# Introduction to stochastic calculus

Lecture notes
Rough preliminary version

Master of mathematics
Université Paris-Dauphine
PSL Université Paris

2017 – 2018
Version compiled September 20, 2018

These are the lecture notes of an introduction course on stochastic calculus, given at Université Paris-Dauphine, for second year master students in mathematics. The initial version is based on a course given by Halim Doss, largely inspired from the book by Ikeda and Watanabe [7]. The prerequisite is a probability theory course based on measure theory and Lebesgue integral, nothing more, and nothing less!

## Notations.

- $\mathbb{R}_+ = [0, \infty)$;
- e is the exponential;
- $\mathbf{1}_A$ is the indicator of $A$;
- d is the differential element;
- $d, i, j, k, m, n, \ell$ are integer numbers;
- $p, q, r, s, t, u, v, \alpha, \beta, \varepsilon$ are real numbers;
- $i$ can be sometimes the complex number $(0, 1)$;
- $X \sim \mu$ $X$ means the random variable $X$ has law $\mu$;
- $\langle x, y\rangle_{\mathrm{H}}$ scalar product in the Hilbert space H;
- $x \cdot y = x_1 y_1 + \cdots + x_d y_d$ for all $x, y \in \mathbb{R}^d$;
- $|x| = \sqrt{x_1^2 + \cdots + x_d^2}$ for all $x \in \mathbb{R}^d$.

# Contents

# Chapter 1

# Preliminaries

## 1.1 Independence and conditioning

**Definition 1.1** (Independence). *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space.*

1. *We say that a family $(\mathcal{A}_i)_{i\in I}$ of sub-$\sigma$-algebras of $\mathcal{A}$ is independent when for all finite $J \subset I$ and for all $A_i \in \mathcal{A}_i$ we have*

$$\mathbb{P}(\cap_{i\in J} A_i) = \prod_{i\in J} \mathbb{P}(A_i).$$

2. *We say that a family $(X_i)_{i\in I}$ of random variables is independent, $X_i : (\Omega, \mathcal{A}) \mapsto (E_i, \mathcal{B}_i)$, when the family of sub-$\sigma$-algebras $(\sigma(X_i))_{i\in I}$ is independent, where $\sigma(X_i) = \{X_i^{-1}(B) : B \in \mathcal{B}_i)$ is the $\sigma$-algebra generated by $X_i$. It follows that $(X_i)_{i\in I}$ is independent if and only if pour all $J \subset I$ finite, we have*

$$\mathbb{P}_{X_i : i\in J} = \otimes_{i\in J} \mathbb{P}_{X_i} \quad \textit{on} \quad (\prod_{i\in J} E_i, \otimes_{i\in J} \mathcal{B}_i).$$

*[If $Z$ is a random variable then we denote its law by $\mathbb{P}_Z$]. It follows that if $X_1, X_2, \ldots, X_n$ are real random variables integrable and independent then*

$$\prod_{i=1}^n X_i \in \mathrm{L}^1 \quad \textit{and} \quad \mathbb{E}\Big(\prod_{i=1}^n X_i\Big) = \prod_{i=1}^n \mathbb{E}(X_i).$$

### 1.1.1 Conditioning

1. We start by recalling some material on Hilbert spaces. Let $H$ be a Hilbert space and $F \subset H$ be a closed sub-space. For all $x \in H$ there exists a unique $y \in F$, called the orthogonal projection of $x$ on $F$, which satisfies one (and thus all) the following equivalent properties:

   (a) for all $z \in F$, $x - y \perp z$ *i.e.* $\langle x, z\rangle = \langle y, z\rangle$

   (b) for all $z \in F$, $\|x - y\| \leq \|x - z\|$ *i.e.* $\|x - y\| = \min_{z\in F} \|x - z\|$.

2. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $\mathcal{F}$ be a sub-$\sigma$-algebra of $\mathcal{A}$. Let us consider the Hilbert space $H = \mathrm{L}^2(\Omega, \mathcal{A}, \mathbb{P})$. The set $F = \mathrm{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ is a closed sub-space of $H$. If $X$ is a square integrable random variable *i.e.* an element of $H$, it is natural to consider the best (least squares) approximation of $X$ by an element of $F$, denoted $Y$. The random variable $Y$ is the orthogonal projection of $X$ on $F$, characterized by the following:

   (a) $Y \in \mathrm{L}^2(\Omega, \mathcal{F}, \mathbb{P})$;

(b) for all $Z \in \mathrm{L}^2(\Omega, \mathcal{F}, \mathbb{P})$, $\mathbb{E}(|X - Y|^2) \leq \mathbb{E}(|X - Z|^2)$.

Using the relation to scalar product, the second property is equivalent to

- for all $Z \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, $\mathbb{E}(XZ) = \mathbb{E}(YZ)$

and also to

- for all $B \in \mathcal{B}$, $\mathbb{E}(X\mathbf{1}_B) = \mathbb{E}(Y\mathbf{1}_B)$.

This gives three characterizations of $Y$. We denote $Y = \mathbb{E}(X \mid \mathcal{F})$ or $Y = \mathbb{E}^{\mathcal{F}}(X)$ and we call it the *conditional expectation of* $Y$ *given* $\mathcal{F}$. It is the best approximation in $\mathrm{L}^2$ (in a sense least squares) of $X$ by an $\mathcal{F}$-measurable square integrable random variable.

3. Suppose now that $X \in \mathrm{L}^1(\Omega, \mathcal{A}, \mathbb{P})$. We can, by extension, define the conditional expectation $Y = \mathbb{E}(X \mid \mathcal{F})$. It is a real random variable characterized by the following properties:

(a) $Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$

(b) for all $Z$ bounded and $\mathcal{F}$ measurable, we have $\mathbb{E}(XZ) = \mathbb{E}(YZ)$. Equivalently, for all $B \in \mathbb{F}$, we have $\mathbb{E}(X\mathbf{1}_B) = \mathbb{E}(Y\mathbf{1}_B)$.

*Proof.* Let $\mu$ be the bounded measure on $(\Omega, \mathcal{F})$ defined by $\mu(B) = \mathbb{E}(X\mathbf{1}_B)$, $B \in \mathcal{F}$. Let us define $\nu = \mathbb{P}_{\mathcal{F}}$. For all $B \in \mathcal{F}$, if $\nu(B) = 0$ then $\mu(B) = 0$. From the Radon–Nicodym theorem, there exists a unique $Y \in L^1(\Omega, \mathcal{F}, \nu)$ such that $\int_B Y \mathrm{d}\nu = \mu(B)$, for all $B \in \mathcal{F}$ *i.e.* $\mathbb{E}(Y\mathbf{1}_B) = \mathbb{E}(X\mathbf{1}_B)$, for all $B \in \mathcal{F}$. ∎

**Remark 1.2** (Variational interpretation). *The expectation and the variance of square integrable random variables have a variational interpretation. Namely if $X \in \mathrm{L}^2$ then its variance* $\mathrm{var}(X)$ *is the square distance in* $\mathrm{L}^2$ *of $X$ to the sub-space of constant random variables:*

$$\mathrm{var}(X) = \inf_{c \in \mathbb{R}} \mathbb{E}((X - c)^2) = \inf_{c \in \mathbb{R}} (\mathbb{E}(X^2) - 2c\mathbb{E}(X) + c^2).$$

*This infimum is a minimum, achieved for $c = \mathbb{E}(X)$, which is therefore the orthogonal projection of $X$ in* $\mathrm{L}^2$ *on the sub-space of constant random variables, and*

$$\begin{aligned}\mathrm{var}(X) &= \mathbb{E}((X - \mathbb{E}(X))^2) \\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X\mathbb{E}(X)) + (\mathbb{E}(X)^2) \\ &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2.\end{aligned}$$

*The following identity is an instance of the Pythagoras theorem in* $\mathrm{L}^2$*:*

$$\begin{aligned}\mathrm{var}(X) &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \\ &= \mathbb{E}(X^2) - \mathbb{E}((\mathbb{E}(X \mid \mathcal{F}))^2) + \mathbb{E}((\mathbb{E}(X \mid \mathcal{F}))^2) - (\mathbb{E}(X))^2 \\ &= \mathbb{E}(\mathrm{var}(X \mid \mathcal{F})) + \mathrm{var}(\mathbb{E}(X \mid \mathcal{F}))\end{aligned}$$

*where*

$$\mathrm{var}(X \mid \mathcal{F}) = \mathbb{E}(X^2 \mid \mathcal{F}) - (\mathbb{E}(X \mid \mathcal{F}))^2.$$

*Note that by definition of* $\mathbb{E}(X \mid \mathcal{F})$*,*

$$\begin{aligned}\inf_{Y:\sigma(Y) \subset \mathcal{F}} \mathbb{E}((X - Y)^2) &= \mathbb{E}((X - \mathbb{E}(X \mid \mathcal{F}))^2) \\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X\mathbb{E}(X \mid \mathcal{F})) + \mathbb{E}((\mathbb{E}(X \mid \mathcal{F}))^2) \\ &= \mathbb{E}(X^2) - \mathbb{E}((\mathbb{E}(X \mid \mathcal{F}))^2) \\ &= \mathbb{E}(\mathrm{var}(X \mid \mathcal{F})).\end{aligned}$$

### 1.1.2 Properties of the conditional expectation

The expectation operator $\mathbb{E}$ is nothing else but the conditional expectation associated to the trivial sub-$\sigma$-field $\{\varnothing, \Omega\}$, hence a particular case. Actually the conditional expectation operator have all the properties of an expectation, and more. Namely, for all sub-$\sigma$-algebra $\mathcal{F}$ of $\mathcal{A}$:

- **Linearity.** for all $\alpha, \beta \in \mathbb{R}$ and $X, Y \in \mathrm{L}^1$, $\mathbb{E}(\alpha X + \beta Y \mid \mathcal{F}) = \alpha\mathbb{E}(X \mid \mathcal{F}) + \beta\mathbb{E}(Y \mid \mathcal{F})$;
- **"Projection".** If $X$ is $\mathcal{F}$-measurable, $Y \in \mathrm{L}^1$, $XY \in \mathrm{L}^1$, then $\mathbb{E}(XY \mid \mathcal{F}) = X\mathbb{E}(Y \mid \mathcal{F})$, in particular $\mathbb{E}(X \mid \mathcal{F}) = X$ if $X \in \mathrm{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ which is the case when $X$ is constant;
- **Composed "projections".** For all sub-$\sigma$-algebras $\mathcal{F}, \mathcal{G}$ with $\mathcal{G} \subset \mathcal{F}$ and all $X \in \mathrm{L}^1$,

  $$\mathbb{E}(\mathbb{E}(X \mid \mathcal{F}) \mid \mathcal{G}) = \mathbb{E}(\mathbb{E}(X \mid \mathcal{G}) \mid \mathcal{F}) = \mathbb{E}(X \mid \mathcal{G}),$$

  and in particular for all $X \in \mathrm{L}^1$

  - $\mathbb{E}(\mathbb{E}(X \mid \mathcal{F})) = \mathbb{E}(X)$;
  - if $X$ is independent of $\mathcal{F}$ then $\mathbb{E}(X \mid \mathcal{F}) = \mathbb{E}(X)$;
  - if $X$ is constant then $\mathbb{E}(X \mid \mathcal{F}) = X$;
- **Normalization.** $\mathbb{E}(\mathbf{1}_\Omega \mid \mathcal{F}) = \mathbf{1}_\Omega$ (follows from projection properties);
- **Positivity or monotonicity.** For all $X, Y \in \mathrm{L}^1$, if $X \leq Y$ then $\mathbb{E}(X \mid \mathcal{F}) \leq \mathbb{E}(Y \mid \mathcal{F})$, or equivalently for all $X \in \mathrm{L}^1$ if $X \geq 0$ then $\mathbb{E}(X \mid \mathcal{F}) \geq 0$. In particular for all $X \in \mathrm{L}^1$,

  $$|\mathbb{E}(X \mid \mathcal{F})| \leq \mathbb{E}(|X| \mid \mathcal{F});$$
- **Convexity.** Jensen inequality: for all non-negative convex $\varphi : \mathbb{R} \to \mathbb{R}$ and all $X \in \mathrm{L}^1$,

  $$\varphi(\mathbb{E}(X \mid \mathcal{F})) \leq \mathbb{E}(\varphi(X) \mid \mathcal{F}).$$

  In particular, for all $p \in [1, \infty)$, $|\mathbb{E}(X \mid \mathcal{F})|^p \leq \mathbb{E}(|X|^p \mid \mathcal{F})$. Moreover for all $X \in \mathrm{L}^p$ and $Y \in \mathrm{L}^q$ with $1 \leq p, q < \infty$, $1/p + 1/q = 1$ ($q = p/(p-1)$), we have the Hölder inequality

  $$|\mathbb{E}(XY \mid \mathcal{F})| \leq (\mathbb{E}(|X|^p \mid \mathcal{F})^{1/p}\mathbb{E}(|Y|^q \mid \mathcal{F})^{1/q}.$$

  The Cauchy–Schwarz inequality corresponds to the special case $p = q = 1/2$;
- **Stability by monotone convergence.** If $X_n \geq 0$, $X_n \nearrow X$, $X \in \mathrm{L}^1$, then $\mathbb{E}(X_n \mid \mathcal{F}) \nearrow \mathbb{E}(X \mid \mathcal{F})$. This allows to define $\mathbb{E}(X \mid \mathcal{F})$ for all non-negative random variable $X$.

**Theorem 1.3** (Transfer)**.** *Let $T : \Omega \to (F, \mathcal{F})$ and $Y : \Omega \to (\mathbb{R}, \mathcal{B}_\mathbb{R})$. Then $Y$ is $\sigma(T)$ measurable if and only if there exists $g : (F, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}_\mathbb{R})$ measurable such that $Y = g \circ T$.*

*Proof.* If $Y = \mathbf{1}_A$ for some $A \in \sigma(T)$. We have $A = T^{-1}(B)$ for some $B \in \mathcal{F}$, and therefore $Y = \mathbf{1}_B \circ T$. If $Y = \sum_{i \in I} a_i \mathbf{1}_{A_i}$ with $I$ finite and $A_i = T^{-1}(B_i)$, $B_i \in \mathcal{F}$, then $Y = (\sum_{i \in I} a_i \mathbf{1}_{B_i}) \circ T$. The property is thus satisfied when $Y$ is a step function. Now, if $Y$ is non-negative and $\sigma(T)$ measurable, then there exists a sequence $(Y_n)_n$ of step functions, $\sigma(T)$ measurable, such that $Y_n \nearrow Y$, and $Y_n = g_n \circ T$. By setting $g = \overline{\lim}\, g_n$, we get $Y = g \circ T$. Finally, if $Y$ is just $\sigma(T)$ measurable, then it suffices to write $Y = Y_+ - Y_-$. ■

Let $X \in \mathrm{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ and let $T : (\Omega, \mathcal{A}) \to (F, \mathcal{F})$ be a random variable. The conditional expectation of $X$ given $T$, denoted $\mathbb{E}(X \mid T)$, is defined by

$$\mathbb{E}(X \mid T) = \mathbb{E}(X \mid \sigma(T)).$$

It is characterized by the following properties:

1. There exists $g : (F, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ with $\mathbb{E}(X \mid T) = g(T)$ and $g(T) \in \mathrm{L}^1$;

2. For all $h : (F, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ measurable and bounded,

$$\mathbb{E}(Xh(T)) = \mathbb{E}(g(T)h(T)).$$

If $X \in \mathrm{L}^2$ then, thanks to the transfer theorem, the conditional expectation $\mathbb{E}(X \mid T)$ is the best approximation in $\mathrm{L}^2$ (least squares!) of $X$ by a measurable function of $T$.

**Exercise 1.4.**

1. *Give a meaning to* $\mathbb{P}(A \mid \mathcal{F})$*;*

2. *Compute* $\mathbb{E}(X \mid T)$ *when* $X \in \mathrm{L}^1$ *and* $T$ *takes its values in an at most countable set* $F$*;*

3. *Compute* $\mathbb{E}(h(X) \mid T)$ *when* $(X, Y)$ *takes its values in* $\mathbb{R}^n \times \mathbb{R}^m$*, with density* $(x, t) \mapsto f(x, t)$*, and* $h : \mathbb{R}^n \to \mathbb{R}$ *is bounded and measurable;*

4. *if* $X$ *and* $Y$ *are random variables with* $X$ *independent of* $\mathcal{F}$ *and* $Y$ $\mathcal{F}$ *measurable and if* $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *is measurable and bounded or positive then*

$$\mathbb{E}(f(X, Y) \mid \mathcal{F}) = g(Y) \quad \textit{where} \quad g(y) = \mathbb{E}(f(X, y) \mid \mathcal{F}).$$

*This suggests to interpret the conditional expectation as an averaging of residual randomness, and not only as the best approximation in the sense of least squares.*

### 1.1.3 Conditional law

Let $(F, \mathcal{F})$ and $(G, \mathcal{G})$ be two measurable spaces. A transition kernel from $G$ to $F$ is a family of probability measures $(N(t, \cdot))_{t \in G}$ on $(F, \mathcal{F})$ such that for all $A \in \mathcal{F}$, the map $t \in G \mapsto N(t, A) \in [0, 1]$ is measurable for $(G, \mathcal{G})$.

Let $X$ and $T$ be two random variables taking values in $(F, \mathcal{F})$ and $(G, \mathcal{G})$ respectively. The conditional law of $X$ given $T$ is a transition kernel $(N(t, \cdot))_{t \in G}$ from $G$ to $F$ such that for all $h : F \to \mathbb{R}$ $\mathcal{F}$ measurable we have

$$\mathbb{E}(h(X) \mid T) = \int_F h(x) N(T, \cdot).$$

In particular $\mathbb{P}(X \in A \mid T) = N(T, A)$ for all $A \in \mathbb{F}$. We say then that $N(t, \cdot)$ is the condition law of $X$ given $T = t$.

**Remark 1.5.** *If* $X$ *and* $T$ *are independent then* $N(t, \cdot)$ *does not depend on* $t$*, namely for all* $t \in G$*,* $N(t, \cdot) = \mathbb{P}_X$ *where* $\mathbb{P}_X$ *is the law of* $X$*.*

**Exercise 1.6.** *Compute the conditional law of* $X$ *given* $T = t$ *for each of the previous examples.*

## 1.2 Generalities on processes

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathcal{F}_t)_{t \geq 0}$ be an increasing family of sub-$\sigma$-algebras of $\mathcal{F}$. We say that $(\mathcal{F}_t)_{t \geq 0}$ is right continuous if, for all $t \in \mathbb{R}_+$,

$$\mathcal{F}_t = \mathcal{F}_{t+} \quad \text{where} \quad \mathcal{F}_{t+} = \cap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon}.$$

In the sequel, we suppose in general that $(\mathrm{F}_t)_{t \geq 0}$ is right continuous and complete (*i.e.* $\mathcal{F}_0$ contains all the negligible sets of $\mathcal{F}$). The family $(\mathcal{F}_t)_{t \geq 0}$ is called filtration or reference family.

Let $(G, \mathcal{B})$ be a measurable space where $G$ is a metric space, complete and separable, and where $\mathcal{B}$ is its Borel $\sigma$-algebra.

A $G$-valued stochastic process or process is a family of random variables $X = (X_t)_{t \geq 0}$ indexed by a parameter interpreted as a time and taking values in some space $(G, \mathcal{B})$.

We say that a process $X$ is...

- $(\mathcal{F}_t)_{t \geq 0}$-adapted when $X_t$ is $\mathcal{F}_t$ measurable for all $t \geq 0$;
- measurable if the map $(t, \omega) \in \mathbb{R}_+ \times \Omega \mapsto X_t(\omega) \in G$ is $(\mathcal{B}_{\mathbb{R}_+} \otimes \mathcal{F})$-measurable.
- right continuous (respectively left continuous, respectively continuous) if for almost all $\omega \in \Omega$, the sample path $t \in \mathbb{R}_+ \mapsto X_t(\omega) \in G$ is right continuous (respectively left continuous, respectively continuous);
- $(\mathcal{F}_t)_{t \geq 0}$-predictable if the map $(t, \omega) \in \mathbb{R}_+ \times \Omega \mapsto X_t(\omega) \in G$ is measurable for the $\sigma$-algebra $\mathcal{P}$ on $\mathbb{R}_+ \times \Omega$ generated by all $(\mathcal{F}_t)_{t \geq 0}$-adapted left continuous processes;
- $(\mathcal{F}_t)_{t \geq 0}$-optional if the map $(t, \omega) \in \mathbb{R}_+ \times \Omega \mapsto X_t(\omega) \in G$ is measurable for the $\sigma$-algebra $\mathcal{O}$ on $\mathbb{R}_+ \times \Omega$ generated by all $(\mathcal{F}_t)_{t \geq 0}$-adapted right continuous processes.

**Theorem 1.7.**

1. *Every predictable process is optional;*
2. *If $X = (X_t)_{t \geq 0}$ is optional then for all $T > 0$ the map $(t, \omega) \in [0, T] \times \Omega \mapsto X_t(\omega) \in G$ is measurable for $\mathcal{B}_{[0,T]} \times \mathcal{F}_T$, and we say then that $X$ is progressively mesurable;*
3. *Every optional process is mesurable and adapted.*

*Proof.*

1. Let $X = (X_t)_{t \geq 0}$ be a left continuous adapted process. For all $n \geq 1$, we set $X_t^{(n)} = X_{k/2^k}$ if $t \in [k/2^n, (k+1)/2^n)$, $k \geq 0$. Then $X^{(n)}$ is right continuous and adapted, and $\lim_{n \to \infty} X^{(n)} = X$, therefore $X$ is optional.
2. Let $X = (X_t)_{t \geq 0}$ be a right continuous adapted process. For all $n \geq 1$, we set, for a fixed $T > 0$: $X_t^{(n)} = X_{((k+1)/2^n) \wedge T}$ if $t \in [k/2^n, (k+1)/2^n) \cap [0, T]$, $k \geq 0$. Then $X^{(n)}$ satisfies the desired property.
3. The third item is immediate.

■

Two processes $X = (X_t)_{t \geq 0}$ and $Y = (Y_t)_{t \geq 0}$ taking values in $G$ are indistinguishable if for almost all $\omega \in \Omega$ the sample paths $t \mapsto X_t(\omega)$ and $t \mapsto Y_t(\omega)$ coincide. We say that $Y$ is a modification of $X$ if for all $t \geq 0$ the set $\{\omega \in \Omega : X_t(\omega) \neq Y_t(\omega)\}$ is negligible.

### 1.2.1 Stopping time

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(F_t)_{t \geq 0}$ be a filtration. A map $T : \Omega \to [0, \infty]$ is an $(F_t)_{t \geq 0}$ stopping time or optional time when for all $t \geq 0$ we have $\{T \leq t\} \in \mathcal{F}_t$.

For a stopping time $T$, we define the stopping $\sigma$-algebra by

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

**Example 1.8.** *Let $X = (X_t)_{t\geq 0}$ be a continuous adapted process taking its values in $\mathbb{R}$. For all $E \subset \mathbb{R}$, let us define*

$$T_E = \inf\{t \geq 0 : X_t \in E\}$$

*with the convention $\inf \emptyset = \infty$. This is called the hitting time of $E$. Then if $E$ is closed then $T_E$ is a stopping time. Indeed, for all $t \geq 0$, we have*

$$\{T_E > t\} = \bigcup_{\varepsilon\in\mathbb{Q}\cap(0,\infty)} \bigcap_{r\in\mathbb{Q}\cap[0,t)} \{X_r \notin V_\varepsilon(E)\} \quad \textit{where} \quad V_\varepsilon(E) = \{y \in \mathbb{R} : d(y, E) < \varepsilon\}.$$

*More generally, if $(F_t)_{t\geq 0}$ is right continuous and complete, then it is possible to show that for all optional process $X = (X_t)_{t\geq 0}$ et for all Borel set $E \in \mathcal{B}_\mathbb{R}$, the hitting time $T_E$ is a stopping time. See for instance [5, 4, 3].*

**Theorem 1.9** (Stopping times games)**.** *Let $S$, $T$, and $S_n$, $n \geq 0$ be stopping times. Then:*

1. *$S \wedge T$ and $S \vee T$ are stopping times and $\mathcal{F}_{S\wedge T} \subset \mathcal{F}_{S\vee T}$;*
2. *if $(F_t)_{t\geq 0}$ is right continuous then $\underline{\lim}_n S_n$ and $\overline{\lim}_n S_n$ are stopping times and*
$$\cap_n \mathcal{F}_{S_n} = \mathcal{F}_{\inf_n S_n};$$
3. *$S$ is $\mathcal{F}_S$ mesurable;*
4. *if $X = (X_t)_{t\geq 0}$ is optional then the random variable $Z = X_S \mathbf{1}_{S<\infty}$ is $\mathcal{F}_S$-mesurable.*

*Proof.*

1. For all $t \geq 0$ we have
$$\{S \wedge T > t\} = \{S > t\} \cap \{T > t\} \in \mathcal{F}_t \quad \text{and} \quad \{S \vee T \leq t\} = \{S \leq t\} \cap \{T \leq t\} \in \mathcal{F}_t.$$
If $S \leq T$ and $A \in \mathcal{F}_S$ then, for all $t \geq 0$, we have $A \cap \{T \leq t\} = A \cap \{S \leq t\} \cap \{T \leq t\} \in \mathcal{F}_t$, and thus $A \in \mathcal{F}_T$.
2. It suffice to show that $\sup_n S_n$ and $\inf_n S_n$ are stopping times. But
$$\{\sup_n S_n \leq t\} = \cap_n \{S_n \leq t\} \in \mathcal{F}_t \quad \text{and} \quad \{\inf_n S_n < t\} = \cup_n \{S_n < t\} \in \mathcal{F}_t$$
and therefore
$$\{\inf_n S_n \leq t\} = \cap_{\varepsilon>0} \{\inf_n S_n < t + \varepsilon\} \in \mathcal{F}_{t+} = \mathcal{F}_t.$$
Let $A \in \cap_n \mathcal{F}_{S_n}$. Then
$$A \cap \{\inf_n S_n < t\} = \cup_n A \cap \{S_n < t\} \in \mathcal{F}_t.$$
Therefore
$$A \cap \{\inf_n S_n \leq t\} \in \mathcal{F}_{t+} = \mathcal{F}_t.$$
3. Obvious;
4. Let $B \in \mathcal{B}_\mathbb{R}$ and $t \geq 0$. Then we have:
$$\{Z \in B\} \cap \{S \leq t\} = \{X_{S\wedge t} \in B\} \cap \{S \leq t\}.$$
Now we consider the composition of measurable maps:
$$\omega \in (\Omega, \mathcal{F}_t) \mapsto (\sigma(\omega) \wedge t, \omega) \in ([0, t] \times \Omega, \mathcal{B}_{[0,1]} \otimes \mathcal{F}_t) \mapsto X_{\sigma(\omega)\wedge t}(\omega) \in (\mathbb{R}, \mathcal{B}_\mathbb{R})$$
and we use the fact that $X$ is progressively measurable.

■

### 1.2.2 Martingales

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space. A stochastic real process $X = (X_t)_{t\geq 0}$ is an $(F_t)_{t\geq 0}$-martingale (respectively super-martingale, respectively sub-martingale) when:

1. for all $t \geq 0$, $X_t$ is integrable;
2. for all $t \geq 0$, $X_t$ is $\mathcal{F}_t$ measurable;
3. for all $0 \leq s \leq t$, $\mathbb{E}(X_t \mid \mathcal{F}_t) = X_s$ (respectively $\leq$, respectively $\geq$).

**Example 1.10.**

1. *Let $Y \in \mathrm{L}^1$. If we set $X_t = \mathbb{E}(Y \mid \mathcal{F}_t)$ then we can check immediately that $(X_t)_{t\geq 0}$ is an $(\mathcal{F}_t)_{t\geq 0}$ martingale. Conversely, we can prove that if $(X_t)_{t\geq 0}$ is a right continuous $(\mathcal{F}_t)_{t\geq 0}$ martingale such that $\sup_{t\geq 0} \mathbb{E}(|X_t|^p) < \infty$ for some $p \in (1, \infty)$ then there exists $Y \in \mathrm{L}^p$ such that $\lim_{t\to\infty} X_t = Y$ almost surely and in $\mathrm{L}^p$ and $X_t = \mathbb{E}(Y \mid \mathcal{F}_t)$ for all $t \geq 0$;*
2. *Let $(X_t)_{t\geq 0}$ be a martingale. Then for all convex $f : \mathbb{R} \to \mathbb{R}$, such that $f(X_t) \in \mathrm{L}^1$ for all $t \geq 0$, $(Y_t)_{t\geq 0} = (f(X_t)_{t\geq 0}$ is a sub-martingale. In particular $(|X_t|)_{t\geq 0}$ is a sub-martingale;*
3. *Let $X = (X_t)_{t\geq 0}$ be an $(F_t)_{t\geq 0}$-martingale. For all $t \geq 0$, let us define the $\sigma$-algebra $\mathcal{G}_t = \sigma(X_s : s \leq t) \subset \mathcal{F}_t$. The family $\mathcal{G} = (\mathcal{G}_t)_{t\geq 0}$ is called the natural filtration associated to $X$. We can check immediately that $X$ is also a $G$ martingale.*

From now on, most of the martingales that we encounter in this course are continuous. We admit the following theorems proved in most textbooks on stochastic processes.

**Theorem 1.11** (Doob[1] maximal inequality)**.** *Let $M = (M_t)_{t\geq 0}$ be a continuous $(\mathcal{F}_t)_{t\geq 0}$-martingale.*

1. *For all $t > 0$, $\lambda > 0$, and $p \in [1, \infty)$,*

$$\mathbb{P}\Big(\sup_{t\in[0,t]} |M_t| \geq \lambda\Big) \leq \frac{\mathbb{E}(|M_t|^p)}{\lambda^p}.$$

2. *For all $t > 0$ and $p \in (1, \infty)$,*

$$\mathbb{E}\Big(\sup_{t\in[0,T]} |M_t|^p\Big) \leq \Big(\frac{p}{p-1}\Big)^p \mathbb{E}(|M_t|^p)$$

*in other words, denoting $M_t^* = \sup_{0\leq s\leq t} |M_t|$,*

$$\|M_t^*\|_p \leq \frac{p}{p-1} \|M_t\|_p.$$

Note that $q = p/(p-1)$ is the conjugate of $p$ in the sense that $1/p + 1/q = 1$.

**Theorem 1.12** (Doob optional stopping)**.** *Let $X = (X_t)_{t\geq 0}$ be a continuous $(\mathcal{F}_t)_{t\geq 0}$-martingale and let $S$ and $T$ be stopping times such that $0 \leq S \leq T \leq C$ where $C$ is a finite constant. Then*

$$\mathbb{E}(X_T \mid \mathcal{F}_S) = X_S.$$

*In particular if $T$ is a bounded stopping time then*

$$\mathbb{E}(X_T) = \mathbb{E}(X_0).$$

[1] Named after Joseph L. Doob (1910 – 2004), American mathematician.

### 1.2.3 Bounded variation, Doob–Meyer decomposition, quadratic variation

Let $I \subset \mathbb{R}$ be a finite interval and let $(S, d)$ be a metric space.

**Definition 1.13** ($p$-variation). *For all $p \in [1, \infty)$, the p-variation of $f : I \to S$ is*

$$\|f\|_{p\text{-}var} = \Big( \sup_{t_k} \sum_k d(f(t_k), f(t_{k+1}))^p \Big)^{1/p}$$

*where the supremum runs over the finite partitions or sub-divisions of the interval $I$ namely the finite sequences of the form $\min(I) = t_0 < t_1 < \cdots < t_{n+1} = \max(I)$, $n \geq 0$.*

- the $p$-variation decreases with $p$;
- $\|f\|_{\text{1-var}}$ is called sometimes the total variation of $f$;
- $f : I \to S$ has finite variation or is of bounded variation when it has finite 1-variation;
- if $f$ is of bounded variation then $f$ is bounded.
- if $f$ if of bounded variation and is differentiable with Riemann integrable derivative then

$$\|f\|_{\text{1-var}} = \int_a^b |f'(x)| \mathrm{d}x.$$

- if $f$ is continuously differentiable then $f$ has bounded variation and the latter holds true.

**Theorem 1.14** (Bounded variation). *For all $f : I \mapsto S$, the following properties are equivalent:*

1. *$f$ is of bounded variation;*
2. *$f$ is the difference of two positive bounded increasing functions.*

*In this case, such a decomposition is not unique, the number of discontinuity points of $f$ is at most countable, and $f$ is differentiable almost everywhere, and is integrable.*

*Proof.* FIXME: Take a look at a textbook in analysis. ■

We say that a process $X = (X_t)_{t \geq 0}$ is...

- issued from the origin when $X_0 = 0$;
- square integrable when $\mathbb{E}(X_t^2) < \infty$ for all $t \geq 0$;
- of finite variation when almost surely $t \mapsto X_t$ is of bounded variation on all finite intervals of $\mathbb{R}_+$. Equivalently $X$ is the difference of two positive increasing processes.

**Theorem 1.15** (Doob–Meyer[2] decomposition).

1. *For all square integrable continuous $(\mathcal{F}_t)_{t \geq 0}$-martingale $M = (M_t)_{t \geq 0}$, there exists a unique continuous and increasing process $\langle M \rangle = (\langle M_t \rangle)_{t \geq 0}$ issued from the origin such that $(M_t^2 - \langle M_t \rangle)_{t \geq 0}$ is an $(\mathcal{F}_t)_{t \geq 0}$-martingale;*

[2] Named after Paul-Anré Meyer (1934 – 2003), French mathematician.

2. *For all square integrable continuous $(\mathcal{F}_t)_{t\geq 0}$-martingales $(M_t)_{t\geq 0}$ and $N = (N_t)_{t\geq 0}$, there exists a unique continuous with finite variation process denoted*
$$\langle M, N\rangle = (\langle M, N\rangle)_{t\geq 0}$$
*such that the following process is an $(\mathcal{F})_{t\geq 0}$-martingale:*
$$(M_t N_t - \langle M_t, N_t\rangle)_{t\geq 0}.$$

Theorem 2.11 states that if $B = (B^1, \dots, B^d)$ is a $d$-dimensional Brownian motion then $\langle B^j, B^k\rangle_t = t\mathbf{1}_{j=k}$ for all $t \geq 0$ and $1 \leq j, k \leq d$.

*Proof.* FIXME: ■

**Definition 1.16** (Quadratic variation). *The quadratic variation $[X] = ([X]_t)_{t\geq 0}$ of a real process $X = (X_t)_{t\geq 0}$ is defined for all $t \geq 0$ by the following limit when it exists:*
$$[X]_t = \lim_{|\delta|\to}^{\mathbb{P}} \sum_k |X_{t_{k+1}} - X_{t_k}|^2$$
*where the convergence takes place in probability, where $\delta : 0 = t_0 < t_1 < \cdots < t_n = t$ runs over all the partitions or sub-divisions of $[0, t]$ and where $|\delta| = \max_k |t_{k+1} - t_k|$ is the mesh of $\delta$. Similarly the quadratic variation of a couple $X = (X_t)_{t\geq 0}$ and $Y = (Y_t)_{t\geq 0}$ of real processes is defined for all $t \geq 0$ by the following limit when it exists:*
$$[X, Y]_t = \lim_{|\delta|\to}^{\mathbb{P}} \sum_k |X_{t_{k+1}} - X_{t_k}||Y_{t_{k+1}} - Y_{t_k}|.$$

We have $[X] = [X, X]$. The set of processes with quadratic variation is a vector space. The operator $[\cdot]$ is bilinear on this space and we have by polarization $4[X, Y] = [X + Y] - [X - Y]$.

If a process $X = (X_t)_{t\geq 0}$ is continuous and has finite variation then it has zero quadratic variation. Indeed, we have, for all $t > 0$ and all partition $\delta : 0 = t_0 < \cdots < t_n = t$ of $[0, t]$, $n \geq 1$,
$$\sum_k |X_{t_{k+1}} - X_{t_k}|^2 \leq \max_k |X_{t_{k+1}} - X_{t_k}| \sum_k |X_{t_{k+1}} - X_{t_k}| \underset{|\delta|\to 0}{\longrightarrow} 0.$$
The max part of the right hand side tends to zero since $X$ is continuous, while the $\sum$ part of the right hand side tends to the 1-variation of $X$ on $[0, t]$ which is finite since $X$ has finite variation.

Theorem 2.13 states that a $d$-dimensional Brownian motion $B$ issued from the origin is a continuous process with infinite variation on every interval and with quadratic variation given by $[B]_t = t$ for all $t \geq 0$. We also have $\langle B\rangle_t = t$ for all $t \geq 0$. More generally Theorem 4.1 states that for all continuous local martingale $M$ issued from the origin we have $[M] = \langle M\rangle$.

## 1.3 Monotone class theorem

The monotone class theorem is a sort of Stone–Weierstrass theorem of measure theory.

**Definition 1.17** ($\pi$-systems and $\lambda$-systems).

- *We say that $\mathcal{C} \subset \mathcal{P}(\Omega)$ is a $\pi$-system when $A \cap B \in \mathcal{C}$ for all $A, B \in \mathcal{C}$;*
- *We say that $\mathcal{S} \subset \mathcal{P}(\Omega)$ is a $\lambda$-system (or a monotone class or a Dynkin[3] system) when*

[3] Named after Eugene Dynkin (1924 – 2014), Soviet and American mathematician.

– $\cup_n A_n \in \mathcal{C}$ *for all* $(A_n)_n$ *such that* $A_n \subset A_{n+1}$ *and* $A_n \in \mathcal{C}$ *for all* $n$*;*
– $A \setminus B \in \mathcal{C}$ *for all* $A, B \in \mathcal{C}$ *such that* $B \subset A$.

Basic examples of $\pi$-systems are given by the class of singletons $\{\{x\} : x \in \mathbb{R}\} \cup \{\varnothing\}$, the class of product subsets $\{A \times B : A, B \in \mathcal{P}(\Omega)\}$, and the class of intervals $\{(-\infty, x] : x \in \mathbb{R}\}$.

A basic yet important example of $\lambda$-system is given by $\{A \in \mathcal{A} : \mathbb{P}(A) = \mathbb{Q}(A)\}$ where $\mathbb{P}$ and $\mathbb{Q}$ are probability measures on $(\Omega, \mathcal{A})$, see Corollary 1.20 for an application.

**Lemma 1.18** ($\sigma$-algebras)**.** *A* $\lambda$*-system that contains* $\Omega$ *and which is a* $\pi$*-system is a* $\sigma$*-algebra.*

Note that conversely, a $\sigma$-algebra is always a $\pi$-system, but not a $\lambda$-system in general.

*Proof.* If a $\lambda$-system $\mathcal{S} \subset \mathbb{P}(\Omega)$ contains $\Omega$ and is a $\pi$-system then for all $A, B \in \mathcal{S}$ we have

$$A \cup B = \Omega \setminus ((\Omega \setminus A) \cap (\Omega \setminus B)),$$

which means that $\mathcal{S}$ is table by finite union. This allows to drop the non-decreasing condition in the stability of $\mathcal{S}$ by countable union, which simply means finally that $\mathcal{S}$ is a $\sigma$-algebra. ∎

**Theorem 1.19** (Dynkin $\pi$-$\lambda$ Theorem)**.** *If* $\mathcal{S} \subset \mathcal{P}(\Omega)$ *is a* $\lambda$*-system containing* $\Omega$ *and including a* $\pi$*-system* $\mathcal{C}$*, then* $\mathcal{S}$ *contains also the* $\sigma$*-algebra* $\sigma(\mathcal{C})$ *generated by* $\mathcal{C}$*.*

*Proof.* The $\lambda$-system generated by a subset of $\mathcal{P}(\Omega)$ is by definition the intersection of all $\lambda$-systems which include this subset. This makes sense, indeed this intersection is not empty since it contains $\mathcal{P}(\Omega)$, and it not difficult to check that it is a $\lambda$-system. It is the smallest (for the inclusion) $\lambda$-system containing the initial subset of $\mathcal{P}(\Omega)$.

Let $\mathcal{S}'$ be the $\lambda$-system generated by $\mathcal{C}$ and $\Omega$. It suffices to show that $\mathcal{S}'$ is a $\sigma$-algebra. For that, and thanks to lemma 1.18, it suffices to show that $\mathcal{S}'$ is a $\pi$-system. To do so, let us define

$$\mathcal{S}_1 = \{A \in \mathcal{S}' : A \cap B \in \mathcal{S}' \text{ for all } B \in \mathcal{C}\},$$

which is a $\lambda$-system including $\Omega$ and containing $\mathcal{C}$, hence $\mathcal{S}_1 \subset \mathcal{S}'$, and thus $\mathcal{S}_1 = \mathcal{S}'$. Now,

$$\mathcal{S}_2 = \{A \in \mathcal{S}' : A \cap B \in \mathcal{S}' \text{ for all } B \in \mathcal{S}'\}$$

is a $\lambda$-system containing $\Omega$ and including $\mathcal{S}$ and thus $\mathcal{S}_2 = \mathcal{S}'$, hence $\mathcal{S}'$ is a $\pi$-system. ∎

**Corollary 1.20** (Sierpiński[4]–Dynkin monotone class theorem)**.**

1. *For all probability measures* $\mathbb{P}$ *and* $\mathbb{Q}$ *on a measurable space* $(\Omega, \mathcal{A})$*, if* $\mathbb{P}(A) = \mathbb{Q}(A)$ *for all* $A \in \mathcal{C}$ *where* $\mathcal{C}$ *is a* $\pi$*-system such that* $\sigma(\mathcal{C}) = \mathcal{A}$*, then* $\mathbb{P} = \mathbb{Q}$*;*

2. *Let* $H$ *be a vector space of bounded functions* $\Omega \to \mathbb{R}$ *such that*

   (a) $H$ *is stable by monotone convergence: if* $f_n \in H \nearrow f$ *with* $f$ *bounded then* $f \in H$*;*

   (b) $H$ *contains constant functions and is stable by product:* $\mathbf{1}_\Omega \in H$ *and* $\mathbf{1}_A \in H$ *for all* $A \in \mathcal{C}$ *for a* $\pi$*-system* $\mathcal{C}$ *such that* $\sigma(\mathcal{C}) = \mathcal{A}$*;*

   *then* $H$ *contains all* $\mathcal{A}$*-measurable bounded functions* $\Omega \to \mathbb{R}$*.*

Note that $H$ is an algebra in the sense that it is a vector space stable by product.

*Proof.*

1. Take $\mathcal{S} = \{A \in \mathcal{A} : \mathbb{P}(A) = \mathbb{Q}(A)\}$ and use Theorem 1.19.

[4] Named after Wacław Sierpiński (1882 – 1969), Polish mathematician.

2. Take $\mathcal{S} = \{A \in \mathcal{A} : \mathbf{1}_A \in H\}$ and use Theorem 1.19.

■

**Corollary 1.21** (Monotone class theorem for stochastic processes)**.** *Let $\mathcal{L}$ be a vector space of real processes such that the following properties hold true:*

1. *$\mathcal{L}$ contains real left-continuous (respectively right-continuous) $(\mathcal{F}_t)_{t\geq 0}$-adapted processes;*
2. *if $\varphi_n \nearrow \varphi$ with $\varphi_n \in \mathcal{L}$ for all $n$ and with $\varphi$ bounded then $\varphi \in \mathcal{L}$;*

*then $\mathcal{L}$ contains all bounded $(\mathcal{F}_t)_{t\geq 0}$-predictable (respectively optional) processes.*

*Proof.* Exercise! ■

# Chapter 2

# Brownian motion

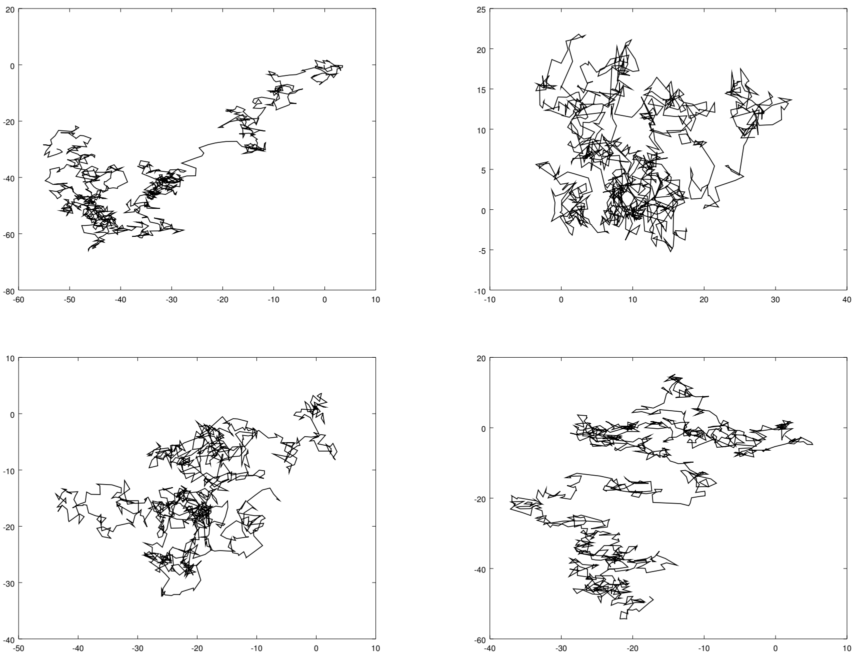

Figure 2.1: First steps of four sample paths of 2-dimensional Brownian motion issued from the origin, numerically simulated with a programming code like `plot(cumsum(randn(2,1000)))`.

For all $t > 0$, $d \geq 1$, the density of the Gaussian distribution $\mathcal{N}(0, tI_d)$ on $\mathbb{R}^d$ is

$$x \in \mathbb{R}^d \mapsto p_t(x) = \frac{\mathrm{e}^{-\frac{|x|^2}{2t}}}{(\sqrt{2\pi t})^d} \quad \text{where} \quad |x|^2 = x_1^2 + \cdots + x_d^2.$$

We have, for all $s, t > 0$,

$$p_{t+s}(x) = (p_t * p_s)(x) = \int_{\mathbb{R}^d} p_t(x - z) p_s(z) \mathrm{d}z.$$

**Definition 2.1** (Brownian motion)**.** *A $d$-dimensional Brownian motion is a continuous $d$-dimensional process $B = (B_t)_{t \geq 0}$ such that:*

1. *For all $0 \leq s \leq t$, the random variable $B_t - B_s$ follows the Gaussian law $\mathcal{N}(0, (t-s)I_d)$;*
2. *$(B_t)_{t \geq 0}$ has independent increments* i.e. *for all $t_0 = 0 < t_1 < \cdots < t_n$, $n \geq 0$, the random variables $B_{t_1} - B_{t_0}, \ldots, B_{t_n} - B_{t_{n-1}}$ are independent.*

**Remark 2.2** (Gaussian processes and Lévy[1] processes)**.** *For all $n \geq 1$ and all $0 \leq t_1 < \cdots < t_n$ the random vector $(B_{t_1}, \ldots, B_{t_n})$ is Gaussian, and we say then that Brownian motion is a Gaussian process. On the other hand, for all $n \geq 1$ and $t_0 = 0 < t_1 < \cdots < t_n$ the increments $B_{t_1} - B_{t_0}, \ldots, B_{t_n} - B_{t_{n-1}}$ are independent and stationary in the sense that their law depends only on the differences $t_1 - t_0, \ldots, t_n - t_{n-1}$ between successive times. Also Brownian motion has independent and stationary increments and such processes are called Lévy processes.*

**Remark 2.3** (Basic properties)**.** *The first two items below are easy consequences of the definition and show that the study of a d-dimensional Brownian motion reduces to the study of the one-dimensional Brownian motion issued from the origin.*

1. *Let $B = (B_t)_{t \geq 0}$ be a Brownian motion issued from the origin* i.e. *$B_0 = 0$, and let $H$ be a random variable independent of $B$. Then according to the definition above the process $(H + B_t)_{t \geq 0}$ is also a Brownian motion;*

2. *Let $X = (X_t)_{t \geq 0}$ be a d-dimensional process and let $X_t = (X_t^1, \ldots, X_t^d)$ be the coordinates of $X_t$ in $\mathbb{R}^d$. Then $X$ is a Brownian motion issued from the origin if and only if the following two properties hold true:*

   *(a) for all $1 \leq i \leq d$, the process $(X_t^i)_{t \geq 0}$ is a Brownian motion issued from the origin;*

   *(b) the processes $(X_t^1)_{t \geq 0}, \ldots, (X_t^d)_{t \geq 0}$ are independent.*

3. *Let $X = (X_t)_{t \geq 0}$ be a continuous process on $\mathbb{R}^d$, issued from $x \in \mathbb{R}^d$, then $X$ is a Brownian motion if and only if for all $n \geq 0$, $0 < t_1 \cdots < t_n$, all $A_i \in \mathcal{B}_{\mathbb{R}^d}$, $1 \leq i \leq n$, we have*

$$\begin{aligned}&\mathbb{P}(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n)\\ &\qquad = \int_{A_1 \times \cdots \times A_n} p_{t_1}(x_1 - x) p_{t_2 - t_1}(x_2 - x_1) \cdots p_{t_n - t_{n-1}}(x_n - x_{n-1}) \mathrm{d}x_1 \cdots \mathrm{d}x_n.\end{aligned}$$

**Theorem 2.4** (Characterization of BM by Gaussianity and covariance)**.** *Let $X = (X_t)_{t \geq 0}$ be a continuous real process, issued from the origin. Then $X$ is a Brownian motion if and only if $X$ is a Gaussian process, centered, with covariance given by $\mathbb{E}(X_t X_s) = s \wedge t$ for all $s, t \geq 0$.*

*Proof.*

1. Suppose that $X = (X_t)_{t \geq 0}$ is a Brownian motion issued from the origin, then for all $0 < t_1 < \cdots <_n$ the random variables $X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ are Gaussian, centered, and independent, and $X_0 = 0$, and $(X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}})$ and $(X_{t_1}, \ldots, X_{t_n})$ are (centered) Gaussian random vectors in the sense that all linear combinations of their coordinates are Gaussian. Moreover, for all $0 \leq s \leq t$, we have

$$\mathbb{E}(X_s X_t) = \mathbb{E}(X_s(X_t - X_s)) + \mathbb{E}(X_s^2) = 0 + s = s.$$

2. Conversely, suppose that $X = (X_t)_{t \geq 0}$ is a Gaussian process, centered, such that $\mathbb{E}(X_s X_t) = s \wedge t$, for all $s, t \geq 0$. It is easy to deduce that for all $0 < t_1 < \cdots < t_n$, the random vector $(X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}})$ is Gaussian, centered, with diagonal covariance $\mathrm{diag}(t_1, t_2 - t_1, \ldots, t_n - t_{n-1})$ which implies that $(X_t)_{t \geq 0}$ is a Brownian motion.

■

**Corollary 2.5.** *If $X = (X_t)_{t \geq 0}$ is a Brownian motion on $\mathbb{R}$, issued form the origin, then for all $c \in \mathbb{R} \setminus \{0\}$, the process $X^{(c)} = (\frac{1}{c} X_{c^2 t})_{t \geq 0}$ is a Brownian motion.*

[1] Named after Paul Lévy (1886 – 1971), French mathematician.

**Theorem 2.6.** *If $X = (X_t)_{t\geq 0}$ is a Brownian motion on $\mathbb{R}$ issued from the origin then the process $Y = (tX_{1/t})_{t\geq 0}$, with $Y_0 = 0$, is a Brownian motion.*

*Proof.* The process $Y$ is Gaussian, centered, with $\mathbb{E}(Y_s Y_t) = s \wedge t$ for all $s, t \geq 0$. It remains to prove that $Y$ is continuous. By definition $Y$ is continuous on $(0, \infty)$. To prove the continuity at $t = 0$, we note that the processes $(X_t)_{t>0}$ and $(Y_t)_{t>0}$ are continuous with same law, and $\overline{\lim}_{t\to 0, t>0} |X_t| = 0$ almost surely, hence $\overline{\lim}_{t\to 0, t>0} |Y_t| = 0$ almost surely. ∎

**Remark 2.7** (Strong law of large numbers)**.** *The continuity of $Y$ at $t = 0$ gives*

$$\lim_{t\to\infty} \frac{X_t}{t} = 0 \quad \textit{almost surely}$$

*which is a sort of strong law of large numbers for the increments.*

Note: a construction of Brownian motion is considered later on.

**Theorem 2.8** (Fourier and Laplace martingales characterizations of Brownian motion)**.** *Let $X = (X_t)_{t\geq 0}$ be a $d$-dimensional continuous process issued from the origin. Let us define the $\sigma$-algebra $\mathcal{G}_t = \sigma(X_s : s \leq t)$ for all $t \geq 0$. The following properties are equivalent:*

1. *$X$ is a Brownian motion;*
2. *For all $\lambda \in \mathbb{R}^d$, $(M_t^\lambda)_{t\geq 0} = (\mathrm{e}^{i\lambda\cdot X_t + \frac{|\lambda|^2 t}{2}})_{t\geq 0}$ is a $(\mathcal{G}_t)_{t\geq 0}$-martingale;*
3. *For all $\lambda \in \mathbb{R}^d$, $(N_t)_{t\geq 0} = (\mathrm{e}^{\lambda\cdot X_t - \frac{|\lambda|^2 t}{2}})_{t\geq 0}$ is a $(\mathcal{G}_t)_{t\geq 0}$-martingale.*

*Proof.* The process $X$ is a Brownian motion if and only if for all $0 \leq s < t$, $X_t - X_s$ is independent of $\mathcal{G}_s$ and $X_t - X_s \sim \mathcal{N}(0, (t-s)I_d)$, in other words if and only if for all $0 \leq s < t$ and $\lambda \in \mathbb{R}^d$,

$$\mathbb{E}(\mathrm{e}^{i\lambda\cdot(X_t - X_s)} \mid \mathcal{G}_s) = \mathrm{e}^{-\frac{|\lambda|^2(t-s)}{2}}.$$

This proves the equivalence between the first two properties. For the third property, it suffices to use an argument of analytic continuation. ∎

Let $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, equipped with a filtration $(\mathcal{F}_t)_{t\geq 0}$.

**Definition 2.9** (Brownian motion with respect to a filtration)**.** *We say that a continuous $d$-dimensional process $X = (X_t)_{t\geq 0}$ is an $(\mathcal{F}_t)_{t\geq 0}$ Brownian motion when it is $(\mathcal{F}_t)_{t\geq 0}$ adapted and for all $0 \leq s < t$, $X_t - X_s$ is independent of $\mathcal{F}_s$ and follows the Gaussian law $\mathcal{N}(0, (t-s)I_d)$, which means that for all $\lambda \in \mathbb{R}^d$, the following process is an $(\mathcal{F}_t)_{t\geq 0}$-martingale:*

$$(\mathrm{e}^{i\lambda\cdot X_t + \frac{|\lambda|^2 t}{2}})_{t\geq 0}.$$

**Remark 2.10** (Definitions of Brownian motion)**.** *We see that if $X = (X_t)_{t\geq 0}$ is an $(\mathcal{F}_t)_{t\geq 0}$ Brownian motion, then $X$ is a Brownian motion in the sense of Definition 2.1. Conversely, any Brownian motion $(X_t)_{t\geq 0}$ in the sense of Definition 2.1 is an $(\mathcal{G}_t)_{t\geq 0}$ Brownian motion where $\mathcal{G}_t = \sigma(X_s : s \leq t)$ for all $t \geq 0$ is the natural filtration associated to $X$ (see Theorem 2.8).*

**Theorem 2.11** (Martingale property and Doob–Meyer decomposition)**.** *Let $B = (B_t)_{t\geq 0}$ be an $(\mathcal{F}_t)_{t\geq 0}$ $d$-dimensional Brownian motion and let $B_t = (B_t^1, \ldots, B_t^d)$ be the coordinates of the random vector $B_t$. Then for all $0 \leq s < t$, we have*

- $\mathbb{E}(B_t^j - B_s^j \mid \mathrm{F}_s) = 0$ *for all $1 \leq j \leq d$;*

- $\mathbb{E}((B_t^j - B_s^j)(B_t^k - B_t^k) \mid \mathcal{F}_s) = (t-s)\mathbf{1}_{j=k}$ *for all* $1 \leq j, k \leq d$.

*In particular, if* $\mathbb{E}(|B_0|^2) < \infty$ *then,*

- *for all* $1 \leq j \leq d$, $(B_t^j)_{t\geq 0}$ *is a continuous* $(\mathcal{F}_t)_{t\geq 0}$*-martingale;*
- *for all* $1 \leq j, k \leq d$, $(B_t^j B_t^k - \mathbf{1}_{j=k}t)_{t\geq 0}$ *is a continuous* $(\mathcal{F}_t)_{t\geq 0}$*-martingale, and*

$$(\langle B^j, B^k\rangle_t)_{t\geq 0} = (t\mathbf{1}_{j=k})_{t\geq 0}.$$

Actually Theorem 4.5 states that these properties characterize Brownian motion.

*Proof.* Left as an exercise. ■

## 2.1 Markov property of Brownian motion

Let $X = (X_t)_{t\geq 0}$ is an $(\mathcal{F}_t)_{t\geq 0}$ $d$-dimensional Brownian motion. We easily check that fir all fixed $T > 0$ the process $(X_{t+T} - X_T)_{t\geq 0}$ is a Brownian motion, issued form the origin, <u>independent</u> of $\mathcal{F}_T$. This is the simple Markov property, which extends to all $(\mathcal{F}_t)_{t\geq 0}$ stopping time $T$, namely:

**Theorem 2.12** (Strong Markov[2] property)**.** *Let $T$ be an $(\mathcal{F}_t)_{t\geq 0}$ stopping time, almost surely finite, let $\mathcal{F}_T$ be its stopping $\sigma$-algebra, and let $X = (X_t)_{t\geq 0}$ be an $(\mathcal{F}_t)_{t\geq 0}$ $d$-dimensional Brownian motion. Then the following properties hold true:*

1. $X^* = (X_{t+T} - X_T)_{t\geq 0}$ *<u>is a Brownian motion</u> issued from the origin, <u>independent</u> of* $\mathcal{F}_T$*;*
2. *For all measurable and bounded* $f : \mathbb{R}^d \to \mathbb{R}$*, we have, for all* $t > 0$*,*

$$\mathbb{E}(f(X_{t+T}) \mid \mathcal{F}_T) = P_t(f)(X_T) \quad \textit{where} \quad P_t(f)(x) = \frac{1}{(\sqrt{2\pi t})^d}\int_{\mathbb{R}^d} \mathrm{e}^{-\frac{|x-y|^2}{2t}} f(y)\mathrm{d}y.$$

*Proof.* For all $n \geq 1$, let us define

$$T_n = \sum_{k\geq 0} \frac{k+1}{2^n}\mathbf{1}_{[\frac{k}{2^n}, \frac{k+1}{2^n})}(T).$$

We check easily that $T_n$ is a stopping time, and that $T_n \searrow T$ as $n \to \infty$. Let $A \in \mathcal{F}_T$ and $0 = t_0 < t_1 < \cdots < t_m < \infty$, We check easily that $T_n$ is a stopping time, and that $T_n \searrow T$ as $n \to \infty$. Let $A \in \mathcal{F}_T$ and let $0 = t_0 < t_1 < \cdots < t_m < \infty$, $m \geq 0$. We have, for all continuous and bounded $\varphi : (\mathbb{R}^d)^m \to \mathbb{R}$,

$$\begin{aligned}\mathbb{E}(\mathbf{1}_A\varphi(X^*_{t_1}, \ldots, X^*_{t_m})) &= \mathbb{E}(\mathbf{1}_A\varphi(X_{t_1+T} - X_T, \ldots, X_{t_m+T} - X_T)) \\ &= \lim_{n\to\infty} \mathbb{E}(\mathbf{1}_A\varphi(X_{t_1+T_n} - X_{T_n}, \ldots, X_{t_m+T_n} - X_{T_n})).\end{aligned}$$

Moreover, for all $n \geq 1$, we have $A \in \mathcal{F}_{T_n}$ since $T \leq T_n$ and

$$\begin{aligned}\mathbb{E}(\mathbf{1}_A\varphi(X_{t_1+T_n} - X_{T_n}, \ldots, X_{t_m+T_n})) &= \sum_{r\in D_n} \mathbb{E}(\mathbf{1}_{A\cap\{T_n=r\}}\varphi(X_{t_1+r} - X_r, \ldots, X_{t_m+r} - X_r)) \\ &= \sum_{r\in D_n} \mathbb{P}(A\cap\{T_n = r\})\mathbb{E}(\varphi(X_{t_1+r} - X_r, \ldots, X_{t_m+r} - X_r)) \\ &= \mathbb{P}(A)\mathbb{E}(\varphi(X_{t_1} - X_0, \ldots, X_{t_m} - X_0))\end{aligned}$$

where $D_n = \{k/2^n : r \geq 0\}$ is the set of dyadics. This implies the first property since $(X_t - X_0)_{t\geq 0}$ is a Brownian motion issued from the origin. The second property is a consequence of the first one, and this is left as an exercise. ■

[2]Named after Andrey Markov (1856 – 1922), Russian mathematician.

Recall that the notions of finite variation of a function and quadratic variation of a process are considered in Section 1.2.3.

**Theorem 2.13** (Variation of Brownian motion sample paths). *Let $B=(B_t)_{t\geq 0}$ be a Brownian motion issued for the origin, let $[u,v]$ be a finite interval, $0\leq u<v$, and let $\delta$ be a partition or sub-division of $[u,v]$, $\delta : t_0=u<t_1<\cdots<t_n=v$, $n\geq 1$. Let us consider the quantities*

$$r_1(\delta)=\sum_{i=1}^{n-1}|B_{t_{i+1}}-B_{t_i}| \quad \textit{and} \quad r_2(\delta)=\sum_{i=0}^{n-1}|B_{t_{i+1}}-B_{t_i}|^2.$$

*Then the following properties hold true:*

1. $\lim_{|\delta|\to 0} r_2(\delta)=v-u$ *in* $\mathrm{L}^2$ *and thus in* $\mathbb{P}$*, where* $|\delta|=\sup_{i\in\{0,1,\ldots,n-1\}}(t_{i+1}-t_i)$*. In other words the quadratic variation of* $B$ *is given by* $[B]_t=t$ *for all* $t\geq 0$*;*
2. $\sup_{\delta\in\mathcal{P}} r_1(\delta)=+\infty$ *almost surely, where* $\mathcal{P}$ *is the set of subdivision of* $[u,v]$*. In other words the sample paths of* $B$ *are almost surely of infinite variation on all interval* $[u,v]$*.*

*Proof.* If $Z\sim\mathcal{N}(0,1)$ then $\mathbb{E}(Z^4)=3$. Therefore

$$\begin{aligned}
\mathbb{E}((r_2(\delta))^2)&=\mathbb{E}((\sum_i|B_{t_{i+1}}-B_{t_i}|^2)^2)\\
&=\sum_i\mathbb{E}(|B_{i+1}-B_{t_i}|^4)+2\sum_{i<j}\mathbb{E}(|B_{t_{i+1}}-B_{t_i}|^2|B_{t_{j+1}}-B_{t_j}|^2)\\
&=3\sum_i(t_{i+1}-t_i)^2+2\sum_{i<j}(t_{i+1}-t_i)(t_{j+1}-t_j)\\
&=2\sum_i(t_{i+1}-t_i)^2+(\sum_i(t_{i+1}-t_i))^2\\
&=2\sum_i(t_{i+1}-t_i)^2+(v-u)^2.
\end{aligned}$$

Moreover $\mathbb{E}(r_2(\delta))=\sum_i(t_{i+1}-t_i)=v-u$. Thus

$$\mathbb{E}((r_2(\delta)-(v-u))^2)=2\sum_i(t_{i+1}-t_i)^2\leq 2\sup_i(t_{i+1}-t_i)(v-u)\underset{|\delta|\to 0}{\longrightarrow}0.$$

Therefore there exists a sequence of subdivisions $(\delta^k)_k$ of $[u,v]$ such that

$$\lim_{k\to\infty} r_2(\delta^k)=v-u \quad \text{almost surely.}$$

Moreover we see that

$$\sup_\delta r_1(\delta)\geq r_1(\delta^k)=\sum_i|B_{t^k_{i+1}}-B_{t^k_i}|\geq\frac{\sum_i|B_{t^k_{i+1}}-B_{t^k_i}|}{\sup_i|B_{t^k_{i+1}}-B_{t^k_i}|}\underset{k\to\infty}{\longrightarrow}+\infty.$$

■

**Theorem 2.14** (Law of the iterated logarithm). *If $B=(B_t)_{t\geq 0}$ is a Brownian motion on $\mathbb{R}$ issued from the origin then*

$$\mathbb{P}\left(\overline{\lim_{t\searrow 0}}\frac{B_t}{\sqrt{2t\log(\log(1/t))}}=1\right)=1$$

*and*

$$\mathbb{P}\left(\underline{\lim_{t\searrow 0}}\frac{B_t}{\sqrt{2t\log(\log(1/t))}}=-1\right)=1$$

*and*

$$\mathbb{P}\left(\overline{\lim_{t\to\infty}}\frac{B_t}{\sqrt{2t\log(\log(t))}}=1,\quad \underline{\lim_{t\to\infty}}\frac{B_t}{\sqrt{2t\log(\log(t))}}=-1\right)=1.$$

*Proof.* The first property is essentially a consequence of the Borel–Cantelli lemma, see for instance [11]. The second (respectively third) property follows by replacing $B_t$ by $-B_t$ (respectively $tB_{1/t}$) which is also a Brownian motion. ■

**Theorem 2.15** (Regularity of the sample paths of Browian motion)**.** *If $B = (B_t)_{t\geq 0}$ is a Brownian motion issued from the origin then*

$$\mathbb{P}\left(\overline{\lim_{t\searrow 0}} \frac{1}{\sqrt{2t\log(\log(1/t))}} \max_{\substack{0\leq t_1<t_2\leq 1\\ |t_1-t_2|\leq t}} |B_{t_2}-B_{t_1}| = 1\right) = 1.$$

*In particular almost surely the sample paths of $B$ are Hölder continuous of order $\alpha$ for all $\alpha \in (0,1/2)$ and are nowhere differentiable on $\mathbb{R}_+$.*

*Proof.* See for instance [11]. ■

## 2.2 A construction of Brownian motion

The mathematical existence of Brownian motion is not obvious. Norbert Wiener[3] was the first to give a rigorous construction in 1923, and for this reason, the Brownian motion is sometimes called the Wiener process.

**Lemma 2.16.** *Let $X = (X_n)_{n\geq 0}$ be a sequence of Gaussian random variables with $X_n \sim \mathcal{N}(m_n, \sigma_n)$. We suppose that $\lim_{n\to\infty} X_n = X$ in probability. Then $X$ is a Gaussian random variable and moreover $\lim_{n\to\infty} X_n = X$ in $\mathrm{L}^2$.*

*Proof.* Left as an exercise (hint: use characteristic functions). ■

**Theorem 2.17** (Brownian measures)**.** *Let us consider the Hilbert space $\mathrm{G} = \mathrm{L}^2(\mathbb{R}, \mathrm{d}x)$, and let $\langle f,g\rangle_{\mathrm{G}} = \int f(x)g(x)\mathrm{d}x$ for all $f,g \in \mathrm{G}$. Then there exists a <u>centered Gaussian process</u> $\widetilde{B} = (\widetilde{B}_g)_{g\in\mathrm{G}}$ defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ such that for all $f,g \in \mathrm{G}$ and $\alpha,\beta \in \mathbb{R}$ the following properties hold true:*

1. *$\langle f,g\rangle_{\mathrm{G}} = \mathbb{E}(\widetilde{B}_f \widetilde{B}_g)$;*
2. *$\widetilde{B}_{\alpha f+\beta g} = \alpha\widetilde{B}_f + \beta\widetilde{B}_g$.*

*The isometry $g \in \mathrm{G} \mapsto \widetilde{B}_g \in \mathrm{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ is called the* Brownian measure*.*

Beware that the Brownian measure defined above is not a true random measure because the negligible event behind the equality in the second property (equality behind two random variables) depends on $f$ and $g$ and G is not countable. Indeed, two uncountable families of random variables are not necessarily equal when they are pointwise equal as random variables!

*Proof.* Let $(X_n)_{n\geq 0}$ be a sequence of real Gaussian random variables independent and identically distributed with law $\mathcal{N}(0,1)$, defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, and let $(e_n)_{n\geq 0}$ be an orthonormal sequence of the Hilbert space $\mathrm{G} = \mathrm{L}^2(\mathbb{R}, \mathrm{d}x)$. For all $g \in \mathrm{G}$, the series $\sum_{n\geq 0} X_n(\omega)\langle g, e_n\rangle_{\mathrm{G}}$ converges in $\mathrm{L}^2(\Omega, \mathcal{A}, \mathbb{P})$. Indeed

$$\mathbb{E}\Big(\Big(\sum_{n=p}^{p+q} X_n\langle g,e_n\rangle_{\mathrm{G}}\Big)^2\Big) = \sum_{n=p}^{p+q}\langle g,e_n\rangle_{\mathrm{G}}^2 \underset{p,q\to\infty}{\longrightarrow} 0.$$

[3]Named after Norbert Wiener (1894 – 1964), American mathematician.

Let us define now

$$\widetilde{B}_g = \sum_{n\geq 0} X_n \langle g, e_n\rangle_{\mathrm{G}}.$$

We see from Lemma 2.16 that $\widetilde{B}$ is a centered Gaussian random variable and that

$$\|\widetilde{B}_g\|_{\mathrm{G}}^2 = \mathbb{E}((\widetilde{B}_g)^2) = \langle g, g\rangle_{\mathrm{G}} = \|g\|_{\mathrm{G}}^2$$

hence $g \mapsto \widetilde{B}_g$ is an isometry. The linearity of $\widetilde{B}_g$ with respect to $g$ is immediate. By polarization we get, for all $f, g \in \mathrm{G}$,

$$\begin{aligned}
4\mathbb{E}(\widetilde{B}_f\widetilde{B}_g) &= \mathbb{E}((\widetilde{B}_f + \widetilde{B}_g)^2) - \mathbb{E}((\widetilde{B}_f - \widetilde{B}_g)^2) \\
&= \mathbb{E}(\widetilde{B}_{f+g}^2) - \mathbb{E}(\widetilde{B}_{f-g}^2) \\
&= \|f+g\|_{\mathrm{G}}^2 - \|f-g\|_{\mathrm{G}}^2 \\
&= \langle f, g\rangle_{\mathrm{G}}.
\end{aligned}$$

Finally, let $\mathrm{H} \subset \mathrm{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ be the closed sub-space generated by the family of random variables $\{\widetilde{B}_g : g \in \mathrm{G}\}$. Then we easily check that H is isomorphic, thanks to the correspondence $g \mapsto \widetilde{B}_g$, to the Hilbert space $\mathrm{G} = \mathrm{L}^2(\mathbb{R}, \mathrm{d}x)$. ■

For all $t \in \mathbb{R}_+$, let us define

$$B_t = \widetilde{B}_{\mathbf{1}_{[0,t]}}.$$

We see that the process $B = (B_t)_{t\geq 0}$ is a centered Gaussian process, with covariance given for all $s, t \in \mathbb{R}_+$ by $\mathbb{E}(B_sB_t) = \langle \mathbf{1}_{[0,s]}, \mathbf{1}_{[0,t]}\rangle_{\mathrm{L}^2(\mathbb{R},\mathrm{d}x)} = s \wedge t$. However the process $B$ has no reason to be continuous. Let us remark however that for all $0 \leq s < t$,

$$\frac{B_t - B_s}{\sqrt{t-s}} \sim \mathcal{N}(0,1) \quad \text{and} \quad \mathbb{E}((B_t - B_s)^4) = C(t-s)^2 \text{ where } C = \mathbb{E}(\xi^4),\ \xi \sim \mathcal{N}(0,1).$$

Now Theorem 2.18 below allows then to construct a continuous modification $B^*$ of $B$, which is a Brownian motion on $\mathbb{R}$ issued from the origin.

**Theorem 2.18** (Kolmogorov continuity criterion)**.** *Let $Z = (Z_t)_{t\geq 0}$ be a process on $(\Omega, \mathcal{A}, \mathbb{P})$ taking values in a Banach space, and such that there exist $p \in [1, \infty)$, $\varepsilon \in (0, \infty)$, and $A \in \mathbb{R}_+$ such that for all $s, t \in \mathbb{R}_+$,*

$$\mathbb{E}(\|Z_t - Z_s\|^p) \leq A|t-s|^{1+\varepsilon}.$$

*Then there exists a process $Z^* = (Z_t^*)_{t\geq 0}$ with continous sample paths such that $Z_t = Z_t^*$ (as random variables, in other words almost surely) for all $t \in \mathbb{R}_+$.*

*Proof.* For all $n \geq 1$, let us define $D_n = \{k/2^n : k \geq 0\}$ and

$$B_n = \{(s,t) \text{ with } s, t \leq n; s, t \in D_n; |s-t| \leq 2^{-n}\}.$$

and

$$U_n = \sup_{(s,t)\in B_n} \|Z_t - Z_s\|.$$

*First step.* Let us show that $\sum_n U_n < \infty$ almost surely. Indeed

$$\mathbb{E}(U_n^p) = \mathbb{E}(\sup_{(s,t)\in B_n} \|Z_t - Z_s\|^p) \leq \sum_{(s,t)\in B_n} \mathbb{E}(\|Z_t - Z_s\|^p) \leq 2n2^n A 2^{-n-n\varepsilon} = An2^{1-n\varepsilon}.$$

therefore

$$\mathbb{E}(\sum_n U_n)) = \sum_n \mathbb{E}(U_n) \leq \sum_n (\mathbb{E}(U_n^p))^{1/p} < \infty$$

and thus $\sum_n U_n < \infty$ almost surely.

*Second step.* For all $n \geq 1$ and all $\omega \in \Omega$, let $t \in \mathbb{R}_+ \mapsto Z_t^{(n)}(\omega)$ be piecewise linear such that $Z_t^{(n)}(\omega) = Z_t(\omega)$ for all $t \in D_n$, namely

$$Z_t^{(n)}(\omega) = Z_{k/2^n}(\omega) + 2^n(t - k/2^n)(Z_{(k+1)/2^n}(\omega) - Z_{k/2^n}(\omega))$$

for all $t \in [k/2^n, (k+1)/2^n)$. Moreover we have $\sup_{t \leq n} \|Z_t^{(n+1)} - Z_t^{(n)}\| \leq 4U_{n+1}$. The series $\sum_n \|Z_t^{(n+1)} - Z_t^{(n)}\|$ converges then almost surely uniformly on every compact subset of $\mathbb{R}_+$. Therefore, there exists a continuous process $Z^* = (Z_t^*)_{t \geq 0}$ such that $\lim_{n\to\infty} Z_t^{(n)} = Z_t^*$ almost surely uniformly on every compact of $\mathbb{R}_+$.

The set $D = \cup_{n \geq 1} D_n$ is dense in $\mathbb{R}_+$, and for all $r \in D_n$, by construction, $Z_r = Z_r^*$ almost surely. Now let $t \in \mathbb{R}_+$, then there exists a sequence $(s_n)_n$ in $D$ such that $\lim_{n\to\infty} s_n = t$ and thus $\lim_{n\to\infty} Z_{s_n} = Z_t$ in $\mathrm{L}^p$. Thus there exists a subsequence $(s_{n_k})_k$ such that $\lim_{k\to\infty} Z_{n_k} = Z_t$ almost surely. Therefore, using the continuity of $Z^*$, we get $Z_{s_k} = Z_{s_k}^* \to Z_t^* = Z_t$ almost surely as $k \to \infty$. ■

## 2.3 Wiener integral

**Theorem 2.19** (Wiener integral). *Let $B = (B_t)_{t\geq 0} = ((B_t^1, \ldots, B_t^n))_{t \geq 0}$ be a d-dimensional Brownian motion issued form the origin, defined on $(\Omega, \mathcal{A}, \mathbb{P})$. Let $\mathrm{G}$ be the Gaussian sub-space of $\mathrm{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ generated by the real random variables $\{B_t^i : t \geq 0, 1 \leq i \leq d\}$. Then there exists a unique linear and continuous bijection $I$ between $\mathrm{L}^2_{\mathbb{R}^d}(\mathbb{R}_+, \mathrm{d}x)$ and $\mathrm{G}$ such that:*

1. *If $g = a\mathbf{1}_{(s,t]}$ with $0 \leq s \leq t$ and $a \in \mathbb{R}^d$ then $I(g) = a \cdot (B_t - B_s)$;*
2. *If $f$ and $g$ belong to $\mathrm{L}^2_{\mathbb{R}^d}(\mathbb{R}_+, \mathrm{d}x)$ then*

$$\int_{\mathbb{R}_+} f(s) \cdot g(s)\mathrm{d}s = \mathbb{E}(I(f)I(g)).$$

*The Wiener integral of $g$ is the random variable $I(g)$ and we denote*

$$I(g)(\omega) = \int_{\mathbb{R}_+} g(s)\mathrm{d}B_s(\omega).$$

*Proof.* The following sub-space

$$S = \left\{ f \in \mathrm{L}^2_{\mathbb{R}^d}(\mathbb{R}_+, \mathrm{d}x) : f = \sum_{i=0}^{n} a_i \mathbf{1}_{(t_i, t_{i+1}]}, t_0 = 0 < t_1 < \cdots < t_n, n \geq 0, a_i \in \mathbb{R}^d \right\}$$

of $\mathrm{L}^2_{\mathbb{R}^d}(\mathbb{R}_+, \mathrm{d}x)$ is dense. If $f \in S$, then $f = \sum_{\text{finite}} a_i \mathbf{1}_{(t_i, t_{i+1}]}$, and we define

$$I(f) = \sum_{\text{finite}} a_i \cdot (B_{t_{i+1}} - B_{t_i}).$$

We check easily that this definition does not depend on the decomposition chosen for $f$, and that the map $f \mapsto I(f)$ is linear. Moreover, we remark that thanks to the properties of Brownian motion, we have

$$\begin{aligned}
\mathbb{E}((I(f))^2) &= \sum_{i,j} \mathbb{E}((a_i \cdot (B_{t_{i+1}} - B_{t_i}))(a_j \cdot (B_{t_{j+1}} - B_{t_j}))) \\
&= \sum_i \mathbb{E}((a_i \cdot (B_{t_{i+1}} - B_{t_i}))^2) \\
&= \sum_i |a_i|^2 (t_{i+1} - t_i) \\
&= \int_{\mathbb{R}_+} |f(x)|^2 \mathrm{d}x.
\end{aligned}$$

Since $S$ is dense, the map $I$ can be extended by continuity to the whole $\mathrm{L}^2(\mathbb{R}_+, \mathrm{d}x)$. For all $f \in \mathrm{L}^2_{\mathbb{R}^d}$, there exists a sequence $(f_n)_n$ is $S$ such that $\|f_n - f\| \to 0$. Therefore $\|f_n - f_m\| = \|I(f_n) - I(f_m)\|_{\mathrm{L}^2(\Omega,\mathcal{A},\mathbb{P})} \to$ as $n, m \to \infty$. Set $I(f) = \lim_{n\to\infty} I(f_n)$. This limit does not depend on the sequence $(f_n)_n$ used to approximate $f$. Moreover $\|f\|_2^2 = \mathbb{E}((I(f))^2)$, and, by polarization, for all $f, g \in \mathrm{L}^2_{\mathbb{R}^d}$,

$$\int_{\mathbb{R}_+} f(s)g(s)\mathrm{d}s = \mathbb{E}(I(f)I(g)).$$

The uniqueness of the map $I$ defined this way is obvious. The map $I$ is an isometry from $\mathrm{L}^2_{\mathbb{R}^d}$ to G since on one hand $F = I(\mathrm{L}^2_{\mathbb{R}^d})$ is a closed sub-space of G, and on the other hand, for all $t \geq 0$ and $1 \leq i \leq d$, $B^i_t \in F$ and therefore $F$ is dense in G. ∎

**Remark 2.20.** *We see that for all $f, g \in \mathrm{L}^2_{\mathbb{R}^d}$, $I(f) \sim \mathcal{N}(0, \|f\|_2^2)$ and the real random variables $I(f)$ and $I(g)$ are independent if and only if $\int_{\mathbb{R}_+} f(s)g(s)\mathrm{d}s = 0$.*

**Corollary 2.21** (Properties of the Wiener integral)**.**

1. *For all $t \geq 0$ and $1 \leq i \leq d$ and $f \in \mathrm{L}^2_{\mathbb{R}^d}$, we have*

$$\mathbb{E}\left(B^i_t \int_{\mathbb{R}_+} f(s)\mathrm{d}B_s\right) = \int_0^t f^i(s)\mathrm{d}s$$

   *where $f^i(s)$ is the $i$-th coordinate of $f(s) = (f^1(s), \dots, f^d(s))$;*

2. *Let $(f_n)_{n\geq 0}$ be an orthonormal basis of $\mathrm{L}^2_{\mathbb{R}^d}$, then $(I(f_n))_{n\geq 0}$ is a sequence of independent Gaussian real random variables with mean zero and unit variance and for all $t \geq 0$, we have the following expansion in $\mathrm{L}^2(\Omega, \mathcal{A}, \mathbb{P})$:*

$$B_t(\omega) = \sum_{n\geq 0} \left(\int_{\mathbb{R}_+} f_n(s)\mathrm{d}B_s\right) \int_0^t f_n(s)\mathrm{d}s.$$

*Proof.*

1. Take $g = e_i \mathbf{1}_{(0,t]}$ then by definition of $I$ we have $I(g) = B^i_t$ and

$$\mathbb{E}\left(B^i_t \int_{\mathbb{R}_+} f(s)\mathrm{d}B_s\right) = \mathbb{E}(I(g)I(f)) = \int_{\mathbb{R}_+} g(s) \cdot f(s)\mathrm{d}s = \int_0^t f^i(s)\mathrm{d}s.$$

2. If $(f_n)_{n\geq 0}$ is an orthonormal basis of $\mathrm{L}^2_{\mathbb{R}^d}$ then $(I(f_n))_{n\geq 0}$ is an orthonormal basis of the Gaussian space G and moreover $\langle B^i_t, I(f_n)\rangle_{\mathrm{G}} = \int_0^t f^i_n(s)\mathrm{d}s$.

∎

## 2.4 Wiener measure and canonical Brownian motion

As a random variable on trajectories, Brownian motion is not unique. We can construct an infinite number of versions of it. What is unique is its law. It is known as the Wiener measure. There exists however a special realization of Brownian motion as a random variable, which is called the canonical Brownian motion, defined on a canonical space.

The canonical space is $\mathrm{W} = \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$, equipped with the topology of uniform convergence on every compact subset of $\mathbb{R}_+$, and with its Borel $\sigma$-algebra $\mathcal{B}_{\mathrm{W}}$.

Let $B = (B_t)_{t\geq 0}$ be an arbitrary $d$-dimensional Brownian motion issued from the origin, and defined on $(\Omega, \mathcal{A}, \mathbb{P})$. We can consider $B$ as a random variable defined on $(\Omega, \mathcal{A})$ and taking values in $(\mathrm{W}, \mathcal{B}_{\mathrm{W}})$: for almost all $\omega \in \Omega$, $B(\omega) \in \mathrm{W}$ is the continuous sample path $t \mapsto B_t(\omega)$.

**Theorem 2.22** (Wiener measure)**.** *There exists a unique probability measure $\mu$ on the canonical space $(\mathrm{W}, \mathcal{B}_{\mathrm{w}})$ such that for all $n \geq 1$ and $0 < t_1 < t_2 < \cdots < t_n$ and all $A_1, A_2, \ldots, A_n \in \mathbb{B}_{\mathbb{R}^d}$,*

$$\mu(\{w \in \mathrm{W} : w_{t_1} \in A_1, w_{t_2} \in A_2, \ldots, w_{t_n} \in A_n\}) \\ = \int_{A_1 \times A_2 \times \cdots \times A_n} p_{t_1}(x_1) p_{t_2 - t_1}(x_2 - x_1) \cdots p_{t_n - t_{n-1}}(x_n - x_{n-1}) \mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_n$$

*where $p$ is the heat or Gaussian kernel defined for all $t > 0$ and $x \in \mathbb{R}^d$ by*

$$p_t(x) = \frac{1}{(\sqrt{2\pi t})^d} \mathrm{e}^{-\frac{|x|^2}{2t}}.$$

*Moreover for all d-dimensional Brownian motion $B = (B_t)_{t \geq 0}$ issued from the origin, we have, for all measurable and bounded or positive $\Phi : \mathrm{W} \to \mathbb{R}$,*

$$\mathbb{E}(\Phi(B)) = \int_{\mathrm{W}} \Phi(w) \mu(\mathrm{d}w).$$

*We say that $\mu$ is the Wiener measure.*

*Proof.* We know how to construct a $d$-dimensional Brownian motion $B = (B_t)_{t \geq 0}$ issued form the origin. If $\mu$ is the law of $B$ seen as a random variable taking values on the canonical space $(\mathrm{W}, \mathbb{B}_{\mathrm{W}})$, then it is immediate to get the first desired property since

$$\mu(B_{t_1} \in A_1, \ldots, B_{t_n} \in A_n) = \mu(\{w \in \mathrm{W} : w_{t_1} \in A_1, \ldots, w_{t_n} \in A_n\}).$$

The uniqueness of $\mu$ is a consequence of the fact that it is entirely determined on the family $\mathcal{C}$ of cylindrical subsets of W, which is stable by finite intersections and generates $\mathbb{B}_{\mathrm{W}}$, see the monotone class theorem (Corollary 1.20). ■

### 2.4.1 Cameron–Martin formula

On the probability space $(\mathrm{W} = \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d), \mathcal{B}_w, \mu)$, where $\mu$ is the Wiener measure, let us consider the coordinates process $\pi = (\pi_t)_{t \geq 0}$ defined by

$$\pi_t(w) = w_t$$

for all $t \geq 0$ and $w \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$. We see that under $\mu$, the process $\pi$ is a $d$-dimensional Brownian motion issued from the origin. It is called the canonical Brownian motion.

The Cameron–Martin space is defined by

$$\mathrm{H} = \left\{ h \in \mathrm{W} : \forall t \geq 0, h(t) = \int_0^t \dot{h}(s) \mathrm{d}s, \dot{h} \in \mathrm{L}^2_{\mathbb{R}^d}(\mathbb{R}_+, \mathrm{d}x) \right\}.$$

This is a subspace of $\mathrm{W} = \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$. We equip H with the scalar product

$$\langle h_1, h_2 \rangle_{\mathrm{H}} = \int_{\mathbb{R}_+} \dot{h}_1(s) \cdot \dot{h}_2(s) \mathrm{d}s.$$

This makes H a Hilbert space isomorphic to $\mathrm{L}^2_{\mathbb{R}^d}(\mathbb{R}_+, \mathrm{d}x)$, and the canonical injection from H to W is continuous. For every $h \in \mathrm{H}$ we denote

$$|h|^2_{\mathrm{H}} = \int_{\mathbb{R}_+} |\dot{h}(s)|^2 \mathrm{d}s.$$

Let $(h_n)_{n\geq 0}$ be an orthonormal basis of H and for all $n \geq 0$ and $t \geq 0$,

$$h_n(t) = \int_0^t \dot{h}_n(s)\mathrm{d}s.$$

The sequence $(\dot{h}_n)_{n\geq 0}$ is an orthonormal basis of $\mathrm{L}^2_{\mathbb{R}^d}(\mathbb{R}_+, \mathrm{d}x)$. Let $\pi = (\pi_t)_{t\geq 0}$ be the canonical Brownian motion and let us define, for all $w \in \mathrm{W}$ and $h \in \mathrm{H}$, using the Wiener integral,

$$(w, h) = \int_{\mathbb{R}_+} \dot{h}(s)\mathrm{d}\pi_s(w) = \int_{\mathbb{R}_+} \dot{h}(s)\mathrm{d}w_s.$$

Then the second property provided by Corollary 2.21 gives that the real random variables $(w, h_n), n \geq 0$, are independent, Gaussian, with zero mean and unit variance, and for all $t \geq 0$ we have the following expansion in $\mathrm{L}^2(\mathrm{W}, \mathcal{B}_\mathrm{W}, \mu)$:

$$\pi_t(w) = w_t = \sum_{n\geq 0}(w, h_n)h_n(t).$$

**Remark 2.23.** *We can prove by using a convergence theorem for vector maringales that for all $T > 0$ and for almost all $w$ for $\mu$, we have*

$$\lim_{N\to\infty} \sup_{t\in[0,T]} \left| w_t - \sum_{n=0}^{N}(w, h_n)h_n(t) \right| = 0.$$

**Theorem 2.24** (Cameron[4]–Martin[5] formula)**.** *If $\Phi : \mathrm{W} \to \mathbb{R}$ is measurable and bounded then for all $h$ in the Cameron–Martin space* H*, we have*

$$\mathbb{E}_\mu(\Phi(w + h)) = \mathbb{E}_\mu\Big(\Phi(w) \exp\Big((w, h) - \frac{|h|^2_\mathrm{H}}{2}\Big)\Big)$$

*where* W *is the Wiener space and $\mu$ is the Wiener measure.*

*Proof.* We can assume without loss of generality that $\Phi$ is such that

$$\Phi(w) = f(w_{t_1}, \ldots, w_{t_n})$$

where $n \geq 1$ and $0 \leq t_1 < \cdots < t_n$ and $f \in \mathcal{C}^\infty(\mathbb{R}^n)$ with compact support. Let $h \in \mathrm{H}$, $h \neq 0$. There exists an orthonormal basis $(h_m)_{m\geq 0}$ of H such that $h_0 = h/|h|_\mathrm{H}$. For all $m \geq 1$, let $w^{(m)} \in W$ be defined $\mu$ almost surely by

$$w^{(m)}(t) = \sum_{\ell=0}^{m}(w, h_\ell)h_\ell(t), \quad t \geq 0.$$

We can check easily that

$$\lim_{m\to\infty} w_m = w = \sum_{n\geq 0}(w, h_n)h_n, \quad \text{and} \quad \lim_{m\to\infty} \mathbb{E}_\mu(\Phi(w^{(m)} + h)) = \mathbb{E}_\mu(\Phi(w + h)),$$

and similarly that

$$\lim_{m\to\infty} \mathbb{E}_\mu\Big(\Phi(w^{(m)} \exp\Big((w, h) - \frac{|h|^2_\mathrm{H}}{2}\Big)\Big) = \mathbb{E}_\mu\Big(\Phi(w) \exp\Big((w, h) - \frac{|h|^2_\mathrm{H}}{2}\Big)\Big).$$

[4]Named after Robert Horton Cameron (1908 – 1989), American mathematician.
[5]Named after William Ted Martin (1911 – 2004), American mathematician.

Also, to prove the desired formula, it suffices to show that for all $m \geq 1$,

$$\mathbb{E}_\mu(\Phi(w^{(m)} + h)) = \mathbb{E}_\mu\Big(\Phi(w^{(m)} \exp\Big((w,h) - \frac{|h|_{\mathrm{H}}^2}{2}\Big)\Big).$$

This boils down to a simple computation in finite dimension. Namely, since

$$w^{(m)} + h = ((w,h_0) + |h|_{\mathrm{H}})\frac{h}{|h|_{\mathrm{H}}} + \sum_{\ell=1}^{m}(w,h_\ell)h_\ell,$$

where $(w,h_\ell)$, $\ell \geq 0$ are independent and identically distributed with law $\mathcal{N}(0,1)$, we have

$$\begin{aligned}\mathbb{E}_\mu(\Phi(w^{(m)} + h)) &= \frac{1}{(\sqrt{2\pi})^{m+1}}\int_{\mathbb{R}^{m+1}}\Phi\Big((x_0 + |h|_{\mathrm{H}})h_0 + \sum_{\ell=0}^{m} x_\ell h_\ell\Big)\mathrm{e}^{-\frac{1}{2}\sum_{\ell=0}^m x_\ell^2}\mathrm{d}x_0\cdots\mathrm{d}x_m\\ &= \frac{1}{(\sqrt{2\pi})^{m+1}}\int_{\mathbb{R}^{m+1}}\Phi\Big(x_0'h_0 + \sum_{\ell=0}^{m} x_\ell' h_\ell\Big)\mathrm{e}^{x_0'|h|_{\mathrm{H}} - \frac{1}{2}|h|_{\mathrm{H}}^2 - \frac{1}{2}\sum_{\ell=0}^m x_\ell'^2}\mathrm{d}x_0'\cdots\mathrm{d}x_m'\\ &= \mathbb{E}_\mu(\Phi(w^{(m)})\mathrm{e}^{(w,\frac{h}{|h|_{\mathrm{H}}}) - \frac{|h|_{\mathrm{H}}^2}{2}}).\end{aligned}$$

■

**Corollary 2.25** (Density of Cameron–Martin translations)**.** *If $B = (B_t)_{t\geq 0}$ is a $d$-dimensional Brownian motion issued from the origin defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ then,*

$$\textit{for all } h \in \mathrm{H}, \quad \mathbb{E}(Z_h) = 1, \quad \textit{where} \quad Z_h = \exp\Big(\int_{\mathbb{R}_+}\dot{h}_s\mathrm{d}B_s - \frac{1}{2}\int_{\mathbb{R}_+}|\dot{h}_s|^2\mathrm{d}s\Big).$$

*Moreover, for all $h \in \mathrm{H}$, the law of the translated process*

$$\widetilde{B}^{(h)} = (B_t - h_t)_{t\geq 0}$$

*is absolutely continuous with respect to the law of $B$, with density given by $Z_h$, in other words the translation process $\widetilde{B}^{(h)}$ is a $d$-dimensional Brownian motion issued from the origin with respect to the probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{A})$ given by $\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P} = Z_h$.*

*Proof.* We know that $\int_{\mathbb{R}_+}\dot{h}_s\mathrm{d}B_s \sim \mathcal{N}(0, |h|_{\mathrm{H}}^2)$, and thus

$$\mathbb{E}\Big(\exp\Big(\int_{\mathbb{R}_+}\dot{h}_s\mathrm{d}B_s\Big)\Big) = \exp\Big(\frac{|h|_{\mathrm{H}}^2}{2}\Big),$$

which gives $\mathbb{E}(Z_h) = 1$. Since the Wiener measure $\mu$ is the law of $B$ (under $\mathbb{P}$), the Cameron–Martin formula of Theorem 2.24 writes, for all measurable and bounded $\Phi : \mathrm{W} \to \mathbb{R}$,

$$\mathbb{E}(\Phi(B+h)) = \mathbb{E}\Big(\Phi(B)\exp\Big(\int_{\mathbb{R}_+}\dot{h}_s\mathrm{d}B_s - \frac{|h|_{\mathrm{H}}^2}{2}\Big)\Big),$$

and therefore

$$\mathbb{E}(\Phi(B)) = \mathbb{E}(\Phi(B-h)Z) = \mathbb{E}_{\mathbb{Q}}(\Phi(B-h)).$$

The law of $\mathbb{Q}$ under $\mathbb{P}$ is therefore the law of $\widetilde{B}^{(h)} = B - h$ under $\mathbb{Q}$. ■

# Chapter 3

# Stochastic integral

## 3.1 Stochastic integral with respect to Brownian motion

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probabilty space, and let $B = (B_t)_{t\geq 0}$ be a $d$-dimensional $(\mathcal{F}_t)_{t\geq 0}$ Brownian motion issued from the origin. We suppose in the sequel that the filtration $(\mathcal{F}_t)_{t\geq 0}$ is right continuous and complete.

### 3.1.1 Stochastic integral of step functions

We say that a $d$-dimensional process $\varphi = (\varphi_t)_{t\geq 0}$ is a step process when there exists a subdivision $t_0 = 0 < t_1 < \cdots < t_n < \cdots$ such that $\lim_{n\to\infty} t_n = +\infty$ and a sequence $(U_n)_{n\geq 0}$ of bounded random vectors of $\mathbb{R}^d$ such that $U_n$ is $\mathcal{F}_{t_n}$-measurable for all $n \geq 0$ and for all $t \geq 0$,

$$\varphi_t = U_0 \mathbf{1}_0(t) + \sum_{n=0}^{\infty} U_n \mathbf{1}_{(t_n,t_{n+1}]}(t).$$

The process $\varphi$ is adapted, mesurable, and left-continuous and in particular predictable.

The set $\mathcal{C}$ of all step processes is a vector space. If $\varphi \in \mathcal{C}$ then we define the stochastic integral of $\varphi$ with respect to $B$ by the following formula for all $t \geq 0$:

$$J(\varphi)_t = \int_0^t \varphi_s \mathrm{d}B_s = \begin{cases} 0 & \text{if } t = 0; \\ U_0 \cdot (B_{t_1} - B_{t_0}) + \cdots + U_k \cdot (B_t - B_{t_k}) & \text{if } t \in (t_k, t_{k+1}]. \end{cases}$$

The process $J(\varphi) = (J(\varphi)_t)_{t\geq 0}$ is real, $(\mathcal{F}_t)_{t\geq 0}$-adapted, and issued from the origin. It does not depend on the decomposition chosen for $\varphi$. Since $B$ has continuous sample paths, we see immediately that $J(\varphi)$ has also continuous sample paths.

The map $J$ from $\mathcal{C}$ to the space of processes is clearly linear.

**Theorem 3.1** (Properties of the stochastic integral)**.**

1. *For all $\varphi \in \mathcal{C}$, the process $J(\varphi) = (\int_0^t \varphi_s \mathrm{d}B_s)_{t\geq 0}$ is a real continuous martingale, centered, square integrable, and issued from the origin;*

2. *If $\varphi, \psi \in \mathcal{C}$ then the following process is an is an $(\mathcal{F}_t)_{t\geq 0}$-martingale:*

$$\left( \left( \int_0^t \varphi_s \mathrm{d}B_s \right) \left( \int_0^t \psi_s \mathrm{d}B_s \right) - \int_0^t (\varphi_s \cdot \psi_s) \mathrm{d}s \right)_{t\geq 0}.$$

*In particular the following process is an $(\mathcal{F}_t)_{t\geq 0}$-martingale:*

$$\left( \left( \int_0^t \varphi_s \mathrm{d}B_s \right)^2 - \int_0^t |\varphi_s|^2 \mathrm{d}s \right)_{t\geq 0}.$$

*In particular, the conservation of mean gives for all $t \geq 0$ the isometry*

$$\mathbb{E}\left(\left(\int_0^t \varphi_s \mathrm{d}B_s\right)^2\right) = \mathbb{E}\left(\int_0^t |\varphi_s|^2 \mathrm{d}s\right).$$

*Proof.*

1. This follows easily from the properties of Brownian motion, and is left to the reader;

2. By polarization one can assume without loss of generality that $\varphi = \psi$, namely, since $J$ is linear, we have $4J(\varphi)J(\psi) = J(\varphi+\psi)^2 - J(\varphi-\psi)^2$. Let us assume now that $\varphi = \psi$. Let $0 \leq s < t$ and $A \in \mathcal{F}_s$, and set $X_t = J(\varphi)_t$ to lightweight notation. It suffices to show that

$$\mathbb{E}\left(\mathbf{1}_A\left(X_t^2 - X_s^2 - \int_s^t |\varphi_u|^2 \mathrm{d}u\right)\right) = 0.$$

But since $X$ is a square integrable martingale, we have

$$\mathbb{E}(\mathbf{1}_A(X_t - X_s)^2) = \mathbb{E}(\mathbf{1}_A(X_t^2 - 2X_tX_s + X_s^2)) = \mathbb{E}(\mathbf{1}_A(X_t^2 - X_s^2)).$$

Let us write

$$\varphi_t = U_0\mathbf{1}_0(t) + \sum_i U_i \mathbf{1}_{(t_i,t_{i+1}]}(t)$$

and, denoting $\Delta_{u,v} = B_v - B_u$,

$$X_t = U_0 \cdot B_{t_1} + U_1 \cdot \Delta_{t_2,t_1}) + \cdots + U_k \cdot \Delta_{t,t_k} \text{ if } t \in (t_k, t_{k+1}].$$

Since $s < t$, there exists $\ell \leq k$ such that $s \in [t_\ell, t_{\ell+1}]$, and thus

$$\mathbf{1}_A(X_t - X_s) = \widetilde{U}_\ell \cdot (B_{t_{\ell+1}} - B_s) + \widetilde{U}_{\ell+1}(B_{t_{\ell+2}} - B_{t_{\ell+1}}) + \cdots + \widetilde{U}_k \cdot (B_t - B_{t_k})$$

where $\widetilde{U}_i = \mathbf{1}_A U_i$ for all $\ell \leq i \leq k$. Let $i, i' \in \{\ell, \ell+1, \ldots, k\}$. If $i < i'$ then

$$\mathbb{E}((\widetilde{U}_i \cdot \Delta_{t_{i+1},t_i})(\widetilde{U}_{i'} \cdot \Delta_{t_{i'+1},t_{i'}})) = \mathbb{E}((\widetilde{U}_i \cdot \Delta_{t_{i+1},t_i})(\widetilde{U}_{i'} \cdot \Delta_{t_{i'+1},t_{i'}}) \mid \mathcal{F}_{i'}) = 0,$$

while if $i = i'$ then

$$\mathbb{E}((\widetilde{U}_i \cdot \Delta_{t_{i+1},t_i})^2) = \mathbb{E}(|\widetilde{U}_i|^2)(t_{i+1} - t_i)$$

where we used the fact that $\mathbb{E}(\Delta^i_{t_{i+1},t_i} \cdot \Delta^j_{t_{i+1},t_i}) = (t_{i+1} - t_i)\mathbf{1}_{i=j}$.

Now it remains to write

$$\begin{aligned}
\mathbb{E}(\mathbf{1}_A(X_t - X_s)^2) &= \sum_{i,i'=\ell}^k \mathbb{E}((\widetilde{U}_i \cdot \Delta_{t_{i+1},t_i})(\widetilde{U}_{i'} \cdot \Delta_{t_{i'+1},t_{i'}})) \\
&= \sum_{i=\ell}^k \mathbb{E}(|\widetilde{U}_i|^2)(t_{i+1} - t_i) \\
&= \mathbb{E}\left(\mathbf{1}_A \sum_{i=\ell}^k |U_\ell|^2 (t_{i+1} - t_i)\right) \\
&= \mathbb{E}\left(\mathbf{1}_A \int_s^t |\varphi_u|^2 \mathrm{d}u\right).
\end{aligned}$$

■

### 3.1.2 Extension to $\Lambda^2$ processes

Let $\Lambda^2$ be the of processes $\varphi = (\varphi_t)_{t\geq 0}$ on $\mathbb{R}^d$, adapted, mesurable, such that for all $t \geq 0$,

$$\mathbb{E}\int_0^t |\varphi_s|^2 \mathrm{d}s < \infty.$$

**Lemma 3.2** (Approximation). *For all $\varphi \in \Lambda^2$, $\varepsilon > 0$ and $t > 0$, there exists $\psi \in \mathcal{C}$ such that*

$$\mathbb{E}\int_0^t |\varphi_s - \psi_s|^2 \mathrm{d}s < \varepsilon.$$

*Proof.* We can assume without loss of generality that $\varphi$ is bounded. Indeed, if not then taking $\varphi_s^{(n)} = \varphi_s \mathbf{1}_{[-n,n]}(\varphi_s)$, $n \geq 0$, the Lebesgue dominated convergence theorem gives

$$\lim_{n\to\infty} \mathbb{E}\int_0^t |\varphi_s - \varphi_s^{(n)}|^2 \mathrm{d}s = 0.$$

From now we assume that $\varphi$ is bounded. For all $s \geq 0$ and $k \geq 1$, set $\alpha_k(s) = \ell/2^k$ if $s \in (\ell/2^k, (\ell+1)/2^k]$ and $\widetilde{\varphi}_s = \varphi_s \mathbf{1}_{0\leq s\leq t}$. We have then

$$\lim_{k\to\infty} \int_{\mathbb{R}} |\widetilde{\varphi}_{\alpha_k(s)+u} - \widetilde{\varphi}_{s+u}|^2 \mathrm{d}u = 0,$$

where we have used the "mean continuity" namely the fact that if $f \in \mathrm{L}^2(\mathbb{R}, \mathrm{d}u)$ then

$$\lim_{h\to 0} \int_{\mathbb{R}} |f(u) - f(u+h)|^2 \mathrm{d}u = 0,$$

which can be proved by approximating $f$ in $\mathrm{L}^2$ by compactly supported continuous functions.

The Lebesgue dominated convergence theorem gives then

$$\lim_{k\to\infty} \mathbb{E}\int_{-t}^t \left(\int_{\mathbb{R}} |\widetilde{\varphi}_{\alpha_k(s)+u} - \widetilde{\varphi}_{s+u}|^2 \mathrm{d}u\right) \mathrm{d}s = 0,$$

therefore, by the Fubini–Tonelli theorem,

$$\lim_{k\to\infty} \int_{\mathbb{R}} \mathbb{E}\int_{-t}^t \left(|\widetilde{\varphi}_{\alpha_k(s)+u} - \widetilde{\varphi}_{s+u}|^2 \mathrm{d}s\right) \mathrm{d}u = 0.$$

Thus, there exists a sub-sequence $(k_i)_i$ such that for almost all $u$,

$$\lim_{p\to\infty} \mathbb{E}\int_{-t}^t |\widetilde{\varphi}_{\alpha_{k_i}(s)+u} - \widetilde{\varphi}_{s+u}|^2 \mathrm{d}s = 0.$$

It follows that for all $\varepsilon > 0$, there exists $u \in (0, t)$ and $\ell \geq 1$ such that

$$\mathbb{E}\int_{-t}^t |\widetilde{\varphi}_{\alpha_\ell(s)+u} - \widetilde{\varphi}_{s+u}|^2 \mathrm{d}s < \varepsilon$$

in other words

$$\mathbb{E}\int_{-t+u}^{t+u} |\widetilde{\varphi}_{\alpha_\ell(s-u)+u} - \widetilde{\varphi}_s|^2 \mathrm{d}s < \varepsilon$$

and in particular

$$\mathbb{E}\int_0^t |\widetilde{\varphi}_{\alpha_\ell(s-u)+u} - \varphi_s|^2 \mathrm{d}s < \varepsilon.$$

The process $(\widetilde{\varphi}_{\alpha_\ell(s-u)+u})_{s\geq 0}$ is an adapted step process *i.e.* belongs to $\mathcal{C}$, as desired. ∎

Let $\mathcal{M}_c^2$ be the set of martingales which are real, continuous, and square integrable.

**Theorem 3.3** (Extension of the stochastic integral to $\Lambda^2$)**.** *There exists a unique linear map $I : \Lambda^2 \to \mathcal{M}_c^2$ such that the following properties hold true:*

1. *for all $\varphi \in \mathcal{C}$, we have $I(\varphi) = J(\varphi)$;*
2. *for all $\varphi \in \Lambda^2$ and all $t \geq 0$,*

$$\mathbb{E}((I(\varphi)_t)^2) = \mathbb{E}\int_0^t |\varphi_s|^2 \mathrm{d}s.$$

This leads, for all $\varphi \in \mathcal{M}_c^2$ and $t \geq 0$, to the notation

$$I(\varphi)_t = \int_0^t \varphi_s \mathrm{d}B_s.$$

*Proof.* Let us prove first uniqueness. Suppose that $I$ and $\widetilde{I}$ are two linear maps from $\Lambda^2$ to $\mathcal{M}_c^2$ satisfying the two properties of the theorem. Let $\varphi\Lambda^2$. From Lemma 3.2, for all $n \geq 1$, there exists $\psi^{(n)} \in \mathcal{C}$ such that

$$\mathbb{E}\int_0^n |\varphi_s - \psi_s^{(n)}|^2 \mathrm{d}s \leq \frac{1}{2^n}.$$

It follows that for all $t \geq 0$, in $\mathrm{L}^2$,

$$I(\varphi)_t = \lim_{n\to\infty} J(\psi_n)_t = \widetilde{I}(\varphi)_t.$$

Therefore, since $I(\varphi)$ and $\widetilde{I}(\varphi)$ are continuous, they are indistinguishable, and therefore $I = \widetilde{I}$.

Let us prove now the existence. Let $\varphi \in \Lambda^2$ and $\psi^{(n)} \in \mathcal{C}$ be as above. For all $t \geq 0$, let us set $X_t^{(n)} = J(\psi^{(n)})_t$. The martingale property of $J$ and the definition of $\psi^{(n)}$ give, for all $n \geq t$,

$$\mathbb{E}(|X_t^{(n)} - X_t^{(n+1)}|^2) = \mathbb{E}\int_0^t |\psi_s^{(n)} - \psi_s^{(n+1)}|^2 \mathrm{d}s \leq \frac{4}{2^n}.$$

Next, using the Doob maximal inequality of Theorem 1.11 we get

$$\mathbb{E}(\sup_{s\in[0,t]} |X_s^{(n)} - X_s^{(n+1)}|^2) \leq 4\mathbb{E}(|X_t^{(n)} - X_t^{(n+1)}|^2) \leq \frac{16}{2^n}.$$

Therefore,

$$\mathbb{E}\sum_{n\geq 0}\sup_{s\in[0,t]} |X_s^{(n)} - X_s^{(n+1)}| = \sum_{n\geq 0}\mathbb{E}\sup_{s\in[0,t]} |X_s^{(n)} - X_s^{(n+1)}| \leq \sum_{n\geq 0}\Big\|\sup_{s\in[0,t]} |X_s^{(n)} - X_s^{(n+1)}|\Big\|_2 < \infty$$

and thus, almost surely,

$$\sum_{n\geq 0}\sup_{s\in[0,t]} |X_s^{(n)} - X_s^{(n+1)}| < \infty.$$

It follows that the sequence $(X^{(n)})_n$ of continuous martingales converges almost surely and uniformly on every compact subset of $\mathbb{R}_+$, as $n \to \infty$, towards a continuous process $X = (X_t)_{t\geq 0}$. This process is an $(\mathcal{F}_t)_{t\geq 0}$-martingale since for all $0 \leq s < t$ and all $A \in \mathcal{F}_s$,

$$\mathbb{E}(\mathbf{1}_A(X_t - X_s)) = \lim_{n\to\infty} \mathbb{E}(\mathbf{1}_A(X_t^{(n)} - X_s^{(n)})) = 0.$$

The process $X$ depends only on $\varphi$ and does not depend on the particular sequence $(\psi^{(n)})_n$ chosen to construct it. We define $I(\varphi) = X$.

From the preceding estimates, it follows moreover that for all $t \geq 0$,

$$\mathbb{E}(X_t^2) = \lim_{n\to\infty} \mathbb{E}((X_t^{(n)})^2) = \lim_{n\to\infty} \mathbb{E}\int_0^t |\psi_s^{(n)}|^2 \mathrm{d}s = \mathbb{E}\int_0^t |\varphi_s|^2 \mathrm{d}s.$$

The construction of $I$ shows that $I$ is linear: for all $\varphi, \widetilde{\varphi} \in \Lambda^2$ and all $\alpha \in \mathbb{R}$,

$$I(\alpha\varphi + \widetilde{\varphi}) = \alpha I(\varphi) + I(\widetilde{\varphi}).$$

■

**Theorem 3.4** (Martingale properties of the stochastic integral $I$ on $\Lambda^2$)**.** *For all $\varphi, \psi \in \Lambda^2$ the following process is a continuous martingale, centered and square integrable:*

$$\left(\left(\int_0^t \varphi_s \mathrm{d}B_s\right)\left(\int_0^t \psi_s \mathrm{d}B_s\right) - \int_0^t (\varphi_s \cdot \psi_s)\mathrm{d}s\right)_{t\geq 0}$$

*and in particular for all bounded stopping time $T$,*

$$\mathbb{E}\left(\left(\int_0^T \varphi_s \mathrm{d}B_s\right)\left(\int_0^T \psi_s \mathrm{d}B_s\right)\right) = \mathbb{E}\int_0^T \varphi_s\psi_s \mathrm{d}s \quad \textit{and} \quad \mathbb{E}\left(\left(\int_0^T \varphi_s \mathrm{d}s\right)^2\right) = \mathbb{E}\int_0^T |\varphi_s|^2 \mathrm{d}s.$$

*Furthermore, for all $\varphi \in \Lambda^2$ and stopping time $T$, we have $(\varphi_t \mathbf{1}_{t\leq T})_{t\geq 0} \in \Lambda^2$ and*

$$\int_0^\bullet \varphi_s \mathbf{1}_{s\leq T} \mathrm{d}s = \int_0^{\bullet\wedge T} \varphi_s \mathrm{d}B_s.$$

*Proof.* The first statement can be obtain by a passage to the limit by considering first the case where $\varphi$ and $\widetilde{\varphi}$ are step processes. The second statement is an application of the Doob optional stopping theorem (Theorem 1.12). For the third and last statement, we set $X_t = \int_0^t \varphi_s \mathrm{d}B_s$ and $Y_t = \int_0^t \varphi_s \mathbf{1}_{s\leq T} \mathrm{d}B_s$. Thanks to the first property, for all $t \geq 0$,

$$\begin{aligned}
\mathbb{E}((Y_t - X_{t\wedge T})^2) &= \mathbb{E}(Y_t^2) - 2\mathbb{E}(Y_t X_{t\wedge T}) + \mathbb{E}(X_{t\wedge T}^2) \\
&= \mathbb{E}(Y_t^2) - 2\mathbb{E}(Y_{t\wedge T} X_{t\wedge T}) + \mathbb{E}(X_{t\wedge T}^2) \\
&= \mathbb{E}\int_0^t |\varphi_s|^2 \mathbf{1}_{s\leq T} \mathrm{d}s - 2\mathbb{E}\int_0^{t\wedge T} |\varphi_s|^2 \mathrm{d}s + \mathbb{E}\int_0^{t\wedge T} |\varphi_s|^2 \mathrm{d}s \\
&= 0
\end{aligned}$$

which gives $Y_t = X_{t\wedge T}$ (as random variables, in other words almost surely). ■

### 3.1.3 Extension to $\Lambda^0$ processes

Let $\Lambda^0$ be the set of $d$-dimensional adapted and progressively measurable processes $\varphi = (\varphi_t)_{t\geq 0}$ such that for all $t \geq 0$, almost surely

$$\int_0^t |\varphi_s|^2 \mathrm{d}s < \infty.$$

Let us extend to the elements of $\Lambda^0$ the stochastic integral $I$ that we have already defined on $\Lambda^2$, which is itself and extension of the stochastic integral $J$ on $\mathcal{C}$.

**Theorem 3.5** (Extension of the stochastic integral to $\Lambda^0$)**.** *There exists a unique linear map $\widetilde{I} : \Lambda_0 \to \mathcal{M}_{\mathrm{c,loc}}$ such that for all $\varphi \in \Lambda^0$ and for all stopping time $T$ such that*

$$\mathbb{E}\int_0^T |\varphi_s|^2 \mathrm{d}s < \infty,$$

*we have, for all $t \geq 0$,*

$$\widetilde{I}(\varphi)_{t\wedge T} = \int_0^t \varphi_s \mathbf{1}_{0\leq s\leq T} \mathrm{d}B_s$$

This leads, for all $\varphi \in \Lambda^0$ and $t \geq 0$, to the notation

$$\widetilde{I}(\varphi)_t = \int_0^t \varphi_s \mathrm{d}B_s.$$

*Proof.* Let $\varphi \in \Lambda^0$. For all $n \geq 0$, we define

$$T_n = \inf\left\{t \geq 0 : \int_0^t |\varphi_s|^2 \mathrm{d}s \geq n\right\}$$

with as usual the convention $\inf \varnothing = +\infty$. We can check easily that $(T_n)_{n\geq 0}$ is a sequence of stopping times such that $T_n \nearrow +\infty$ almost surely. Let us define $\varphi_s^{(n)} = \varphi_s \mathbf{1}_{0\leq s\leq T_n}$. We see that $\varphi^{(n)} \in \Lambda^2$, and thus for all $t \geq 0$ the following stochastic integral is well defined:

$$X_t^{(n)} = \int_0^t \varphi_s^{(n)} \mathrm{d}B_s = \int_0^t \varphi_s \mathbf{1}_{0\leq s\leq T_n} \mathrm{d}B_s.$$

Moreover, the last property of Theorem 3.4 gives, for all $0 \leq k \leq m$, almost surely, for all $t \geq 0$,

$$X_{t\wedge T_m}^{(k)} = X_t^{(m)}.$$

Now since $T_m \to +\infty$ almost surely, the sequence $(X_t^{(m)})_{m\geq 0}$ is almost surely stationary for $m$ sufficiently large. Now we define the process $X = (X_t)_{t\geq 0}$ outside a negligible event by $X_t = \lim_{m\to\infty} X_t^{(m)}$, and therefore, almost surely, $X_{t\wedge T_m} = X_t^{(m)}$ for all $m \geq 0$ and $t \geq 0$.

Let $T$ be a stopping time such that $\mathbb{E}\int_0^T |\varphi_s|^2 \mathrm{d}s < \infty$. Let us denote $Y_t = \int_0^t \varphi_s \mathbf{1}_{s\leq T} \mathrm{d}B_s$. Thanks to the last property of Theorem 3.4, we have, for all $m \geq 0$,

$$Y_{t\wedge T_m} = \int_0^t \varphi_s \mathbf{1}_{s\leq T\wedge T_m} \mathrm{d}B_s = X_{t\wedge T}^{(m)}$$

and thus

$$Y_t = \lim_{m\to\infty} Y_{t\wedge T_m} = \lim_{m\to\infty} X_{t\wedge T}^{(m)} = X_{t\wedge T}.$$

By defining $\widetilde{I}(\varphi)_t = X_t$ for all $t \geq 0$, we obtain a map from $\Lambda^0$ to $\mathcal{M}_{c,\mathrm{loc}}$ with the desired property. The uniqueness and the linearity are obvious. ∎

## 3.2 Stochastic integral with respect to continuous martingales

### 3.2.1 Continuous local martingales

**Lemma 3.6** (Filtrations). *For all optional process $M = (M_t)_{t\geq 0}$ defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and all stopping time $T$, the following properties are equivalent:*

1. *$(M_{t\wedge T})_{t\geq 0}$ is an $(\mathcal{F}_t)_{t\geq 0}$-martingale;*
2. *$(M_{t\wedge T})_{t\geq 0}$ is an $(\mathcal{F}_{t\wedge T})_{t\geq 0}$-martingale.*

*Proof.* Proof of $\Rightarrow$. For all $0 \leq s < t$ we have

$$M_{s\wedge T} = \mathbb{E}(M_{t\wedge T} \mid \mathcal{F}_s) = \mathbb{E}(\mathbb{E}(M_{t\wedge T} \mid \mathcal{F}_{s\wedge T}) \mid \mathcal{F}_s) = \mathbb{E}(M_{t\wedge T} \mid \mathcal{F}_{s\wedge T}).$$

Proof of $\Leftarrow$. For all $0 \leq s < t$ and $A \in \mathcal{F}_s$, we have, using the fact that $A \cap \{T > s\} \in \mathcal{F}_{s\wedge T}$,

$$\begin{aligned}
\mathbb{E}(M_{t\wedge T}\mathbf{1}_A) &= \mathbb{E}(M_{t\wedge T}\mathbf{1}_{A\cap\{T\leq s\}}) + \mathbb{E}(M_{t\wedge T}\mathbf{1}_{A\cap\{T>s\}}) \\
&= \mathbb{E}(M_T\mathbf{1}_{A\cap\{T\leq s\}}) + \mathbb{E}(\mathbb{E}(M_{t\wedge T} \mid \mathcal{F}_{s\wedge T})\mathbf{1}_{A\cap\{T>s\}}) \\
&= \mathbb{E}(M_{s\wedge T}\mathbf{1}_{A\cap\{T\leq s\}}) + \mathbb{E}(M_{s\wedge T}\mathbf{1}_{A\cap\{T>s\}}) \\
&= \mathbb{E}(M_{s\wedge T}\mathbf{1}_A).
\end{aligned}$$

∎

**Definition 3.7** (Continuous local martingale). *A process $M = (M_t)_{t\geq 0}$ is a continuous local martingale if it is real, continuous, $(\mathcal{F}_t)_{t\geq 0}$-adapted, and such that there exists a sequence $(T_n)_{n\geq 0}$ of stopping times such that $\lim_{n\to\infty} T_n = +\infty$ almost surely and for all $n \geq 0$ the stopped process $(M_{t\wedge T_n})_{t\geq 0}$ is a continuous and square integrable $(\mathcal{F}_t)_{t\geq 0}$-martingale. We denote by $\mathcal{M}_{c,\mathrm{loc}}$ the set of continuous local martingales.*

Lemma 3.6 implies that $\mathcal{M}_{c,\mathrm{loc}}$ is a real vector space.

**Remark 3.8** (Localization and boundedness). *If $M \in \mathcal{M}_{c,\mathrm{loc}}$ then we can always chose a sequence $(\widetilde{T}_n)_{n\geq 0}$ of stopping times such that for all $n \geq 0$, the random variable $M_{t\wedge\widetilde{T}_n}$ is bounded. Indeed, taking $S_n = \inf\{t \geq 0 : |M_t| \geq n\}$ then $S_n \nearrow +\infty$ almost surely and $|M_{t\wedge S_n}| \leq n$ for all $t \geq 0$ and $n \geq 0$ since $M$ is continuous. Now if $(T_n)_{n\geq 0}$ is the sequence of stopping times associated by definition to $M$ then we can take $\widetilde{T}_n = S_n \wedge T_n$, which satisfies $\widetilde{T}_n \nearrow +\infty$ almost surely and $(M_{t\wedge\widetilde{T}_n})_{n\geq 0}$ is a continuous $(\mathcal{F}_t)_{\geq 0}$-martingale and $\sup_{t\geq 0} |M_{t\wedge\widetilde{T}_n}| \leq n$ for all $n \geq 0$.*

### 3.2.2 Definition of the stochastic integral on $\Lambda^2(M)$

Let $M \in \mathcal{M}_{c,\mathrm{loc}}$ and let $\langle M\rangle = (\langle M\rangle_t)_{t\geq 0}$ be the increasing process associated to $M$ provided by the Doob–Meyer decomposition of Theorem 1.15. Let $\Lambda^2(M)$ be the set of real predictable processes $(\varphi_t)_{t\geq 0}$ such that for all $t \geq 0$,

$$\mathbb{E}\int_0^t \varphi_s^2 \mathrm{d}\langle M\rangle_s < \infty.$$

Note that the set $\mathcal{C}$ of step functions considered in Section 3.1.1 is included in $\Lambda^2(M)$.

Note that we integrate now with respect to a real process, because we would like to avoid the subtleties of vector valued martingales. Also we integrate real valued test processes.

**Lemma 3.9** (Approximation). *For all $\varphi \in \Lambda^2(M)$, $t > 0$, $\varepsilon > 0$, there exists $\psi \in \mathcal{C}$ such that*

$$\mathbb{E}\int_0^t |\varphi_s - \psi_s|^2 \mathrm{d}\langle M\rangle_s < \varepsilon.$$

The proof can be skipped in a first reading.

*Proof.* Let $t > 0$. We consider the vector space $\mathcal{L}$ formed by all $\varphi \in \Lambda^2(M)$ such that there exists a sequence $(\psi^{(n)})_{n\geq 0}$ in $\mathcal{C}$ such that

$$\lim_{n\to\infty} \mathbb{E}\int_0^t |\varphi_s - \psi_s^{(n)}|^2 \mathrm{d}\langle M\rangle_s = 0.$$

Let $(\varphi_n)_{n\geq 0}$ be a sequence in $\mathcal{L}$ such that $0 \leq \varphi_n \leq \varphi_{n+1}$ for all $n \geq 0$ and $\varphi_n \nearrow \varphi$ where $\varphi$ is bounded. The Lebesgue monotone convergence theorem implies that $\varphi \in \mathcal{C}$. Moreover, if $(\varphi_t)_{t\geq 0}$ is a left continuous bounded $(\mathcal{F}_t)_{t\geq 0}$-adapted process then

$$\lim_{n\to\infty} \int_0^t |\varphi_s - \psi_s^{(n)}|^2 \mathrm{d}\langle M\rangle_s = 0$$

where

$$\psi_s^{(n)} = \varphi_0 \mathbf{1}_{\{0\}}(s) + \sum_{k\geq 0} \varphi_{k/2^n} \mathbf{1}_{(k/2^n,(k+1)/2^n]}(s).$$

Therefore $\varphi \in \mathcal{L}$. The monotone class theorem (Corollary 1.21) implies that $\mathcal{L}$ contains all bounded $(\mathcal{F}_t)_{t\geq 0}$-predictable processes. But if $\varphi \in \Lambda^2(M)$ then

$$\lim_{n\to\infty} \mathbb{E}\int_0^t |\varphi_s - \varphi_s^{(n)}|^2 \mathrm{d}\langle M\rangle_s = 0$$

where $(\varphi_t^{(n)})_{t\geq 0} = (\varphi_t \mathbf{1}_{|\varphi_t|\leq n})_{t\geq 0}$ is a bounded $(\mathcal{F}_t)_{t\geq 0}$-predictable process, thus $\mathcal{L} = \Lambda^2(M)$. ■

If $\varphi$ is a step process defined by $\varphi_t = U_0 \mathbf{1}_0(t) + \sum_i U_i \mathbf{1}_{(t_i,t_{i+1}]}(t)$ with $0 = t_0 < t_1 < \cdots < t_n < \cdots \to \infty$ and $U_i$ is an $\mathcal{F}_i$-measurable bounded random variable for all $i$, then we define the stochastic integral of $\varphi$ with respect to $M$ by, for all $t \geq 0$,

$$J(\varphi)_t = \int_0^t \varphi_s \mathrm{d}M_s = U_0 \cdot (M_{t_1} - M_{t_0}) + U_1 \cdot (M_{t_2} - M_{t_1}) + \cdots + U_k \cdot (M_t - M_{t_k})$$

if $t \in (t_k, t_{k+1}]$, for all $k$. The martingale $M$ being continuous and $(\mathcal{F}_t)_{t\geq 0}$-adapted, it follows that that $J(\varphi)$ is a continuous and $(\mathcal{F}_t)_{t\geq 0}$-adapted process.

**Theorem 3.10** (Properties of the stochastic integral on $\mathcal{C}$). *For all $\varphi \in \mathcal{C}$, the following process is a continuous and square integrable $(\mathcal{F}_t)_{t\geq 0}$-martingale issued from the origin:*

$$J(\varphi) = \left(\int_0^t \varphi_s \mathrm{d}M_s\right)_{t\geq 0}.$$

*For all $\varphi, \psi \in \mathcal{C}$, the following process is a continuous centered $(\mathcal{F}_t)_{t\geq 0}$-martingale:*

$$\left(\left(\int_0^t \varphi_s \mathrm{d}M_s\right)\left(\int_0^t \psi_s \mathrm{d}M_s\right) - \int_0^t (\varphi_s\psi_s)\mathrm{d}\langle M\rangle_s\right)_{t\geq 0}.$$

*In particular, the conservation of mean gives the isometry: for all $t \geq 0$*

$$\mathbb{E}\left(\left(\int_0^t \varphi_s \mathrm{d}M_s\right)^2\right) = \mathbb{E}\int_0^t |\varphi_s|^2 \mathrm{d}\langle M\rangle_s.$$

*Proof.* Similar to the proof that we gave when $M$ is a Brownian motion. Left as an exercise. ■

**Theorem 3.11** (Extension of the stochastic integral to $\Lambda^2(M)$). *There exists a unique linear map $I$ from $\Lambda^2(M)$ to $\mathcal{M}_{\mathrm{c,loc}}$ such that:*

1. *For all $\varphi \in \mathcal{C}$, $I(\varphi) = J(\varphi)$;*
2. *For all $\varphi \in \Lambda^2(M)$ and all $t \geq 0$,*

$$\mathbb{E}((I(\varphi)_t)^2) = \mathbb{E}\int_0^t \varphi_s^2 \mathrm{d}\langle M\rangle_s.$$

This leads to denote for all $\varphi \in \Lambda^2(M)$ and all $t \geq 0$

$$I(\varphi)_t = \int_0^t \varphi_s \mathrm{d}M_s.$$

*Proof.* Similar to the proof that we gave when $M$ is a Brownian motion. Left as an exercise. ■

**Theorem 3.12** (Properties of the stochastic integral). *For all $\varphi$ and $\psi$ in $\Lambda^2(M)$, the following process is a continuous $(\mathcal{F}_t)_{t\geq 0}$-martingale, centered and issued from the origin:*

$$\left(\left(\int_0^t \varphi_s \mathrm{d}M_s\right)\left(\int_0^t \psi_s \mathrm{d}M_s\right) - \int_0^t \varphi_s\psi_s \mathrm{d}\langle M\rangle_s\right)_{t\geq 0}.$$

*In particular the following process is a martingale*

$$\left(\left(\int_0^t \varphi_s \mathrm{d}M_s\right)^2 - \int_0^t \varphi_s^2 \mathrm{d}\langle M\rangle_s\right)_{t\geq 0}$$

*in other words*

$$\left\langle\int_0^\bullet \varphi_s \mathrm{d}M_s\right\rangle = \int_0^\bullet \varphi_s^2 \mathrm{d}\langle M\rangle_s.$$

*and in particular for all bounded stopping time $T$,*

$$\mathbb{E}\left(\left(\int_0^T \varphi_s \mathrm{d}B_s\right)\left(\int_0^T \psi_s \mathrm{d}B_s\right)\right) = \mathbb{E}\int_0^T \varphi_s\psi_s \mathrm{d}\langle M\rangle_s$$

*and in particular*

$$\mathbb{E}\left(\left(\int_0^T \varphi_s \mathrm{d}B_s\right)^2\right) = \mathbb{E}\int_0^T \varphi_s^2 \mathrm{d}\langle M\rangle_s$$

*Furthermore, for all $\varphi \in \Lambda^2(M)$ and stopping time $T$, we have $(\varphi_t \mathbf{1}_{t\leq T})_{t\geq 0} \in \Lambda^2(M)$ and*

$$\int_0^\bullet \varphi_s \mathbf{1}_{s\leq T} \mathrm{d}M_s = \int_0^{\bullet\wedge T} \varphi_s \mathrm{d}M_s.$$

*Proof.* Similar to the proof that we gave when $M$ is a Brownian motion. Left as an exercise. ■

**Lemma 3.13** (Cauchy–Schwarz like inequality for stochastic integrals). *For all square integrable continuous martingales $M$ and $N$ and all $\varphi \in \Lambda^2(M)$ and $\psi \in \Lambda^2(N)$ and all $t \geq 0$,*

$$\int_0^t |\varphi_s\psi_s| \mathrm{d}|\langle M,N\rangle|_s \leq \left(\int_0^t \varphi_s^2 \mathrm{d}\langle M\rangle_s\right)^{1/2}\left(\int_0^t \psi_s^2 \mathrm{d}\langle N\rangle_s\right)^{1/2}.$$

*Proof.* Let us consider $0 \leq u < v$ and let us define

$$\Delta\langle M,N\rangle = \langle M,N\rangle_v - \langle M,N\rangle_u, \quad \text{and} \quad \Delta\langle M\rangle = \Delta\langle M,M\rangle.$$

Then for all $\alpha \in \mathbb{R}$,

$$\Delta\langle M+\alpha N, M+\alpha N\rangle = \Delta\langle M\rangle + 2\alpha\langle M,N\rangle + \alpha^2\Delta\langle N\rangle \geq 0$$

which gives

$$|\Delta\langle M,N\rangle|^2 \leq (\Delta\langle M\rangle)(\Delta\langle N\rangle).$$

The desired inequality, when $\varphi$ and $\psi$ are step functions, follows via the Cauchy–Schwarz inequality and by considering sub-divisions of $[0,t]$. The general case is obtained by approximation using Lemma 3.9. ■

**Theorem 3.14.** *For all square integrable continuous martingales $M, N$ and all $\varphi \in \Lambda^2(M)$ and $N \in \Lambda^2(N)$, the process*

$$\left(\left(\int_0^t \varphi_s \mathrm{d}M_s\right)\left(\int_0^t \psi_s \mathrm{d}M_s\right) - \int_0^t \varphi_s\psi_s \mathrm{d}\langle M,N\rangle_s\right)_{t\geq 0}$$

*is a continuous and centered martingale issued from the origin. In particular*

$$\langle\int_0^\bullet \varphi_s \mathrm{d}M_s, \int_0^\bullet \psi_s \mathrm{d}N_s\rangle = \int_0^\bullet \varphi_s\psi_s \mathrm{d}\langle M,N\rangle_s$$

*or*

$$\mathrm{d}\langle\int_0^\bullet \varphi_s \mathrm{d}M_s, \int_0^\bullet \psi_s \mathrm{d}N_s\rangle_\bullet = \varphi_\bullet\psi_\bullet \mathrm{d}\langle M,N\rangle_\bullet.$$

*Proof.* Following by approximation from the case where $\varphi$ and $\psi$ are step functions, by using Lemma 3.9 and Lemma 3.13. ■

### 3.2.3 Extension of the stochastic integral to $\Lambda^2_{\text{loc}}$

Let $M, N \in \mathcal{M}_{c,\text{loc}}$. We can find a sequence $(T_n)_{n\geq 0}$ of stopping times such that $T_n \nearrow +\infty$ as $n \to \infty$ and for all $n \geq 0$, $M^{T_n} = (M_{t\wedge T_n})_{t\geq 0}$ and $N^{T_n} = (N_{t\wedge T_n})_{t\geq 0}$ are square integrable continuous martingales. A consequence of the uniqueness of the Doob–Meyer decomposition of Theorem 1.15 is that for all $0 \leq n \leq m$ and all $t \geq 0$,

$$\langle M^{T_m}, N^{T_m}\rangle_{t\wedge T_n} = \langle M^{T_n}, N^{T_n}\rangle_t.$$

It follows that there exists a unique continuous process with finite variations and issued from the origin, denoted $\langle M, N\rangle$ such that for all $t \geq 0$ and all $n \geq 0$,

$$\langle M, N\rangle_{t\wedge T_n} = \langle M^{T_n}, N^{T_n}\rangle.$$

We set

$$\langle M\rangle = \langle M, M\rangle.$$

Let $\Lambda^2_{\text{loc}}(M)$ be the set of predictable real processes $\varphi = (\varphi_t)_{t\geq 0}$ such that for all $t \geq 0$,

$$\int_0^t \varphi_s^2 \mathrm{d}\langle M\rangle_s < \infty.$$

If $M \in \mathcal{M}_{c,\text{loc}}$ and $\varphi \in \Lambda^2_{\text{loc}}(M)$, we can find a sequence $(T_n)_{n\geq 0}$ of stopping times such that $T_n \nearrow +\infty$ and for all $n \geq 0$, $M^{T_n} = (M_{t\wedge T_n})_{t\geq 0}$ is a square integrable continuous martingale and for all $n \geq 0$ and $t \geq 0$

$$\mathbb{E}\int_0^{t\wedge T_n} \varphi_s^2 \mathrm{d}\langle M\rangle_s < \infty.$$

For all $n \geq 0$ and $t \geq 0$, the stochastic integral

$$I_n(\varphi)_t = \int_0^t \varphi_s \mathbf{1}_{s\leq T_n} \mathrm{d}M_s^{T_n}$$

is then well defined and we can easily check that, for all $0 \leq n \leq m$ and all $t \geq 0$,

$$I_m(\varphi)_{t\wedge T_n} = I_n(\varphi)_t.$$

Therefore there exists a unique local martingale $(I(\varphi)_t)_{t\geq 0}$ such that for all $n \geq 0$ and $t \geq 0$,

$$I(\varphi)_{t\wedge T_n} = I_n(\varphi)_t.$$

For all $t \geq 0$, we denote

$$\int_0^t \varphi_s \mathrm{d}\langle M\rangle_s = I(\varphi)_t.$$

This is the stochastic integral of $\varphi \in \Lambda^2_{\text{loc}}(M)$ with respect to $\mathcal{M}_{c,\text{loc}}$.

**Remark 3.15** (Properties of the stochastic integral)**.** *The following properties hold true.*

1. ***Linearity.*** *For all $M \in \mathcal{M}_{c,\text{loc}}$, all $\varphi, \psi \in \Lambda^2_{\text{loc}}(M)$, and all $t \geq 0$,*

$$\int_0^t (\varphi + \psi)_s \mathrm{d}M_s = \int_0^t \varphi_s \mathrm{d}M_s + \int_0^t \psi_s \mathrm{d}M_s.$$

2. ***Linearity.*** *For all $M, N \in \mathrm{M}_{c,\text{loc}}$, $\varphi \in \Lambda^2_{\text{loc}}(M) \cap \Lambda^2_{\text{loc}}(N)$, and all $t \geq 0$,*

$$\int_0^t \varphi_s \mathrm{d}(M + N)_s = \int_0^t \varphi_s \mathrm{d}M_s + \int_0^t \varphi_s \mathrm{d}N_s.$$

3. ***Substitution.*** *For all $M \in \mathcal{M}_{c,\text{loc}}$ and all $\varphi \in \Lambda^2_{\text{loc}}(M)$, if one defines*

$$N = \left(\int_0^t \varphi_s \mathrm{d}M_s\right)_{t\geq 0}$$

*then for all $\psi \in \Lambda^2_{\text{loc}}(N)$, we have $\varphi\psi \in \Lambda^2_{\text{loc}}(M)$ and for all $t \geq 0$*

$$\int_0^t \varphi_s \psi_s \mathrm{d}M_s = \int_0^t \psi_s \mathrm{d}N_s.$$

# Chapter 4

# Itô formula and applications

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space, with $(\mathcal{F}_t)_{t\geq 0}$ complete and right continuous.

## 4.1 Quadratic variation and semi-martingales

Recall that the notion of quadratic variation of a process is considered in Section 1.2.3. The quadratic variation of Brownian motion is considered in Theorem 2.13. For a Brownian motion $B$, Theorem 2.13 already gives $[B]_t = \langle B\rangle_t = t$ for all $t \geq 0$.

**Theorem 4.1** (Quadratic variation of local martingales)**.** *Let $M = (M_t)_{t\geq 0}$ and $N = (N_t)_{t\geq 0}$ be continuous local martingales issued from the origin. For all $t > 0$ and all partition or subdivision $\delta : 0 = t_0 < t_1 < \cdots < t_n = s$ of $[0, t]$, we define $S(\delta) = \sum_{i=0}^{n-1}(M_{t_{i+1}} - M_{t_i})(N_{t_{i+1}} - N_{t_i})$. Then, denoting $|\delta| = \max_{0\leq i\leq n-1}(t_{i+1} - t_i)$, we have, for all $t > 0$,*

$$[M, N]_t = \lim_{|\delta|\to 0}^{\mathbb{P}} S(\delta) = \langle M, N\rangle_t.$$

*In particular $[M] = [M, M] = \langle M, M\rangle = \langle M\rangle$. In other words, for continuous local martingales, the quadratic variation matches the increasing process given by the Doob–Meyer decomposition.*

*Proof.* We can assume without loss of generality that $M = N$. Suppose first that for all $t \geq 0$ we have $|M_t| + \langle M\rangle_t \leq C$ almost surely where $C$ is a fixed constant. We have

$$\begin{aligned}
\mathbb{E}(S(\delta)) &= \sum_i \mathbb{E}((M_{t_{i+1}} - M_{t_i})^2) \\
&= \sum_i \mathbb{E}(M_{t_{i+1}}^2 - M_{t_i}^2) \\
&= \sum_i \mathbb{E}(\langle M_{t_{i+1}}\rangle - \langle M_{t_i}\rangle) \\
&= \mathbb{E}(\langle M\rangle_t) \\
&\leq C.
\end{aligned}$$

Moreover, denoting $\Delta_i = M_{t_{i+1}} - M_{t_i}$ and $\langle\Delta\rangle_i = \langle M\rangle_{t_{i+1}} - \langle M\rangle_{t_i}$,

$$\begin{aligned}
\mathbb{E}((S(\delta) - \langle M\rangle_t)^2) &= \mathbb{E}\Big(\Big(\sum_i(\Delta_i^2 - \langle\Delta\rangle_i)\Big)^2\Big) \\
&= \mathbb{E}\sum_{i,j}(\Delta_i^2 - \langle\Delta\rangle_i)(\Delta_j^2 - \langle\Delta\rangle_j) \\
&\leq 2\sum_i \mathbb{E}(\Delta_i^4) + \mathbb{E}(\langle\Delta\rangle_i^2) \\
&\overset{\star}{\leq} 2\mathbb{E}(S(\delta)\max_i \Delta_i^2) + 2\mathbb{E}(\langle M\rangle_t \max_i\langle\Delta\rangle_i) \\
&\leq 2 \times 4 \times C \times C^2 + 2C^2.
\end{aligned}$$

Therefore

$$\sup_\delta(\mathbb{E}(S(\delta)^2))^{1/2} \le (8C^3+2C^2)^{1/2}+C<\infty.$$

Back to the $\star$ inequality above, we have

$$\mathbb{E}((S(\delta)-\langle M\rangle_t)^2) \le 2\sup_\delta \|S(\delta)\|_2\,(\mathbb{E}(\max_i \Delta_i^4))^{1/2} + 2C\mathbb{E}\max_i\langle\Delta\rangle_i \underset{|\delta|\to 0}{\longrightarrow} 0$$

where we used the uniform continuity of $s\in[0,t]\mapsto M_s$ and $s\in[0,t]\mapsto\langle M\rangle_s$ (Heine theorem).

Let us pass to the general case by localization with stopping times. For all $n\ge 0$ we define

$$T_n=\inf\{t\ge 0: |M_t|+\langle M\rangle_t\ge n\}.$$

Then $(T_n)_{n\ge 0}$ is a sequence of stopping times such that $T_n\le T_{n+1}$ for all $n\ge 0$ and $T_n\nearrow+\infty$ almost surely. We know that $\langle M^{T_n}\rangle_t=\langle M\rangle_{t\wedge T_n}$ for all $n\ge 0$ and all $t\ge 0$, hence

$$\sup_t(|M_t^{T_n}|+\langle M^{T_n}\rangle_t)\le n.$$

From the first part of the proof, we have

$$S^{T_n}(\delta)=\sum_i(M^{T_n}_{t_{i+1}}-M^{T_n}_{t_i})\underset{|\delta|\to 0}{\overset{\mathrm{L}^2}{\longrightarrow}}\langle M\rangle_t.$$

Now for all $\varepsilon>0$ and all $n\ge 0$,

$$\begin{aligned}\mathbb{P}(|S(\delta)-\langle M\rangle_t|>\varepsilon)&\le\mathbb{P}(T_n\le t)+\mathbb{P}(|S(\delta)-\langle M\rangle_t|>\varepsilon, t<T_n)\\&\le\mathbb{P}(T_n\le t)+\mathbb{P}(|S^{T_n}(\delta)-\langle M\rangle_t|>\varepsilon).\end{aligned}$$

This gives easily that $\lim_{|\delta|\to 0}\mathbb{P}(|S(\delta)-\langle M\rangle_t|>\varepsilon)=0$. ■

**Definition 4.2** (Semi-martingales). *A real $(\mathcal{F}_t)_{t\ge0}$-adapted process $X=(X_t)_{t\ge0}$ is a continuous semi-martingale when it admits a decomposition of the form*

$$X=X_0+M+V$$

*where $M$ and $V$ are $(\mathcal{F}_t)_{t\ge0}$-adapted continuous processes issued from the origin such that*

- *$M=(M_t)_{t\ge0}$ is a local martingale;*
- *$V=(V_t)_{t\ge0}$ has finite variation on every compact subset of $\mathbb{R}_+$.*

**Remark 4.3** (Uniqueness of the decomposition). *The decomposition of a semi-martingale is unique. Indeed, if $X=X_0+M+V=X_0+\widetilde{M}+\widetilde{V}$ then, with $W=\widetilde{V}-V=M-\widetilde{M}$, Theorem 4.1 gives, for all $t>0$, $n\ge0$, $\delta:0=t_0<t_1<\cdots<t_n=t$,*

$$\langle M-\widetilde{M}\rangle_t=\lim_{|\delta|\to0}\sum_k(W_{t_{k+1}}-W_{t_k})^2,$$

*and thus, by using the uniform continuity of $s\in[0,t]\mapsto W_s$ (Heine theorem) we get*

$$\langle M-\widetilde{M}\rangle_t\le\lim_{|\delta|\to0}\max_k|W_{t_{k+1}}-W_{t_k}|(|V|_t+|\widetilde{V}_t|)=0.$$

*Therefore $\langle M-\widetilde{M}\rangle=0$, which implies $M-\widetilde{M}=0$ and thus $V=\widetilde{V}$.*

## 4.2 Itô formula

Classical integral calculus has a fundamental formula expressing a regular function as the integral of its derivative. For the stochastic integral, the analogue is the Itô formula. The lack of regularity of the process used for integration produces an additional term of second order.

**Theorem 4.4** (Itô[1] formula)**.** *If $X = (X_t)_{t\geq 0}$ is a $d$-dimensional continuous process such that for all $1 \leq i \leq d$ its $i$-th coordinate $(X_t^i)_{t\geq 0}$ is a semi-martingale with decomposition $X^i = X_0^i + M^i + V^i$ then for all $f : \mathbb{R}^d \to \mathbb{R}$ of class $\mathcal{C}^2$ and all $t \geq 0$,*

$$f(X_t) = f(X_0) + \sum_{i=1}^{d} \int_0^t \frac{\partial f}{\partial x_i}(X_s)\mathrm{d}M_s^i + \sum_{i=1}^{d} \int_0^t \frac{\partial f}{\partial x_i}(X_s)\mathrm{d}V_s^i + \frac{1}{2}\sum_{i,j=1}^{n} \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s)\mathrm{d}\langle M^i, M^j\rangle_s.$$

*(equality holds as random variables, in other words almost surely).*

*Proof.* We suppose firs that there exists a constant $C > 0$ such that

$$\sup_{\substack{t\geq 0\\ 1\leq i,j\leq d}} \Big(|X_0| + |M_t^i| + |V_t^i| + |\langle M^i, M^j\rangle_t|\Big) \leq C$$

and that $f$ has compact support. The Taylor formula gives, for all $x, y \in \mathbb{R}^d$,

$$\begin{aligned} f(y) - f(x) &= \langle \nabla f(x), y - x\rangle + \frac{1}{2}\langle \mathrm{Hess}(f)(x)(y-x), y - x\rangle + r(x,y)|x-y|^2 \\ &= \sum_i \frac{\partial f}{\partial x_i}(a)(y_i - x_i) + \frac{1}{2}\sum_{i,j}\frac{\partial^2 f}{\partial x_i \partial x_j}(a)(y_i - x_i)(y_j - x_j) + r(x,y)|y-x|^2. \end{aligned}$$

Since $f$ is $\mathcal{C}^2$ with compact support, by Heine theorem, $x \mapsto f''(x) = \mathrm{Hess}(f)(x) = (\frac{\partial^2 f}{\partial x_i \partial x_j})_{1\leq i,j\leq d}$ is uniformly continuous, and therefore there exists a bounded continuous non-decreasing function $g : \mathbb{R}^d \to \mathbb{R}$ such that $\lim_{x\to 0} g(x) = 0$ and $|r(x,y)| \leq g(|x-y|)$ for all $x, y \in \mathbb{R}$.

Now, for all $n \geq 0$ and $t > 0$ and all sub-division $\delta : 0 = t_0 < t_1 < \cdots < t_{n+1} = t$ of $[0,t]$, denoting $\Delta X_k = X_{t_{k+1}} - X_{t_k}$ for all $0 \leq k \leq n$, we have

$$\begin{aligned} f(X_t) - f(X_0) &= \sum_k (f(X_{t_{k+1}}) - f(X_{t_k})) \\ &= \sum_k \langle \nabla f(X_{t_i}), \Delta X_k\rangle + \frac{1}{2}\sum_k \langle \mathrm{Hess}(f)(X_{t_k})\Delta X_k, \Delta X_k\rangle + \sum_k r(X_{t_k}, X_{t_{k+1}})|\Delta X_k|^2 \\ &= S_1 + S_2 + S_3. \end{aligned}$$

Denoting $\Delta V_k = V_{t_{k+1}} - V_{t_k}$, $\Delta M_k = M_{t_{k+1}} - M_{t_k}$, and $|\delta| = \max_{0\leq k\leq n}(t_{k+1} - t_k)$, we have

$$\begin{aligned} S_1 &= \sum_k \langle \nabla f(X_{t_k}), \Delta V_k\rangle + \sum_k \langle \nabla f(X_{t_k}), \Delta M_k\rangle \\ &\xrightarrow[|\delta|\to 0]{\mathrm{L}^2} \int_0^t \nabla f(X_s)\mathrm{d}V_s + \int_0^t \nabla f(X_s)\mathrm{d}M_s \\ &= \sum_{i=1}^{d}\int_0^t \frac{\partial f}{\partial x_i}(X_s)(\mathrm{d}V_s^i + \mathrm{d}M_s^i). \end{aligned}$$

[1]Named after Kiyosi Itô (1915 – 2008), Japanese mathematician. He used the notation "Kiyosi Itô" for his name (Kunrei-shiki romanization), instead of the more standard "Kiyoshi Itō" (Hepburn romanization).

Next, denoting $\Delta X_k^i = X_{t_{k+1}}^i - X_{t_k}^i$, $\Delta M_k^i = M_{t_{k+1}}^i - M_{t_k}^i$, $\Delta V_k^i = V_{t_{k+1}}^i - V_{t_k}^i$,

$$\begin{aligned}
S_2 &= \frac{1}{2}\sum_k \langle \mathrm{Hess}(f)(X_s)\Delta X_k, \Delta X_k\rangle \\
&= \frac{1}{2}\sum_k\sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_{t_k})\Delta X_k^i \Delta X_k^j \\
&= \frac{1}{2}\sum_k\sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_{t_k})\Delta M_k^i \Delta M_k^j \\
&\quad + \sum_k\sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_{t_k})\Delta M_k^i \Delta V_k^j \\
&\quad + \frac{1}{2}\sum_k\sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_{t_k})\Delta V_k^i \Delta V_k^j \\
&= S_2' + S_2'' + S_2'''.
\end{aligned}$$

Now we have

$$|S_2''| \leq \sum_{i,j} \max_k |\Delta M_k^i| \sum_k \left|\frac{\partial^2 f}{\partial x_i \partial x_j}(X_{t_k})\right| |\Delta V_k^j| \underset{|\delta|\to 0}{\longrightarrow} \sum_{i,j} 0 \times \int_0^t \left|\frac{\partial^2 f}{\partial x_i \partial x_j}(X_s)\right| \mathrm{d}|V^j|_s = 0,$$

and similarly

$$|S_2'''| \leq \sum_{i,j} \max_k |\Delta V_k^i| \sum_k \left|\frac{\partial^2 f}{\partial x_i \partial x_j}(X_{t_k})\right| |\Delta V_k^j| \underset{|\delta|\to 0}{\longrightarrow} \sum_{i,j} 0 \times \int_0^t \left|\frac{\partial^2 f}{\partial x_i \partial x_j}(X_s)\right| \mathrm{d}|V^j|_s = 0.$$

For $S_2'$, denoting $\langle \Delta M^{i,j}\rangle_k = \langle M^i, M^j\rangle_{t_{k+1}} - \langle M^i, M^j\rangle_{t_k}$, we have, using Theorem 4.1,

$$\begin{aligned}
&\mathbb{E}\Big(\Big(\sum_k\sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_{t_k})(\Delta M_k^i \Delta M_k^j - \langle \Delta M^{i,j}\rangle_k)\Big)^2\Big) \\
&\qquad = \sum_k \mathbb{E}\Big(\Big(\sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_{t_k})(\Delta M_k^i \Delta M_k^j - \langle \Delta M^{i,j}\rangle_k)\Big)^2\Big) \\
&\qquad \leq C\mathbb{E}\Big(\sum_{i,j}\Big(\sum_k (\Delta M_k^i \Delta M_k^j - \langle \Delta^{i,j}\rangle_k)\Big)^2\Big) \underset{|\delta|\to 0}{\longrightarrow} 0.
\end{aligned}$$

It follows that in $\mathrm{L}^2$,

$$\begin{aligned}
\lim_{|\delta|\to 0} S_2' &= \lim_{|\delta|\to 0} \frac{1}{2}\sum_{i,j=1}^d \sum_{k=0}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(X_{t_k})(\langle M^i, M^j\rangle_{t_{k+1}} - \langle M^i, M^j\rangle_{t_k}) \\
&= \frac{1}{2}\sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s)\mathrm{d}\langle M^i, M^j\rangle_s.
\end{aligned}$$

Regarding $S_3$, we have, using Theorem 4.1,

$$\begin{aligned}
&\sum_{k=0}^n |r(X_{t_k}, X_{t_{k+1}})||X_{t_{k+1}} - X_{t_k}|^2 \\
&\qquad \leq 2\,\underbrace{g(\max_{0\leq k\leq n} |X_{t_{k+1}} - X_{t_k}|)}_{\to 0 \text{ as } |\delta|\to 0} \sum_{i=1}^d \Big(\underbrace{\sum_{k=0}^n (M_{t_{k+1}}^i - M_{t_k}^i)^2}_{\to \langle M^i\rangle_t \text{ in } \mathrm{L}^2} + \underbrace{\sum_{k=0}^n (V_{t_{k+1}}^i - V_{t_k}^i)^2}_{\to 0}\Big).
\end{aligned}$$

This achieves the proof under the assumptions of boundedness of $M$ and $V$ and compactness of the support of $f$. To prove the general case, we consider the sequence $(T_n)_{n\geq 0}$ of stopping times defined for all $n \geq 0$ by

$$T_n = \inf\Big\{t \geq 0 : \sum_{i,j=1}^{d} |X_0| + |M_t^i| + |V_t^i| + |\langle M^i, M^j\rangle_t| \geq n\Big\}$$

Then $T_n \nearrow +\infty$ almost surely and from the first part of the proof

$$f(X_{t\wedge T_n}) = f(X_0) + \sum_{i=1}^{d}\int_0^{t\wedge T_n} \frac{\partial f}{\partial x_i}(X_s)(\mathrm{d}M_s^i + \mathrm{d}V_s^i) + \frac{1}{2}\sum_{i,j=1}^{d}\int_0^{t\wedge T_n} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s)\mathrm{d}\langle M^i, M^j\rangle_s.$$

It suffices then to let $n \to \infty$. ■

## 4.3 Applications of the Itô formula

Let us start with two immediate consequences of the Itô formula of Theorem 4.4 is that for all continuous local martingales $M, N \in \mathcal{M}_{\mathrm{c,loc}}$ and for all continuous process $V$ issued from the origin and of bounded variation on every compact subset of $\mathbb{R}_+$, we have, for all $t \geq 0$,

$$M_tN_t = M_0N_0 + \int_0^t M_s\mathrm{d}N_s + \int_0^t N_s\mathrm{d}M_s + \langle M, N\rangle_t$$

and

$$M_tV_t = M_0V_0 + \int_0^t M_s\mathrm{d}V_s + \int_0^t V_s\mathrm{d}M_s,$$

obtained with $f(x_1, x_2) = x_1x_2$ and $X = (M, N)$ for the first and $X = (M, V)$ for the second.

### 4.3.1 Lévy characterization of Brownian motion and Dubins–Schwarz theorem

**Theorem 4.5** (Lévy[2] characterization of Brownian motion)**.** *If $M = (M_t)_{t\geq 0}$ is a $d$-dimensional $(\mathcal{F}_t)_{t\geq 0}$-adapted process such that for $1 \leq i \leq d$ its $i$-th coordinate $(M_t^i)_{t\geq 0}$ is a continuous local martingale issued from the origin and if $\langle M^i, M^j\rangle_t = \mathbf{1}_{i=j}t$ for all $1 \leq i, j \leq d$ and all $t \geq 0$, then $M$ is an $(\mathcal{F}_t)_{t\geq 0}$ Brownian motion.*

*Proof.* Let $\lambda \in \mathbb{R}^d$ and $N_t^\lambda = \mathrm{e}^{i\lambda\cdot M_t + \frac{1}{2}|\lambda|^2 t}$ and let us reserve the notation $i$ for the complex number $(0, 1)$ in this proof. Thanks to Theorem 2.8, it suffices to show that $(N_t^\lambda)_{t\geq 0}$ is an $(\mathcal{F}_t)_{t\geq 0}$-martingale. From the Itô formula of Theorem 4.4 with $f(x) = \mathrm{e}^{i\lambda\cdot x + \frac{1}{2}|x|^2 t}$, we have

$$\begin{aligned}
N_t^\lambda &= 1 + \int_0^t N_s^\lambda\Big(i\mathrm{d}(\lambda\cdot M_s) + \frac{1}{2}|\lambda|^2\mathrm{d}s\Big) + \frac{1}{2}i^2\sum_{j,k=1}^{d}\lambda_j\lambda_k\int_0^t N_s^\lambda \mathbf{1}_{j=k}\mathrm{d}s \\
&= 1 + i\int_0^t N_s^\lambda\mathrm{d}(\lambda\cdot M_s) + \frac{1}{2}|\lambda|^2\int_0^t N_s^\lambda\mathrm{d}s - \frac{1}{2}|\lambda|^2\int_0^t N_s^\lambda\mathrm{d}s \\
&= 1 + i\int_0^t N_s^\lambda\mathrm{d}(\lambda\cdot M_s).
\end{aligned}$$

The stochastic integral $\Big(\int_0^t N_s^\lambda\mathrm{d}(\lambda\cdot M_s)\Big)_{t\geq 0}$ is an $(\mathcal{F}_t)_{t\geq 0}$-martingale since for all $t \geq 0$,

$$\mathbb{E}\int_0^t |N_s^\lambda|^2\mathrm{d}\langle\lambda\cdot M, \lambda\cdot M\rangle_s = \int_0^t \mathrm{e}^{|\lambda|^2 s}|\lambda|^2\mathrm{d}s < \infty.$$

We used the criterion for being a martingale given by Remark 4.6. ■

[2] See footnote 1 page 14.

**Remark 4.6** (Martingale criterion)**.** *For all $X \in \mathcal{M}_{\mathrm{c,loc}}$, a necessary (and sufficient!) condition for $X$ to be a square integrable martingale is that $\mathbb{E}(\langle X\rangle_t) < \infty$ for all $t \geq 0$. Proof: exercise!*

**Corollary 4.7** (Dubins–Schwarz theorem)**.** *If $M = (M_t)_{t\geq 0}$ is a continuous local martingale issued from the origin, with $\langle M\rangle_\infty = \infty$, and if we define, for all $t \geq 0$, the stopping time*

$$T_t = \inf\{s \geq 0 : \langle M\rangle_s > t\},$$

*then $M_{T_t} = M_{\langle M\rangle_t}$ for all $t \geq 0$ and the process $(M_{\langle M\rangle_t})_{t\geq 0}$ is an $(\mathcal{F}_{T_t})_{t\geq 0}$-Brownian motion.*

*Proof.* FIXME: ■

**Theorem 4.8** (Martingale criterion)**.** *Let $M = (M_t)_{t\geq 0}$ and $V = (V_t)_{t\geq 0}$ be continuous adapted real processes issued from the origin, with $V$ increasing. For all $\lambda \in \mathbb{R}$ let us define the process*

$$X^\lambda = (X_t^\lambda)_{t\geq 0} = (\mathrm{e}^{\lambda M_t - \frac{\lambda^2}{2}V_t})_{t\geq 0}.$$

*Then the following properties are equivalent.*

1. *$M$ is a local martingale and $\langle M\rangle = V$;*
2. *$X^\lambda$ is a local martingale for all $\lambda \in \mathbb{R}$.*

*In this case $X^\lambda$ is a super-martingale, and is a martingale if and only if $\mathbb{E}X_t^\lambda = 1$ for all $t \geq 0$.*

*Moreover if $\mathbb{E}\int_0^t \mathrm{e}^{2\lambda M_s}\mathrm{d}V_s < \infty$ for all $t \geq 0$ then $X^\lambda$ is a martingale. Furthermore if $X^\lambda$ is a martingale for all $\lambda \in \mathbb{R}$ and if $\mathbb{E}\mathrm{e}^{\lambda M_t} < \infty$ for all $t \geq 0$ then $M$ is a martingale.*

*Proof.* Suppose that $M$ is a local martingale and that we have $\langle M\rangle = V$. The Itô formula gives

$$X_t^\lambda = 1 + \lambda \cdot \int_0^t X_s^\lambda \mathrm{d}M_s$$

for all $\lambda \in \mathbb{R}$ and all $t \geq 0$, and therefore $X^\lambda$ is a non-negative local martingale. On the other hand, the Fatou lemma gives that $X^\lambda$ is a non-negative local super-martingale: indeed since $X^\lambda$ is a local martingale, there exists a sequence of stopping times $(T_n)_{n\geq 0}$ such that $T_n \nearrow +\infty$ almost surely and such that $(X_{t\wedge T_n}^\lambda)_{t\geq 0}$ is a martingale for all $n \geq 0$ then for all $0 \leq s \leq t$,

$$X_s^\lambda = \lim_{n\to\infty} X_{s\wedge T_n}^\lambda = \lim_{n\to\infty} \mathbb{E}(X_{t\wedge T_n}^\lambda \mid \mathcal{F}_s) \geq \mathbb{E}(\varliminf_{n\to\infty} X_{t\wedge T_n}^\lambda \mid \mathcal{F}_s) = \mathbb{E}(X_t^\lambda \mid \mathcal{F}_s).$$

Also $\mathbb{E}X_t^\lambda \leq \mathbb{E}X_0^\lambda = 1$ and $X^\lambda$ is martingale if and only if $\mathbb{E}X_t^\lambda = 1$ for all $t \geq 0$.

Moreover, for all $t \geq 0$,

$$\mathbb{E}\langle X^\lambda\rangle_t = \lambda^2 \mathbb{E}\int_0^t \mathrm{e}^{2\lambda M_s - \lambda^2 V_s}\mathrm{d}V_s \leq \lambda^2 \mathbb{E}\int_0^t \mathrm{e}^{2\lambda M_s}\mathrm{d}V_s.$$

If this last term is finite, then thanks to Remark 4.6, the process $X^\lambda$ is a martingale.

Conversely, suppose first that $X^\lambda$ is a martingale and that $\mathbb{E}\mathrm{e}^{\lambda M_t} < \infty$ for all $\lambda \in \mathbb{R}$. Then for all $0 \leq s < t$ and all $A \in \mathcal{F}_s$,

$$\mathbb{E}(\mathbf{1}_A \mathrm{e}^{\lambda M_t - \frac{\lambda^2}{2}V_t}) = \mathbb{E}(\mathbf{1}_A X_t^\lambda) = \mathbb{E}(\mathbf{1}_A X_s^\lambda) = \mathbb{E}(\mathbf{1}_A \mathrm{e}^{\lambda M_s - \frac{\lambda^2}{2}V_s}).$$

Taking the derivative with respect to $\lambda$, which is allowed here by dominated convergence, gives

$$\mathbb{E}(\mathbf{1}_A \mathrm{e}^{\lambda M_t - \frac{\lambda^2}{2}V_t}(M_t - \lambda V_t)) = \mathbb{E}(\mathbf{1}_A \mathrm{e}^{\lambda M_s - \frac{\lambda^2}{2}V_s}(M_s - \lambda V_s)),$$

and additionally by taking the derivative with respect to $\lambda$ again,

$$\mathbb{E}(\mathbf{1}_A X_t^\lambda((M_t - \lambda V_t)^2 - V_t)) = \mathbb{E}(\mathbf{1}_A X_s^\lambda((M_s - \lambda V_s)^2 - V_s)).$$

Taking $\lambda = 0$ gives that $M$ and $(M_t^2 - V_t)_{t\geq 0}$ are martingales, and in particular $\langle M\rangle = V$. Finally, to address the general case, we consider the sequence $(T_n)_{n\geq 0}$ of stopping times defined by $T_n = \inf\{t \geq 0 : |M_t| \geq n\}$ for all $n \geq 0$, then $(X_{t\wedge T_n}^\lambda)_{t\geq 0}$ is a bounded local martingale, and thus a martingale, and we are in the conditions above. ■

### 4.3.2 Girsanov theorem for Itô integrals

The following result is a generalization to random translations of the result due to Cameron–Martin about the density of translated processes (Theorem 2.25).

**Theorem 4.9** (Girsanov)**.** *Let $B = (B_t)_{t\geq 0}$ be an $(\mathcal{F}_t)_{t\geq 0}$ d-dimensional Brownian motion issued from the origin and $\varphi = (\varphi_t)_{t\geq 0}$ be a d-dimensional $(\mathcal{F}_t)_{t\geq 0}$-adapted and mesurable process, bounded on each compact subset of $\mathbb{R}_+$. Then the process $M = (M_t)_{t\geq 0}$ defined for all $t \geq 0$ by*

$$M_t = \exp\Big(\int_0^t \varphi_s \mathrm{d}B_s - \frac{1}{2}\int_0^t |\varphi_s|^2 \mathrm{d}s\Big)$$

*is a martingale and $\mathbb{E}M_t = 1$ for all $t \geq 0$. Moreover for all fixed $T \geq 0$ the law of the process*

$$\widetilde{B} = \Big(B_t - \int_0^t \varphi_s \mathrm{d}s\Big)_{0\leq t\leq T}$$

*is absolutely continuous with respect to the law of $(B_t)_{0\leq t\leq T}$ with density $M_T$. In other words $\widetilde{B}$ is an $(\mathcal{F}_t)_{0\leq t\leq T}$ Brownian motion on $(\Omega, \mathcal{F}, \mathbb{Q})$ where $\mathbb{Q}$ is defined by $\mathrm{d}\mathbb{Q} = M_T \mathrm{d}\mathbb{P}$.*

We recover Theorem 2.25 of Cameron–Martin when $\varphi$ is deterministic.

*Proof.* From Theorem 4.8, the process $M$ is a non-negative local martingale and thus a non-negative super-martingale. For all $t \geq 0$, let us consider

$$N_t = \int_0^t \varphi_s \mathrm{d}B_s \quad \text{and} \quad \langle N\rangle_t = \int_0^t |\varphi_s|^2 \mathrm{d}s.$$

For all $0 \leq t < T$ we have, denoting $C = \sup_{s\in[0,T]} |\varphi_s|^2$,

$$\begin{aligned}
\mathbb{E}(\langle M\rangle_t) &= \mathbb{E}\int_0^t \mathrm{e}^{2N_s - \langle N\rangle_s} |\varphi_s|^2 \mathrm{d}s \\
&\leq C\int_0^t \mathbb{E}(\mathrm{e}^{2N_s - 2\langle N\rangle_s}\mathrm{e}^{\langle N\rangle_s})\mathrm{d}s \\
&\leq C\mathrm{e}^{Ct}\int_0^t \mathbb{E}\mathrm{e}^{2N_s - 2\langle N\rangle_s}\mathrm{d}s \\
&\leq C\mathrm{e}^{Ct} t < \infty.
\end{aligned}$$

Therefore, $M$ is a martingale thanks to the criterion given by Remark 4.6.

In order to check that $\widetilde{B}$ is a Brownian motion under $\mathbb{Q}$, we use Theorem 2.8 which reduces the problem to show that for all $\lambda \in \mathbb{R}^d$ and all fixed $T \geq 0$, the process

$$\Big(\mathrm{e}^{\lambda\cdot\widetilde{B}_t - \frac{|\lambda|^2}{2}t}\Big)_{0\leq t\leq T}$$

is a martingale under $\mathbb{Q}$. Indeed, for all $0 \leq s < t \leq T$ and $A \in \mathcal{F}_s$,

$$\begin{aligned}
\mathbb{E}_Q\Big(\mathbf{1}_A \mathrm{e}^{\lambda\cdot\widetilde{B}_t - \frac{|\lambda|^2}{2}t}\Big) &= \mathbb{E}\Big(\mathbf{1}_A \mathrm{e}^{\lambda\cdot B_t - \lambda\cdot\int_0^t \varphi_s \mathrm{d}s - \frac{|\lambda|^2}{2}t} M_t\Big) \\
&= \mathbb{E}\Big(\mathbf{1}_A \mathrm{e}^{\int_0^t (\lambda+\varphi_s)\cdot \mathrm{d}B_s - \frac{1}{2}\int_0^t |\lambda+\varphi_s|^2\cdot \mathrm{d}s}\Big) \\
&\overset{\star}{=} \mathbb{E}\Big(\mathbf{1}_A \mathrm{e}^{\int_0^s (\lambda+\varphi_u)\cdot \mathrm{d}B_u - \frac{1}{2}\int_0^s |\lambda+\varphi_u|^2\cdot \mathrm{d}u}\Big) \\
&= \mathbb{E}_Q\Big(\mathbf{1}_A \mathrm{e}^{\lambda\cdot\widetilde{B}_s - \frac{|\lambda|^2}{2}s}\Big)
\end{aligned}$$

where we used in $\star$ the fact that $M$ is a martingale with $\lambda + \varphi$ instead of $\varphi$. ■

### 4.3.3 Sub-Gaussian deviation inequality

**Theorem 4.10** (Sub-Gaussian deviation inequality)**.** *For all continuous local martingale $M = (M_t)_{t\geq 0}$ issued from the origin and for all $t > 0$, $K \geq 0$, $r \geq 0$, we have*

$$\mathbb{P}\Big(\langle M\rangle_t \leq K, \sup_{0\leq s\leq t}|M_s| \geq r\Big) \leq 2\mathrm{e}^{-\frac{r^2}{2K}}.$$

*In particular, if $\langle M\rangle_s \leq Ks$ for all $0 \leq s \leq t$ then*

$$\mathbb{P}\Big(\sup_{0\leq s\leq t}|M_s| \geq r\Big) \leq 2\mathrm{e}^{-\frac{r^2}{2Kt}}.$$

The condition on $\langle M\rangle_s$ is a comparison to Brownian motion, for which $s \mapsto \langle B\rangle_s$ is linear.

*Proof.* For all $\lambda, t \geq 0$, by Theorem 4.8, the process

$$X^\lambda = \Big(\mathrm{e}^{\lambda M_t - \frac{\lambda^2}{2}\langle M\rangle_t}\Big)_{t\geq 0}.$$

is a non-negative super-martingale and $\mathbb{E}X_t^\lambda \leq 1$ for all $t, \lambda \geq 0$. Therefore, for all $t, \lambda, r, K \geq 0$,

$$\begin{aligned}\mathbb{P}\Big(\langle M\rangle_t \leq K, \sup_{0\leq s\leq t} M_s \geq r\Big) &\leq \mathbb{P}\Big(\langle M\rangle_t \leq K, \sup_{0\leq s\leq t} X_s^\lambda \geq \mathrm{e}^{\lambda r - \frac{\lambda^2}{2}K}\Big)\\ &\leq \mathbb{P}\Big(\sup_{0\leq s\leq t} X_s^\lambda \geq \mathrm{e}^{\lambda r - \frac{\lambda^2}{2}K}\Big)\\ &\leq \mathrm{e}^{-\lambda r + \frac{\lambda^2}{2}K}\end{aligned}$$

where the last step comes from the maximal inequality (Theorem 1.11). Taking $\lambda = r/K$ gives

$$\mathbb{P}\Big(\langle M\rangle_t \leq K, \sup_{0\leq s\leq t} M_s \geq r\Big) \leq \mathrm{e}^{-\frac{r^2}{2K}}.$$

The same reasoning provides

$$\mathbb{P}\Big(\langle M\rangle_t \leq K, \sup_{0\leq s\leq t}(-M_s) \geq r\Big) \leq \mathrm{e}^{-\frac{r^2}{2K}}.$$

The desired result follows now by the union bound, hence the factor 2 in the right hand side. ■

**Exercise 4.11** (Sub-Gaussian exponential integrability)**.** *Let $M$ be as in Theorem 4.10 and for all $t \geq 0$. Show that if $\langle M\rangle_t \leq Kt$ for all $t \geq 0$ then for all $\alpha < 1/(2Kt)$,*

$$\mathbb{E}\mathrm{e}^{\alpha\|M\|_t^2} < \infty \quad \textit{where} \quad \|M\|_t = \sup_{0\leq s\leq t}|M_s|.$$

### 4.3.4 Burkholder–Davis–Gundy inequalities

**Theorem 4.12** (Burkholder–David–Gundy inequalities)**.** *For all $p \in (0, +\infty)$ there exists universal constants $c_p > 0$ and $C_p > 0$ such that for all continuous local martingale $M = (M_t)_{t\geq 0}$ issued form the origin, we have, for all fixed $T \geq 0$, denoting $\|M\|_T = \sup_{0\leq t\leq T}|M_t|$,*

$$c_p\mathbb{E}(\|M\|_T^{2p}) \leq \mathbb{E}(|\langle M\rangle_T|^p) \leq C_p\mathbb{E}(\|M\|_T^{2p}).$$

*Proof.* Let $(T_n)_{n\geq 0}$ be the sequence of stopping times defined for all $n \geq 0$ by

$$T_n = \inf\{t \geq 0 : |M_t| \geq n \text{ or } \langle M\rangle_t \geq n\}.$$

Then if the desired inequalities are satisfied for $M^{T_n} = (M_{t\wedge T_n})_{t\geq 0}$ with constants $c_p > 0$ and $C_p > 0$, then they will be also satisfied by $M$ by letting $n \to \infty$. This localization method allows in fact to assume without loss of generality that $M$ is a bounded continuous martingale.

Let us fix $T > 0$. The maximal inequality of Theorem 1.11 writes, for all $r \in (1, +\infty)$,

$$\mathbb{E}(\|M\|_T^r) \leq \left(\frac{r}{r-1}\right)^r \mathbb{E}(|M_T|^r)$$

*Case* $p = 1$. In this case $\mathbb{E}(\langle M\rangle_T) = \mathbb{E}(M_T^{2p})$ and the desired BGD inequality is verified with $c_1 = 1/4$ (maximal inequality with $r = 2$) and $C_1 = 1$ (monotony of expectation).

*Case* $p > 1$. We have, from the Itô formula of Theorem 4.4, for all $t \geq 0$,

$$|M_t|^{2p} = 2p\int_0^t |M_s|^{2p-1}\text{sign}(M_s)\mathrm{d}M_s + p(2p-1)\int_0^t |M_s|^{2p-2}\mathrm{d}\langle M\rangle_s$$

and thus, for all $0 \leq t \leq T$, using the Hölder inequality with $p$ and $q = 1/(1-1/p) = p/(p-1)$,

$$\begin{aligned}\mathbb{E}(|M_t|^{2p}) &= p(2p-1)\mathbb{E}\int_0^t |M_s|^{2p-2}\mathrm{d}\langle M\rangle_s \\ &\leq p(2p-1)\mathbb{E}(\|M\|_T^{2(p-1)}\langle M\rangle_T) \\ &\leq p(2p-1)\mathbb{E}(\|M\|_T^{2p})^{1-1/p}(\mathbb{E}(\langle M\rangle_T^p)^{1/p}.\end{aligned}$$

Combined with the maximal inequality above used with $r = 2p$, we obtain the second BGD inequality. To prove the first BGD inequality, we write, using the Itô formula of Theorem 4.4,

$$M_t\langle M\rangle_t^{(p-1)/2} = \int_0^t \langle M\rangle_s^{(p-1)/2}\mathrm{d}M_s + \int_0^t M_s\mathrm{d}(\langle M\rangle_s^{(p-1)/2}).$$

If we set $N_t = \int_0^t \langle M\rangle_s^{(p-1)/2}\mathrm{d}M_s$, we have, for all $t \in [0, T]$,

$$|N_t| \leq 2\|M\|_T\langle M\rangle_T^{(p-1)/2},$$

which gives, using the Hölder inequality with $p$ and $q = 1/(1-1/p) = p/(p-1)$,

$$\mathbb{E}(N_t^2) \leq 4\mathbb{E}(\|M\|_T^2\langle M\rangle_T^{p-1}) \leq 4(\mathbb{E}(\|M\|_T^{2p}))^{1/p}(\mathbb{E}(\langle M\rangle_T^p)^{1-1/p}.$$

Combined with

$$\mathbb{E}(N_t^2) = \mathbb{E}\int_0^t \langle M\rangle_s^{p-1}\mathrm{d}\langle M\rangle_s = \frac{1}{p}\mathbb{E}(\langle M\rangle_t^p)$$

we obtain

$$\mathbb{E}(\langle M\rangle_T^p) \leq (4p)^p\mathbb{E}(\|M\|_T^{2p}),$$

which is the first BGD inequality.

*Case* $0 < p < 1$. Let us define $N_t = \int_0^t \langle M\rangle_s^{(p-1)/2}\mathrm{d}M_s$. We have

$$M_t = \int_0^t \langle M\rangle_s^{(1-p)/2}\mathrm{d}N_s$$

and

$$\begin{aligned}N_t\langle M\rangle_t^{(1-p)/2} &= \int_0^t \langle M\rangle_s^{(1-p)/2}\mathrm{d}N_s + \int_0^t N_s\mathrm{d}(\langle M\rangle_s^{(1-p)/2}) \\ &= M_t + \int_0^t N_s\mathrm{d}(\langle M\rangle_s^{(1-p)/2}).\end{aligned}$$

Therefore, for all $t \in [0,T]$,

$$|M_t| \le 2\|N\|_T \langle M\rangle_T^{(1-p)/2} \quad \text{and} \quad \|M\|_T \le 2\|N\|_T \langle M\rangle_T^{(1-p)/2},$$

thus, using the Hölder inequality with $1/p$ and its conjugate exponent $1/(1-p)$,

$$\begin{aligned}
\mathbb{E}(\|M\|_T^{2p}) &\le 4^p \mathbb{E}(\|N\|_T^{2p} \langle M\rangle_T^{p(1-p)}) \\
&\le (4^p)^2 (\mathbb{E}(\|N_T\|^2))^p (\mathbb{E}(\langle M\rangle_T^p)^{1-p} \\
&\le (4^p)^2 (\mathbb{E}(N_T^2))^p (\mathbb{E}(\langle M\rangle_T^p)^{1-p} \\
&= 16^p (p^{-1}\mathbb{E}(\langle M\rangle_T^p))^p (\mathbb{E}(\langle M\rangle_T^p))^{1-p} \\
&= \left(\frac{16}{p}\right)^p \mathbb{E}(\langle M\rangle_T^p).
\end{aligned}$$

This proves the first BGD inequality. To prove the second BGD inequality, let $\alpha > 0$. The reason for $\alpha > 0$ is to avoid the singularity at 0 of $x \mapsto x^{p-1}$ due to the fact that $p-1 < 0$. Now write, using the Itô formula of Theorem 4.4,

$$\begin{aligned}
M_t(\alpha + \|M\|_t)^{p-1} &= \int_0^t (\alpha + \|M\|_s)^{p-1} \mathrm{d}M_s + \int_0^t M_s \mathrm{d}(\alpha + \|M\|_s)^{p-1} \\
&= N_t + (p-1) \int_0^t M_s (\alpha + \|M\|_s)^{p-2} \mathrm{d}\|M\|_s
\end{aligned}$$

where $N_t = \displaystyle\int_0^t (\alpha + \|M\|_s)^{p-1} \mathrm{d}M_s$. We have then (taking $\alpha \to 0$)

$$|N_t| \le \|M_t\|^p + (1-p) \int_0^t \|M\|_s^{p-1} \mathrm{d}\|M\|_s = \frac{1}{p}\|M\|_t^p$$

and thus

$$\mathbb{E}\int_0^t (\alpha + \|M\|_s)^{2(p-1)} \mathrm{d}\langle M\rangle_s = \mathbb{E}(N_t^2) \le \frac{1}{p^2}\mathbb{E}(\|M\|_t^{2p}),$$

which gives finally the inequality (recall that $2(1-p) < 0$)

$$\mathbb{E}((\alpha + \|M\|_t)^{2(p-1)} \langle M\rangle_t) \le \frac{1}{p^2}\mathbb{E}(\|M\|_t^{2p}).$$

But the identity

$$\langle M\rangle_t^p = (\langle M\rangle_t^p (\alpha + \|M\|_t)^{2p(p-1)})(\alpha + \|M\|_t)^{2p(1-p)}$$

gives, using the Hölder inequality with $1/p$ and its conjugate exponent $1/(1-p)$, that

$$\begin{aligned}
\mathbb{E}(\langle M\rangle_t^p) &\le (\mathbb{E}(\langle M\rangle_t (\alpha + \|M\|_t)^{2(p-1)}))^p (\mathbb{E}((\alpha + \|M\|_t)^{2p}))^{1-p} \\
&\le \left(\frac{1}{p^2}\right)^p (\mathbb{E}(\|M\|_t^{2p}))^p (\mathbb{E}((\alpha + \|M\|_t)^{2p}))^{1-p}.
\end{aligned}$$

Taking the limit as $\alpha \to 0$, we obtain

$$\mathbb{E}(\langle M\rangle_t^p) \le \frac{1}{p^{2p}}\mathbb{E}(\|M\|_t^{2p})$$

which is the second BGD inequality. ■

# Chapter 5

# Stochastic differential equations

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space, with $(\mathcal{F}_t)_{t\geq 0}$ complete and right continuous.

Let $B = (B_t)_{t\geq 0}$ be a $d$-dimensional $(\mathcal{F}_t)_{t\geq 0}$ Brownian motion issued from the origin.

## 5.1 Stochastic differential equations with Lipschitz coefficients

Let us denote $|M| = (\sum_{j,k} M_{jk}^2)^{1/2}$ for all matrix $M \in \mathcal{M}_{q,d}(\mathbb{R})$ with $q$ rows and $d$ columns.

Let us consider two maps

$$\sigma : \mathbb{R}_+ \times \Omega \times \mathbb{R}^q \to \mathcal{M}_{q,d}(\mathbb{R}) \quad \text{and} \quad b : \mathbb{R}_+ \times \Omega \times \mathbb{R}^q \to \mathbb{R}^q$$

such that the following properties hold true:

1. the maps $\sigma, b$ are Lipschitz in the space variable, namely there exists a constant $c > 0$ such that for all $(u, \omega) \in \mathbb{R}_+ \times \Omega$ and all $x, y \in \mathbb{R}^q$,

$$|\sigma(u, \omega, x) - \sigma(u, \omega, y)| \leq c|x - y| \quad \text{and} \quad |b(u, \omega, x) - b(u, \omega, y)| \leq c|x - y|;$$

2. the maps $\sigma, b$ are measurable for the time/random variables, namely for all $t > 0$, $x \in \mathbb{R}^q$,

$$(u, \omega) \in [0, t] \times \Omega \mapsto \sigma(u, \omega, x) \quad \text{and} \quad (u, \omega) \in [0, t] \times \Omega \mapsto b(u, \omega, x)$$

   are $(\mathcal{B}_{[0,t]} \otimes \mathcal{F}_t)$-measurable;

3. the maps $\sigma, b$ are locally square integrable, namely for all $t > 0$ and $x \in \mathbb{R}^q$,

$$\mathbb{E} \int_0^t |\sigma|^2(u, \cdot, x) \mathrm{d}u < \infty \quad \text{and} \quad \mathbb{E} \int_0^t |b|^2(u, \cdot, x) \mathrm{d}u < \infty.$$

   Note that hanks to the first property (Lipschitz regularity), if this square integrability property is satisfied for some $x \in \mathbb{R}^q$ then it will be satisfied for all $x \in \mathbb{R}^q$.

**Theorem 5.1** (Solving stochastic differential equations). *For all $s \geq 0$ and all $\mathcal{F}_s$-measurable and square integrable random vector $\eta$ of $\mathbb{R}^q$, there exists an adapted and continuous $q$-dimensional process $X = (X_t)_{t\geq s}$ such that the following properties hold true:*

1. *for all $t \geq s$, $\mathbb{E} \int_s^t |X_u|^2 \mathrm{d}u < \infty$;*
2. *$X$ solves the stochastic differential equation (SDE)*

$$X_t = \eta + \int_s^t \sigma(u, X_u) \mathrm{d}B_u + \int_s^t b(u, X_u) \mathrm{d}u \quad a.s., \quad t \geq s,$$

   *in other words for all $1 \leq j \leq q$,*

$$X_t^j = \eta_j + \sum_{k=1}^d \int_s^t \sigma_{jk}(u, X_u) \mathrm{d}B_u^k + \int_s^t b_j(u, X_u) \mathrm{d}u \quad a.s., \quad t \geq s.$$

*Moreover this process $X$ is unique up to indistinguishablility.*

Note that the first property ensures that the second property has a meaning since

$$|\sigma(u, X_u)|^2 \leq 2(|\sigma(u,0)|^2 + c^2|X_u|^2).$$

The process $X$ solves the stochastic differential equation

$$X_s = \eta, \quad \mathrm{d}X_t = \sigma(u, X_t)\mathrm{d}B_t + b(t, X_t)\mathrm{d}t \quad \text{a.s.,} \quad t \geq s.$$

*Proof.* Let $\mathcal{D}$ be the set of continuous adapted $q$-dimensional processes $(Y_t)_{t\geq s}$ with, for all $t \geq 0$,

$$\|Y\|_t^2 = \mathbb{E}(\sup_{s\leq u\leq t} |Y_u|^2) < \infty.$$

For all $Y \in \mathcal{D}$, we define, for all $t \geq s$,

$$SY(t) = \eta + \int_s^t \sigma(u, Y_u)\mathrm{d}B_u + \int_s^t b(u, Y_u)\mathrm{d}u.$$

For all $Y^1$ and $Y^2$ in $\mathcal{D}$ we have, for all $t \geq s$,

$$SY^1(t) - SY^2(t) = \int_s^t (\sigma(u, Y_u^1) - \sigma(u, Y_u^2))\mathrm{d}B_u + \int_s^t (b(u, Y_u^1) - b(u, Y_u^2))\mathrm{d}u,$$

and then

$$\begin{aligned}
&\sup_{s\leq u\leq t} |SY^1(u) - SY^2(u)|^2 \\
&\qquad \leq 2 \sup_{s\leq u\leq t} \left|\int_s^u (\sigma(v, Y_v^1) - \sigma(v, Y_v^2))\mathrm{d}B_v\right|^2 + 2(t-s)\int_s^t |b(u, Y_u^1) - b(u, Y_u^2)|^2\mathrm{d}u.
\end{aligned}$$

By using the Doob maximal inequality of Theorem 1.11 with $p = 2$, we get

$$\begin{aligned}
\|SY^1 - SY^2\|_t^2 &\leq 8\int_s^t \mathbb{E}(|\sigma(u, Y_u^1) - \sigma(u, Y_u^2)|^2)\mathrm{d}u + 2(t-s)\int_s^t \mathbb{E}(|b(u, Y_u^1) - b(u, Y_u^2)|^2)\mathrm{d}u \\
&\leq 2c^2(4 + (t-s))\int_s^t \mathbb{E}(|Y_u^1 - Y_u^2|^2)\mathrm{d}u.
\end{aligned}$$

Taking $Y^2 \equiv 0$, this shows that $SY \in \mathcal{D}$ when $Y \in \mathcal{D}$. If we set $C_t = 2c^2(4 + (t-s))$ and $\varphi(u) = \mathbb{E}(|Y_u^1 - Y_u^2|^2)$, we get, for all $n \geq 1$, denoting $S^n = S \circ \cdots \circ S$ the $n$-th iteration of $S$,

$$\begin{aligned}
\|S^nY^1 - S^nY^2\|_t^2 &\leq (C_t)^2 \int_s^t \mathrm{d}u \int_s^u \mathbb{E}(|S^{n-2}Y^1 - S^{n-2}Y^2(v)|^2)\mathrm{d}v \qquad (\star) \\
&\;\;\vdots \\
&\leq (C_t)^n \int \mathbf{1}_{t\geq u_1\geq\cdots\geq u_n\geq s}\varphi(u_n)\mathrm{d}u_1\ldots\mathrm{d}u_n \\
&\leq (C_t)^n\|Y^1 - Y^2\|_t^2\frac{(t-s)^n}{n!}.
\end{aligned}$$

Let us show now that $S$ admits a fixed point. We start from an arbitrary $Y \in \mathcal{D}$, and we set $X^0 = Y$, and $X^n = S^nY$ for all $n \geq 1$. Then we have

$$\mathbb{E}\Big(\sup_{s\leq u\leq t} |X_u^n - X_u^{n+1}|^2\Big) \leq \frac{(C_t(t-s))^n}{n!}\|Y - SY\|_t^2. \qquad (\star\star)$$

It follows that

$$\mathbb{E}\sum_{n\geq 0}\sup_{s\leq u\leq t}|X_u^n - X_v^{n+1}| \leq \sum_{n\geq 0}\Big(\frac{(C_t(t-s))^n}{n!}\|Y-SY\|_t^2\Big)^{1/2} < \infty.$$

Thus, for all $t > s$, almost surely

$$\sum_{n\geq 0}\sup_{s\leq u\leq t}|X_u^n - X_u^{n+1}| < \infty.$$

Therefore, the sequence of continuous processes $(X_u^n)_{u\geq s}$ converges almost surely uniformly on every compact subset of $[s,\infty)$ towards a continuous adapted process denoted $X = (X_u)_{u\geq 0}$ and from the inequality $(\star\star)$ we get

$$\Big(\mathbb{E}\Big(\sup_{s\leq u\leq t}|X_u^n - X_u|^2\Big)\Big)^{1/2} \leq \sum_{m\geq n}\|X^n - X^{n+1}\|_t \underset{n\to\infty}{\longrightarrow} 0.$$

It follows that $X\in\mathcal{D}$, that $X^n\to X$ in $\mathcal{D}$, and that

$$\|X - SX\|_t \leq \|X - X^{n+1}\|_t + \|SX^n - SX\|_t \underset{n\to\infty}{\longrightarrow} 0.$$

It follows that $X = SX$. Now let $X$ and $\widetilde{X}$ two fixed points of $S$. We have, for all $n\geq 0$, $X - \widetilde{X} = S^nX - S^n\widetilde{X}$ and from $(\star)$, for all $t\geq 0$,

$$\|X-\widetilde{X}\|_t^2 \leq \frac{(C_t(t-s))^n}{n!}\|X-\widetilde{X}\|_t^2 \underset{n\to\infty}{\longrightarrow} 0$$

and therefore $X = \widetilde{X}$, hence the uniqueness. ■

**Lemma 5.2** (Grönwall lemma)**.** *If $f : [s,+\infty)\to\mathbb{R}_+$ is measurable and $h : [s,\infty)\to\mathbb{R}$ is continuous and increasing and $a > 0$ and $b > 0$ are constants such that for all $t\geq s$,*

$$f(t) \leq a + b\int_s^t f(u)\mathrm{d}h_u.$$

*Then, for all $t\geq s$,*

$$f(t) \leq a\mathrm{e}^{b(h_t-h_s)}.$$

*Proof.* Let $\varepsilon > 0$ and $g(t) = (a+\varepsilon)\mathrm{e}^{b(h_t-h_s)}$. We have, for all $t\geq s$,

$$g(t) = (a+\varepsilon) + b\int_s^t g(u)\mathrm{d}h_u,$$

and therefore, for all $t\geq s$,

$$f(t) - g(t) \leq -\varepsilon + b\int_s^t (f(u)-g(u))\mathrm{d}h_u.$$

Set $T = \inf\{t\geq s : f(t)-g(t)\geq 0\}\in[0,+\infty]$. If $T<\infty$ then, for some $T_n\searrow T$,

$$0 \leq f(T_n) - g(T_n) \leq -\varepsilon + b\int_s^{T_n}(f(u)-g(u))\mathrm{d}h_u,$$

thus, using the definition of $T$,

$$0 \leq \overline{\lim_{n\to\infty}}(f(T_n)-g(T_n)) \leq -\varepsilon + b\int_s^T (f(u)-g(u))\mathrm{d}h_u \leq -\varepsilon$$

which is a contradiction. Therefore $T = +\infty$ and thus $f(t)\leq g(t)$ for all $t\geq s$. ■

**Theorem 5.3** (Dependency over initial condition)**.** *For all $s \geq 0$, for all $\mathcal{F}_s$-measurable square integrable random vectors $\eta$ and $\widetilde{\eta}$ of $\mathbb{R}^q$, if $X$ and $\widetilde{X}$ are the solutions of*

$$X_t = \eta + \int_s^t \sigma(u, X_u)\mathrm{d}B_u + \int_s^t b(u, X_u)\mathrm{d}u \quad a.s., \quad t \geq s,$$

*and*

$$\widetilde{X}_t = \widetilde{\eta} + \int_s^t \sigma(u, \widetilde{X}_u)\mathrm{d}B_u + \int_s^t b(u, \widetilde{X}_u)\mathrm{d}u \quad a.s., \quad t \geq s,$$

*then, for all $t \geq s$, there exists a constant $C_t > 0$ such that*

$$\mathbb{E}\Big(\sup_{s\leq u\leq t} |X_u - \widetilde{X}_u|^2\Big) \leq C_t \mathbb{E}(|\eta - \widetilde{\eta}|^2).$$

*Proof.* We have

$$X_t - \widetilde{X}_t = \eta - \widetilde{\eta} + \int_s^t (\sigma(u, X_u) - \sigma(u, \widetilde{X}_u))\mathrm{d}B_u + \int_s^t (b(u, X_u) - b(u, \widetilde{X}_u))\mathrm{d}u.$$

Setting $f(t) = \mathbb{E}\Big(\sup_{s\leq u\leq t} |X_u - \widetilde{X}_u|^2\Big)$, by the maximal inequality (Theorem 1.11), for all $t \geq s$,

$$\begin{aligned} f(t) &\leq 3\mathbb{E}(|\eta - \widetilde{\eta}|^2) + 12\mathbb{E}\int_s^t |\sigma(u, X_u) - \sigma(u, \widetilde{X}_u)|^2\mathrm{d}u + 3(t-s)\int_s^t \mathbb{E}(|b(u, X_u) - b(u, \widetilde{X}_u)|^2)\mathrm{d}u \\ &\leq 3\mathbb{E}(|\eta - \widetilde{\eta}|^2) + c^2(12 + 3(t-s))\int_s^t f(u)\mathrm{d}u. \end{aligned}$$

The desired result follows now from the Grönwall lemma (Lemma 5.2). ∎

**Theorem 5.4** (Regular solution of the stochastic differential equation)**.** *For all $s \geq 0$, there exists a family $(X_t^s(x, \omega) : x \in \mathbb{R}^q, \omega \in \Omega, s \leq t)$ of random variables such that:*

1. *for all $t \geq s$, the map $(x, \omega) \in \mathbb{R}^q \times \Omega \mapsto X_t^s(x, \omega) \in \mathbb{R}^q$ is $(\mathcal{B}_{\mathbb{R}^q} \otimes \mathcal{F}_t)$-measurable;*
2. *for all square integrable $\mathcal{F}_s$-measurable random vector $\eta$ of $\mathbb{R}^q$, the random variable $Y_t(\omega) = X_t^s(\eta(\omega), \omega)$ solves the stochastic differential equation*

$$Y_t = \eta + \int_s^t \sigma(u, Y_u)\mathrm{d}B_u + \int_s^t b(u, Y_u)\mathrm{d}u \quad a.s., \quad t \geq s. \tag{$\star$}$$

*Proof.* For all $n \geq 0$, let $(T_k)_{k\geq 0}$ be an at most countable partition of $\mathbb{R}^q$ such that for all $k \geq 0$, $\mathrm{diam}(T_k) \leq 2^{-n}$. For each $k \geq 0$, we select $z_k \in T_k$, and we define, for all $x \in \mathbb{R}^q$,

$$g_n(x) = \sum_k z_k \mathbf{1}_{T_k}(x).$$

Let $z \in \mathbb{R}^q$. We consider the solution $\widetilde{X}_t(z, \omega)$ of

$$\widetilde{X}_t = z + \int_s^t \sigma(u, \widetilde{X}_u)\mathrm{d}B_s + \int_s^t b(u, \widetilde{X}_u)\mathrm{d}u,$$

for all $t \geq s$ and all $\omega \notin N_z$ where $N_z$ is a negligible set. Let us define

$$N_n = \cup_k N_{z_k} \quad \text{and} \quad X_t^n(x, \omega) = \widetilde{X}_t(g_n(x), \omega)\mathbf{1}_{\Omega\setminus N_n}(\omega).$$

The map $(x, \omega) \mapsto X_t^n(x, \omega)$ is $(\mathcal{B}_{\mathbb{R}^q} \otimes \mathcal{F}_t)$-measurable and

$$\mathbb{E}\Big(\sup_{s\leq u\leq t} |X_u^n - \widetilde{X}_u(x)|^2\Big) \leq C_t|x - g_n(x)|^2 \leq C_t\Big(\frac{1}{2^n}\Big)^2.$$

Thus

$$\mathbb{E}\sum_{n\geq 0}\sup_{s\leq u\leq t}|X_u^n-\widetilde{X}_u(x)|<\infty$$

and therefore, for all $t\geq s$, almost surely,

$$\sup_{s\leq u\leq t}|X_u^n-\widetilde{X}_u|\underset{n\to\infty}{\longrightarrow}0.$$

Now we define

$$X_t^s(x,\omega)=\overline{\lim_{n\to\infty}}\,X_t^n(x,\omega).$$

Let $\eta$ be a square integrable $\mathcal{F}_s$-measurable random vector $\eta$ of $\mathbb{R}^q$. We have that the random variable $Y_t^n(\omega)=X_t^n(\eta(\omega),\omega)$ solves

$$Y_t^n=g_n(\eta)+\int_s^t\sigma(u,Y_u^n)\mathrm{d}B_u+\int_s^t b(u,Y_u^n)\mathrm{d}u,\quad t\geq s.$$

This follows from the fact that, for all $k$, almost surely,

$$\mathbf{1}_{\eta\in T_k}\int_s^t\sigma(u,X_u^n(z_k))\mathrm{d}B_u=\int_s^t\mathbf{1}_{\eta\in T_k}\sigma(u,X_u^n)\mathrm{d}B_u.$$

Finally, for all $t\geq s$, almost surely,

$$\sup_{s\leq u\leq t}|Y_u^n-Y_u|\underset{n\to\infty}{\longrightarrow}0,$$

where $Y=(Y_t)_{t\geq 0}$ is the solution of $(\star)$. It follows that for all $t\geq s$, almost surely,

$$X_t^s(\eta(\omega),\omega)=Y_t(\omega).$$

■

**Corollary 5.5.** *For all $0\leq s\leq t\leq u$, $x\in\mathbb{R}^q$, with the notions of Theorem 5.4, almost surely,*

$$X_u^s(x,\omega)=X_u^t(X_t^s(s,\omega),\omega)).$$

*Proof.* We have

$$\begin{aligned}X_u^s(x)&=x+\int_s^u\sigma(u,X_u^s)\mathrm{d}B_v+\int_s^u b(v,X_v^s)\mathrm{d}v\\&=x+\int_s^t\sigma(u,X^s\ u)\mathrm{d}B_v+\int_s^t b(v,X_v^s)\mathrm{d}v+\int_t^u\sigma(u,X_u^s)\mathrm{d}B_v+\int_t^u b(v,X_v^s)\mathrm{d}v\\&=X_t^s(x)+\int_t^u\sigma(u,X_u^s)\mathrm{d}B_v+\int_t^u b(v,X_v^s)\mathrm{d}v\end{aligned}$$

where the last equality holds almost surely, therefore, almost surely, from Theorem 5.4,

$$X_u^s(x,\omega)=X_u^t(X_t^s(x,\omega),\omega).$$

■

## 5.2 Deterministic case

### Important remark about the assumptions

In this section, we assume that $\sigma$ and $b$ do not depend on $\omega$ in other words that $\sigma$ and $b$ are measurable maps from $\mathbb{R}_+ \times \mathbb{R}^q$ to $\mathcal{M}_{q,d}(\mathbb{R})$ and $\mathbb{R}^q$ respectively, and, for simplicity, that

1. there exists a constant $C > 0$ such that for all $x, y \in \mathbb{R}^q$ and all $u \in \mathbb{R}_+$,
$$|\sigma(u,x) - \sigma(u,y)| + |b(u,x) - b(u,y)| \leq C|x-y|;$$
2. for all $t > 0$ and all $x \in \mathbb{R}^q$,
$$\int_0^t (|\sigma(u,x)|^2 + |b|^2(u,x))\mathrm{d}u < \infty.$$

We denote by $(X_t^s(x,\omega))_{0\leq s\leq t}$ the regular solution, given by Theorem 5.4, of

$$X_s^s(x) = x, \quad X_t^s(x) = x + \int_s^t \sigma(u, X_u^s(x))\mathrm{d}B_u + \int_s^t b(u, X_u^s(x))\mathrm{d}u, \quad t \geq s, \qquad \text{(SDE)}$$

in other words

$$X_s^s(x) = x, \quad \mathrm{d}X_t^s(x) = \sigma(t, X_t^s(x))\mathrm{d}B_s + b(t, X_t^s(x))\mathrm{d}t, \quad t \geq s.$$

**Theorem 5.6** (Markov property)**.** *For all $s \geq 0$ and $x \in \mathbb{R}^q$, let $(X_t^s(x))_{t\geq s}$ be the solution of* (SDE)*. Then for all bounded and measurable $f : \mathbb{R}^q \to \mathbb{R}$, and for all $u \geq t \geq s$,*

$$\mathbb{E}(f(X_u^s(x)) \mid \mathcal{F}_t) = \mathbb{E}(f(X_u^s(x)) \mid X_t^s(x)) = \Pi_{t,u}(f)(X_t^s(x))$$

*almost surely, where for all $z \in \mathbb{R}^q$,*

$$\Pi_{t,u}(f)(z) = \mathbb{E}(f(X_u^t(z))).$$

*Proof.* Let $z \in \mathbb{R}^q$ and $0 \leq t \leq u$. We have, almost surely,

$$\begin{aligned} X_u^t(z) &= z + \int_t^u \sigma(v, X_v^t(z))\mathrm{d}B_v + \int_t^u b(v, X_v^t(z))\mathrm{d}v \\ &= z + \int_0^{u-t} \sigma(t+v, X_{t+v}^t(z))\mathrm{d}B_v^t + \int_0^{u-t} b(t+v), X_{t+v}^t(z))\mathrm{d}v, \end{aligned}$$

where $B^t = (B_v^t)_{v\geq 0} = (B_{t+v} - B_t)_{v\geq 0}$ is a translated Brownian motion, independent of $\mathcal{F}_t$.

Now, let $(\mathcal{S}_u^t)_{u\geq t}$ be the complete filtration given by $\mathcal{S}_u^t = \sigma(B_{t+v} - B_t : 0 \leq v \leq u-t)$, for all $u \geq t$. Thanks to Theorem 5.4, the following map is $(\mathcal{B}_{\mathbb{R}^q} \otimes \mathcal{S}_u^t)$-measurable:

$$(z,\omega) \in \mathbb{R}^q \times \Omega \mapsto X_u^t(z,\omega) \in \mathbb{R}^q.$$

By considering the process $B^t$ as a random variables taking values in $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$, we see that there exists a measurable map $\Theta_u^t : \mathbb{R}^q \times \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^q$ such that $X_u^t(z,\omega) = \Theta_u^t(z, B^t(\omega))$ for all $(z,\omega) \in \mathbb{R}^q \times \Omega$. Let $\mu$ be the Wiener measure on $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$, the law of $B^t$. We have

$$\mathbb{E}(f(X_u^t(z))) = \int_{\mathcal{C}(\mathbb{R}_+,\mathbb{R}^d)} f(\Theta_u^t(z,\omega))\mu(\mathrm{d}\omega) = \Pi_{t,u}(f)(z).$$

Moreover, for all $u \geq t \geq s \geq 0$, from Corollary 5.5,

$$\begin{aligned} \mathbb{E}(f(X_u^s) \mid \mathcal{F}_t) &= \mathbb{E}(f(X_u^t(X_t^s(x))) \mid \mathcal{F}_t)) \\ &= \mathbb{E}(f(\Theta_u^t(X_t^s(x), B^t)) \mid \mathcal{F}_t) \\ &= \int_{\mathcal{C}(\mathbb{R}_+,\mathbb{R}^d)} f(\Theta_u^t(X_t^s(x), w))\mu(\mathrm{d}w) \\ &= \Pi_{t,u}(f)(X_t^s(x)) \end{aligned}$$

almost surely since $X_t^s(x)$ is $\mathcal{F}_t$-measurable and since $B^t$ is independent of $\mathcal{F}_t$. ■

**Remark 5.7** (Markov transition kernel and Markov semi-group)**.** *For all* $0 \leq s \leq t$*, let* $\Pi_{s,t}(x, \mathrm{d}y)$ *be the* *Markov transition kernel* *on* $\mathbb{R}^q$ *defined for all* $x \in \mathbb{R}^q$ *and* $A \in \mathcal{B}_{\mathbb{R}^q}$ *by*

$$\Pi_{s,t}(x, A) = \mathbb{P}(X_t^s(x) \in A).$$

*It acts on bounded measurable* $f : \mathbb{R}^q \to \mathbb{R}$ *as*

$$\Pi_{s,t}(f)(x) = \int_{\mathbb{R}^q} f(y)\Pi_{s,t}(x, \mathrm{d}y), \quad x \in \mathbb{R}^q.$$

*Theorem 5.6 gives, of* $u \geq t$*,*

$$\mathbb{E}(f(X_t^s(x))) = \mathbb{E}(\Pi_{t,u}(f)(X_t^s(x))) = \Pi_{s,u}(f)(x).$$

*This gives the (non-homogeneous)* *Markov semi-group property:*

$$\Pi_{s,u}(x, \mathrm{d}y) = \int_{\mathbb{R}^q} \Pi_{s,t}(x, \mathrm{d}z)\Pi_{t,u}(z, \mathrm{d}y), \quad u \geq t \geq s \geq 0, \quad x \in \mathbb{R}^q.$$

*Moreover Theorem 5.6 shows that the following process is an* $(\mathcal{F}_t)_{t\geq 0}$*-martingale:*

$$(\Pi_{t,u}(f)(X_t^s(x)))_{s \leq t \leq u}.$$

*Conversely, the Markov semi-group* $(\Pi_{s,t}(x, \mathrm{d}y))_{0 \leq s \leq t}$ *determines entirely the law of process* $(X_t^s(x))_{t \geq s}$*. Indeed, for all* $n \geq 1$*, all* $0 \leq s \leq t_1 \leq t_2 \leq \cdots \leq t_n$*, and all bounded and measurable* $f_1, \ldots, f_n$ *from* $\mathbb{R}^q$ *to* $\mathbb{R}$*, we have*

$$\begin{aligned}
&\mathbb{E}(f_1(X_{t_1}^s(x)) \cdots f_n(X_{t_n}^s(x))) \\
&\qquad = \mathbb{E}(f_1(X_{t_1}^s(x)) \cdots f_{n-1}(X_{t_{n-1}}^s)\Pi_{t_{n-1},t_n}(f_n)(X_{t_{n-1}}^s(x))) \\
&\qquad = \int_{\mathbb{R}^q} \Pi_{s,t_1}(x, \mathrm{d}y_1)\Pi_{t_1,t_2}(y_1, \mathrm{d}y_2) \cdots \Pi_{t_{n-1},t_n}(y_{n-1}, \mathrm{d}y_n) f_1(y_1) \cdots f_n(y_n).
\end{aligned}$$

**Theorem 5.8** (Uniqueness in law)**.** *Let* $(\widetilde{\Omega}, \widetilde{F}, (\widetilde{F}_t)_{t\geq 0}, \widetilde{P})$ *be another filtered probability space with a complete and right continuous filtration, on which is defined a d-dimensional* $(\widetilde{F}_t)_{t\geq 0}$ *Brownian motion* $\widetilde{B} = (\widetilde{B}_t)_{t\geq 0}$ *issued from the origin. Let* $x \in \mathbb{R}^q$*, and let* $X = (X_t(x, \omega))_{t\geq 0}$ *and* $\widetilde{X} = (\widetilde{X}_t(x, \omega)_{t\geq 0}$ *be the solutions of the respective stochastic differential equations:*

$$X_t(x) = x + \int_0^t \sigma(x, X_u(x))\mathrm{d}B_u + \int_0^t b(u, X_u(x))\mathrm{d}u \quad a.s. \quad t \geq 0,$$

*and*

$$\widetilde{X}_t(x) = x + \int_0^t \sigma(x, \widetilde{X}_u(x))\mathrm{d}\widetilde{B}_u + \int_0^t b(u, \widetilde{X}_u(x))\mathrm{d}u \quad a.s. \quad t \geq 0.$$

*Then these processes* $X$ *and* $\widetilde{X}$ *have the same law on* $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^q)$*.*

*Proof.* We consider the canonical Brownian motion $\pi = (\pi_t(\omega))_{t\geq 0}$ defined on the Wiener space $(W = \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d), \mathcal{B}_W, (\mathcal{B}_t)_{t\geq 0}, \mu)$ where $\mu$ is the Wiener measure. Let $(Y_t(x, w))_{t\geq 0, w\in W}$ be the regular solution provided by Theorem 5.4 of the stochastic differential equation

$$Y_t(x) = x + \int_0^t \sigma(u, Y_u(x))\mathrm{d}\pi_u + \int_0^t b(u, Y_u(x))\mathrm{d}u$$

$\mu$ almost-surely. We can check easily that the processes $Z_t(x, \omega) = Y_t(x, B(\omega))$, $t \geq 0$, $\omega \in \Omega$, and $\widetilde{Z}_t(x, \widetilde{\omega}) = Y_t(x, \widetilde{B}(\widetilde{\omega}))$, $t \geq 0$, $\widetilde{\omega} \in \widetilde{\Omega}$ are respectively solutions of the stochastic differential equations satisfied by $X$ and $\widetilde{X}$. The sample path uniqueness of these solutions give that $(Y_t(x, B(\omega))_{t\geq 0} = (X_t(x, \omega)_{t\geq 0}$ $\mathbb{P}$-a.s. and $(Y_t(x, \widetilde{B}(\widetilde{\omega}))_{t\geq 0} = (\widetilde{X}_t(x, \widetilde{\omega}))_{t\geq 0}$ $\widetilde{\mathbb{P}}$-a.s. But the Brownian motions $B$ and $\widetilde{B}$ have same law on $W = \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$, which is the Wiener measure $\mu$, and therefore the processes $X$ and $\widetilde{X}$ and $(Y_t)_{t\geq 0}$ have the same law on $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^q)$. ■

## 5.3 Homogeneous case

In this section, we assume that $\sigma$ and $b$ do not depend on the randomness $\omega$ and on time $u$, in other words $\sigma$ and $b$ are two maps from $\mathbb{R}^q$ to $\mathcal{M}_{q,d}(\mathbb{R})$ and $\mathbb{R}^q$ respectively. We also assume for simplicity that there exists a constant $C > 0$ such that for all $x, y \in \mathbb{R}^q$,

$$|\sigma(x) - \sigma(y)| + |b(x) - b(y)| \leq C|x - y|.$$

**Theorem 5.9** (Simple Markov property). *Let $(X_t(x))_{t\geq 0}$ be the regular solution of the SDE*

$$X_t(x) = x + \int_0^t \sigma(X_u(x))\mathrm{d}B_u + \int_0^t b(X_u(x))\mathrm{d}u \quad a.s., \quad t \geq 0,\ x \in \mathbb{R}^q,$$

*provided by Theorem 5.4. Then for all $t \geq u \geq 0$ and all measurable and bounded $f : \mathbb{R}^q \mapsto \mathbb{R}$,*

$$\mathbb{E}(f(X_u(x)) \mid \mathcal{F}_t) = \mathbb{E}(f(X_u(x)) \mid X_t(x)) = \Pi_{u-t}(f)(X_t(x)) \quad a.s.$$

*where for all $s \geq 0$ and $x \in \mathbb{R}^q$,*

$$\Pi_s(f)(x) = \mathbb{E}(f(X_s(x))).$$

*Proof.* Thanks to Theorem 5.6 with $s = 0$ it suffices to show that for all $u \geq t \geq 0$,

$$\mathbb{E}(f(X_u^t(x))) = \mathbb{E}(f(X_{u-t}^0(x))).$$

But

$$X_u^t(x) = x + \int_0^{u-t} \sigma(X_{t+s}^t(x))\mathrm{d}B_s^t + \int_0^{u-t} b(X_{t+s}^t)\mathrm{d}s,$$

where $B_s^t = B_{t+s} - B_t$ for all $s \geq 0$, in other words, setting $Y_s(x) = X_{t+s}^t(x)$,

$$Y_s(x) = x + \int_0^s \sigma(Y_u(x))\mathrm{d}B_u^t + \int_0^s b(Y_u(x))\mathrm{d}u \quad \text{a.s.}, \quad s \geq 0,$$

Thus the process $Y(x)$ solves a stochastic differential equation similar to the one solves by $X(x)$, obtained by replacing the Brownian motion $B$ by the translation Brownian motion $B^t$. From Theorem 5.8, it follows that the processes $X(x)$ and $Y(x)$ have same law, and thus, for all $s \geq 0$,

$$\mathbb{E}(f(X_{t+s}^t(x))) = \mathbb{E}(f(X_s^0(x))).$$

■

For all $t \geq 0$, let $\Pi_t(\cdot, \mathrm{d}y)$ be the Markov transition kernel on $\mathbb{R}^q$ defined by

$$\Pi_t(x, A) = \mathbb{P}(X_t(x) \in A), \quad x \in \mathbb{R}^q,\ A \in \mathcal{B}_{\mathbb{R}^q}.$$

It acts on bounded measurable functions $f : \mathbb{R}^q \to \mathbb{R}$ as

$$\Pi_t(f)(x) = \int_{\mathbb{R}^q} f(y)\Pi(x, \mathrm{d}y) = \mathbb{E}(f(X_t(x))), \quad x \in \mathbb{R}^q.$$

It defines a <u>homogeneous Markov semi-group</u> $(\Pi_t(x, \mathrm{d}y))_{t\geq 0}$,

$$\Pi_0 = \mathrm{Id}, \quad \Pi_s \circ \Pi_t = \Pi_{t+s},\ s, t \geq 0.$$

In other words for all $s, t \geq 0$ and all $x \in \mathbb{R}^q$,

$$\Pi_{s+t}(x, \mathrm{d}y) = \int_{\mathbb{R}^q} \Pi_t(x, \mathrm{d}z)\Pi_s(z, \mathrm{d}y) = (\Pi_t\Pi_s)(x, \mathrm{d}y)).$$

**Theorem 5.10** (Markov semi-group properties). *For all $t \geq 0$ the operator $\Pi_t$ preserves globally*

1. *the set $\mathcal{M}(\mathbb{R}^q, \mathbb{R})$ of bounded and measurable functions $\mathbb{R}^q \to \mathbb{R}$;*
2. *the set $\mathcal{C}_b(\mathbb{R}^q, \mathbb{R})$ of bounded and continuous functions $\mathbb{R}^q \to \mathbb{R}$;*
3. *the set $\mathcal{C}_0(\mathbb{R}^q, \mathbb{R})$ of bounded and continuous functions $\mathbb{R}^q \to \mathbb{R}$ tending to zero at infinity when the coefficients $\sigma$ and $b$ are bounded.*

*Proof.*

1. Immediate for a Markov transition kernel;
2. Let $t \geq 0$, $f \in \mathcal{C}_b(\mathbb{R}^q, \mathbb{R})$, $x = \lim_{n\to\infty} x_n \in \mathbb{R}^q$. We have,
$$|\Pi_t(f)(x_n) - \Pi_t(f)(x)||\mathbb{E}(f(X_t(x_n))) - \mathbb{E}(f(X_t(x)))|.$$
Since $\mathbb{E}(|X_t(x_n)-X_t(x)|^2) \leq C|x_n-x|^2$, it follows that $\lim_{n\to\infty} X_t(x_n) = X_t(x)$ in $\mathrm{L}^2$, and thus in law, and therefore $\lim_{n\to\infty} \Pi_t(f)(x_n) = \Pi_t(f)(x)$, which implies $\Pi_t(f) \in \mathcal{C}_b(\mathbb{R}^q, \mathbb{R})$.
3. Let $f \in \mathcal{C}_0(\mathbb{R}^q, \mathbb{R})$ and $\varepsilon > 0$. There exists $A > 0$ such that for all $y \in \mathbb{R}^q$ such that $|y| > A$, we have $|f(y)| < \varepsilon$. Let $(X_t(x))_{t\geq 0}$ be the solution of the stochastic differential equation associated to the semi-group, namely
$$X_t(x) = x + \int_0^t \sigma(X_s(x))\mathrm{d}B_s + \int_0^t b(X_s(x))\mathrm{d}s.$$
We have, for all $x \in \mathbb{R}^q$ such that $|x| > B > A$, using the Markov inequality,
$$\begin{aligned}|\mathbb{E}(f(X_t(x))| &\leq \mathbb{E}(|f(X_t(x))|\mathbf{1}_{|X_t(x)|>A}) + \|f\|_\infty \mathbb{P}(|X_t(x)| \leq A)\\ &\leq \varepsilon + \|f\|_\infty \mathbb{P}\left(\left|\int_0^t \sigma(X_s(x))\mathrm{d}B_s + \int_0^t b(X_s(x))\mathrm{d}s\right| \geq B - A\right)\\ &\leq \varepsilon + \frac{\|f\|_\infty}{(B-A)^2}\mathbb{E}\left(\left|\int_0^t \sigma(X_s(x))\mathrm{d}B_s + \int_0^t b(X_s(x))\mathrm{d}s\right|^2\right)\\ &\leq \varepsilon + 2\frac{\|f\|_\infty}{(B-A)^2}(\||\sigma|\|_\infty^2 t + (\||b|\|_\infty t)^2)\\ &\leq 2\varepsilon \text{ for } B \text{ sufficiently large.}\end{aligned}$$

■

Let $\mathcal{C}^2(\mathbb{R}^q, \mathbb{R})$ be the space of functions $\mathbb{R}^q \to \mathbb{R}$ of class $\mathcal{C}^2$ *i.e.* twice differentiable with continuous second derivative (Hessian). We define the second order linear differential operator without constant term $L : \mathcal{C}^2(\mathbb{R}^q, \mathbb{R}) \to \mathcal{C}(\mathbb{R}^q, \mathbb{R})$, by, for all $f \in \mathcal{C}^2(\mathbb{R}^2, \mathbb{R})$ and all $x \in \mathbb{R}^q$,

$$L(f)(x) = \frac{1}{2}\sum_{i,j=1}^q a_{i,j}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum_{i=1}^q b_i(x)\frac{\partial f}{\partial x_i}(x), \tag{L}$$

where $b(x) = (b_1(x), \ldots, b_q(x))$ and $a(x) = \sigma(x)(\sigma(x))^\top$ in other words

$$a_{i,j}(x) = \sum_{k=1}^d \sigma_{i,k}(x)\sigma_{j,k}(x).$$

For all $x \in \mathbb{R}^q$, the matrix $a(x)$ is symmetric positive semi-definite, *i.e.* for all $y \in \mathbb{R}^q$,

$$a(x)y \cdot y = |(\sigma(x)^\top)y|^2 \geq 0.$$

**Theorem 5.11** (Martingale and Duhamel[1] formula)**.** *Let $x \in \mathbb{R}^q$ and let $(X_t(x))_{t\geq 0}$ be the regular solution of the SDE as in Theorem 5.9. Then for all $f \in \mathcal{C}^2(\mathbb{R}^q, \mathbb{R})$, the following process is an $(\mathcal{F}_t)_{t\geq 0}$ local martingale issued from the origin:*

$$M^f = (M_t^f)_{t\geq 0} = \Big(f(X_t(x)) - f(x) - \int_0^t Lf(X_s(x))\mathrm{d}s\Big)_{t\geq 0}.$$

*Moreover if $f \in \mathcal{C}_b^2(\mathbb{R}^q, \mathbb{R})$* i.e. *has bounded first and second order derivatives then $M^f$ is an $(\mathcal{F}_t)_{t\geq 0}$ martingale and we have the Duhamel formula:*

$$\Pi_t(f)(x) = f(x) + \int_0^t \Pi_s(L(f))(x)\mathrm{d}s, \quad t \geq 0.$$

*Proof.* Let $f \in \mathcal{C}^2(\mathbb{R}^q, \mathbb{R})$. The Itô formula of Theorem 4.4 applied to $f(X_t(x))$ gives

$$\begin{aligned} f(X_t(x)) = f(x) & \\ &+ \sum_{i=1}^q \int_0^t \frac{\partial f}{\partial x_i}(X_s(x))\mathrm{d}M_s^x \\ &+ \sum_{i=1}^q \int_0^t \frac{\partial f}{\partial x_i}(X_s(x))b_i(X_s(x))\mathrm{d}s \\ &+ \frac{1}{2}\sum_{i,j=1}^q \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s(x))\mathrm{d}\langle M^i, M^j\rangle_s, \end{aligned}$$

where

$$M_t^x = \sum_{k=1}^d \int_0^t \sigma_{i,k}(X_s(x))\mathrm{d}B_s^k.$$

We have

$$\begin{aligned} \langle M^i, M^j\rangle &= \sum_{k,\ell=1}^d \int_0^t \sigma_{i,k}(X_s(x))\sigma_{j,\ell}(X_s(x))\mathrm{d}\langle B^k, B^\ell\rangle_s \\ &= \sum_{k=1}^d \int_0^t \sigma_{i,k}(X_s(x))\sigma_{j,k}(X_s(x))\mathrm{d}s \\ &= \int_0^t a_{i,j}(X_s(x))\mathrm{d}s. \end{aligned}$$

Therefore

$$\begin{aligned} M_t^f &= f(X_t(x)) - f(x) - \int_0^t L(f)(X_s(x))\mathrm{d}s \\ &= \sum_{i=1}^q \int_0^t \frac{\partial f}{\partial x_i}(X_s(x))\mathrm{d}M_s^i \end{aligned}$$

is an $(\mathcal{F}_t)_{t\geq 0}$ local martingale. Moreover if $f \in \mathcal{C}_b^2(\mathbb{R}^q, \mathbb{R})$ then we can easily check that

$$\mathbb{E}\int_0^t \Big(\frac{\partial f}{\partial x_i}(X_s(x))\Big)^2 \mathrm{d}\langle M^i\rangle_s < \infty, \quad 1 \leq i \leq q,$$

and then $M^f$ is a martingale. The Duhamel formula is immediate. ■

**Corollary 5.12** (Infinitesimal generator of Markov semi-group)**.** *Let us equip the space $\mathcal{C}_0(\mathbb{R}^q, \mathbb{R})$ with the uniform norm $\|f\| = \|f\|_\infty = \sup_{x\in\mathbb{R}^q} |f(x)|$. The following properties hold true:*

[1]Named after Jean-Marie Duhamel (1797 – 1872), French mathematician.

1. ***Continuity.*** *For all* $f \in \mathcal{C}_0(\mathbb{R}^q, \mathbb{R})$, $\lim_{t\to 0^+} \|\Pi_t(f) - f\| = 0$*;*
2. ***Differentiability.*** *For all* $f \in \mathcal{C}_c^2(\mathbb{R}^q, \mathbb{R})$ i.e. $\mathcal{C}^2(\mathbb{R}^q, \mathbb{R})$ *with compact support,*

$$\lim_{t\to 0^+} \left\| \frac{\Pi_t(f) - f}{t} - Lf \right\| = 0.$$

*Proof.*

1. Suppose first that $f \in \mathcal{C}_c(\mathbb{R}^q, \mathbb{R})$. We have then

$$\Pi_t(f)(x) - f(x) = \int_0^t \Pi_s(L(f))(x)\mathrm{d}s = \mathbb{E}\int_0^t L(f)(X_s(x))\mathrm{d}s$$

and thus $\|\Pi(f) - f\| \leq t\|Lf\| \to 0$ as $t \to \infty$. Now if $f \in \mathcal{C}_0(\mathbb{R}^q, \mathbb{R})$, then, for all $\varepsilon > 0$ and $g \in \mathcal{C}_c(\mathbb{R}^q, \mathbb{R})$ such that $\|f - g\| \leq \varepsilon$, we have, for sufficiently small $t > 0$,

$$\|\Pi_t(f) - f\| \leq \|\Pi_t(f - g)\| + \|\Pi_t(g) - g\| + \|g - f\| \leq 2\varepsilon + \|\Pi_t(g) - g\| \leq 3\varepsilon.$$

2. For all $f \in \mathcal{C}_c^2(\mathbb{R}^q, \mathbb{R})$ and all $t > 0$, we have, using the first part for the last step,

$$\left\| \frac{\Pi_t(f) - f}{t} - Lf \right\| = \left\| \frac{1}{t}\int_0^t (\Pi_s(L(f)) - L(f))\mathrm{d}s \right\| \leq \sup_{s\in[0,t]} \|\Pi_s(L(f)) - L(f)\| \underset{t\to 0^+}{\longrightarrow} 0.$$

■

**Theorem 5.13** (Strong Markov property)**.** *Let* $x \in \mathbb{R}^q$ *and* $(X_t(x))_{t\geq 0}$ *be the regular solution of the SDE as in Theorem 5.9. Let* $T$ *be an* $(\mathcal{F}_t)_{t\geq 0}$ *stopping time and let* $\mathcal{F}_T$ *be its stopping* $\sigma$*-algebra. The following properties hold true:*

1. *for all bounded measurable* $f : \mathbb{R}^q \to \mathbb{R}$*, and all* $t \geq 0$*,*

$$\mathbb{E}(f(X_{T+t}(x))\mathbf{1}_{T<\infty} \mid \mathcal{F}_T) = \Pi_t(f)(X_T(x))\mathbf{1}_{T<\infty} \quad a.s.;$$

2. *for all bounded measurable* $\Phi : \mathcal{C}(\mathbb{R}_+, \mathbb{R}^q) \to \mathbb{R}$*,*

$$\mathbb{E}(\Phi((X_{T+s}(x))_{s\geq 0})\mathbf{1}_{T<\infty} \mid \mathcal{F}_T) = \Psi(X_T(x))\mathbf{1}_{T<\infty} \quad a.s.$$

*where the measurable function* $\Psi : \mathbb{R}^q \to \mathbb{R}$ *is defined for all* $y \in \mathbb{R}^q$ *by*

$$\Psi(y) = \mathbb{E}\Phi((X_s(y))_{s\geq 0}).$$

*Proof.*

1. Suppose first that $T$ takes its values in an at most countable set $\Im(T) \subset [0, \infty]$. We have to show that for all $A \in \mathcal{F}_T$ and for all $t \geq 0$,

$$\mathbb{E}(f(X_{T+t})\mathbf{1}_{A\cap\{T<\infty\}}) = \mathbb{E}(\Pi_t(f)(X_T(x))\mathbf{1}_{A\cap\{T<\infty\}}).$$

Indeed, using the simple Markov property of Theorem 5.9, the left hand side is equal to

$$\begin{aligned}\sum_{r\in\Im(T)\setminus\{\infty\}} \mathbb{E}(f(X_{r+t}(x))\mathbf{1}_{A\cap\{T=r\}}) &= \sum_{r\in\Im(T)\setminus\{\infty\}} \mathbb{E}(\Pi_t(f)(X_r(x))\mathbf{1}_{A\cap\{T=r\}}) \\ &= \mathbb{E}(\Pi_t(f)(X_T(x))\mathbf{1}_{A\cap\{T<\infty\}}).\end{aligned}$$

Suppose now that $T$ takes arbitrary values in $[0,\infty]$. It suffices to prove the desired property for all bounded continuous $f$. Let us define, for all $n \geq 0$, the stopping time

$$T_n = \sum_{k\geq 0} \frac{k+1}{2^n} \mathbf{1}_{[k/2^n,(k+1)/2^n[}(T) + \infty \mathbf{1}_{T=\infty}.$$

We have $T_n \searrow T$. For all $n \geq 0$ and all $A \in \mathcal{F}_{T_n}$, we get, from the first part of the proof,

$$\mathbb{E}(f(X_{T_n+t}(x))\mathbf{1}_{A\cap\{T_n<\infty\}}) = \mathbb{E}(\Pi_t(f)(X_{T_n}(x))\mathbf{1}_{A\cap\{T_n<\infty\}}).$$

By letting $n \to \infty$ and using the Lebesgue dominated convergence theorem, we obtain,

$$\mathbb{E}(f(X_{T+t}(x))\mathbf{1}_{A\cap\{T<\infty\}}) = \mathbb{E}(\Pi_t(f)(X_T(x))\mathbf{1}_{A\cap\{T<\infty\}}),$$

where we used Theorem 5.10 to get

$$\Pi_t(f)(X_{T_n}(x))\mathbf{1}_{T_n<\infty} \underset{n\to\infty}{\longrightarrow} \Pi_t(f)(X_T(x))\mathbf{1}_{T<\infty} \quad \text{a.s.}$$

2. Suppose that $\Phi$ is cylindrical, in the sense that for some $n \geq 1$, some $s_n \geq \cdots \geq s_1 \geq 0$, and some bounded measurable $f_1, \ldots, f_n : \mathbb{R}^q \mapsto \mathbb{R}$, we have, for all $w \in W$,

$$\Phi(w) = f_1(w_{s_1}) \cdots f_n(w_{s_n}).$$

We have in this case to show that

$$\mathbb{E}(f_1(X_{T+s_1}(x)) \cdots f_n(X_{T+s_n}(x))\mathbf{1}_{T<\infty} \mid \mathcal{F}_T) = \Psi(X_T(x))\mathbf{1}_{T<\infty} \quad \text{a.s.}$$

where $\Psi : \mathbb{R}^q \to \mathbb{R}$ is the function defined for all $y \in \mathbb{R}^q$ by

$$\Psi(y) = \mathbb{E}(f_1(X_{s_1}(y)) \cdots f_n(X_{s_n}(y))).$$

Indeed, for $n = 1$, this is the first property of the Theorem that we have already proved. We then proceed by induction on $n$, and suppose that it is already proved for some $n \geq 1$. Let us prove it for $n+1$. We have, denoting for short $Y_i = f_i(X_{T+s_i}(x))$,

$$\begin{aligned}\mathbb{E}(Y_1 \cdots Y_n Y_{n+1}\mathbf{1}_{T<\infty} \mid \mathcal{F}_T) &= \mathbb{E}(Y_1 \cdots Y_n \mathbb{E}(Y_{n+1}\mathbf{1}_{T<\infty} \mid \mathcal{F}_{s_n+T}) \mid \mathcal{F}_T)\\ &= \mathbb{E}(Y_1 \cdots Y_n \Pi_{s_{n+1}-s_n}(f_{n+1})(X_{T+s_n}(y))\mathbf{1}_{T<\infty} \mid \mathcal{F}_T)\\ &= \Psi(X_T(x))\mathbf{1}_{T<\infty}\end{aligned}$$

where

$$\Psi(y) = \mathbb{E}(f_1(X_{s_1}(y)) \cdots f_n(X_{s_n}(y))\Pi_{s_{n+1}-s_n}(f_{n+1})(X_{s_n}(y))).$$

But using the induction hypothesis and the simple Markov property of Theorem 5.9,

$$\begin{aligned}\Psi(y) &= \mathbb{E}(f_1(X_{s_1}(y)) \cdots f_n(X_{s_n}(y))\mathbb{E}(f_{n+1}(X_{s_{n+1}}(y)) \mid \mathcal{F}_{s_n}))\\ &= \mathbb{E}(f_1(X_{s_1}(y)) \cdots f_n(X_{s_n}(y))f_{n+1}(X_{s_{n+1}}(y))).\end{aligned}$$

This proves the property for all cylindrical $\Phi$. It remains to use a monotone class argument (see Section 1.3) to address the case of a general $\Phi$.

■

**Theorem 5.14** (Heat equation)**.** *Assume that $\sigma$ and $b$ are moreover $\mathcal{C}_b^2$, in other words $\mathcal{C}^2$ with bounded first and second derivatives. Let $L$ be the differential operator* (L)*. Then, for all $f \in \mathcal{C}_b^2(\mathbb{R}^q, \mathbb{R})$, there exists a unique $\Psi = (\Psi(t,x))_{(t,x)\in\mathbb{R}_+\times\mathbb{R}^q}$ solution of the following problem:*

1. $(t,x) \mapsto \Psi(t,x)$ *is* $\mathcal{C}^1$ *in* $t$ *and* $\mathcal{C}_b^2$ *in* $x$*;*
2. *for all* $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^q$*,*

$$\frac{\partial \Psi}{\partial t}(x) = L(\Psi(t,\cdot))(x) = \frac{1}{2}\sum_{i,j=1}^{q} a_{i,j}(x)\frac{\partial^2 \Psi}{\partial x_i \partial x_j}(t,x) + \sum_{i=1}^{q} b_i(x)\frac{\partial \Psi}{\partial x_i}(t,x);$$

3. *for all* $x \in \mathbb{R}^q$*,* $\Psi(0,x) = f(x)$*.*

*Moreover* $\Psi(x,t)$ *is given by the following formula:*

$$\Psi(t,x) = \mathbb{E}(f(X_t(x))) = \Pi_t(f)(x)$$

*where* $(X_t(x))_{t\geq 0}$ *is the solution of the stochastic differential equation as in Theorem 5.9.*

Theorem 5.14 states that the infinitesimal generator $L$ in (L) determines the Markov semigroup $(\Pi_t)_{t\geq 0}$ which characterizes the law of the Markov diffusion process $(X_t(x))_{t\geq 0}$.

*Proof.* We admit the following result, which follows from the assumptions made on $\sigma$ and $b$, see [6]: for all $f \in \mathcal{C}_b^2(\mathbb{R}^q,\mathbb{R})$, the quantity $\Pi_t(f)(x)$ is $\mathcal{C}_b^2$ in $x$, while $\Pi_t(f)(x)$ is $\mathcal{C}^1$ in $t$ as we can check on the Duhamel formula of Theorem 5.11:

$$\Pi_t(f)(x) = f(x) + \int_0^t \Pi_s(L(f))(x)\mathrm{d}s.$$

Let $t \geq u > 0$. The Itô formula of Theorem 4.4 gives

$$\begin{aligned}\mathbb{E}(f(X_u(x)) \mid \mathcal{F}_t) &= \Pi_{u-t}(f)(X_t(x))\\ &= \Pi_u(f)(x) + N_t + \int_0^t \Big(-\frac{\partial}{\partial u}\Pi_{u-s}(f)(X_s(x)) + L(\Pi_{u-s}(f))(X_s(x))\Big)\mathrm{d}s\end{aligned}$$

where $(N_t)_{t\geq 0}$ is a continuous local martingale. It follows that

$$\Big(\int_0^t \Big(-\frac{\partial}{\partial u}\Pi_{u-s}(f)(X_s(x)) + L(\Pi_{u-s}(f))(X_s(x))\Big)\mathrm{d}s\Big)_{t\geq 0}$$

is a continuous local martingale, with finite variation, issued from the origin, and therefore identically equal to zero, and thus for all $s \in [0,u]$,

$$-\frac{\partial}{\partial u}\Pi_{u-s}(f)(X_s(x)) + L(\Pi_{u-s}(f))(X_s(x)) = 0.$$

In particular, for $s = 0$, we get

$$\frac{\partial}{\partial u}\Pi_u(f)(x) = L(\Pi_u(f))(x).$$

Therefore the formula for $\Psi$ which appears at the end of the statement of the Theorem provides a solution to the problem (heat equation) considered in the statement of the Theorem, since $\Pi_0(f)(x) = f(x)$. Conversely, if $\Psi(t,x)$ is a solution to this problem then for all $u > 0$ the Itô formula of Theorem 4.4 for the process $(\Psi(u-t,X_t(x)))_{t\in[0,u]}$ gives

$$\begin{aligned}\Psi(0,X_u(x)) &= f(X_u(x))\\ &= \Psi(u,x) + \widetilde{N}_u + \int_0^u \Big[-\frac{\partial}{\partial u}\Psi(u-t,X_t(x)) + L(\Psi(u-t,\cdot))(X_t(x))\Big]\mathrm{d}t\\ &= \Psi(u,x) + \widetilde{N}_u\end{aligned}$$

where $(\widetilde{N}_u)_{u\geq 0}$ is a stochastic integral with zero expectation, and therefore

$$\Psi(u,x) = \mathbb{E}(f(X_u(x))) = \Pi_u(f)(x).$$

■

## 5.4 Locally Lipschitz coefficients and explosion time

We assume in this section that $\sigma$ and $b$ are non-random and constant in time, defined on $\mathbb{R}^q$ and taking values in $\mathcal{M}_{q,d}(\mathbb{R})$ and $\mathbb{R}^q$ respectively, and locally Lipschitz in the sense that for all bounded subset $K$ of $\mathbb{R}^q$ there exists a constant $C_K > 0$ such that for all $x, y \in K$,

$$|\sigma(x) - \sigma(y)| + |b(x) - b(y)| \leq C_K |x - y|.$$

This is the case for instance when $\sigma$ and $b$ are $\mathcal{C}^1$ on $\mathbb{R}^q$.

Under this regularity assumption, for all $n \geq 1$, we can find two applications $\sigma_n$ and $b_n$ from $\mathbb{R}^q$ to $\mathcal{M}_{q,d}(\mathbb{R})$ and $m\mathbb{R}^q$ respectively such that the following properties hold true:

1. for all $x \in \mathbb{R}^q$ with $|x| \leq n$, we have $\sigma_n(x) = \sigma(x)$ and $b_n(x) = b(x)$;
2. there exists a constant $D_n > 0$ such that for all $x, y \in \mathbb{R}^q$,

$$|\sigma_n(x) - \sigma(x)| + |b_n(x) - b(x)| \leq D_n |x - y|.$$

Beware that these assumptions are not a specialization of the general assumptions made at the beginning of the chapter, since locally Lipschitz is strictly more general than Lipschitz. The main problem with these assumptions on $\sigma$ and $b$ is that the stochastic differential equation

$$X_t(x) = x + \int_0^t \sigma(X_s(x)) \mathrm{d}B_s + \int_0^t b(X_s(x)) \mathrm{d}s$$

may not have a solution $X_t(x)$ for all time $t$, and an explosion may occur in finite (random) time. In this case a simple way to still define a solution for all time is to use a localization procedure in order to define the solution process before explosion, and then to stick the process to an extra point at infinity after explosion. This suggests to consider the Alexandroff[2] compactification $\widehat{\mathbb{R}^q} = \mathbb{R}^q \cup \{\infty\}$ of $\mathbb{R}^q$ obtained by adding to $\mathbb{R}^q$ a point at infinity denoted $\infty$. The neighborhoods of $\infty$ in $\widehat{\mathbb{R}^q}$ are the complements of the closed proper subsets of $\mathbb{R}^q$.

**Theorem 5.15** (Solving SDE with locally Lipschitz coefficients). *For all $x \in \mathbb{R}^q$, there exists a unique couple $(X^x, \xi^x)$ where $\xi^x$ is a stopping time taking values in $(0, \infty]$ called explosion time and where $X^x = (X_t(x))_{t\geq 0}$ is an adapted process such that the following properties hold true:*

1. *a.s. the path $t \mapsto X_t(x)$ is continuous from $[0, \xi^x)$ to $\mathbb{R}^q$ and $X_t(x) = \infty$ for all $t \geq \xi^x$;*
2. *almost surely, on the event $\{\xi^x < \infty\}$,*

$$\lim_{t \overset{<}{\to} \xi^x} |X_t(x)| = +\infty;$$

3. *for all stopping time $T$ such that $\{T < \xi^x\}$ almost surely on $\{\xi^x < \infty\}$,*

$$X_{t\wedge T}(x) = x + \int_0^t \mathbf{1}_{s\leq T} \sigma(X_s(x)) \mathrm{d}B_s + \int_0^t \mathbf{1}_{s\leq T} b(X_s(x)) \mathrm{d}s \quad a.s., \quad t \geq 0.$$

Before giving the proof of Theorem 5.15, let us prepare some ingredients.

For all $n \geq 1$ and $x \in \mathbb{R}^q$, let $(X_t^n(x))_{t\geq 0}$ be the solution of the stochastic differential equation

$$X_t^n(x) = x + \int_0^t \sigma_n(X_s^n(x)) \mathrm{d}B_s + \int_0^t b_n(X_s^n(x)) \mathrm{d}s \quad \text{a.s.,} \quad t \geq 0.$$

For all $m \geq 1$, let us define the stopping time

$$T_n^m = \inf\{t \geq 0 : |X_t^n(x)| \geq m\}.$$

[2]Named after Pavel Alexandrov (1896 – 1982), Russian mathematician.

**Lemma 5.16.** *For all $m \geq n \geq 1$ and $x \in \mathbb{R}^q$, we have*

$$T_n^n(x) = T_m^n(x) \leq T_m^m(x) \quad a.s.,$$

*and, for all $t \in [0, T_n^n(x)]$, we have*

$$X_t^n(x) = X_t^m(x) \quad a.s.$$

*Proof of Lemma 5.16.* Let us define $T = T_n^n(x) \wedge T_m^n(x)$. We have

$$X_{t\wedge T}^n = x + \int_0^t \mathbf{1}_{s\leq T}\sigma_n(X_{s\wedge T}^n)\mathrm{d}B_s + \int_0^t \mathbf{1}_{s\leq T}b_n(X_{s\wedge T}^n)\mathrm{d}s.$$

and

$$X_{t\wedge T}^m = x + \int_0^t \mathbf{1}_{s\leq T}\sigma_m(X_{s\wedge T}^m)\mathrm{d}B_s + \int_0^t \mathbf{1}_{s\leq T}b_m(X_{s\wedge T}^m)\mathrm{d}s.$$

By definition of $T$, the processes $(X_{t\wedge T}^n)_{t\geq 0}$ and $(X_{t\wedge T}^m)_{t\geq 0}$ solve the same SDE

$$Z_t = x + \int_0^t \mathbf{1}_{s\leq T}\sigma(Z_s)\mathrm{d}B_s + \int_0^t \mathbf{1}_{s\leq T}b(Z_s)\mathrm{d}s,$$

and thus $X_{t\wedge T}^n(x) = X_{t\wedge T}^m(x)$ a.s. for all $t \geq 0$. On the event $\{0 < T < \infty\}$, then for all $t \in [0, T)$,

$$|X_t^m(x)| = |X_t^n(x)| < n \quad \text{and} \quad [X_T^m(x)| = |X_T^n(x)| = n$$

therefore $T = T_n^n(x = T_m^n(x)$. Now, if $T = 0$, then $|x| \geq n$ and $T = T_n^n(x) = T_m^n(x) = 0$, while if $T = \infty$, then $T_n^n(x) = T_m^n(x) = \infty$. ■

*Proof of Theorem 5.15. Proof of existence.* We set $\xi^x = \sup_{n\geq 0} T_n(x)$ where $T_n(x) = T_n^n(x)$. We check immediately that if $|x| < n$ then $T_n(x) > 0$ and thus $\xi^x > 0$. Let $t \in [0, \xi^x)$. There exists $n$ such that $T_n(x) > t$ and for all $m \geq n$, we have $X_t^m(x) = X_t^n(x)$ almost surely. From Lemma 5.16, we can then define

$$X_t(x) = \lim_{n\to\infty} X_t^n(x)\mathbf{1}_{[0,\xi^x(t)} + \infty\mathbf{1}_{[\xi^x,\infty)}(t).$$

This process $(X_t(x))_{t\geq 0}$ verifies the first property stated by the Theorem. Moreover, on the event $\{\xi^x < \infty\}$, we have $T_n(x) < T_{n+1}(x) < \cdots < \xi^x$ and $|X_{T_n(x)}| = n$, and therefore, almost surely, on the event $\{\xi^x < \infty\}$,

$$\overline{\lim_{t \overset{<}{\to} \xi^x}} |X_t(x)| = +\infty.$$

Suppose that

$$\mathbb{P}(\underline{\lim_{t \overset{<}{\to} \xi^x}} |X_t(x)| < \infty; \xi^x < \infty) > 0,$$

then we can find real numbers $r$ and $R$ such that $0 < r < R < \infty$ and

$$\mathbb{P}(\underline{\lim_{t \overset{<}{\to} \xi^x}} |X_t(x)| < r; \overline{\lim_{t \overset{<}{\to} \xi^x}} |X_t(x)| > R; \xi^x < \infty) > 0. \qquad (\star)$$

Let $f \in \mathcal{C}_c^2(\mathbb{R}^q, \mathbb{R})$ be such that $f(x) = 0$ if $|x| = r$ and $f(x) = 1$ if $|x| = R$. If $L$ is the differential operator (L) then, for $n$ sufficiently large,

$$\left(f(X_t^n(x)) - \int_0^t L(f)(X_s^n(x))\mathrm{d}s\right)_{t\geq 0}$$

is a martingale, and thus

$$\left(f(X^n_{t\wedge T_m}(x)) - \int_0^{t\wedge T_m} L(f)(X^n_s(x))\mathrm{d}s\right)_{t\geq 0}$$

is a martingale. By letting $n\to\infty$, the Lebesgue dominated convergence theorem gives that

$$\left(f(X_{t\wedge T_m}(x)) - \int_0^{t\wedge T_m} L(f)(X_s(x))\mathrm{d}s\right)_{t\geq 0}$$

is a martingale. Now by letting $m\to\infty$, we see similarly that

$$\left(f(X_t(x))\mathbf{1}_{t<\xi^x} - \int_0^{t\wedge \xi^x} L(f)(X_s(x))\mathrm{d}s\right)_{t\geq 0}$$

is a martingale since $t<\xi^x$ implies $X_{t\wedge T_m(x)}(x)\to X_t(x)$ as $m\to\infty$ while $t\geq\xi^x$ implies $f(X_{t\wedge T_m(x)}(x)) = f(X_{T_m(x)}(x))\to 0$ as $m\to\infty$. We can check (exercise) that the above martingale is continuous, and it follows then that the process $(f(X_t(x))(x))\mathbf{1}_{t<\xi^x}{}_{t\geq 0}$ is continuous martingale, which contradicts $(\star)$ and the definition of $f$. It follows that the second property stated in the Theorem holds true: almost surely, on $\{\xi^x<\infty\}$,

$$\lim_{t\overset{<}{\to}\xi^x}|X_t(x)| = +\infty.$$

Furthermore, the third and last property stated in the Theorem can be deduced as follows: if $T$ is a stopping time such that $T<\xi^x$ almost surely on $\{\xi^x<\infty\}$ then for all $n\geq 1$ we have

$$X^n_{t\wedge T\wedge T_n} = x + \int_0^t \mathbf{1}_{s\leq T\wedge T_n}\sigma(X^n_s)\mathrm{d}B_s + \int_0^t \mathbf{1}_{s\leq T\wedge T_n} b(X^n_s)\mathrm{d}s,$$

thus,

$$X_{t\wedge T\wedge T_n} = x + \int_0^t \mathbf{1}_{s\leq T\wedge T_n}\sigma(X_s)\mathrm{d}B_s + \int_0^t \mathbf{1}_{s\leq T\wedge T} b(X_s)\mathrm{d}s,$$

and the desired result follows by letting $n\to\infty$.

*Proof of uniqueness.* Exercise. ■

## 5.5 Girsanov theorem

This section provides an analogue of Theorem 4.9 for stochastic differential equations. More precisely, it gives the mutual density of the law of the solution of a stochastic differential equation for different drifts and same diffusion coefficient, in particular it gives the density of the law of the solution with respect to the law of the driving Brownian motion.

Let $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq 0},\mathbb{P})$ be a filtered probability space, with $(\mathcal{F}_t)_{t\geq 0}$ complete and right continuous. Let $B=(B_t)_{t\geq 0}$ be a $d$-dimensional $(\mathcal{F}_t)_{t\geq 0}$ Brownian motion issued from the origin. Let us consider three Lipschitz maps

$$\sigma:\mathbb{R}^q\to\mathcal{M}_{q,d}(\mathbb{R}),\quad b:\mathbb{R}^q\to\mathbb{R}^q,\quad \widetilde{b}:\mathbb{R}^q\to\mathbb{R}^q.$$

For all $x\in\mathbb{R}^q$, let $(X_t(x))_{t\geq 0}$ and $(\widetilde{X}_t(x))_{t\geq 0}$ be the solutions of the SDE

$$X_t(x) = x + \int_0^t \sigma(X_s(x))\mathrm{d}B_s + \int_0^t b(X_s(x))\mathrm{d}s$$

and

$$\widetilde{X}_t(x) = x + \int_0^t \sigma(\widetilde{X}_s(x))\mathrm{d}B_s + \int_0^t \widetilde{b}(\widetilde{X}_s(x))\mathrm{d}s.$$

Let $T>0$ be a fixed real number. Let $Q_{x,\sigma,b}$ and $Q_{x,\sigma,\widetilde{b}}$ be the respective laws of the processes $(X_t)_{t\in[0,T]}$ and $(\widetilde{X}_t)_{t\in[0,T]}$ on the canonical space $(W=\mathcal{C}([0,T],\mathbb{R}^q),\mathcal{B}_T,(\pi_t)_{t\in[0,T]})$ where we have $\mathcal{B}_T=\sigma(\pi_s:s\in[0,T])$ and $\pi_s(w)=w_s$ for all $w\in W$ and all $s\in[0,T]$.

**Theorem 5.17** (Girsanov theorem for SDE solutions)**.** *If for all $x \in \mathbb{R}^q$, $b(x)-\widetilde{b}(x) = \sigma(x)\varphi(x)$ where $\varphi : \mathbb{R}^q \to \mathbb{R}^q$ is measurable and bounded, then the probability measures $Q_{x,\sigma,b}$ and $Q_{x,\sigma,\widetilde{b}}$ are* <u>*equivalent*</u> *(*i.e. *mutually absolutely continuous) on the measurable space $(\Omega, \mathcal{B}_T)$, and*

$$\mathrm{d}Q_{x,\sigma,b} = D\mathrm{d}Q_{x,\sigma,\widetilde{b}}$$

*where the density $D$ is given for all $w \in W$ by*

$$D(w) = \exp\Big(\int_0^T (a(w_s))^{-1}(b(w_s) - \widetilde{b}(w_s)) \cdot \mathrm{d}w_s \\ - \frac{1}{2}\int_0^T (a(w_s))^{-1}(b(w_s) - \widetilde{b}(w_s)) \cdot (b(w_s) - \widetilde{b}(w_s))\mathrm{d}s\Big) \quad a.s.$$

*where $a(x) = \sigma(x)(\sigma(x))^\top \in \mathcal{M}_{q,q}(\mathbb{R})$ for all $x \in \mathbb{R}^q$.*

*Proof.* Note first that the formula for $D$ makes sense since under $Q_{x,\sigma,\widetilde{b}}$ or $Q_{x,\sigma,b}$, the canonical process $(\pi_s(w))_{s\geq 0} = (w_s)_{s\geq 0}$ is a semi-martingale. Note also that $a(x)|_{\mathrm{image}(a(x))}$ is an isomorphism on image$(\sigma(x))$ and thus $a^{-1}(x)(b(x) - \widetilde{b}(x))$ is well defined for all $x \in \mathbb{R}^q$.

We can write

$$\begin{aligned}\widetilde{X}_t &= x + \int_0^t \sigma(\widetilde{X}_s(x))\mathrm{d}B_s + \int_0^t b(\widetilde{X}_s(x))\mathrm{d}s - \int_0^t (b(\widetilde{X}_s(x)) - \widetilde{b}(\widetilde{X}_s(x)))\mathrm{d}s \\ &= x + \int_0^t \sigma(\widetilde{X}_s(x))(\mathrm{d}B_s - \varphi(\widetilde{X}_s(x))\mathrm{d}s) + \int_0^t b(\widetilde{X}_s(x))\mathrm{d}s.\end{aligned}$$

Let $\widetilde{\mathbb{P}}$ be the probability measure on $(\Omega, \mathcal{F}_T)$ given by

$$\frac{\mathrm{d}\widetilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}} = \exp\Big(\int_0^T \varphi(\widetilde{X}_s(x))\mathrm{d}B_s - \frac{1}{2}\int_0^T |\varphi(\widetilde{X}_s(x))|^2\mathrm{d}s\Big) = Z.$$

Then, under $\widetilde{\mathbb{P}}$, the process $\Big(B_t - \int_0^t \varphi(\widetilde{X}_s(x))\mathrm{d}s\Big)_{t\in[0,T]}$ is an $(\mathcal{F}_t)_{t\in[0,T]}$ Brownian motion. Therefore, from the uniqueness in law provided by Theorem 5.8, the law of $\widetilde{X}$ under $\widetilde{\mathbb{P}}$ is equal to the law of $X$ under $\mathbb{P}$ which is $Q_{x,\sigma,b}$. Moreover the law of $\widetilde{X}$ under $\mathbb{P}$ is $Q_{x,\sigma,\widetilde{b}}$. For all bounded measurable $\Psi : W \to \mathbb{R}$, we have

$$\mathbb{E}_{\widetilde{\mathbb{P}}}(\Psi(\widetilde{X})) = \mathbb{E}_{\mathbb{P}}(\Psi(\widetilde{X})Z) = \int_W \Psi(w)\mathrm{d}Q_{x,\sigma,b}(w).$$

It remains to "compute" $Z$. Suppose first that $d = q$ and that $\sigma(x)$ is a invertible matrix. Then $b(x) - \widetilde{b}(x) = \sigma(x)\varphi(x)$ implies $\varphi(x) = (\sigma(x))^{-1}(b(x) - \widetilde{b}(x))$ and

$$|\varphi(x)|^2 = |(\sigma(x))^{-1}(b(x) - \widetilde{b}(x))|^2 = (a(x))^{-1}(b(x) - \widetilde{b}(x)) \cdot (b(x) - \widetilde{b}(x)),$$

and moreover

$$\mathrm{d}\widetilde{X}_t = \sigma(\widetilde{X}_t)\mathrm{d}B_t + \widetilde{b}(\widetilde{X}_t)\mathrm{d}t$$

gives

$$\mathrm{d}B_t = (\sigma(\widetilde{X}_t))^{-1}(\mathrm{d}\widetilde{X}_t - \widetilde{b}(\widetilde{X}_r)\mathrm{d}t)$$

and

$$\begin{aligned}\varphi(\widetilde{X}_t)\mathrm{d}B_t &= (\sigma^{-1}(\widetilde{X}_t)(b(\widetilde{X}_t) - \widetilde{b}(\widetilde{X}_t)) \cdot (\sigma(\widetilde{X}_t))^{-1}(\mathrm{d}\widetilde{X}_t - \widetilde{b}(\widetilde{X}_r)\mathrm{d}t) \\ &= ((a(\widetilde{X}_t))^{-1}(b(\widetilde{X}_t) - \widetilde{b}(\widetilde{X}_t)) \cdot \mathrm{d}\widetilde{X}_t \times (a(\widetilde{X}_t))^{-1}(b - \widetilde{b})(\widetilde{X}_t) \cdot \widetilde{b}(\widetilde{X}_t),\end{aligned}$$

hence

$$Z = \exp\Big( \int_0^T (a(\widetilde{X}_s))^{-1}(b(\widetilde{X}_s) - \widetilde{b}(\widetilde{X}_s)) \cdot \mathrm{d}\widetilde{X}_s - \frac{1}{2}\int_0^T (a(\widetilde{X}_s))^{-1}(b(\widetilde{X}_s) - \widetilde{b}(\widetilde{X}_s)) \cdot (b(\widetilde{X}_s) - \widetilde{b}(\widetilde{X}_s))\mathrm{d}s\Big)$$

which gives the desired formula by considering the image laws, namely

$$\int_W \Psi(w)\mathrm{d}Q_{x,\sigma,b} = \int_W \Psi(w)\exp\Big( \int_0^T (a(w_s))^{-1}(b(w_s) - \widetilde{b}(w_s)) \cdot \mathrm{d}w_s - \frac{1}{2}\int_0^T (a(w_s))^{-1}(b(w_s) - \widetilde{b}(w_s)) \cdot (b(w_s) - \widetilde{b}(w_s))\mathrm{d}s\Big)\mathrm{d}Q_{s,\sigma,\widetilde{b}}(w).$$

To address the general case on $\sigma$, we can assume that $\varphi(x) \in \mathrm{kernel}((\sigma(x))^\top)^\perp = \mathrm{image}((\sigma(x))^\top)$ and thus $\varphi(x) = (\sigma(x))^\top\widetilde{\varphi}(x)$ where $\widetilde{\varphi}(x) \in \mathrm{kernel}((\sigma(x))^\top)^\perp$. This gives

$$b(x) - \widetilde{b}(x) = \sigma(x)(\sigma(x))^\top\widetilde{\varphi}(x) = a(x)\widetilde{\varphi}(x)$$

and

$$|\varphi(x)|^2 = |(\sigma(x))^\top\widetilde{\varphi}(x)|^2 = a(x)\widetilde{\varphi}(x)\cdot\widetilde{\varphi}(x) = (b(x) - \widetilde{b}(x))\cdot(a(x))^{-1}(b(x) - \widetilde{b}(x)),$$

and

$$\begin{aligned}\varphi(x)\cdot\mathrm{d}B_t &= (\sigma(x))^\top\widetilde{\varphi}(x)\cdot\mathrm{d}B_t\\ &= \widetilde{\varphi}(x)\cdot\sigma(x)\mathrm{d}B_t\\ &= (a(x))^{-1}(b(x) - \widetilde{b}(x))\cdot(\mathrm{d}\widetilde{x}_t - \widetilde{b}(x)\mathrm{d}t).\end{aligned}$$

■

# Chapter 6

# Probabilistic formulation of Dirichlet problems

The simplest instance of the Dirichlet[1] problem on an open domain $D \subset \mathbb{R}^d$ consists in giving a function $\Psi : \partial D \to \mathbb{R}$ at the boundary $\partial D$ of $D$ and seeking for a function $f : \overline{D} = D \cup \partial D \to \mathbb{R}$ such that $\Delta f = 0$ on $D$ and $f = \Psi$ on $\partial D$.

More generally, the Dirichlet problem on an open domain $D \subset \mathbb{R}^d$ with boundary $\partial D$ writes

$$\begin{cases} \mathrm{L}u(x) - c(x)u(x) = f(x) & \text{for all } x \in D, \\ \lim\limits_{\substack{x \to x_0 \\ x \in D}} u(x) = \Psi(x) & \text{for all } x_0 \in \partial D, \end{cases} \tag{DirP}$$

where

$$\mathrm{L}u(x) = \frac{1}{2} \sum_{i,j=1}^{d} a_{i,j}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{d} b_i(x) \frac{\partial u}{\partial x_i}(x),$$

where for all $x \in \mathbb{R}^d$,

- $a(x) = (a_{i,j}(x))_{1 \le i,j \le d} = \sigma(x)\sigma^*(x)$ where $\sigma(x)$ is a $d \times q$ matrix, Lipschitz in $x$;
- $b(x) = (b_i(x))_{1 \le i \le d}$ is a vector field, Lipschitz in $x$;
- $f : D \mapsto \mathbb{R}$ and $\Psi : \partial D \mapsto \mathbb{R}$ are continuous and bounded;
- $c : D \mapsto \mathbb{R}$ is continuous and non-negative.

Let us consider the stochastic differential equation

$$X_t^x = x + \int_0^t \sigma(X_s^x) \mathrm{d}B_s + \int_0^t b(X_s^x) \mathrm{d}s, \quad t \ge 0, x \in \mathbb{R}^d \tag{SDE}$$

where $B = (B_s)_{s \ge 0}$ is an $(\mathcal{F}_s)_{s \ge 0}$ $d$-dimensional Brownian motion defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_s)_{s \ge 0}, \mathbb{P})$ satisfying the "usual" assumptions. We define the stopping time

$$T = T_D^x = \inf\{t \ge 0 : X_t^x \in \partial D\},$$

with the usual convention $\inf \varnothing = +\infty$.

**Theorem 6.1** (Kakutani[2] probabilistic representation of the Dirichlet problem)**.** *If $\mathbb{E}(T_D^x) < \infty$ for all $x \in D$ and if $x \in D \mapsto u(x)$ is $\mathcal{C}_b^2(D)$ solution of* (DirP)*, then, for all $x \in D$,*

$$u(x) = \mathbb{E}\Big( \Psi(X_{T_D^x}^x) \exp\Big( \int_0^{T_D^x} c(X_s^x) \mathrm{d}s \Big) \Big) - \mathbb{E}\Big( \int_0^{T_D^x} \Big( f(X_s^x) \exp\Big( - \int_0^s c(X_u^x) \mathrm{d}u \Big) \Big) \mathrm{d}s \Big).$$

[1]Named after Peter Gustav Lejeune Dirichlet (1805 – 1859), German mathematician.
[2]Named after Shizuo Kakutani (1911 – 2004), Japanese mathematician.

*Proof.* Let us suppose first that $u$ can be extended to the whole space $\mathbb{R}^d$ as a $\mathcal{C}_b^2$ function. Let us consider the process $u(X_t^x)\exp(Y_t^x)$ where $(X_t^x)_{t\geq 0}$ is a solution of (SDE) and where $Y_t^x = -\int_0^t c(X_s^x)\mathrm{d}s$. The Itô formula of Theorem 4.4 gives, for all $t \geq 0$, almost surely,

$$
\begin{aligned}
u(X_t^x)\exp(Y_t^x) - u(x) = \int_0^t \langle \nabla u(X_s^x)\exp(Y_s^x), \sigma(X_s^x)\mathrm{d}B_s\rangle \\
+ \int_0^t f(X_s^x)\exp(Y_s^x)\mathrm{d}s \\
+ \int_0^t (\mathrm{L}u(X_s^x) - f(X_s^x) - c(X_s^x)u(X_s^x))\exp(Y_s^x)\mathrm{d}s\rangle.
\end{aligned}
$$

This equality remains valid if one replaces $t$ by the stopping time $t \wedge T_D^x$, which gives, thanks to the assumptions,

$$
\mathbb{E}\Big(u(X_{t\wedge T_D^x}^x)\exp\Big(-\int_0^{t\wedge T_D^x} c(X_s^x)\mathrm{d}s\Big)\Big) = u(x) + \mathbb{E}\Big(\int_0^{t\wedge T_D^x} f(X_s^x)\exp(Y_s^x)\mathrm{d}s\Big). \tag{6.1}
$$

The desired formula follows then by letting $t \to \infty$ and using dominated convergence.

Then general case on $u$ can be addressed as follows. We consider an increasing sequence $(D_n)_n$ of open domains with smooth boundary $\partial D_n$ such that $\overline{D_n} \subset D_{n+1} \subset D$ and $\cup_n D_n = D$. The solution $u_n$ of the Dirichlet problem on $D_n$ can be extended on the whole $\mathbb{R}^d$ into a $\mathcal{C}_b^2$ function, and $u_n = u|_{\overline{D_n}}$. We have then, for all $x \in D_n$,

$$
u_n(x) = u(x) = \mathbb{E}(u(X_{T_n^x}^x)\exp(Y_{T_n^x}^x)) - \mathbb{E}\Big(\int_0^{T_0^x} f(X_s^x)\exp(Y_s^x)\mathrm{d}s\Big)
$$

where $T_n^x = \inf\{t \geq 0 : X_t^x \in \partial D_n\} = \inf\{t \geq 0 : X_t^x \notin D_n\}$. We have $T_n^x \leq T_{n+1}^x \nearrow T^x$, $\mathbb{E}(T^x) < \infty$, and it suffices to let $n \to \infty$ to get the desired result. ∎

**Remark 6.2.**

1. *If there exists $a > 0$ such that $c(x) > a > 0$ for all $x \in D$ then the assumption $\mathbb{E}(T_D^x) < \infty$ for all $x \in D$ is useless. Namely, by letting $t \to \infty$ in* (6.1), *we get*

$$
u(x) = \mathbb{E}\Big(u(X_{t\wedge T_D^x}^x)\exp\Big(-\int_0^{t\wedge T_D^x} c(X_s^x)\mathrm{d}s\Big)\mathbf{1}_{T_D^x<\infty}\Big) - \mathbb{E}\Big(\int_0^{t\wedge T_D^x} f(X_s^x)\exp(Y_s^x)\mathrm{d}s\Big);
$$

2. *If $f = 0$ then the condition $\mathbb{E}(T_D^x) < \infty$ for all $x \in D$ can be replaced by $\mathbb{P}(T_D^x < \infty) = 1$ for all $x \in D$, and, from* (6.1),

$$
u(x) = \mathbb{E}\Big(\Psi(X_{T_D^x}^x)\exp\Big(-\int_0^{T_D^x} c(X_s^x)\mathrm{d}s\Big)\Big).
$$

3. *If, for some $a > 0$, we have, for all $x \in D$, $\mathbb{E}(\exp(aT_D^x)) < \infty$, then the probabilistic representation provided by Theorem 6.1 remains valid for all coefficient $c$ such that $\inf_x c(x) \geq -a$. In particular, if, for all $x \in D$, $\mathbb{P}(T_D^x \leq T(x)) = 1$ where $T(x)$ is deterministic, then it remains valid for all coefficients bounded below by c.*

4. *Suppose that $D$ is a open, bounded, with regular boundary and that* L *is non degenerate, then $\mathbb{E}(T_D^x) \leq c < \infty$ for all $x \in D$. For all continuous function $\Psi$ on $\partial D$, the solution $u$ of* (DirP) *exists with the properties required by Theorem 6.1 provided for example that $c$ and $f$ are Hölder and $c \geq 0$. In particular, for $f = c = 0$, the solution writes*

$$
u(x) = \mathbb{E}(\Psi(X_{T_D^x}^x)) = \int_{\partial D} \Psi(y)\Pi(x, \mathrm{d}y)
$$

*where* $\Pi(x,y)$*,* $x \in D$*,* $y \in \partial D$ *is the* *Poisson kernel* *of* $L$*, see [12]. It follows that the exit law* $\Pi(x, \mathrm{d}y) = \mathbb{P}(X^x_{T^x_D} \in \mathrm{d}y)$*, also known as the* *harmonic measure**, has density* $\Pi(x, \mathrm{d}y) = \Pi(x,y)\mathrm{d}y$*. The function* $(x,y) \mapsto \Pi(x,y)$ *is strictly positive,* $\mathcal{C}^2$ *in* $x \in D$*, and*

$$\begin{cases} \lim\limits_{\substack{x \to y_0 \\ x \in D}} \Pi(x,y) = 0 & \textit{for all } y, y_0 \in \partial D, y \neq y_0, \\ \lim\limits_{\substack{x \to y_0 \\ x \in D}} |\Pi(x,y_0)| = +\infty & \textit{for all } y_0 \in \partial D. \end{cases}$$

5. *In the proof of Theorem 6.1, we can simply assume that* (SDE) *admits a (weak) solution for all* $x \in D$*, the coefficients* $\sigma$ *and* $b$ *being supposed mesurable (not necessarily Lipschitz!). The hypothesis of continuity for* $f$*,* $c$*, and* $\Psi$ *are superfluous as well, and one can assume that they are just mesurable,* $f$ *and* $\Psi$ *being bounded and non-negative. This remark is useful for certain problems in stochastic control theory.*

6. *The probabilistic representation provided by Theorem 6.1 shows that it suffices to give* $\Psi$ *on a subset* $\partial_R D$ *of* $\partial D$ *such that* $\mathbb{P}(X^x_{T^x_D} \in \partial_R D) = 1$ *for all* $x \in D$*. The elements of such as subset are called* *regular points of the boundary**. This same representation shows also that the Dirichlet problem is ill posed if* $\Psi$ *is arbitrary outside* $\partial_R D$*.*

7. *The Itô formula of Theorem 4.4 can be generalized to functions which are not* $\mathcal{C}^2$ *but are differentiable in a weak sense, and it follows that the probabilistic representation provided by Theorem 6.1 remains valid in this general case.*

# Appendix A

# More material

FIXME:TODO

## A.1 Fokker–Planck equation

## A.2 Feynman–Kac formula

## A.3 Hamilton–Jacobi–Bellman equation

## A.4 Statistical inference of diffusion processes

## A.5 Euler–Maruyama numerical scheme

## A.6 Stratonovich integral

## A.7 Local time and Tanaka formula

## A.8 Exemples of stochastic processes

## A.9 Infinitesimal generator = simulation algorithm

### A.9.1 Ornstein–Uhlenbeck and Bakry–Émery processes

### A.9.2 Laguerre and Jacobi processes

### A.9.3 Bessel and Cox–Ingersoll–Ross processes

### A.9.4 Dyson Brownian motion and Dyson–Ornstein–Uhlenbeck process

### A.9.5 Fisher–Wright processes

### A.9.6 Diffusions with jumps and piecewise deterministic Markov processes

## A.10 Additive functionals, ergodic theorem, central limit theorem

# Bibliography


[1] G. W. Anderson, A. Guionnet & O. Zeitouni – *An introduction to random matrices*, Cambridge Studies in Advanced Mathematics, vol. 118, Cambridge University Press, Cambridge, 2010.

[2] D. Bakry, I. Gentil & M. Ledoux – *Analysis and geometry of Markov diffusion operators*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 348, Springer, Cham, 2014.

[3] C. Dellacherie & P.-A. Meyer – *Probabilities and potential*, North-Holland Mathematics Studies, vol. 29, North-Holland Publishing Co., Amsterdam-New York; North-Holland Publishing Co., Amsterdam-New York, 1978. 6

[4] — , *Probabilities and potential. B*, North-Holland Mathematics Studies, vol. 72, North-Holland Publishing Co., Amsterdam, 1982, Theory of martingales, Translated from the French by J. P. Wilson. 6

[5] — , *Probabilities and potential. C*, North-Holland Mathematics Studies, vol. 151, North-Holland Publishing Co., Amsterdam, 1988, Potential theory for discrete and continuous semigroups, Translated from the French by J. Norris. 6

[6] I. I. Gihman & A. V. Skorohod – *The theory of stochastic processes. III*, Springer-Verlag, Berlin-New York, 1979, Translated from the Russian by Samuel Kotz, With an appendix containing corrections to Volumes I and II, Grundlehren der Mathematischen Wissenschaften, 232. 57

[7] N. Ikeda & S. Watanabe – *Stochastic differential equations and diffusion processes*, second ed., North-Holland Mathematical Library, vol. 24, North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, 1989. ii

[8] I. Karatzas & S. E. Shreve – *Brownian motion and stochastic calculus*, second ed., Graduate Texts in Mathematics, vol. 113, Springer-Verlag, New York, 1991.

[9] D. Lamberton & B. Lapeyre – *Introduction to stochastic calculus applied to finance*, second ed., Chapman & Hall/CRC Financial Mathematics Series, Chapman & Hall/CRC, Boca Raton, FL, 2008.

[10] J.-F. Le Gall – *Brownian motion, martingales, and stochastic calculus*, french ed., Graduate Texts in Mathematics, vol. 274, Springer, [Cham], 2016.

[11] H. P. McKean – *Stochastic integrals*, AMS Chelsea Publishing, Providence, RI, 2005, Reprint of the 1969 edition, with errata. 18

[12] C. Miranda – *Partial differential equations of elliptic type*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 2, Springer-Verlag, New York-Berlin, 1970, Second revised edition. Translated from the Italian by Zane C. Motteler. 65

[13] D. Revuz & M. Yor – *Continuous martingales and Brownian motion*, third ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293, Springer-Verlag, Berlin, 1999.

[14] D. W. Stroock & S. R. S. Varadhan – *Multidimensional diffusion processes*, Classics in Mathematics, Springer-Verlag, Berlin, 2006, Reprint of the 1997 edition.

[15] B. Øksendal – *Stochastic differential equations*, sixth ed., Universitext, Springer-Verlag, Berlin, 2003, An introduction with applications.