Advanced continuous processes

2023 - 2024

François Simenhaus
Office B 640
simenhaus@ceremade.dauphine.fr

Contents

Translated by ChatGPT with almost no human control so far...

One lecture of 1.5 hours per week and a tutorial session of the same duration. The grade is determined by

$$\text{Grade} = \max(\text{Exam}\ ;\ 0.4\ \text{Midterm} + 0.6\ \text{Exam}).$$

These lecture notes are largely inspired by the notes and books of Fabrice Baudoin [1], Francis Comets [5], Jean-François Legall [7], Marc Yor and Daniel Revuz [8]. I also invite you to consult the lecture notes of Djalil Chafaï [4], Philippe Bougerol [3], and Nadine Guillotin [6], which you can easily find online.

Throughout this course, we consider a probability space $(\Omega, \mathcal{F}, \mathrm{P})$. All random objects considered are constructed on this space unless explicitly stated. We will denote E as the associated expectation. If X is a random variable and $A \in \mathcal{F}$ an event, we will sometimes use the notation (standard)

$$\mathrm{E}(X, A) := \mathrm{E}(X 1_A).$$

Definition 1. *A random process indexed by a set of indices T and taking values in a measurable space $(E, \mathcal{E})$ is a family $(X_t)_{t\in\mathsf{T}}$ of random variables defined on $(\Omega, \mathcal{F}, \mathrm{P})$ and taking values in $(E, \mathcal{E})$.*

Examples of common index sets:

1. $\mathsf{T} = \mathbb{N}$ and we then talk about *discrete processes.* This is the framework of the course you took in the first semester. The most important examples of such random processes are, of course, discrete-time martingales and Markov chains.

2. $\mathsf{T} = \mathbb{R}^2$ and we then talk about *random fields.* [1]

In this course, we will mainly focus on **continuous-time processes** taking values in $\mathbb{R}$:

$$\mathsf{T} = \mathbb{R}^+ \text{ and } (E, \mathcal{E}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

[1] See, for example, the Wikipedia page on the Gaussian free field for a famous example.

(sometimes $\mathsf{T} = \mathbb{R}$ and $(E, \mathcal{E}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ with $d \geq 1$ an integer).

We will also mainly focus on **random processes with continuous paths**. We will, in particular, define, construct, and study the **Brownian motion** $(B_t)_{t\geq 0}$. One of our objectives is to give meaning to the **stochastic differential equation**:

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t.$$

This equation can be understood as follows: the infinitesimal displacement of the particle X at time t is the sum of two terms,

1. a term corresponding to the velocity field b: $b(X_t)dt$. If we only had this term, our equation would be an ordinary differential equation $X_t' = b(X_t)$ similar to those you studied last year in your third year;

2. a term accounting for microscopic collisions due to thermal agitation σ: $\sigma(X_t)dB_t$.

To make sense of this equation, we will need to construct the Brownian motion and also construct the **stochastic integral** with respect to this Brownian motion.

1 Revisions

In this chapter, we propose some review points that will be very useful in the rest of the course. There are two main sections: the first one allows us to review the concept of the *"monotone class argument" (in frecnh) or Dynkin's lemma*, which is often invoked in the study of random processes; the second one includes some reviews on Gaussian variables and vectors, which will be very useful since Brownian motion is a Gaussian process. Some results from measure theory, which we will also use during this semester, complete this part.

1.1 Dynkin's Lemma

This lemma is particularly useful in the context of random process theory. In general, in what context is it used?

To show that a property $\mathcal{P}$ is true for any event A in a sigma-algebra $\mathcal{F}$, we often proceed as follows:

1. We find a class (i.e., a collection of subsets of Ω) I generating $\mathcal{F}$ (i.e., $\sigma(I) = \mathcal{F}$) such that $\mathcal{P}(A)$ is true for all $A \in I$;

2. We show that the class $\{A \text{ such that } \mathcal{P}(A) \text{ is true}\}$ forms a sigma-algebra;

and we conclude that $\mathcal{P}$ is true (at least) on $\sigma(I) = \mathcal{F}$. The problem is that in some cases, the second point is false because the class we consider is not closed under countable union but only under countable *increasing* union. This is the case, for example, when considering μ and ν as two probabilities on $(\Omega, \mathcal{F})$ and focusing on the class $\{A \text{ such that } \mu(A) = \nu(A)\}$. It is in these cases that Dynkin's lemma is useful.

Definition 2. *A class $\mathcal{M} \subset \mathcal{P}(\Omega)$ is a $\lambda-$**system** if*

1. $\Omega \in \mathcal{M}$

2. *For all $A, B \in \mathcal{M}$ with $A \subset B$, $B \setminus A \in \mathcal{M}$*

3. *For any **increasing** sequence $(A_n)_{n\geq 0}$ in $\mathcal{M}$,*

$$\bigcup_{n\geq 0} \uparrow A_n \in \mathcal{M}.$$

There are other equivalent definitions of a λ-system in the literature[2]. We define $\lambda(\mathcal{C})$ as the $\lambda-$system generated by a class I using the same idea as for the σ-algebra generated by I, denoted as $\sigma(I)$. Recall that:

$$\sigma(I) = \bigcap_{\mathcal{T} \text{ sigma-algebra s.t. } I \subset \mathcal{T}} \mathcal{T}.$$

Thus, we define

$$\lambda(I) = \bigcap_{\mathcal{M} \text{ monotone class s.t. } I \subset \mathcal{M}} \mathcal{M}.$$

To ensure that this definition makes sense, it is necessary to verify that any arbitrary intersection of $\lambda-$system is still a $\lambda-$system and that the intersection is non-empty since $\mathcal{P}(\Omega)$ is $\lambda-$system and contains I. Note that a sigma-algebra is also a $\lambda-$system, and thus, we always have

$$\lambda(I) \subset \sigma(I). \tag{1}$$

Returning to our initial problem: we find ourselves in a case where the second part of the proof does not work because the class

$$\{A \text{ such that } \mathcal{P}(A) \text{ is true}\}$$

is not a sigma-algebra but only a $\lambda-$system. Therefore, we obtain that property A is true on $\lambda(I)$. The purpose of Dynkin's lemma is to show, with an additional assumption, the converse of (1), and thus $\lambda(I) = \sigma(I)$.

Lemma 1 (Dynkin's Lemma)**.** *Let $I \subset \mathcal{P}(\Omega)$ be a class* ***closed under finite intersection****. Then,*

$$\sigma(I) = \lambda(I).$$

Proof. The idea of the proof is to show that with the additional assumption of closure under finite intersection, $\lambda(I)$ is, in fact, a sigma-algebra. This is the subject of Exercise 1 in Tutorial 1. □

Here are some important applications:

1. Let μ and ν be two probability measures that coincide on a class $I \subset \mathcal{P}(\Omega)$ closed under finite intersection and such that $\sigma(I) = \mathcal{F}$ (which means $\mu(A) = \nu(A)$ for all $A \in I$). Then $\mu = \nu$, i.e.,

$$\mu(A) = \nu(A), \quad \text{for all } A \in \mathcal{F}.$$

[2] See, for example, the English Wikipedia page on the monotone class lemma

Indeed, the class $\mathcal{M} = \{A \text{ such that } \mu(A) = \nu(A)\}$ is a $\lambda-$system (verify this!) and contains I. Therefore, $\mathcal{M} \supset \lambda(I)$. Moreover, I is closed under finite intersection, so according to Dynkin's lemma, $\lambda(I) = \sigma(I)$. Thus, $\mu = \nu$.

It is important to note that μ and ν do not necessarily coincide if we remove the assumption "*I is closed under finite intersection*" as demonstrated by the following example: consider $\Omega = \{1, 2, 3, 4, 5\}$, the sigma-algebra $\mathcal{F} = \mathcal{P}(\Omega)$, and the class $I = \{\{1, 2, 3\}, \{2, 4\}, \{3, 4, 5\}\}$, which is not closed under finite intersection. It can be easily verified that $\sigma(I) = \mathcal{F}$. Let μ be the uniform probability and ν be the probability such that $\nu(\{1\}) = 3/10,\ \nu(\{2\}) = 1/10,\ \nu(\{3\}) = 1/5,\ \nu(\{4\}) = 3/10,\ \nu(\{5\}) = 1/10$. It can be checked that μ and ν coincide on I but not on $\mathcal{P}(\Omega)$.

Here are two important examples where we use this property:

(a) Consider $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. If two measures μ and ν coincide on open bounded intervals (i.e., $\mu(]a, b[) = \nu(]a, b[)$ for all $a < b$), then they are equal. This is how the uniqueness of the Lebesgue measure is ultimately proved, although the existence is a different story!

(b) The cumulative distribution function characterizes the distribution because the class $\{]-\infty, a],\ a \in \mathbb{R}\}$ is closed under finite intersection and also generates the Borel sigma-algebra.

2. Consider a family $(X_t)_{t\in T}$ of random variables. By definition, $\sigma(X_t, t \in T)$ is the smallest sigma-algebra making all $X_t,\ t \in T$, measurable [verify that this definition makes sense and that, when $|T| < +\infty$, $\sigma(X_t, t \in T) = \{(X_t)_{t\in T} \in A,\ A \in \mathcal{B}(\mathbb{R})^{|T|}\}$. See Exercise 5 in Tutorial 1 on this point]. Also, consider $\mathcal{G}$ a sub-sigma-algebra of $\mathcal{F}$. Then the following two propositions are equivalent:

   (a) $\mathcal{G}$ and $\sigma(X_t, t \in T)$ are independent.

   (b) For every $S \subset T$ **finite**, $\mathcal{G}$ and $\sigma(X_t, t \in S)$ are independent.

   Proof. Fix $A \in \mathcal{G}$ and introduce

$$\mathcal{M} = \{B \in \mathcal{F},\ \mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)\}.$$

   We want to show that $\mathcal{M} = \sigma(X_t, t \in T)$. To do this, we show that $\mathcal{M}$ is a $\lambda-$system. According to the assumption, this class contains

$$\mathcal{C} = \bigcup_{|J|<+\infty} \sigma(X_j,\ j \in J).$$

This class is closed under finite intersection because if $A \in \sigma(X_j,\ j \in J_1)$ and $B \in \sigma(X_j,\ j \in J_2)$ where J_1 and J_2 are two finite subsets of J, then $A \cap B \in \sigma(X_j,\ j \in J_1 \cup J_2)$.

Therefore, $\lambda(\mathcal{C}) = \sigma(\mathcal{C})$, and we deduce that $\mathcal{M}$ contains $\sigma(\mathcal{C})$. Since $\sigma(\mathcal{C})$ makes all $X_t,\ t \in T$, measurable, we deduce that $\mathcal{M}$ contains $\sigma(X_t, t \in T)$. □

As a corollary of this property, it can also be shown (Exercise 3 Tutorial 1) that two families of random variables $(X_t)_{t\in T}$ and $(Y_s)_{s\in S}$ are independent if and only if for all finite families $t_1, \ldots, t_n \in T$ and $s_1, \ldots, s_m \in S$, the vectors $(X_{t_1}, \ldots, X_{t_n})$ and $(Y_{s_1}, \ldots, Y_{s_m})$ are independent.

In this course, we will encounter many other applications of this Dynkin's lemma. I recommend mastering this concept thoroughly.

1.2 Other Measure Theory Reminders

Here are some results that you need to know how to prove and use.

Lemma 2. *Let X be a function taking values in a set E, and $\mathcal{C} \subset \mathcal{P}(E)$. Then, X is a random variable from $(\Omega, \mathcal{F})$ to $(E, \sigma(\mathcal{C}))$ if and only if, for every $B \in \mathcal{C}$, $\{X \in B\} \in \mathcal{F}$.*

Proof. Exercise! □

For any random variable X, we denote ϕ_X as the characteristic function of X, defined for any $\xi \in \mathbb{R}$ by

$$\phi_X(\xi) = \mathrm{E}(e^{i\xi X}).$$

Now, we recall the fundamental theorem due to Paul Lévy:

Theorem 1 (Lévy's Theorem)**.**

1. *A sequence $(X_n)_{n\geq 1}$ of random variables converges in law to the random variable X if and only if, for every $\xi \in \mathbb{R}$, $(\phi_{X_n}(\xi))_{n\geq 1}$ converges to $\phi_X(\xi)$.*

2. *Suppose there exists a function ϕ such that for every $\xi \in \mathbb{R}$, $(\phi_{X_n}(\xi))_{n\geq 1}$ converges to $\phi(\xi)$. Then, the following three points are equivalent:*

   (a) $(X_n)_{n\geq 1}$ converges in law;

   (b) ϕ is a characteristic function;

   (c) ϕ is continuous;

(d) ϕ is continuous at 0.

This is a challenging theorem (which does not mean you should give up studying its proof).

1.3 Gaussian Variables and Vectors

It is essential to review all concepts related to Gaussian variables and vectors as they will be very useful in the study of Brownian motion.

A **Gaussian random variable** with parameters $m \in \mathbb{R}$ and $\sigma^2 > 0$, denoted as $\mathcal{N}(m, \sigma^2)$ (also called normal distribution), is either (when $\sigma^2 = 0$) a constant random variable equal to m or (when $\sigma^2 > 0$) a real-valued variable with the density function

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad x \in \mathbb{R},$$

with respect to the Lebesgue measure. It is called **centered** if $m = 0$ and **standardized** if $\sigma^2 = 1$ (it follows a $\mathcal{N}(0,1)$), in which case

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}.$$

It can be verified that if X follows a $\mathcal{N}(m, \sigma^2)$, then $\mathrm{E}(X) = m$ and $\mathrm{Var}(X) = \sigma^2$. If U follows a $\mathcal{N}(0,1)$, then for any $m, \sigma \in \mathbb{R}$, the variable $m + \sigma U$ follows a $\mathcal{N}(m, \sigma^2)$.

The **characteristic function** of a centered standardized variable U is given by

$$\phi_U(\xi) = e^{-\xi^2}, \quad \xi \in \mathbb{R},$$

(this can be proven easily, for instance by solving a simple ODE). More generally, for any $z \in \mathbb{C}$,

$$\mathrm{E}(e^{zU}) = e^{z^2/2}.$$

It can be easily derived from the expression of ϕ_X when X follows a $\mathcal{N}(m, \sigma^2)$:

$$\phi_X(\xi) = e^{i\xi m} e^{-\frac{\xi^2\sigma^2}{2}}, \quad \xi \in \mathbb{R}.$$

Consequently, if X_1 and X_2 are two independent Gaussian variables with respective distributions $\mathcal{N}(m_1, \sigma_1^2)$ and $\mathcal{N}(m_2, \sigma_2^2)$, according to Lévy's characterization, $X_1 + X_2$ follows a $\mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$.

Moments of the standard Gaussian can be easily computed from the characteristic function ϕ_U. For any $p \geq 1$, $\mathrm{E}(U^{2p}) = \frac{(2p)!}{2^p p!}$ and $\mathrm{E}(U^{2p-1}) = 0$.

The following proposition is a classic result that one should know and be able to prove (see Exercise 6 Tutorial 1).

Proposition 1 (A limit (even a weak limit !) of a sequence of Gaussians is a Gaussian)**.** *Let $(X_n)_{n\geq 0}$ be a sequence of Gaussian variables such that for every $n \geq 0$, X_n follows a $\mathcal{N}(m_n, \sigma_n^2)$. Suppose also that $(X_n)_{n\geq 0}$ converges* ***weakly*** *to a random variable X. Then:*

1. *X follows a $\mathcal{N}(m, \sigma^2)$ where $m = \lim_{n\to+\infty} m_n$ and $\sigma^2 = \lim_{n\to+\infty} \sigma_n^2$.*

2. *If, moreover, $(X_n)_{n\geq 0}$ converges* ***in probability*** *to X, then $(X_n)_{n\geq 0}$ also converges in L^p for every $1 \leq p < +\infty$.*

Proof. This is Exercise 5 of Worksheet 1. It is a classic, but not an easy exercise. □

We now move on to **Gaussian vectors**.

Definition 3. *A random vector $(X_1, \ldots, X_n)$ is a Gaussian vector if, for any $(t_1, \ldots, t_n) \in \mathbb{R}^n$, the variable*

$$t \cdot X = \sum_{i=1}^{n} t_i X_i$$

is Gaussian.

If the vector $(X_i)_{i=1,\ldots,n}$ is Gaussian, it implies, of course, that for any $i = 1, \ldots, n$, the variable X_i is Gaussian. However, the converse is not true.

Starting from the one-dimensional case, we can easily find the **characteristic function of a Gaussian vector**: for any vector ξ in $\mathbb{R}^n$,

$$\phi_X(\xi) = \mathrm{E}(e^{i\,{}^t\xi X}) = \phi_{{}^t\xi X}(1).$$

Now, ${}^t\xi X$ follows a $\mathcal{N}({}^t\xi m, {}^t\xi\Gamma\xi)$ with

$$m = \mathrm{E}(X) \quad \text{ and } \Gamma = \mathrm{E}((X-m)^t(X-m)),$$

the mean and variance-covariance matrix of X. Consequently, we finally obtain

$$\phi_X(\xi) = e^{i\,{}^t\xi m - \frac{{}^t\xi\Gamma\xi}{2}},$$

and note that the distribution of a Gaussian vector is characterized by its mean vector and variance-covariance matrix.

An important consequence is the **characterization of independence** between the coordinates of a Gaussian vector:

Proposition 2. *If X is a Gaussian vector, then its components are independent* ***if and only if*** *Γ is diagonal, i.e., if the covariances are zero.*

More generally, if $(X_i)_{i\in I}$ is a Gaussian vector and $I = \cup_{j=1,\dots,m}^{(disjoint)} I_j$, then the subvectors $(X_i)_{i\in I_j}$, $j = 1,\dots,m$, are independent if and only if, for any $1 \le k < \ell \le m$ and any $a, b \in I_k \times I_\ell$, $Cov(X_a, X_b) = 0$ (i.e., Γ is block-diagonal with blocks corresponding to the partition of I into I_j, $j = 1,\dots,m$).

Proof. The direct implication is always true: there is nothing specific in the Gaussian case. For the converse, it is necessary to show that the characteristic function factors. □

As in the one-dimensional case, any centered Gaussian vector can be written as a **linear transformation of a centered standard normal gaussian vector** (i.e., $m = 0$ and $\Gamma = I_n$). Indeed, for any positive symmetric matrix Γ, there exists an orthogonal matrix P (i.e., ${}^tPP = P^tP = I_n$) and $\lambda_1, \dots, \lambda_d > 0$ ($d \le n$) such that

$$ {}^tP\Gamma P = \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_d & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}. $$

It follows that $\Gamma = A^tA$ with

$$ A = P \begin{pmatrix} \sqrt{\lambda_1} & & & & & \\ & \ddots & & & & \\ & & \sqrt{\lambda_d} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}. $$

If U is a centered standard normal vector (i.e., U follows a $\mathcal{N}(0, I_n)$), then we can easily verify that

$$ X \overset{(law)}{=} AU. $$

This is also a simple way to see that for any positive symmetric matrix Γ, there exists a $\mathcal{N}(0,\Gamma)$, as it can be constructed explicitly from U centered reduced (and which is only a vector of independent centred standard Gaussian variables, hence easy to construct).

Density. For any random vector X taking values in $\mathbb{R}^n$, the following two points are equivalent:

1. X follows a $\mathcal{N}(m, \Gamma)$ and $\text{rank}(\Gamma) = n$
2. X has a density with respect to the Lebesgue measure on $\mathbb{R}^n$ given by:

$$f_X(x) = \frac{1}{\sqrt{2\pi}^n \sqrt{\det \Gamma}} e^{-\frac{1}{2} {}^t(x-m)\Gamma^{-1}(x-m)}.$$

It can be proven, for example, by writing X “ cleverly ” as the sum of independent Gaussian vectors.

2 Generalities on random processes

2.1 Definitions. Sigma-algebra and law.

Definition 4. *A random process indexed by a set of indices* T *and taking values in a measurable space* $(E, \mathcal{E})$ *is a family* $(X_t)_{t\in\mathsf{T}}$ *of random variables defined on* $(\Omega, \mathcal{F}, \mathbb{P})$ *and taking values in* $(E, \mathcal{E})$*.*

Examples of common index sets:

1. $\mathsf{T} = \mathbb{N}$, and we then talk about *discrete processes*. This is the framework of the course you took in the first semester. The most important examples of such random processes are, of course, discrete-time martingales and Markov chains.

2. $\mathsf{T} = \mathbb{R}^2$, and we then talk about *random fields*.

In this course, we are interested in **continuous-time processes** taking values in $\mathbb{R}$:

$$\mathsf{T} = \mathbb{R}^+ \text{ and } (E, \mathcal{E}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

(sometimes $\mathsf{T} = \mathbb{R}$ or $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ with $d \geq 1$ an integer).

2.1.1 Random processes as random functions

Remark that, for a fixed $\omega \in \Omega$, a random process $(X_t)_{t\in\mathsf{T}}$ (here $\mathsf{T} = \mathbb{R}^+$) defines a function:

$$\begin{aligned} X(\omega) : \mathbb{R}^+ &\to \mathbb{R} \\ t &\mapsto X_t(\omega). \end{aligned}$$

One way to consider the random process X is to see it as a **random function**. Is this possible? What sigma-algebra $\mathcal{T}$ should be considered on the set $\mathcal{A}(\mathbb{R}^+, \mathbb{R})$ of functions from $\mathbb{R}^+$ to $\mathbb{R}$ so that

$$X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$$

is measurable?

Definition 5 (Cylindrical sigma-algebra)**.** *A cylinder is defined as a subset of* $\mathcal{A}(\mathbb{R}^+, \mathbb{R})$ *of the form*

$$\{f \in \mathcal{A}(\mathbb{R}^+, \mathbb{R}) \text{ such that } f(t_1) \in B_1, \ldots, f(t_n) \in B_n\}$$

where $n \geq 1$ is an integer, $t_1, \ldots, t_n$ are positive real numbers, and $B_1, \ldots, B_n$ are in $\mathcal{B}(\mathbb{R})$. The cylindrical sigma-algebra, denoted by $\mathcal{T}$, is the sigma-algebra on $\mathcal{A}(\mathbb{R}^+, \mathbb{R})$ generated by the cylinders:

$$\mathcal{T} = \sigma(C,\ C \text{ cylinder of } \mathcal{A}(\mathbb{R}^+, \mathbb{R})),$$

meaning it is the smallest sigma-algebra that contains all cylinders.

In fact, the same sigma-algebra is defined by replacing, in the definition of cylinders, the Borel sets B_i, $i = 1, \ldots, n$ with open intervals I_i, $i = 1, \ldots, n$ or even intervals open on only one side (see Exercise 3 of TD 2). Another common and useful definition of this sigma-algebra is noted:

Proposition 3. *The sigma-algebra $\mathcal{T}$ is the smallest sigma-algebra that makes all coordinate maps measurable. That is*

$$\mathcal{T} = \sigma(\pi_t,\ t \geq 0)$$

where for $t \geq 0$, π_t is the coordinate map at t defined by:

$$\begin{aligned} \pi_t : \mathcal{A}(\mathbb{R}^+, \mathbb{R}) &\to \mathbb{R} \\ f &\mapsto f(t). \end{aligned}$$

Proof. For any $t \geq 0$, note that π_t is $\mathcal{T} - \mathcal{B}(\mathbb{R})$-measurable, since for any $B \in \mathcal{B}(\mathbb{R})$,

$$\{\pi_t \in B\} = \{f \in \mathcal{A}(\mathbb{R}^+, \mathbb{R}),\ f(t) \in B\}$$

is a cylinder and thus in $\mathcal{T}$.

Conversely, let $\tilde{\mathcal{T}}$ be a sigma-algebra that makes all coordinate maps measurable, and show that it contains all cylinders. Consider a cylinder

$$C = \{f \text{ such that } f(t_1) \in A_1, \ldots, f(t_n) \in A_n\}.$$

It can be rewritten as

$$C = \bigcap_{i=1}^{n} \{\pi_{t_i} \in A_i\}$$

and it follows that $C \in \tilde{\mathcal{T}}$. Thus, we have shown that $\mathcal{T} \subset \tilde{\mathcal{T}}$. □

We can now consider a random process as a **random function**:

Proposition 4. *If $(X_t)_{t \in \mathbb{R}^+}$ is a random process, then*

$$\begin{aligned} X : (\Omega, \mathcal{F}, \mathbb{P}) &\to (\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T}) \\ \omega &\to \Big(t \mapsto\ X_t(\omega)\Big) \end{aligned}$$

is measurable from $(\Omega, \mathcal{F}, \mathbb{P})$ to $(\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$ and thus defines a random function.

Proof. Since $\mathcal{T} = \sigma(C,\ C$ cylinder of $\mathcal{A}(\mathbb{R}^+, \mathbb{R}))$, it suffices to verify that for any cylinder C, $\{X \in C\} \in \mathcal{F}$. Consider a cylinder

$$C = \bigcap_{i=1}^{n} \{\pi_{t_i} \in B_i\}$$

and verify that

$$\{X \in C\} = \{\omega \text{ such that } \big(t \mapsto\ X_t(\omega)\big) \in C\} = \bigcap_{i=1}^{n} \{X_{t_i} \in B_i\}$$

belongs to $\mathcal{F}$. □

It has just been shown that

$$\sigma(X) = X^{-1}(\mathcal{T}) \subset \sigma(X_t,\ t \in \mathsf{T}).$$

The reverse inclusion is also true (it is easier! see TD 2), and thus, for any random process $(X_t)_{t\in\mathsf{T}}$,

$$\sigma(X_t,\ t \in \mathsf{T}) = X^{-1}(\mathcal{T}). \tag{2}$$

Also note (see again TD 2) that for any finite family $(X_1, \ldots, X_n)$ of real random variables, $\sigma(X_1, \ldots, X_n)$, the smallest sigma-algebra making all X_i $(i = 1, \ldots, n)$ measurable, coincides with $\sigma((X_1, \ldots, X_n))$, the sigma-algebra generated by $(X_1, \ldots, X_n)$ seen as a random vector taking values in $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. By equipping $\mathcal{A}(\mathbb{R}^+, \mathbb{R})$ with the cylindrical sigma-algebra $\mathcal{T}$, this property of finite families of random variables can be extended to random processes.

Furthermore, $\mathcal{T}$ allows us to consider X as a random function. Note also that (2) shows that $\mathcal{T}$ is rich enough for its inverse image to fill the entire $\sigma(X_t, t \in \mathsf{T})$.

2.1.2 Law of a random process

We can define the **law of the random process** $(X_t)_{t\in\mathsf{T}}$ as the image law on $(\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$ of P by the random function X: for all $A \in \mathcal{T}$,

$$\mathrm{P}_X(A) = \mathrm{P}((X_t)_{t\in\mathsf{T}} \in A).$$

Proposition 5. *Probability measures on $(\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$ are characterized by their values on cylinders. In other words, if μ and ν are two probabilities on $(\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$ such that for every cylinder C in $\mathcal{A}(\mathbb{R}^+, \mathbb{R})$, $\mu(C) = \nu(C)$, then $\mu = \nu$.*

Proof. This is another application of Dynkin's lemma. Indeed, since $\mathcal{T} = \sigma(C,\ C$ is a cylinder of $\mathcal{A}(\mathbb{R}^+, \mathbb{R}))$, it is enough to verify that the set of cylinders forms a class stable under finite intersection (which is not very difficult). $\square$

For a random process $(X_t)_{t\in\mathsf{T}}$, a **finite-dimensional marginal** refers to any finite family of marginals (i.e., any finite subset of $(X_t)_{t\in\mathsf{T}}$):

$$(X_{t_i})_{i=1,\cdots,n} \qquad n \geq 1,\ t_i \in \mathsf{T} \text{ for all } i = 1, \cdots, n.$$

And **finite-dimensional laws** refer to the laws of finite-dimensional marginals. From Proposition 5, it follows that two random processes $(X_t)_{t\in\mathsf{T}}$ and $(Y_t)_{t\in\mathsf{T}}$ with the same finite-dimensional laws also have the same law:

If for every $n \geq 1$, all $t_1, \cdots, t_n \in \mathsf{T}$, and all $B_1, \cdots, B_n \in \mathcal{B}(\mathbb{R})$,

$$\mathrm{P}(X_{t_1} \in B_1, \cdots, X_{t_n} \in B_n) = \mathrm{P}(Y_{t_1} \in B_1, \cdots, Y_{t_n} \in B_n),$$

then $\mathrm{P}_X = \mathrm{P}_Y$ (i.e., $X \overset{(law)}{=} Y$).

2.1.3 Independence of two random processes

According to Dynkin's lemma (again! see Exercise 3 TD 1), **two random processes** $(X_t)_{t\in\mathsf{T}}$ **and** $(Y_t)_{t\in\mathsf{T}}$ **are independent** if and only if, for all $n, m \geq 1$, all $t_1, \cdots, t_n \in \mathsf{T}$, and all $s_1, \cdots, s_m \in \mathsf{T}$,

$$(X_{t_1}, \cdots, X_{t_n}) \perp\!\!\!\perp (Y_{s_1}, \cdots, Y_{s_m}).$$

2.1.4 Canonical process

Now, we are interested in the question of **the existence of law** on $\mathcal{A}(\mathbb{R}^+, \mathbb{R})$. First question is: if I consider a probability measure μ on $(\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$, does there exist a random process with law μ? To warm up, you can try to answer the same question for $(\mathbb{R}, \mathcal{B}(R))$ (Exercise 4 TD 2). To answer this question, we introduce the **canonical process**. It consists of the coordinate maps $(\pi_t)_{t\geq 0}$ viewed as random variables on the probability space $(\Omega, \mathcal{F}, \mathrm{P}) = (\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T}, \mu)$: for all $t \geq 0$,

$$\begin{aligned} \pi_t : (\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T}, \mu) &\to (\mathbb{R}, \mathcal{B}(\mathbb{R})) \\ \omega &\to \omega(t) \end{aligned}$$

Note that by the very definition of $\mathcal{T}$ (see Proposition 3), π_t is measurable for all $t \geq 0$. It is important to note that, viewed as a random function, the random process π is actually the identity:

$$\begin{aligned} \pi : (\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T}) &\to (\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T}) \\ \omega &\to (t \to \pi_t(\omega)) \end{aligned}$$

We can verify that for any $A \in \mathcal{T}$, $\mu_\pi(A) = \mu(\pi \in A) = \mu(A)$, which means that π has law μ! Therefore, for any probability on $(\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$, I can construct a random process that has this probability as its law by considering the canonical process.

2.1.5 The Daniell-Kolmogorov Theorem

We have already seen that the finite-dimensional distributions characterize the law of a random process. Is it true that, for any family of finite-dimensional distributions, there exists a random process whose marginals correspond to these distributions? It is clear that we cannot expect this to be true in general: if we consider the finite-dimensional projections of a law on $\mathcal{A}(\mathbb{R}^+, \mathbb{R})$, they clearly have the following **compatibility** property (which is therefore a necessary condition):

Let's consider a probability μ on $(\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$. To make things a bit more concrete, we can imagine that μ is the law of a random process $(X_t)_{t\geq 0}$ (so $\mu = \mathbb{P}_X$, but we will also sometimes use $\mu = X(\mathbb{P})$ for clarity). For any $I \subset \mathbb{R}^+$ finite, let μ_I denote the image law of μ under the restriction application to I:

$$\begin{aligned} \pi_I : (\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T}) &\to (\mathbb{R}^I, \mathcal{B}(\mathbb{R}^I)) \\ \omega &\to (\omega_t)_{t\in I} \end{aligned}$$

Thus, $\mu_I = \mu_{\pi_I}$. Equivalently, we could have defined μ_I as the law of $(X_t)_{t\in I}$ since for any $A \in \mathcal{B}(\mathbb{R}^I)$,

$$\mu_I(A) = \mu(\pi_I \in A) = \mathrm{P}((X_t)_{t\geq 0} \in \{\pi_I \in A\}) = \mathrm{P}(\pi_I \circ (X_t)_{t\geq 0} \in A) = \mathrm{P}((X_t)_{t\in I} \in A).$$

Thus, μ_I is the projection of the law μ onto $\mathbb{R}^I$, and the set of laws μ_I, $I \subset \mathbb{R}^+$ finite, is the set of finite-dimensional laws of $(X_t)_{t\geq 0}$. We also define, for all subsets $J \subset I \subset \mathbb{R}^+$ finite:

$$\begin{aligned} \pi^I_J : (\mathbb{R}^I, \mathcal{B}(\mathbb{R}^I)) &\to (\mathbb{R}^J, \mathcal{B}(\mathbb{R}^J)) \\ (\omega_t)_{t\in I} &\to (\omega_t)_{t\in J}. \end{aligned}$$

We have $\pi^I_J \circ \pi_I = \pi_J$, and thus for all $A \in \mathcal{B}(\mathbb{R}^J)$,

$$\pi^I_J(\mu_I)(A) = \mu_I(\pi^I_J \in A) = \mu(\pi_I \in \{\pi^I_J \in A\}) = \mu(\pi_J \in A) = \mu_J(A).$$

We have just verified that $\pi^I_J(\mu_I)$ is the law of $(X_t)_{t\in J}$. It seems complicated to write, but we haven't said anything very spectacular: I can obtain the law of $(X_t)_{t\in J}$ in two ways: either by directly extracting the vector $(X_t)_{t\in J}$ from the random process $(X_t)_{t\geq 0}$ or by first extracting the vector $(X_t)_{t\in I}$ and then from this vector the sub-vector $(X_t)_{t\in J}$. The family of finite-dimensional laws μ_I therefore satisfies the **compatibility condition**: for all subsets $J \subset I \subset \mathbb{R}^+$ finite

$$\pi^I_J(\mu_I) = \mu_J. \tag{3}$$

We have thus already identified a necessary condition, which is ultimately quite trivial, for a family of finite-dimensional laws to be *derived* from a common law on $\mathcal{A}(\mathbb{R}^+, \mathbb{R})$ (i.e., a family of laws obtained by image measures from a measure μ by the associated restriction applications to finite subsets of $\mathbb{R}^+$). What is remarkable is that the converse is true:

Theorem 2 (Daniell (1918) - Kolmogorov (1933) Theorem)**.** *For any $I \subset \mathbb{R}^+$ finite, let's suppose that a probability μ_I is given on $(\mathbb{R}^I, \mathcal{B}(\mathbb{R}^I))$. We further assume that these probabilities satisfy the compatibility condition* (3)*. Then, there exists a unique probability measure μ on $(\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$ such that for all $I \subset \mathbb{R}^+$ finite*

$$\mu_I = \mu_{\pi_I}.$$

We accept the proof of this theorem for this year (and probably next year too). We deduce that for any compatible family of finite-dimensional probabilities, there exists a random process that has these finite-dimensional laws. Indeed, the Daniell-Kolmogorov theorem provides us with a probability on $\mathcal{A}(\mathbb{R}^+, \mathbb{R})$, and then we construct the desired random process using the canonical process.

An important example of using this theorem is the existence of Gaussian processes.

Definition 6. *A random process $(X_t)_{t \in \mathsf{T}}$ is said to be Gaussian if all its finite-dimensional distributions are Gaussian: for all $n \geq 0$, for all $t_1, \cdots, t_n \in \mathsf{T}$ and all $\lambda_1, \cdots, \lambda_n \in \mathbb{R}$,*

$$\lambda \cdot (X_{t_1}, \cdots, X_{t_n}) = \sum_{i=1}^{n} \lambda_i X_{t_i}$$

is a Gaussian variable.

Proposition 6. *The law of a Gaussian process is characterized by:*

1. *its **mean function**:*

$$\begin{aligned} m : \mathsf{T} &\to \mathbb{R} \\ t &\to \mathbb{E}(X_t). \end{aligned}$$

2. *its **covariance function**:*

$$\begin{aligned} R : \mathsf{T} \times \mathsf{T} &\to \mathbb{R} \\ (s,t) &\to Cov\,(X_s, X_t) = \mathbb{E}\big([X_s - \mathbb{E}(X_s)][X_t - \mathbb{E}(X_t)]\big). \end{aligned}$$

Indeed, according to Proposition 5, the law of a random process is characterized by its finite-dimensional laws, and as here the finite-dimensional laws are Gaussian, they are characterized by their means and covariances (which determine the characteristic function). Note that the covariance function is

1. **symmetric**: for all $s, t \in \mathsf{T}$, $R(s,t) = R(t,s)$,
2. **positive semidefinite** : for all $n \geq 1$ and all $t_1, \cdots, t_n \in \mathsf{T}$, the matrix $(R(t_i,t_j))_{1\leq i,j\leq n}$ is positive semidefinite since for all $\lambda \in \mathbb{R}^n$,

$$ {}^t\lambda R\lambda = \mathbb{E}\Big(\sum_{i=1}^n \lambda_i (X_{t_i} - m(t_i))^2\Big) \geq 0 $$

Proposition 7 (Existence of Gaussian processes)**.** *Let $m : \mathbb{R}^+ \to \mathbb{R}$ and $R : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}$ be positive semidefinite and symmetric. Then there exists a Gaussian process $(X_t)_{t\geq 0}$ with mean m and covariance R. It is uniquely determined in law.*

Proof. Use Theorem 2, see TD2. □

Here is an important example of a Gaussian process that will occupy us for a good part of the semester. We seek to construct a **Gaussian process centered with independent and stationary increments and with variance t at time $t \geq 0$.**

We say that a random process $(X_t)_{t\geq 0}$ has

1. **independent increments** if for all $n \geq 1$ and all real numbers $0 \leq t_1 \leq \cdots \leq t_n$, the random variables $X_{t_1} - X_0, \cdots, X_{t_n} - X_{t_{n-1}}$ are independent.
2. **stationary increments** if for all $0 \leq s < t$ the random variables $X_t - X_s$ and $X_{t-s} - X_0$ have the same distribution.

Let's go back to our problem: if such a random process exists, it satisfies for all $0 \leq s < t$,

$$ R(s,t) = \mathbb{E}(X_s X_t) = \mathbb{E}((X_t - X_s)X_s) + \mathbb{E}(X_s^2) \overset{(PAIS)}{=} \mathbb{E}(X_t - X_s)\mathbb{E}(X_s) + \mathbb{E}(X_s^2) = s. $$

The function $(s,t) \to s \wedge t$ is indeed positive semidefinite and symetric, and therefore we can define the Gaussian process associated with this covariance function R and mean function $m = 0$. We have almost constructed the Brownian motion: what remains is the continuity, which we will address in the next section.

2.2 Continuous Processes

We have worked so far in the very general framework $(\mathcal{A}(\mathbb{R}^+, \mathbb{R}), \mathcal{T})$. Since this space is enormous (or the σ-algebra $\mathcal{T}$ is too small), we note that certain sets we would like to study are not events. We can, for example, show (TD2) that the sets

$$\{f \in \mathcal{A}(\mathbb{R}^+, \mathbb{R}),\ \sup_{t\in[0,1]} f(t) < 1\} \quad \text{and}$$
$$\{f \in \mathcal{A}(\mathbb{R}^+, \mathbb{R}),\ \exists t \in [0,1]\ f(t) = 0\},$$

are not in $\mathcal{T}$ (and the set of continuous functions either... another exercise!).

2.2.1 σ-algebra(s)

In this course, we are mainly interested in **continuous processes**. So we will work in the space of continuous functions $\mathscr{C}(\mathbb{R}^+, \mathbb{R})$. We define on this space the cylindrical σ-algebra $\mathcal{C}$, the smallest σ-algebra making the coordinate applications measurable. In this context, the coordinate applications are defined for all $t \geq 0$ by

$$\begin{aligned} \pi_t : \mathscr{C}(\mathbb{R}^+, \mathbb{R}) &\to \mathbb{R} \\ \omega &\to \omega(t) \end{aligned}$$

Equivalently, we can define the σ-algebra $\mathcal{C}$ as the smallest σ-algebra containing the cylinders of $\mathscr{C}(\mathbb{R}^+, \mathbb{R})$, that is, a set of the form

$$C : \{f \in \mathscr{C}(\mathbb{R}^+, \mathbb{R}) \text{ such that } f(t_1) \in B_1, \cdots, f(t_n) \in B_n\},$$

where $t_1, \cdots, t_n$ are positive real numbers and $B_i, \cdots, B_n$ are Borel sets of $\mathbb{R}$.

Definition 7. *We say that a random process $(X_t)_{t\geq 0}$ is **continuous** if for all $\omega \in \Omega$, the function*

$$t \in \mathbb{R} \to X_t(\omega)$$

is continuous.

Here again we can view X as a random function

$$X : (\Omega, \mathcal{F}, \mathrm{P}) \to (\mathscr{C}(\mathbb{R}^+, \mathbb{R}), \mathcal{C}).$$

However, there is another *natural* σ-algebra on $\mathscr{C}(\mathbb{R}^+, \mathbb{R})$: that of uniform convergence on compacts. We will be content to study these different σ-algebras when $\mathsf{T} = [0,1]$ and not $\mathbb{R}^+$ to bring it back to the topology of uniform norm. We lose little in understanding by limiting ourselves to this framework, and we can study the general framework in exercise 10 of TD2. So we remind you that the uniform norm on $\mathscr{C}([0,1],\mathbb{R})$ is defined for any function f of this space by

$$||f||_\infty = \sup\{|f(t)|;\ t \in [0,1]\}.$$

This norm defines a distance, which itself defines a topology (=a set of open sets). We can therefore consider the Borel σ-algebra $\mathcal{B}$ on $\mathscr{C}([0,1],\mathbb{R})$, the smallest σ-algebra containing all the opens for the uniform norm.

Proposition 8. *The cylindrical and Borel σ-algebras on $\mathscr{C}([0,1],\mathbb{R})$ coincide:*

$$\mathcal{B} = \mathcal{C}.$$

Proof. $\underline{\mathcal{B} \supset \mathcal{C}}$. We will show that $\mathcal{B}$ makes all coordinate applications measurable. Indeed, for all $t \geq 0$,

$$\pi_t : (\mathscr{C}(\mathbb{R}^+,\mathbb{R}), ||\cdot||_\infty) \to (\mathbb{R}, |\cdot|)$$

is continuous, thus measurable.

$\underline{\mathcal{B} \subset \mathcal{C}}$. Since $\mathscr{C}([0,1],\mathbb{R})$ equipped with the uniform norm is separable, the Borel σ-algebra is generated by open balls or, equivalently, by closed balls (Exercise 13 of TD2). It suffices for us to show that all closed balls are in $\mathcal{C}$. We consider a function $f \in \mathscr{C}([0,1],\mathbb{R})$ and $\varepsilon > 0$. Using continuity, we obtain

$$\begin{aligned} B_{\mathrm{f}}(f,\varepsilon) &= \{g \in \mathscr{C}([0,1],\mathbb{R}) \text{ such that } ||g-f||_\infty \leq \varepsilon\} \\ &= \cap_{t\in[0,1]} \{g \in \mathscr{C}([0,1],\mathbb{R}) \text{ such that } g(t) \in [f(t)-\varepsilon, f(t)+\varepsilon]\} \\ &= \cap_{t\in[0,1]\cap\mathbb{Q}} \{\pi_t \in [f(t)-\varepsilon, f(t)+\varepsilon]\}. \end{aligned}$$

We deduce that $B_{\mathrm{f}}(f,\varepsilon) \in \mathcal{C}$, concluding the proof. □

It will be noted that $(\mathscr{C}([0,1],\mathbb{R}), \mathcal{B})$ constitutes a more comfortable space to work with than $(\mathcal{A}(\mathbb{R}^+,\mathbb{R}), \mathcal{T})$: for example, the *sup* function on this space is this time measurable (Exercise 6 TD 2).

2.2.2 Continuous Modification and Kolmogorov Criterion

We consider two random processes $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ (not necessarily continuous). We say that

1. $(Y_t)_{t\geq 0}$ is a **modification** of $(X_t)_{t\geq 0}$ if for all $t \geq 0$

$$\mathrm{P}(X_t = Y_t) = 1.$$

2. $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ are **indistinguishable** if

$$P(\forall\, t \geq 0, X_t = Y_t) = 1.$$

This means that the random processes are equal *as random functions*.

Is the set we consider in the definition of *indistinguishable* indeed an event? We can verify (exercise!) that it is indeed an element of $\mathcal{C}$ but that it is not an event of $\mathcal{T}$. The way this definition is stated is therefore a small abuse: we must understand that the complement of this part is negligible, i.e., included in an event of zero probability or equivalently: there exists $\Omega' \in \mathcal{T}$ such that $P(\Omega') = 1$ and for all $\omega \in \Omega'$ and all $t \geq 0$, $X_t(\omega) = Y_t(\omega)$.

Proposition 9. *Consider two random processes $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$*

1. *If X is a modification of Y then X and Y have the same law.*

2. *If X and Y are indistinguishable then X is a modification of Y.*

3. *If X is a modification of Y and X and Y are continuous (either here or there) then X and Y are indistinguishable.*

Proof. 1. If X is a modification of Y then the two random processes have the same finite dimensional laws and therefore the same law.

2. It's easy: $\cap_{t\geq 0}\{X_t = Y_t\} \subset \{X_t = Y_t\}$ for all $t \geq 0$.

3. Using continuity we get $\cap_{t\in\mathbb{R}^+}\{X_t = Y_t\} = \cap_{t\in\mathbb{Q}^+}\{X_t = Y_t\}$ and therefore

$$P(\exists t \geq 0,\ X_t \neq Y_t) = P(\exists t \in \mathbb{Q}^+,\ X_t \neq Y_t) \leq \sum_{t\in\mathbb{Q}^+} P(X_t \neq Y_t) = 0.$$

□

Consider a random process $(X_t)_{t\geq 0}$. Under what (sufficient) condition does there exist a continuous modification of this random process? In other words, can we make X continuous by changing, for all $t \geq 0$, the values of X_t only for an event of probability zero.

Theorem 3 (Kolmogorov's Criterion)**.** *Let $(X_t)_{0\leq 1}$ be a random process. Suppose there exist $q, \varepsilon, C > 0$ such that for all $s, t \in [0, 1]$,*

$$E(|X_s - X_t|^q) \leq C|t - s|^{1+\varepsilon}.$$

Then there exists a modification Y of X whose paths are Hölder continuous with exponent $\alpha \in]0, \varepsilon/q[$: for all $\alpha \in]0, \varepsilon/q[$ and all ω, there exists a constant $C_\alpha(\omega)$ such that for all $s, t \in [0, 1]$,

$$|Y_s(\omega) - Y_t(\omega)| \leq C_\alpha(\omega)|t - s|^\alpha.$$

In particular Y is a continuous modification of X (unique up to indistinguishability).

Remark 1. *If we work with random processes on $\mathbb{R}^+$ and not $[0, 1]$, we only obtain that the modification is locally Hölder continuous (that is, Hölder continuous on every compact) as the constants could diverge.*

Proof. At the board! We will follow the proof of [7]. □

3 Brownian Motion

[Intro]

3.1 Definitions

In this section, we present several definitions of Brownian motion (and demonstrate their equivalence!). Review the definition of a random process with stationary increments before exploring these definitions.

Definition 8 (B1)**.** *A* ***Brownian motion*** *is defined as any* ***continuous*** *random process* $(B_t)_{t\geq 0}$ *(i.e., for every* $\omega \in \Omega$*, the function* $t \in \mathbb{R}^+ \to B_t(\omega)$ *is continuous) with* ***independent Gaussian increments****, such that* $B_0 = 0$ *almost surely, and for all* $0 \leq s \leq t$

$$B_t - B_s \rightsquigarrow \mathcal{N}(0, t - s).$$

Note that the variance at time t of the Brownian motion reflects the diffusive behavior of this random process.

Definition 9 (B2)**.** *A* ***Brownian motion*** *is any* ***continuous*** *random process* $(B_t)_{t\geq 0}$ *that is* ***centered Gaussian*** *with the variance function defined for all* $0 \leq s \leq t$ *by*

$$R(s, t) = s \wedge t.$$

Definition 10 (B3)**.** *A* ***Brownian motion*** *is any* ***continuous*** *random process* $(B_t)_{t\geq 0}$ *such that* $B_0 = 0$ *almost surely, and for all* $0 \leq s \leq t$

$$\begin{aligned} &B_t - B_s \perp\!\!\!\perp \sigma(B_r;\, 0 \leq r \leq s), \\ &B_t - B_s \rightsquigarrow \mathcal{N}(0, t - s). \end{aligned}$$

Proof of Equivalence of the three Definitions. B1 $\Longrightarrow$ B2. For any $t \geq 0$, $\mathrm{E}(B_t - B_0) = 0$, and since $B_0 = 0$ almost surely, we have that $(B_t)_{t\geq 0}$ is centered. We now show that this random process is Gaussian. For any $n \geq 1$ and $0 \leq t_1 \leq \cdots \leq t_n$, we have

$$\sum_{i=1}^{n} \lambda_i B_{t_i} = \sum_{i=1}^{n} \mu_i (B_{t_i} - B_{t_{i-1}})$$

with $t_0 = 0$ and

$$\mu_i = \sum_{j=i}^{n} \lambda_j, \quad i = 1, \cdots, n.$$

Since the increments are independent and Gaussian, the linear combination is Gaussian, and therefore $(B_t)_{t\geq 0}$ is Gaussian. Finally, for all $0 \leq s \leq t$,

$$R(s,t) = \mathrm{E}(B_s B_t) = \mathrm{E}((B_t - B_s)B_s) + \mathrm{E}(B_s^2) = s.$$

<u>B2 $\Longrightarrow$ B3.</u> Since $(B_t)_{t\geq 0}$ is centered and $\mathrm{E}(B_0^2) = R(0,0) = 0$, we conclude that $B_0 = 0$ almost surely. Let $0 \leq s \leq t$. By a monotone class argument, it suffices to show that for any $n \geq 1$ and any $0 \leq r_1 \leq \cdots \leq r_n \leq s$, $B_t - B_s$ is independent of $(B_{r_1}, \cdots, B_{r_n})$. Since $(B_t)_{t\geq 0}$ is a Gaussian process, we only need to verify that the covariances are zero. For any $1 \leq i \leq n$

$$\mathrm{E}((B_t - B_s)B_{r_i}) = r_i - r_i = 0.$$

Finally, the increment $B_t - B_s$ is indeed a centered Gaussian, and we verify that its variance is

$$\mathrm{E}((B_t - B_s)^2) = t + s - 2(s \wedge t) = t - s.$$

<u>B3 $\Longrightarrow$ B1.</u> We need only show the independence of increments. So, we consider a family $0 = t_0 \leq t_1 \leq \cdots \leq t_n$ and bounded continuous functions $\phi_1, \cdots, \phi_n$. Since $B_{t_n} - B_{t_{n-1}} \perp\!\!\!\perp \sigma(B_r;\, 0 \leq r \leq t_{n-1})$,

$$\mathrm{E}(\phi_n(B_{t_n} - B_{t_{n-1}}) \cdots \phi_1(B_{t_1} - B_{t_0})) = \mathrm{E}(\phi_n(B_{t_n} - B_{t_{n-1}}))\mathrm{E}(\phi_{n-1}(B_{t_{n-1}} - B_{t_{n-2}}) \cdots \phi_1(B_{t_1} - B_{t_0}))$$

and we easily conclude by iteration. □

3.2 Existence

We now address the question of the existence of such a random process. Fortunately, we have already done the work! We use Definition B2. We have already seen how the Daniell-Kolmogorov theorem (Theorem 2) allows us to construct a Gaussian process $(B_t)_{t\geq 0}$ with mean $m = 0$ and variance function $R(s,t) = s \wedge t$ $(s, t \in \mathbb{R}^+)$. We need to ensure the existence of a continuous modification of such a random process (which will therefore have the same distribution!). To do this, we will use the Kolmogorov criterion (Theorem 3): for all $0 \leq s < t$, $B_t - B_s$ follows a Gaussian distribution with variance $t - s$, hence has the same distribution as $\sqrt{t-s}N$ where N is a standard centered Gaussian. So for any $q > 0$,

$$\mathrm{E}(|B_t - B_s|^q) = |t - s|^{q/2}\mathrm{E}(|N|^q) = c_q\ |t - s|^{q/2},$$

where c_q denotes the q-th moment of the standard centered Gaussian. We can therefore apply the Kolmogorov criterion by taking for all $q > 2$, $\varepsilon = q/2 - 1$.

It follows that B has a modification $\tilde{B}$ that is locally Hölder continuous with exponent α for all

$$\alpha < \frac{q/2 - 1}{q} = \frac{1}{2} - \frac{1}{q}.$$

We thus obtain a modification that is locally Hölder continuous with exponent arbitrarily close to $1/2$ by letting q tend to infinity. In particular, $\tilde{B}$ is continuous, which concludes the proof of the existence of the Brownian motion. In passing, we also proved the

Proposition 10. *Let $(B_t)_{t\geq 0}$ be a Brownian motion. Then almost surely, the trajectories of $(B_t)_{t\geq 0}$ are locally Hölder continuous for all $\gamma \in [0, \frac{1}{2}[$.*

Proof. We have just seen that B has a modification X that satisfies this property. Now, B and X are both continuous, so these two random processes are indistinguishable. □

3.3 Wiener Measure and Canonical Process

We call the **Wiener measure** and denote it by W the law of the Brownian motion as a random function on the space $(\mathscr{C}(\mathbb{R}^+, \mathbb{R}), \mathcal{B})$. We call the canonical construction of the Brownian motion or the **canonical Brownian motion**, the canonical process on this space:

$$\begin{aligned} B_t : (\mathscr{C}(\mathbb{R}^+, \mathbb{R}), \mathcal{B}, W) &\to (\mathbb{R}, \mathcal{B}(\mathbb{R})) \\ \omega &\mapsto B_t(\omega) = \omega(t) = \pi_t(\omega) \end{aligned}$$

3.4 First Properties

Proposition 11. *Let $(B_t)_{t\geq 0}$ be a Brownian motion. Then*

1. ***[Symmetry]*** *$(-B_t)_{t\geq 0}$ is also a Brownian motion;*
2. ***[Invariance under Diffusive Scaling]*** *for any $\lambda > 0$, $(B_t^\lambda)_{t\geq 0}$ defined for all $t \geq 0$ by*
$$B_t^\lambda = \frac{1}{\lambda} B_{\lambda^2 t}$$
is also a Brownian motion;
3. ***[Simple Markov Property]*** *for all $s \geq 0$, the random process $(B_t^{(s)})_{t\geq 0}$ defined for $t \geq 0$ by*
$$B_t^{(s)} = B_{s+t} - B_s$$
is a Brownian motion independent of $\sigma(B_r,\ r \leq s)$.

Proof. For all three points, note that continuity of the trajectories poses no problem. Point 1. is straightforward with Definitions B1 or B2. For the second point, we can use the second definition and calculate the covariance function K^λ of $(B_t^\lambda)_{t\geq 0}$: for all $s, t \geq 0$

$$K^\lambda(s,t) = \frac{1}{\lambda^2}\mathrm{E}(B_{\lambda^2 s}B_{\lambda^2 t}) = \frac{1}{\lambda^2}(\lambda^2 s \wedge \lambda^2 t) = s \wedge t.$$

For the last point, we again use Definition B2: the random process is indeed centered Gaussian, and we only need to calculate the covariance $K^{(s)}$: for all $t, u \geq 0$,

$$K^{(s)}(t,u) = \mathrm{E}((B_{s+t} - B_s)(B_{s+u} - B_s)) = (s+t) \wedge (s+u) - s = u \wedge t.$$

□

To continue our study of the Brownian motion, we rely on the following property (see [7] for this part of the course sequence)

Theorem 4 (Blumenthal's $0-1$ Law)**.** *Let $(B_t)_{t\geq 0}$ be a Brownian motion. For all $t \geq 0$, let*

$$\mathcal{F}_t = \sigma(B_s;\ s \leq t) \quad \textit{and}$$
$$\mathcal{F}_{0^+} = \cap_{s>0}\mathcal{F}_s.$$

Then for any $A \in \mathcal{F}_{0^+}$, $\mathrm{P}(A) \in \{0, 1\}$.

Intuitively, $\mathcal{F}_{0^+}$ represents some information: it is what we can infer by observing an arbitrarily small piece of the Brownian motion. For example,

$$\{\exists \varepsilon > 0 \quad \text{such that for all } 0 \leq t \leq \varepsilon,\ B_t \geq 0\} \in \mathcal{F}_{0^+}.$$

[proof?]

Proof. We will show that $\mathcal{F}_{0^+}$ is independent of itself, which implies the result. Let $A \in \mathcal{F}_{0^+}$, $0 < t_1 < \cdots < t_k$ and $g : \mathbb{R}^k \to \mathbb{R}$ bounded and continuous. As g is continuous,

$$\overset{p.s.}{\lim_{\varepsilon\to 0}} 1_A\ g(B_{t_1} - B_\varepsilon, \cdots, B_{t_k} - B_\varepsilon) = 1_A\ g(B_{t_1}, \cdots, B_{t_k})$$

and we have domination of the sequence by $||g||_\infty$, we obtain by the dominated convergence theorem,

$$\lim_{\varepsilon\to 0} \mathrm{E}(1_A\ g(B_{t_1} - B_\varepsilon, \cdots, B_{t_k} - B_\varepsilon)) = \mathrm{E}(1_A\ g(B_{t_1}, \cdots, B_{t_k})).$$

For any $\varepsilon > 0$, $A \in \mathcal{F}_\varepsilon$, and by the simple Markov property, for any $t > \varepsilon$, $B_t - B_\varepsilon$ is independent of $\mathcal{F}_\varepsilon$. So for $\varepsilon > 0$ small enough,

$$\mathrm{E}(1_A \; g(B_{t_1} - B_\varepsilon, \cdots, B_{t_k} - B_\varepsilon)) = \mathrm{P}(A)\mathrm{E}(g(B_{t_1} - B_\varepsilon, \cdots, B_{t_k} - B_\varepsilon)).$$

Using again the dominated convergence theorem, we finally obtain

$$\mathrm{E}(1_A \; g(B_{t_1}, \cdots, B_{t_k})) = \mathrm{P}(A)\mathrm{E}(g(B_{t_1}, \cdots, B_{t_k})).$$

Since this holds for all k and all k-tuples, we have in fact proved that

$$\mathcal{F}_{0^+} \perp\!\!\!\perp \sigma(B_t, \; t > 0).$$

Now, $\sigma(B_t, \; t > 0) = \sigma(B_t, \; t \geq 0)$ because $B_0 = \lim_{t \to 0, \, t > 0} B_t$, and moreover, for any $\varepsilon > 0$,

$$\mathcal{F}_{0^+} \subset \mathcal{F}_\varepsilon \subset \sigma(B_t, \; t \geq 0).$$

We have thus shown that $\mathcal{F}_{0^+}$ is independent of itself. □

This important (and interesting in itself) result allows for the development of new properties for the Brownian motion.

Corollary 1. *Let $(B_t)_{t \geq 0}$ be a Brownian motion. Then the following properties hold:*

1. *a.s. for any $\varepsilon > 0$, $\sup_{0 \leq s \leq \varepsilon} B_s > 0$ and $\inf_{0 \leq s \leq \varepsilon} B_s < 0$.*
2. *a.s. for any $\varepsilon > 0$ $(B_t)_{t \geq 0}$ has a zero on $]0, \varepsilon[$.*
3. *a.s. $(B_t)_{t \geq 0}$ is not monotone on any interval.*

Proof. 1. A remark before starting the proof: the Brownian motion is a continuous process, so $\{\sup_{0 \leq s \leq \varepsilon} B_s > 0\}$ is indeed an event. We consider a decreasing sequence $(\varepsilon_n)_{n \geq 1}$ tending to 0 and define

$$A = \cap_{n \geq 1} \downarrow \{ \sup_{0 \leq s \leq \varepsilon_n} B_s > 0\},$$

and we must show that $\mathrm{P}(A) = 1$. It is clear that for all $t > 0$, $A \in \mathcal{F}_t$ since $A \in \mathcal{F}_{\varepsilon_n}$ for all n and particularly for n large enough such that $\varepsilon_n < t$. We deduce that $A \in \mathcal{F}_{0^+}$ and thus, using Blumenthal's $0 - 1$ Law, $\mathrm{P}(A) \in \{0, 1\}$. Now

$$\mathrm{P}(A) = \lim_{n \to +\infty} \downarrow \mathrm{P}(\sup_{0 \leq s \leq \varepsilon_n} B_s > 0)$$

and since for all $n \geq 1$,

$$\mathrm{P}(\sup_{0 \leq s \leq \varepsilon_n} B_s > 0) \geq \mathrm{P}(B_{\varepsilon_n} > 0) \geq \frac{1}{2},$$

we conclude that $\mathrm{P}(A) = 1$.

2. From the previous point, it follows that a.s. for any $\varepsilon > 0$, there is a zero of $(B_t)_{t\geq 0}$ on $]0, \varepsilon[$. This of course implies that a.s. $(B_t)_{t\geq 0}$ has infinitely many zeros in the neighborhood of 0.

3. From point 1. it follows that a.s. for any $\varepsilon > 0$, $(B_t)_{t\geq 0}$ is not monotone on $[0, \varepsilon]$. We now consider an interval with rational endpoints $[s, t]$. According to the simple Markov property, the random process $(B_t^{(s)})_{t\geq 0}$ is a Brownian motion and thus is not monotone on $[0, t-s]$. We have thus shown

$$\forall\, 0 < s < t \in \mathbb{Q} \quad a.s. \quad (B_t)_{t\geq 0} \text{ is not monotone on } [s, t],$$

and since $\mathbb{Q}$ is countable,

$$a.s. \quad \forall\, 0 < s < t \in \mathbb{Q} \quad (B_t)_{t\geq 0} \text{ is not monotone on } [s, t],$$

finally, by density of $\mathbb{Q}$ in $\mathbb{R}$ and continuity of the trajectories of $(B_t)_{t\geq 0}$,

$$a.s. \quad \forall\, 0 < s < t \in \mathbb{R} \quad (B_t)_{t\geq 0} \text{ is not monotone on } [s, t].$$

□

We note, for any $a \in \mathbb{R}$, T_a the hitting time of a:

$$T_a = \inf\{s \geq 0 \text{ such that } B_s = a\}.$$

Corollary 2 (of the corollary)**.** *Let $(B_t)_{t\geq 0}$ be a Brownian motion. Then a.s. for any $a \in \mathbb{R}$, $T_a < +\infty$.*

Proof. Let's start by showing that $T_1 < +\infty$ a.s. As $\{\sup_{0\leq s\leq 1} B_s > 0\} = \{\exists\, n \geq 1;\ \sup_{0\leq s\leq 1} B_s > 1/n\}$, it follows from Corollary 1

$$\lim_{n\to+\infty} \uparrow \mathrm{P}(\sup_{0\leq s\leq 1} B_s > 1/n) = 1.$$

Now for any $n \geq 1$, using scaling,

$$\mathrm{P}(\sup_{0\leq s\leq 1} B_s > 1/n) = \mathrm{P}(\sup_{0\leq s\leq n^2} nB_{s/n^2} > 1) = \mathrm{P}(\sup_{0\leq s\leq n^2} B_s > 1).$$

Now $\lim_{n\to+\infty} \uparrow \mathrm{P}(\sup_{0\leq s\leq n^2} B_s > 1) = \mathrm{P}(\sup_{s\geq 0} B_s > 1)$ and we deduce that a.s. $\sup_{s\geq 0} B_s > 1$. Using scaling again, for any $a > 0$,

$$\mathrm{P}(\sup_{s\geq 0} B_s > a) = \mathrm{P}(\sup_{s\geq 0} \frac{1}{a} B_{a^2 s} > 1) = \mathrm{P}(\sup_{s\geq 0} B_s > 1) = 1.$$

We obtain a similar result for the inf by considering $(-B_t)_{t\geq 0}$. □

We have thus shown that a.s. $\limsup_{t\to+\infty} B_t = -\liminf_{t\to+\infty} B_t = +\infty$. Indeed, since the trajectories are continuous, $\{\limsup_{t\to+\infty} B_t < +\infty\} \subset \{\sup_{t\to+\infty} B_t < +\infty\}$ which has probability zero. We easily deduce (using trajectory continuity):

Corollary 3 (of the corollary of the corollary)**.** *Let $(B_t)_{t\geq 0}$ be a Brownian motion. Then a.s. the set of zeros of the Brownian motion is unbounded.*

3.5 Strong Markov Property

In this section, we aim to replace, in the strong Markov property, the deterministic time by a random time. Of course, there is no chance that this is true in general (why? Give an example!) and we first introduce the notion of stopping time (in continuous time here... but the definition is similar to what you have seen in the first semester in discrete time). We define for all $t \geq 0$,

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

The family of sigma algebras $(\mathcal{F}_t)_{t\geq 0}$ is a **filtration**, meaning that for all $0 \leq s \leq t$,

$$\mathcal{F}_s \subset \mathcal{F}_t.$$

Definition 11. *A random variable T taking values in $[0, +\infty]$ is a **stopping time** if for all $t \geq 0$,*

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

If T is a stopping time, we define the sigma algebra

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty \text{ such that for all } t \geq 0,\, A \cap \{T \leq t\} \in \mathcal{F}_t\}.$$

It is necessary to show that the above definition is relevant, i.e., that $\mathcal{F}_T$ is indeed a sigma algebra (Exercise!). Now we can state the strong Markov property:

Theorem 5 (Strong Markov Property)**.** *Let $(B_t)_{t\geq 0}$ be a Brownian motion and T a stopping time such that $\mathbb{P}(T < +\infty) > 0$. We define the random process*

$$B_t^{(T)} = 1_{\{T<+\infty\}}(B_{T+t} - B_T), \quad t \geq 0.$$

Then, under $\mathbb{P}(\cdot|T < +\infty)$, $(B_t^{(T)})_{t\geq 0}$ is a Brownian motion independent of $\mathcal{F}_T$.

Remark 2. *Note that on the event $\{T = +\infty\}$, we have $B_t^{(T)} = 0$ for all $t \geq 0$.*

Proof. (see [7]) We content ourselves with the case $T < +\infty$ almost surely. Our goal: to show that for all $A \in \mathcal{F}_T$, all $0 \leq t_1 \leq \cdots \leq t_p$ ($p \geq 1$), and all bounded continuous function $F : \mathbb{R}^p \to \mathbb{R}$,

$$\mathbb{E}(1_A\, F(B_{t_1}^{(T)}, \cdots, B_{t_p}^{(T)})) = \mathbb{P}(A)\mathbb{E}(F(B_{t_1}, \cdots, B_{t_p})). \qquad (\star)$$

This allows us to prove everything we need:

1. by a monotone class argument, that the random process $B^{(T)}$ is independent of the sigma algebra $\mathcal{F}_T$,

2. by taking $A = \Omega$, that the random process $(B^{(T)})_{t\geq 0}$ has the same finite-dimensional laws as $(B_t)_{t\geq 0}$, and thus, again by the monotone class lemma, that $(B^{(T)})_{t\geq 0}$ has the same law as $(B_t)_{t\geq 0}$. This indeed proves that $(B^{(T)})_{t\geq 0}$ is a Brownian motion since the continuity of the trajectories is not a problem.

It remains to prove $(\star)$ and for that, we discretize T to reduce it to simple Markov. We denote by T_n the smallest rational number of the form $k/2^n$ greater than or equal to T (thus $T_n = \lceil T2^n \rceil / 2^n$ and we can show that it is also a stopping time). For all $t \geq 0$, using the continuity of the Brownian motion, we have $B_t^{(T_n)} \to B_t^{(T)}$ and since F is continuous, we obtain

$$\lim_{n\to+\infty} F(B_{t_1}^{(T_n)}, \cdots, B_{t_p}^{(T_n)}) = F(B_{t_1}^{(T)}, \cdots, B_{t_p}^{(T)}),$$

and by the dominated convergence theorem

$$\lim_{n\to+\infty} \mathbb{E}\big(1_A\, F(B_{t_1}^{(T_n)}, \cdots, B_{t_p}^{(T_n)})\big) = \mathbb{E}\big(1_A\, F(B_{t_1}^{(T)}, \cdots, B_{t_p}^{(T)})\big).$$

Now for all $n \geq 1$

$$\mathbb{E}\big(1_A\, F(B_{t_1}^{(T_n)}, \cdots, B_{t_p}^{(T_n)})\big) = \sum_{k=0}^{+\infty} \mathbb{E}\big(1_A 1_{\frac{k-1}{2^n} < T \leq \frac{k}{2^n}}\, F(B_{\frac{k}{2^n}+t_1} - B_{\frac{k}{2^n}}, \cdots, B_{\frac{k}{2^n}+t_p} - B_{\frac{k}{2^n}})\big).$$

Since $A \in \mathcal{F}_T$, $A \cap \{\frac{k-1}{2^n} < T \leq \frac{k}{2^n}\} \in \mathcal{F}_{k/2^n}$ and according to simple Markov, we obtain for all $k \geq 0$,

$$\begin{aligned}&\mathbb{E}\big(1_A 1_{\frac{k-1}{2^n} < T \leq \frac{k}{2^n}}\, F(B_{\frac{k}{2^n}+t_1} - B_{\frac{k}{2^n}}, \cdots, B_{\frac{k}{2^n}+t_p} - B_{\frac{k}{2^n}})\big) \\ &\quad = \mathbb{P}\left(A \cap \left\{\frac{k-1}{2^n} < T \leq \frac{k}{2^n}\right\}\right) \mathbb{E}(F(B_{t_1}, \cdots, B_{t_p})).\end{aligned}$$

Summing over k yields $(\star)$. $\square$

3.6 Donsker's Theorem

An important property of the Brownian motion, and in fact another way to introduce and construct it, is to see it as the limit of a random walk on $\mathbb{Z}$ after a relevant scaling change. This is the subject of Donsker's theorem.

Let's start by recalling the definition of the simple random walk on $\mathbb{Z}$. We consider a family $(X_i)_{i\geq 1}$ of i.i.d. random variables with a common distribution, the uniform distribution on $\{-1,1\}$. We then define the walk by

$$S_0 = 0 \quad \text{and}$$
$$S_n = \sum_{i=1}^{n} X_i, \quad n \geq 1.$$

The variance of S_n is n and the order of magnitude of the distance to the origin of the walk at time n is therefore $\sqrt{n}$. If we want to observe something non-degenerate on a large scale, we must therefore normalize the walk by considering a *diffusive scaling*: contract for large n the time by n (i.e., bring the interval $[0,n]$ to the interval $[0,1]$) and the space by $\sqrt{n}$ (i.e., bring the interval $[-\sqrt{n},\sqrt{n}]$ to the interval $[-1,1]$). This leads us to define for all $n \geq 1$ the random process

$$S_t^{(n)} = \frac{1}{\sqrt{n}} \left\{ \sum_{i=1}^{\lfloor nt \rfloor} X_i + (nt - \lfloor nt \rfloor) X_{\lfloor nt \rfloor + 1} \right\}, \quad t \geq 0.$$

By this operation, we have also transformed the discrete walk into a continuous process by simply connecting the points with straight lines! The Brownian motion then appears as the limit of this sequence (in n) of random processes:

Theorem 6 (Donsker's Theorem)**.** *The sequence of random processes* $(S_t^{(n)})_{0\leq t\leq 1}$*,* $n \geq 1$ *converges in law for the topology of uniform convergence to the random process* $(B_t)_{0\leq t\leq 1}$*, i.e., for any function* $F : \mathscr{C}([0,1],\mathbb{R}) \to \mathbb{R}$*, continuous (for the uniform topology) and bounded,*

$$\mathbb{E}(F(S^{(n)})) \underset{n\to+\infty}{\longrightarrow} \mathbb{E}(F(B)).$$

Proof. The convergence of the finite-dimensional marginals is easy to see since it is nothing but the multidimensional central limit theorem. Another ingredient is missing to obtain the convergence in law of the sequence of random processes, which is the **tightness** of this sequence of laws on $\mathscr{C}([0,1],\mathbb{R})$. This is beyond the scope but you may see it next year! A classic reference on these convergence issues is [2]. □

[Add a picture]

4 Continuous Martingales

This chapter is intentionally short. Its purpose is to learn about continuous martingales only what we need to study stochastic integrals and stochastic differential equations in the next two chapters. The stopping theorem is also often a useful tool for studying Brownian motion (we will see this in many exercises). Many results are assumed, as the proof often involves reducing to the discrete case and using results you have proven in the first semester in the course *Discrete Processes*. For those who wish to learn more, I recommend [7].

4.1 Definition

Definition 12 (Filtration)**.** *A filtration* $(\mathcal{F}_t)_{t\geq 0}$ *is an increasing sequence of sub-*σ*-algebras of* $\mathcal{F}$.

An important example is the **filtration generated by the random process** $(X_t)_{t\geq 0}$: for all $t \geq 0$,

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

The sigma-algebra $\mathcal{F}_t$ contains information about the trajectory of X up to time t. A random process $(X_t)_{t\geq 0}$ is said to be **adapted** to the filtration $(\mathcal{F}_t)_{t\geq 0}$ if for all $t \geq 0$, X_t is $\mathcal{F}_t$-measurable.

Now we need to state a technical condition that will be useful in many theorems:

Definition 13 (Usual Conditions or Habitual Conditions)**.** *A filtration* $(\mathcal{F}_t)_{t\geq 0}$ *satisfies the* ***usual conditions*** *if*

1. *the sigma-algebra* $\mathcal{F}_0$ *contains the negligible sets (which implies that for all* $t \geq 0$, $\mathcal{F}_t$ *contains the negligible sets);*

2. *for all* $t \geq 0$, $\mathcal{F}_{t^+} := \cap_{\epsilon>0}\mathcal{F}_{t+\epsilon}$ *coincides with* $\mathcal{F}_t$ *(we say the filtration is right-continuous).*

It is recalled that a subset of Ω is negligible if it is included in an event (an element of $\mathcal{F}$ hence) of probability zero. In practice, one can always transform a filtration to satisfy the usual conditions by considering (and completing) $\mathcal{F}_{t^+}$ instead of $\mathcal{F}_t$. This condition notably allows us to ensure that a modification of an adapted process is still adapted. This will be particularly useful for regularizations of martingales.

Definition 14 (Sub/Super/$\emptyset$ Martingales)**.** *A random process* $(M_t)_{t\geq 0}$ *is an (sub/super/$\emptyset$)* $(\mathcal{F}_t)$*-martingale (where* $(\mathcal{F}_t)$ *is a filtration) if it is*

1. ***adapted*** *to* $(\mathcal{F}_t)_{t\geq 0}$*,*
2. ***integrable****, i.e., for all* $t \geq 0$*,* $\mathbb{E}(|M_t|) < +\infty$*,*
3. *for all* $0 \leq s \leq t$*,* $\mathbb{E}(M_t|\mathcal{F}_s) = M_s$ *(resp* $\leq$*,* $\geq$*).*

4.2 Convergence and Regularization

In this section, we study asymptotic behavior and regularization.

Proposition 12 (Almost Sure Convergence)**.** *Let* $(M_t)_{t\geq 0}$ *be a martingale with* ***càdlàg*** *paths and bounded in* L^1*. Then there exists* $M_\infty \in \mathrm{L}^1$ *such that*

$$M_t \xrightarrow{a.s.} M_\infty.$$

Some ideas for the proof. We actually look at limsup and liminf for t in $\mathcal{D}$ where $\mathcal{D}$ is a countable dense set in $\mathbb{R}$. We can then apply the convergence theorem for discrete martingales. We return to the continuous case using right-continuity. To show that M_∞ is in L^1, we use Fatou's lemma (exercise!). In general, remember to review the results from the first semester on discrete martingales. □

We recall the two possible (and of course equivalent) definitions of a family of *uniformly integrable* random variables:

Definition 15. *A family* $(X_i)_{i\in I}$ *(where* I *is any set) of random variables is uniformly integrable if*

$$\lim_{a\to+\infty} \sup_{i\in I} \mathbb{E}(|X_i| 1_{|X_i|>a}) = 0.$$

Definition 16. *A family* $(X_i)_{i\in I}$ *(where* I *is any set) of random variables is uniformly integrable if*

1. *the family* $(X_i)_{i\in I}$ *is bounded in* L^1 ***and***
2. *for all* $\epsilon > 0$ *there exists* $\delta > 0$ *such that for all* $A \in \mathcal{F}$ *such that* $\mathbb{P}(A) < \delta$ *and all* $i \in I$*,*

$$\mathbb{E}(|X_i| 1_A) < \epsilon.$$

Proposition 13 (L^1 convergence)**.** *Let* $(M_t)_{t\geq 0}$ *be a martingale with* ***càdlàg*** *paths. The three conditions are equivalent:*

1. $(M_t)_{t\geq 0}$ *converges in* L^1 *and almost surely to* M_∞*,*

2. $(M_t)_{t\geq 0}$ *is closed,*

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

Proof. Recall that a martingale $(M_t)_{t\geq 0}$ is said to be closed if there exists a variable $M_\infty \in \mathrm{L}^1$ such that for all $t \geq 0$, $M_t = \mathbb{E}(M_\infty|\mathcal{F}_t)$. The proof works as in the discrete case you have seen in the first semester. Exercise! □

Theorem 7 (Doob's Regularization Theorem)**.** *Let* $(M_t)_{t\geq 0}$ *be a martingale satisfying the usual conditions. Then there exists a modification* $(\tilde{M}_t)_{t\geq 0}$ *of* $(M_t)_{t\geq 0}$ *such that*

1. $(\tilde{M}_t)_{t\geq 0}$ *is a martingale*

2. $(\tilde{M}_t)_{t\geq 0}$ *has locally bounded and càdlàg paths.*

Proof. The proof is omitted. It relies on the almost sure convergence theorem. As an exercise, you can show that when the filtration satisfies the usual conditions, a modification of a martingale is still a martingale. □

4.3 Doob's Stopping Theorem

Throughout this section, we consider a filtration $(\mathcal{F}_t)_{t\geq 0}$.

Theorem 8 (Doob's Stopping Theorem)**.** *Let* $(M_t)_{t\geq 0}$ *be a martingale with* ***càdlàg paths and uniformly integrable****. Let* S *and* T *be two stopping times such that* $S \leq T$*. Then*

1. M_S *and* M_T *are integrable,*

2. $M_S = \mathbb{E}(M_T|\mathcal{F}_S)$*,*

3. $M_S = \mathbb{E}(M_\infty|\mathcal{F}_S)$*.*

In particular, in this case, $\mathbb{E}(M_S) = \mathbb{E}(M_\infty) = \mathbb{E}(M_0)$*.*

Note that in this theorem, the variable M_T is well-defined even if T is not supposed to be finite almost surely, since the variable M_∞ is well-defined. On the event $\{T = +\infty\}$, we have $M_T = M_\infty$. Again, we will not detail the proof of this theorem which is based on a similar theorem in the discrete case. To reduce to the discrete case, we introduce the sequence of stopping times $T_n = \lceil 2^n T \rceil / 2^n$, which converges almost surely decreasingly to T. We note the other useful forms of the stopping theorem:

Theorem 9. *Let* $(M_t)_{t\geq 0}$ *be a martingale with càdlàg paths and* S *and* T *be two* ***bounded*** *stopping times such that* $S \leq T$*. Then*

1. M_S *and* M_T *are integrable,*

2. $M_S = \mathbb{E}(M_T|\mathcal{F}_S)$*,*

Proof. This theorem is easily deduced from the previous one, since, noting C an upper bound of T, we have that $(M_{t\wedge C})_{t\geq 0}$ is a martingale closed by M_C, therefore uniformly integrable. □

Theorem 10 (Doob's Stopping Theorem)**.** *Let* $(M_t)_{t\geq 0}$ *be a martingale with càdlàg paths and* T *be a stopping time. Then*

1. *the random process* $(M_{t\wedge T})_{t\geq 0}$ *is a martingale,*

2. *if* $(M_t)_{t\geq 0}$ *is uniformly integrable then* $(M_{t\wedge T})_{t\geq 0}$ *is also uniformly integrable and for all* $t \geq 0$

$$M_{t\wedge T} = \mathbb{E}(M_T|\mathcal{F}_t).$$

4.4 Doob's Maximal Inequalities

These inequalities are very useful and important to know. They will be particularly useful for the study of stochastic integrals.

Theorem 11. *Let* $(M_t)_{t\geq 0}$ *be a* ***càd*** *martingale. Let* $M_t^* = \sup_{0\leq s\leq t}|M_s|$*. Then*

1. *For all* $p \geq 1$ *and all* $t \geq 0$*, for all* $\lambda > 0$*,*

$$\mathrm{P}(M_t^* \geq \lambda) \leq \frac{\mathrm{E}(|M_t|^p)}{\lambda^p}.$$

2. *If* $p > 1$ *and* $t > 0$

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

The proof is a classic exercise that you have probably seen in the first semester in the discrete case. To recall the first inequality (which leads to showing the second one), we can notice that it is actually a "super" Markov inequality: instead of controlling the probability that only the final variable M_t is large, we can actually (and this is because $(M_t)_{t\geq 0}$ is a martingale) control the supremum of the entire trajectory between 0 and t. Note also that the control in the second point is valid only for $p > 1$ (or rather, it yields nothing for $p = 1$ since the right-hand side is infinite).

5 Stochastic Integral - Itô (1950)

We aim to integrate with respect to Brownian motion, that is, to give meaning, when $(u_s)_{s\geq 0}$ is a random process and $(B_t)_{t\geq 0}$ is a Brownian motion, to the expression

$$\int u_s \; dB_s.$$

The first idea (which will not work!) is to define this integral ω by ω, that is, to give meaning for every ω to

$$\int u_s(\omega) \; dB_s(\omega).$$

This amounts to asking in what framework we can give a satisfactory definition of a function with respect to another.

5.1 Stieltjes Integral

The Stieltjes-Riemann theory allows us to construct the integral of a continuous function f with respect to another function g, provided that g has bounded variations. Here, we present only a summary of the construction of

$$\int_0^t f(s) \; dg(s),$$

and we refer to [7, 4, 6] for different and much more comprehensive expositions (which we used for this chapter).

1. **The case where g is positive, right-continuous, and non-decreasing on $[0,t]$.** The function g then defines a measure μ on $[0,t]$ by $\mu(\{0\}) = 0$ and $\mu(]a,b]) = g(b) - g(a)$ for all $0 \leq a \leq b \leq t$. In other words, g is the distribution function of μ (although the term is usually used for a probability measure in general). We can then define

$$\int_0^t f(s) \; dg(s) = \int_0^t f(s) \; d\mu(s),$$

for any non-negative function f or in $\mathrm{L}^1(\mu)$. Note that if f is continuous on $[0,t]$, then for any partition sequence $0 = t_0^n \leq t_1^n \leq \cdots \leq t_{p_n}^n = t$ with step size going to 0,

$$\int_0^t f(s) \; dg(s) = \lim_{n\to+\infty} \sum_{i=1}^{p_n} f(t_{i-1}^n)(g(t_i^n) - g(t_{i-1}^n)). \qquad (4)$$

Indeed, for any $n \geq 1$, $\sum_{i=1}^{p_n} f(t_{i-1}^n)(g(t_i^n) - g(t_{i-1}^n)) = \int_0^t f_n(s)\ dg(s)$ where f_n is the piecewise constant function defined by $f(t) = f(t_{i-1}^n)$ if $t \in]t_{i-1}^n, t_i^n]$, and we conclude using the dominated convergence theorem.

2. **The case where g has finite variations.** The total variation of a continuous function $g : [0,t] \to \mathbb{R}$ is defined as:

$$V(t) = \sup_{\substack{n \geq 1 \\ t_0 \leq \cdots \leq t_n}} \left\{ \sum_{i=0}^{n-1} |g(t_{i+1}) - g(t_i)| \right\}$$

where the supremum is taken over all partitions $0 = t_0 \leq \cdots \leq t_n = t$ of $[0,t]$. When $V(t)$ is finite, we say that g has *finite variations*. We will denote $V(g,t)$ when we want to specify the function under consideration. We say that $g : \mathbb{R} \to \mathbb{R}$ has finite variations when $t \to V(g,t)$ is finite. It can be shown that when a function has finite variations, the sup involved in the definition of the variation is also a limit as we consider any sequence of subdivisions with step size approaching 0.

If functions with finite variations are of interest here, it is because of the following result:

Proposition 14. *Let g be a continuous function on $[0,t]$. The following propositions are equivalent:*

(a) g has finite variations;

(b) g is the difference of two continuous, non-decreasing, positive functions.

Proof. The reverse implication $(b) \implies (a)$ is the easiest. Indeed, one can verify that a non-decreasing function indeed has finite variations. Moreover, using the triangle inequality, one can show that the sum of two functions with finite variations is still of finite variations.

Let's prove the converse, first in the case where $g(0) = 0$. It is easy to verify that $V : s \to V(g,s)$ is non-decreasing on $[0,t]$ and non-negative. Let's show that it is also the case for $W = V - g$. So, let $0 \leq s_1 \leq s_2 \leq t$, and we want to show that $V(s_1) - g(s_1) \leq V(s_2) - g(s_2)$. This is indeed the case since

$$V(s_1) + g(s_2) - g(s_1) \leq V(s_1) + |g(s_2) - g(s_1)| \leq V(s_2).$$

Furthermore, for any $0 \leq s \leq t$ and any subdivision $(s_i)_{i=1,\cdots,n}$ of $[0,s]$,

$$g(s) \leq \sum_{i=1}^{n} |g(s_i) - g(s_{i-1})| \leq V(s),$$

so $V - g$ is non-negative. When $g(0)$ is non-zero, we add it according to its sign to V or W. It remains to show that V is continuous (which implies that W is also continuous): an exercise (not so easy)! $\square$

With this proposition, we can construct the integral with respect to a function g with finite variations. We decompose it into $g = g^+ - g^-$ using Proposition 14. We associate g^+ and g^- with two measures μ^+ and μ^-, and we define the signed measure $\mu = \mu^+ - \mu^-$ (a signed measure is the difference of two positive measures). The decomposition of g into g^+ and g^- is not unique, and therefore, the associated pair of measures (μ^+, μ^-) is not unique either, but it can be shown that the measure μ does not depend on this choice (we admit it for this time!). We then define the positive measure $|\mu| = \mu^+ + \mu^-$ and for $f \in \mathrm{L}^1(|\mu|)$, we set

$$\int_0^t f(s)dg(s) = \int_0^t f(s)d\mu(s) = \int_0^t f(s)\mu^+(ds) - \int_0^t f(s)\mu^-(ds).$$

When f is continuous, (4) still remains valid.

The problem is that this theory of integration does not work with Brownian motion, whose trajectories, almost surely, do not have finite variations:

Proposition 15. *Let $(B_t)_{t\geq 0}$ be a Brownian motion and $t \in \mathbb{R}^+$. For any sequence of subdivisions $0 = t_0^n \leq \cdots \leq t_{p_n}^n = t$ with step sizes approaching 0,*

$$\sum_{k=1}^{p_n} (B_{t_k^n} - B_{t_{k-1}^n})^2 \xrightarrow{\mathrm{L}^2} t.$$

It follows that almost surely, the trajectories of Brownian motion have unbounded variations.

Proof (see [5]). Let $W_n = \sum_{k=1}^{p_n} (B_{t_k^n} - B_{t_{k-1}^n})^2$ and Δ_n be the step size of the n-th subdivision (we omit the superscript n to lighten the notation). Since the increments are stationary, for any $n \geq 1$,

$$\mathrm{E}(W_n) = \sum_{k=1}^{p_n} (t_k - t_{k-1}) = t.$$

Thus, as the increments are independent,

$$||W_n - t||_2^2 = \mathrm{Var}(W_n) = \sum_{k=1}^{p_n} \mathrm{Var}[(B_{t_k} - B_{t_{k-1}})^2].$$

Now, for any $1 \leq k \leq p_n$,

$$\text{Var}[(B_{t_k} - B_{t_{k-1}})^2] = (t_k - t_{k-1})^2 \text{Var}(B_1^2) = 2(t_k - t_{k-1})^2,$$

which leads to

$$||W_n - t||_2^2 = 2 \sum_{k=1}^{p_n} (t_k - t_{k-1})^2 \leq 2\Delta_n t,$$

concluding the proof of the first point since $\Delta_n \to 0$ as n tends to infinity.

We thus infer that there exists a subsequence of the subdivision sequence such that $(W_{\phi(n)})_{n \geq 1}$ tends to t almost surely. We then notice that

$$W_{\phi(n)} \leq \sup\{|B_{t_k} - B_{t_{k-1}}|, \ 1 \leq k \leq p_n\} V(B, t).$$

As $\sup\{|B_{t_k} - B_{t_{k-1}}|, \ 1 \leq k \leq p_n\} \to 0$ almost surely since Brownian motion is continuous and thus uniformly continuous on $[0, t]$, we deduce from the first point that almost surely $V(B, t) = +\infty$. □

Remark 3. *It can be shown that the quadratic variation (which consists of taking the square of the increments rather than their absolute values) of a $\mathcal{C}^1$ function is zero (see Exercise ?). The preceding proposition tells us that the variations of Brownian motion over small times are much larger than for smooth functions. Intuitively, for small $\delta > 0$, $B_{t+\delta} - B_t$ is of order $\sqrt{\delta}$, which is much larger than δ when δ is close from* 0.

Thus, it is not possible to define the integral with respect to Brownian motion trajectory by trajectory, and we will need to change strategy!

5.2 Itô Integral

The general idea is to define the integral as the extension of an isometry between two Hilbert spaces. If we consider a "simple" process of the form

$$u_t = \sum_{i=1}^{n} F_i 1_{]t_{i-1}, t_i]}(t) \quad t \geq 0,$$

with the F_i being random variables, it is natural to want

$$\int u_t \, dB_t = \sum_{i=1}^{n} F_i (B_{t_i^n} - B_{t_{i-1}^n}).$$

One could then approximate a general random process by a sequence of simple processes and take the L^2 limit (the previous part tells us that almost sure convergence is not possible). However, be careful: the way we approximate the random process can change the limit, as shown by the following example: what is $\int_0^T B_t \, dB_t$? We can approximate $(B_t)_{0\leq t\leq T}$ in an anticipative way by ϕ^n:

$$\phi_t^n = \sum_{i=0}^{n-1} B_{t_{i+1}^n} 1_{]t_i^n, t_{i+1}^n]}(s)$$

or in a non-anticipative way by ψ_n:

$$\psi_t^n = \sum_{i=0}^{n-1} B_{t_i^n} 1_{]t_i^n, t_{i+1}^n]}.$$

If we're not careful in the definition of the integral, then we obtain

$$\begin{aligned}
\int_0^T \phi_t^n \, dB_t - \int_0^T \psi_t^n \, dB_t &= \sum_{i=1}^{n-1} B_{t_{i+1}^n}(B_{t_{i+1}^n} - B_{t_i^n}) - \sum_{i=1}^{n-1} B_{t_i^n}(B_{t_{i+1}^n} - B_{t_i^n}) \\
&= \sum_{i=1}^{n-1} (B_{t_{i+1}^n} - B_{t_i^n})^2.
\end{aligned}$$

This last quantity converges in L^2 to T according to Proposition 15, so we see that the choice of approximation is not trivial! We will choose a **non-anticipative approximation**, which will ensure, among other things, that the stochastic integral is a martingale.

In this section, we consider a filtration $(\mathcal{F}_t)_{t\geq 0}$ that satisfies the usual conditions (see Definition 13), and $(B_t)_{t\geq 0}$ a $\mathcal{F}$-Brownian motion, i.e., a Brownian motion adapted to $\mathcal{F}$ and such that for all $0 \leq s \leq t$:

$$B_t - B_s \perp\!\!\!\perp \mathcal{F}_s.$$

Definition 17. *A random process $(X_t)_{t\geq 0}$ is said to be:*

1. ***measurable*** *if*

$$X : (\mathbb{R}^+ \times \Omega, \mathcal{B}(\mathbb{R}^+) \otimes \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

is measurable.

2. ***progressively measurable*** *if for all $t \geq 0$*

$$X : ([0,t] \times \Omega, \mathcal{B}([0,t]) \otimes \mathcal{F}_t) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

is measurable.

We note that if $(X_t)_{t\geq 0}$ is progressively measurable, then it is adapted and measurable:

1. *Adapted.* Let $t \geq 0$ and $A \in \mathcal{B}(\mathbb{R})$. Since $(X_s)_{s\geq 0}$ is progressively measurable, $\{X|_{[0,t]} \in A\} = \{(s,\omega) \in [0,t] \times \Omega \text{ such that } X_s(\omega) \in A\} \in \mathcal{B}([0,t]) \otimes \mathcal{F}_t$. We recall that if $\mathcal{E}$ and $\mathcal{F}$ are sigma-algebras on E and F, then for all $A \in \mathcal{E} \otimes \mathcal{F}$ and all $x \in E$, the section of A at x, $A_x = \{y \in F \text{ such that } (x,y) \in A\}$ belongs to $\mathcal{F}$. Here we deduce that $\{X_t \in A\} = \{X|_{[0,t]} \in A\}_t \in \mathcal{F}_t$.

2. *Measurable.* We verify that for all $A \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned}\{X \in A\} &= \{(s,\omega) \in \mathbb{R} \times \Omega \text{ such that } X_s(\omega) \in A\} \\ &= \cup_{n\geq 1}\{(s,\omega) \in [0,n] \times \Omega \text{ such that } X_s(\omega) \in A\}.\end{aligned}$$

   Now, since the random process is progressively measurable, for all $n \geq 1$, $\{(s,\omega) \in [0,n]\times\Omega \text{ such that } X_s(\omega) \in A\} \in \mathcal{B}([0,n])\otimes\mathcal{F}_n \subset \mathcal{B}(\mathbb{R}^+)\otimes\mathcal{F}$.

Proposition 16 (to demystify)**.** *Let $(X_t)_{t\geq 0}$ be an adapted process with càd trajectories. Then $(X_t)_{t\geq 0}$ is progressively measurable.*

Proof. Exercise: see TD 5. □

Now we can define the set of random processes that we will be able to integrate with respect to the Brownian motion:

Definition 18. *We denote by $\mathrm{L}^2(Prog)$ the set of progressively measurable processes $(u_s)_{s\geq 0}$ such that*

$$\mathrm{E}\int u_s^2\,ds < +\infty.$$

Proposition 17. *The space $\mathrm{L}^2(Prog)$ is a Hilbert space.*

Proof. The line of the proof is to show that $\mathrm{L}^2(Prog)$ is in fact a "normal" L^2, and this is a consequence of the following lemma:

Lemma 3. *We define*

$$\mathcal{P} = \{A \in \mathcal{B}(\mathbb{R}^+)\otimes\mathcal{F} \text{ such that } X : (t,\omega) \to 1_A(t,\omega) \text{ is progressively measurable}\}.$$

Then $\mathcal{P}$ is a $\sigma-$algebra (called ***progressive $\sigma-$algebra****) and a random process $(X_t)_{t\geq 0}$ is progressively measurable if and only if the function*

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

is $\mathcal{P}$-measurable.

Proof. Let's first verify that $\mathcal{P}$ is a $\sigma-$algebra. For $A \in \mathcal{B}(\mathbb{R}^+) \otimes \mathcal{F}$, we denote $I(A)|_t$ the function defined on $([0,t]\times\Omega, \mathcal{B}([0,t]\otimes\mathcal{F}_t)$ by $I(A)|_t(s,\omega) = 1_A(s,\omega)$. With these notations, $A \in \mathcal{P}$ if and only if for all $t \geq 0$, $I(A)|_t$ is measurable.

1. We easily deduce that $\mathbb{R}^+ \times \Omega \in \mathcal{P}$.

2. Let $A \in \mathcal{P}$ and $t \geq 0$. Since $I(A^c)|_t = 1 - I(A)|_t$, it is also a measurable function. Hence $A^c \in \mathcal{P}$.

3. Let $(A_n)_{n\geq 1}$ be a sequence in $\mathcal{P}$ and $t \geq 0$. We have $I(\cup_{n\geq 1} A_n)|_t = 1 - \lim_{N\to+\infty} \prod_{n=1}^{N} I(A_n^c)|_t$. We deduce that $\cup_{n\geq 1} A_n \in \mathcal{P}$.

Now we consider a random process $(X_t)_{t\geq 0}$ and $B \in \mathcal{B}(\mathbb{R})$. We assume X is progressively measurable. We must show that $\{X \in B\} \in \mathcal{P}$, i.e., for all $t \geq 0$, $I(X \in B)|_t$ is measurable. Now $\{I(X \in B)|_t = 1\} = \{(s,\omega) \in [0,t] \times \Omega$ such that $X(s,\omega) \in B\} \in \mathcal{B}([0,t] \otimes \mathcal{F}_t$ because X is progressively measurable. Conversely, if X is $\mathcal{P}$-measurable, fixing $t \geq 0$ with the established equality shows that X is progressively measurable. $\square$

In any case, this allows us to see progressively measurable processes as square integrable random variables on $(\mathbb{R}^+ \times \Omega, \mathcal{P})$: $\mathrm{L}^2(Prog)$ is therefore a Hilbert space like any L^2,

$$\mathrm{L}^2(Prog) = \mathrm{L}^2(\mathbb{R}^+ \times \Omega, \mathcal{P}, \mathrm{Leb} \otimes \mathrm{P}).$$

The associated inner product is defined for all random processes $u, v \in \mathrm{L}^2(Prog)$ by

$$< u, v >= \mathrm{E} \int u_t v_t \, dt,$$

and the norm by

$$||u||^2_{\mathrm{L}^2(Prog)} = \mathrm{E} \int u_t^2 \, dt.$$

$\square$

We now focus on a subset of random processes in $\mathrm{L}^2(Prog)$ for which stochastic integration will be *easily* defined.

Definition 19. *A random process $(u_t)_{t\geq 0}$ is called* ***step process or predictable simple process*** *if it is of the form*

$$u_t : \sum_{i=0}^{n-1} F_i \, 1_{[t_i, t_{i+1}[}(t) \quad t \geq 0,$$

where $n \geq 0$ is an integer, $0 = t_0 \leq \cdots \leq t_n$, and for all $i = 1, \cdots, n$, the variable F_i is $\mathcal{F}_{t_i}$-measurable and square integrable. We denote by $\mathcal{E}$ the set of step processes.

Note that a step process is càdlàg. We will be able to define stochastic integration easily on $\mathcal{E}$, but to extend it, we need to ensure that $\mathcal{E}$ is large enough:

Proposition 18. *The set $\mathcal{E}$ is a dense linear subset of $\mathrm{L}^2(Prog)$.*

Proof. It is easy to verify that $\mathcal{E}$ has a vector space structure. We then verify $\mathcal{E} \subset \mathrm{L}^2(Prog)$. Let $0 \leq t_1 \leq t_2$ and F be a $\mathcal{F}_{t_1}$-measurable random variable. We must show that the random process $u : F 1_{[t_1,t_2[}$ is progressively measurable. For all $T \geq 0$,

$$\begin{aligned}\{u|_{[0,T]} \in A\} = \quad & \{(t,\omega) \text{ such that } t < t_1, t \leq T, 0 \in A\} \\ & \cup \{(t,\omega) \text{ such that } t_1 \leq t < t_2, t \leq T, F \in A\} \\ & \cup \{(t,\omega) \text{ such that } t_2 < t, t \leq T, 0 \in A\}\end{aligned}$$

Only the middle set is a bit more delicate: it is equal to $([t_1, t_2[\cap [0,T]) \times \{F \in A\}$, which belongs to $\mathcal{B}([0,T]) \otimes \mathcal{F}_T$. Note that this is not the case if F is only $\mathcal{F}_{t_2}$-measurable.

We now need to prove density. Two methods (at least!) are possible. The first one consists of approximating $u \in \mathrm{L}^2(Prog)$ by a sequence of random processes in $\mathcal{E}$ (passing through bounded processes, then bounded with compact support, then adding continuity and finally arriving at an approximating sequence in $\mathcal{E}$); it can be found in [1, 5, 4] and I recommend referring to it as it is instructive. Here we will follow another approach found in [7] which consists of showing that $\mathcal{E}^\perp = \{0\}$. For this, we need two lemmas (which are interesting on their own).

Lemma 4. *Let $u \in \mathrm{L}^2(prog)$ such that for all $t \geq 0$ and all $\omega \in \Omega$, $\int_0^t |u_s(\omega)|\, ds < +\infty$. Then $t \to \int_0^t u_s\, ds$ is of finite variations and adapted.*

Proof. (see also [7] for this proof) *Adapted.* We fix $t \geq 0$. We recall that $u|_{[0,t]}$ is $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$-measurable and we must show that $\int_0^t u_s\, ds$ is $\mathcal{F}_t$-measurable. If u is of the form $1_A 1_{[\alpha,\beta[}$ with $0 \leq \alpha \leq \beta \leq t$ and $A \in \mathcal{F}_t$, we easily verify that $\int_0^t u_s\, ds = 1_A(\beta - \alpha)$ is $\mathcal{F}_t$ measurable. We now consider the class $\mathcal{M} = \{\Gamma \in \mathcal{B}([0,t]) \otimes \mathcal{F}_t \text{ such that } \int_0^t 1_\Gamma\, ds \text{ is } \mathcal{F}_t\text{-measurable}\}$. It is a $\lambda-$system containing all Γ of the form $[\alpha, \beta[\times A$, and by the Dynkin's lemma, we deduce that $\mathcal{M}$ contains all $\mathcal{B}([0,t]) \otimes \mathcal{F}_t)$. So, we have treated the case of $u = 1_\Gamma$ where $\Gamma \in \mathcal{B}([0,t]) \otimes \mathcal{F}_t$. Next, we easily obtain the case of a step process u, then of any $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$-measurable process by approximation by step processes and the dominated convergence theorem with respect to the Lebesgue measure on $[0,t]$.

Finite variations. Moreover, $\int_0^t u_s, ds = \int_0^t u_s^+, ds - \int_0^t u_s^-, ds$ can be written as the difference of two non-decreasing functions, so it is indeed of finite variations. $\square$

Lemma 5. *Let $(M_t)_{t\geq 0}$ be a continuous martingale, of finite variations, and such that $M_0 = 0$ a.s. Then $(M_t)_{t\geq 0}$ is indistinguishable from 0.*

Proof. (see also [5] for this proof) Let us first suppose that $(M_t)_{t\geq 0}$ has bounded variations almost surely, i.e., almost surely, for all $t \geq 0$, $V(M(\omega),t) < K$. For any subdivision of $[0,t]$,

$$\mathrm{E}(M_t^2) = \mathrm{E}\left(\sum_{i=0}^{n-1} M_{t_{i+1}}^2 - M_{t_i}^2\right) \overset{\text{Mart.}}{=} \sum_{i=0}^{n-1} \mathrm{E}\left[(M_{t_{i+1}} - M_{t_i})^2\right] \leq K\mathrm{E}\left(\max |M_{t_{i+1}} - M_{t_i}|\right).$$

Let δ_n be a sequence of subdivision steps tending to 0. Then, by uniform continuity, we have almost sure convergence of $\max |M_{t_{i+1}} - M_{t_i}|$ to 0. Furthermore, $|M|$ is bounded by K, so we can use the dominated convergence theorem, and we deduce that $\mathrm{E}(M_t^2) = 0$, hence almost surely $M_t = 0$. By using the fact that M is a version of the null process and continuity, we obtain that M is indistinguishable from 0.

If M only has finite variations, we introduce the sequence of stopping times

$$\tau_k = \inf\{s \text{ such that } V(s) \geq k\}, \qquad k \geq 1.$$

The stopped martingales $(M_{\tau_k \wedge t})_{t\geq 0}$ have bounded variations, and using what we have just proven, we deduce that for all $k \geq 1$ almost surely $M_{\tau_k \wedge t} = 0$ for all $t \geq 0$. We then exchange the quantifier that applies to a countable set and almost sure convergence to conclude. $\square$

With these two lemmas in hand, let's go back to our proof of the density of $\mathcal{E}$ in $\mathrm{L}^2(Prog)$. We consider $u \in \mathcal{E}^\perp$ and we want to show that $u = 0$. For this purpose, we define the random process

$$M_t = \int_0^t u_s \; ds \quad t \geq 0.$$

This is a well-defined integral almost surely since

$$\mathrm{E}|M_t| \leq \mathrm{E}\int_0^{+\infty} |u_s| \; 1_{[0,t]}(s) \; ds \overset{\text{C.S.}}{\leq} \sqrt{t}\mathrm{E}\left(\int_0^{+\infty} |u_s|^2 \; ds\right)^{1/2} < +\infty.$$

We even note that $M_t \in \mathrm{L}^1$. Let's show that the random process $(M_t)_{t\geq 0}$ is a martingale. We have just seen that it is integrable, and Lemma 4 assures us that it is adapted. To show the martingale property, we consider $0 \leq s \leq t$ and F a $\mathcal{F}_s$-measurable random variable. We set $v = F1_{[s,t[}$. Since $v \in \mathcal{E}$, we have $< v \cdot u >_{\mathrm{L}^2(Prog)} = 0$. However,

$$< v \cdot u >_{\mathrm{L}^2(Prog)} = \mathrm{E} \int_0^{+\infty} v_r u_r \ dr = \mathrm{E}(F \int_s^t u_r \ dr) = \mathrm{E}(F(M_t - M_s)).$$

We deduce that $(M_t)_{t\geq 0}$ is a martingale.

Moreover, $(M_t)_{t\geq 0}$ is also a random process of finite variation according to Lemma 4. Therefore, by Lemma 5, almost surely, for all $t \geq 0$, $M_t = 0$. So, almost surely, for all $t \geq 0$, $\int_0^t u_r, dr = 0$, and we deduce that $u_r = 0$ almost everywhere and almost surely. To show this last point, we can, for example, show that almost surely u is in the orthogonal of the set of step functions, which is dense in $\mathrm{L}^2(\mathbb{R})$ (but perhaps we can find a shorter argument). □

We now have everything we need to define the stochastic integral.

Theorem 12 (Itô Stochastic Integral)**.** *There exists a unique linear mapping*

$$I : \mathrm{L}^2(Prog) \to \mathrm{L}^2(\Omega, \mathcal{F}, \mathbb{P})$$

such that

1. *For every random process* $u = \sum_{i=0}^{n-1} F_i \ 1_{[t_i,t_{i+1}[}$ *in* $\mathcal{E}$*,*

$$I(u) = \sum_{i=0}^{n-1} F_i \ (B_{t_{i+1}} - B_{t_i}). \tag{5}$$

2. *For every* $u \in \mathrm{L}^2(Prog)$*,*

$$||I(u)||^2_{\mathrm{L}^2(\mathcal{F})} = ||u||^2_{\mathrm{L}^2(Prog)}$$

in other words,

$$\mathrm{E}(I(u)^2) = \mathrm{E} \int u_s^2 \ ds$$

or equivalently, I *is an* ***isometry****.*

Proof. The idea is to use the theorem of extension of isometries between two Hilbert spaces. Since we have already proven the density of $\mathcal{E}$ in $\mathrm{L}^2(Prog)$ in

Proposition 18, it remains to show that the relation (5) defines an isometry from $\mathcal{E}$ into $\mathrm{L}^2(\mathcal{F})$. We verify that

$$
\begin{aligned}
||I(u)||^2_{\mathrm{L}^2(\mathcal{F})} &= \mathrm{E}[(\sum_{i=0}^{n-1} F_i \ (B_{t_{i+1}} - B_{t_i}))^2] \\
&= \mathrm{E}(\sum_{i=0}^{n-1} F_i^2 \ (B_{t_{i+1}} - B_{t_i})^2) + 2 \sum_{0 \le i < j \le n-1} \mathrm{E}(F_i F_j \ (B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})) \\
&= A + B.
\end{aligned}
$$

Now

$$
\begin{aligned}
A &= \sum_{i=0}^{n-1} \mathrm{E}(\underbrace{F_i^2}_{\mathcal{F}_{t_i}-\text{meas.}} \ \underbrace{(B_{t_{i+1}} - B_{t_i})^2)}_{\text{indep. of } \mathcal{F}_{t_i}} \\
&= \sum_{i=0}^{n-1} \mathrm{E}(F_i^2)(t_{i+1} - t_i) \\
&= \mathrm{E} \int_0^t u_s^2 \ ds \\
&= ||u||^2_{\mathrm{L}^2(Prog)}.
\end{aligned}
$$

Moreover, B is zero because for all $i < j$,

$$
\mathrm{E}(\underbrace{F_i F_j \ (B_{t_{i+1}} - B_{t_i})}_{\mathcal{F}_{t_j}-\text{meas.}} \underbrace{(B_{t_{j+1}} - B_{t_j}))}_{\text{indep. of } \mathcal{F}_{t_j}} = 0.
$$

□

We note, in passing, some initial properties of the stochastic integral. Since I is an isometry, for all $u, v \in \mathrm{L}^2(Prog)$

$$
< I(u), I(v) >_{\mathrm{L}^2(\mathcal{F})} = < u, v >_{\mathrm{L}^2(Prog)},
$$

which can be rewritten as

$$
\mathrm{E}\left(\int u_s \ dB_s \int v_s \ dB_s \right) = \mathrm{E} \int u_s v_s \ ds.
$$

Furthermore, for all $u \in \mathcal{E}$,

$$
\mathrm{E}(I(u)) = \sum_{i=0}^{n-1} \mathrm{E}\left(F_i \ (B_{t_{i+1}} - B_{t_i}) \right).
$$

Now for every i, F_i is $\mathcal{F}_{t_i}$-measurable and since $(B_t)_{t\geq 0}$ is an $\mathcal{F}$-Brownian motion, we deduce that $F_i \perp\!\!\!\perp B_{t_{i+1}} - B_{t_i}$, and thus $\mathrm{E}(I(u)) = 0$. We extend this property to all $u \in \mathrm{L}^2(Prog)$ by approximating u by a sequence $(u_n)_{n\geq 1}$ of simple predictable processes. We then obtain that $(I(u_n))_{n\geq 1}$ converges in $\mathrm{L}^2(\mathcal{F})$ and therefore in $\mathrm{L}^1(\mathcal{F})$ to $I(u)$, which implies the convergence of the expectations. Thus, we have successfully shown that for all $u \in \mathrm{L}^2(Prog)$,

$$\mathrm{E} \int u_t \; dB_t = 0.$$

Consider a random process $u \in \mathrm{L}^2(Prog)$. Note that for every $t \geq 0$, the random process $u1_{[0,t]}$ is also in $\mathrm{L}^2(Prog)$ (it is $\mathcal{P}$-measurable as a product of measurable variables, and being square-integrable poses no problem), and thus we can define its integral

$$X_t = \int_0^t u_s \; dB_s = \int u_s 1_{[0,t]}(s) \; dB_s.$$

The choice of a non-anticipative definition of step processes allows us to obtain the following property

Proposition 19. *Let $u \in \mathrm{L}^2(Prog)$. Then the random process*

$$X_t = \int_0^t u_s \; dB_s \qquad t \geq 0,$$

is a bounded martingale in L^2 continuous (i.e. which admits a continuous modification). For all $t \geq 0$,

$$||X_t||_2^2 = \mathbb{E} \int_0^t u_s^2 \; ds.$$

We note that this proposition would be false if we had not made the choice of a non anticipative definition of the stochastic integral.

Proof. The line of the proof to show both continuity and martingale character is to show first these properties when $u \in \mathcal{E}$ then to show that they are preserved by approximation.

Noting to start that for all $u \in \mathrm{L}^2(Prog)$ and all $t \geq 0$,

$$||X_t||_2^2 = ||I(u1_{[0,t]})||_2^2 = \mathbb{E} \int |u_s 1_{[0,t]}(s)|^2 \; ds = \mathbb{E} \int_0^t u_s^2 \; ds.$$

We deduce that for all $t \geq 0$, X_t is in L^2 (thus in L^1) and even that the random process $(X_t)_{t\geq 0}$ is bounded in L^2 by $\mathbb{E} \int u_s^2 \; ds$ which is well finite since $u \in \mathrm{L}^2(Prog)$.

Let's now show that $(X_t)_{t\geq 0}$ is a martingale. We thus assume in a first step that $u = \sum_{i=0}^{n-1} F_i \; 1_{[t_i, t_{i+1}[}$ is a random process of $\mathcal{E}$. We check that for all $t \geq 0$,

$$X_t = \sum_{i=0}^{n-1} F_i \; (B_{t_{i+1}\wedge t} - B_{t_i \wedge t}) \tag{6}$$

and we deduce that X_t is $\mathcal{F}_t$-measurable and thus $(X_t)_{t\geq 0}$ is adapted. Let's now show the martingale property and fix for that $0 \leq s \leq t$. If s and t are in the same interval, i.e. for a certain k, s and t are in $]t_k, t_{k+1}]$ then

$$\int_0^t u_r \; dB_r - \int_0^s u_r \; dB_r = F_k(B_t - B_s).$$

We easily conclude because F_k is $\mathcal{F}_{t_k}$-measurable thus $\mathcal{F}_s$ - measurable thus $\mathbb{E}(F_k(B_t - B_s)|\mathcal{F}_s) = F_k \mathbb{E}(B_t - B_s|\mathcal{F}_s) = 0$. We now move on to the case where s and t are not in the same interval: $t_\ell < s \leq t_{\ell+1} \leq t_k < t \leq t_{k+1}$. We then have

$$\int_0^t u_r \; dB_r - \int_0^s u_r \; dB_r = F_\ell(B_{t_{\ell+1}} - B_s) + \sum_{i=\ell+1}^{k-1} F_i(B_{t_{i+1}} - B_{t_i}) + F_{t_k}(B_t - B_{t_k}).$$

When considering the conditional expectation with respect to $\mathcal{F}_s$, we can handle the first term as in the previous case, and for $i \in \{\ell+1, \cdots, k-1\}$,

$$\begin{aligned}
\mathbb{E}(F_i(B_{t_{i+1}} - B_{t_i})|\mathcal{F}_s) &= \mathbb{E}[\mathbb{E}[F_i(B_{t_{i+1}} - B_{t_i})|\mathcal{F}_{t_i}]|\mathcal{F}_s] \\
&= \mathbb{E}[F_i \mathbb{E}[(B_{t_{i+1}} - B_{t_i})|\mathcal{F}_{t_i}]|\mathcal{F}_s] \\
&= 0,
\end{aligned}$$

where we used that F_i is $\mathcal{F}_{t_i}$-measurable. It should be noted that this calculation is no longer possible if the random process we are integrating is anticipative. We treat the last term similarly.

We no longer assume that $u \in \mathcal{E}$ but only that it is in $\mathrm{L}^2(Prog)$ and we consider an approximating sequence $(u^n)_{n\geq 0}$ in $\mathcal{E}$. For all $t \geq 0$

$$||u^n 1_{[0,t]} - u 1_{[0,t]}||_{\mathrm{L}^2(Prog)} \leq ||u^n - u||_{\mathrm{L}^2(Prog)}$$

thus $(u^n 1_{[0,t]})$ converges in $\mathrm{L}^2(Prog)$ to $u1_{[0,t]}$ and we deduce that $(\int_0^t u_s^n \; dB_s)$ converges in $\mathrm{L}^2(\mathcal{F})$ to $\int_0^t u_s \; dB_s$. As for all $n \geq 1$, $\int_0^t u_s^n \; dB_s$ is $\mathcal{F}_t$-measurable (since $u^n \in \mathcal{E}$) we deduce that $\int_0^t u_s \; dB_s$ is also $\mathcal{F}_t$-measurable. We have well proved that $(X_t)_{t\geq 0}$ is adapted. We move on to the martingale property. We fix $0 \leq s \leq t$. For all $n \geq 1$, as $u^n \in \mathcal{E}$,

$$\mathbb{E}\left(\int_0^t u_r^n \; dB_r|\mathcal{F}_s\right) = \int_0^s u_r^n \; dB_r.$$

We have already seen using the fact that the stochastic integral is an isometry that $(\int_0^s u_r^n \ dB_r)$ converges in $\mathrm{L}^2(\mathcal{F})$ to $\int_0^s u_r \ dB_r$. Furthermore, as conditional expectation is a continuous operator on $\mathrm{L}^2(\mathcal{F})$, we obtain that $(\mathbb{E}(\int_0^t u_r^n \ dB_r|\mathcal{F}_s))$ converges, still in $\mathrm{L}^2(\mathcal{F})$, to $\mathbb{E}(\int_0^t u_r \ dB_r|\mathcal{F}_s)$. We have thus shown the martingale property.

We now need to show the continuity of $(X_t)_{t\geq 0}$. Again, when $u \in \mathcal{E}$, it is easy to conclude using (6) because the paths of the Brownian motion are continuous. For the general case, we fix $T > 0$ and consider an approximating sequence (u^n). According to Doob's inequality (which we can use because we have the continuity of the martingale for processes in $\mathcal{E}$), for all $n, m \geq 0$,

$$\begin{aligned}\mathbb{P}\left(\sup_{0\leq t\leq t} |\int_0^T u_s^n - u_s^m \ dB_s| > \epsilon\right) &\leq \frac{\mathbb{E}(|\int_0^T u_s^n - u_s^m \ dB_s|^2)}{\epsilon^2} \\ &\overset{Iso.}{=} \frac{\mathbb{E}(\int_0^T (u_s^n - u_s^m)^2 \ ds)}{\epsilon^2} \\ &= \frac{||u^n - u^m||^2_{\mathrm{L}^2(Prog)}}{\epsilon^2}.\end{aligned}$$

We can thus extract a subsequence $(n_k)_{k\geq 1}$ such that for all $k \geq 1$

$$\mathbb{P}\left(\sup_{0\leq t\leq T} |\int_0^t u_s^{n_k} - u_s^{n_{k+1}} \ dB_s| > \frac{1}{2^k}\right) \leq \frac{1}{2^k}$$

and by the Borel-Cantelli lemma we obtain that almost surely for large enough k

$$\sup_{0\leq t\leq T} |\int_0^t u_s^{n_k} - u_s^{n_{k+1}} \ dB_s| \leq \frac{1}{2^k}.$$

We deduce that almost surely the sequence of functions $(t \to \int_0^t u_s^{n_k} \ dB_s)$ converges uniformly to a limit which is thus continuous. This limit is indeed $t \to \int_0^t u_s \ dB_s$ because for all $0 \leq t \leq T$, we have L^2 convergence of $(\int_0^t u_s^{n_k} \ dB_s)$ to $\int_0^t u_s \ dB_s$. □

We have made progress... but for now, we can only integrate processes that are in $\mathrm{L}^2(Prog)$ thus satisfying, in addition to the progressive measurability condition, the condition $\mathbb{E}\int u_t^2 \ dt < +\infty$. This is actually quite restrictive and we can do better at a lower cost.

Definition 20. *We denote L^2_{loc} the set of progressively measurable processes such that for all $t \geq 0$,*

$$a.s. \int_0^t u_s^2 \ ds < +\infty.$$

It should be noted that in this definition we could exchange "a.s." and "$\forall$" since $\int_0^t u_s^2\ ds$ is non-decreasing in t and thus it is sufficient to check that these quantities are finite for integer t. It is clear that $\mathrm{L}^2_{loc} \not\subseteq \mathrm{L}^2(Prog)$ but the interest of considering this set lies in the **localization property** as follows:

Lemma 6. *Let $u \in \mathrm{L}^2_{loc}$. Then there exists an non-decreasing sequence of stopping times $(T_n)_{n\geq 1}$ such that*

1. *a.s. $T_n \nearrow +\infty$ $(n \to +\infty)$,*
2. *$\mathbb{E} \int_0^{T_n} u_t^2\ dt < +\infty$ for all $n \geq 1$.*

Proof. We consider for all $n \geq 1$,

$$T_n = \inf\{t \geq 0 \text{ such that } \int_0^t u_s^2\ ds \geq n\}. \tag{7}$$

We check that it is indeed a stopping time because for all $r \geq 0$

$$\{T_n \leq r\} = \left\{\int_0^r u_s^2\ ds \geq n\right\},$$

and, reasoning as in Lemma 4, we check that $\int_0^r u_s^2\ ds$ is $\mathcal{F}_r$-measurable. It is also clear that T_n is non-decreasing and its limit is $+\infty$ by definition of L^2_{loc}. $\square$

We deduce that if $u \in \mathrm{L}^2_{loc}$ then for all $n \geq 1$, $u1_{[0,T_n]}$ is in $\mathrm{L}^2(Prog)$ (exercise: verify the progressively measurability by showing that for any stopping time T the random process $t \to 1_{[0,T]}(t)$ is progressively measurable) and thus we can define

$$\int_0^{T_n} u_s\ dB_s = \int u_s 1_{[0,T_n]}(s)\ dB_s$$

and for all $t \geq 0$,

$$M_t^{(n)} := \int_0^t u_s 1_{[0,T_n]}(s)\ dB_s = \int u_s 1_{[0,T_n]}(s) 1_{[0,t]}(s)\ dB_s.$$

For all $m \geq n$ and $t \geq 0$, we verify that $M^{(m)}_{t\wedge T_n} = M_t^{(n)}$, implying that almost surely for n large enough (such that $T_n \geq t$) $M_t^{(n)}$ is constant and this limit is by definition

$$\int_0^t u_s\ dB_s \overset{\text{a.s.}}{=} \lim_{n\to+\infty} \int_0^t u_s 1_{[0,T_n]}(s)\ dB_s.$$

Here are the first important properties of this stochastic integral defined on L^2_{loc}:

1. **When $u \in \mathrm{L}^2(Prog)$, the two definitions coincide.** Indeed, in this case the sequence of processes $(s \to u_s 1_{[0,t]}(s) 1_{[0,T_n]}(s))_{n\geq 0}$ converges in $\mathrm{L}^2(Prog)$ to $s \to u_s 1_{[0,t]}(s)$ by the dominated convergence theorem, and we deduce that $(\int_0^t u_s 1_{[0,T_n]}(s)\, dB_s)_{n\geq 0}$ converges in $\mathrm{L}^2(\mathcal{F})$ to $\int_0^t u_s\, dB_s$

2. **The random process $(\int_0^t u_s\, dB_s)_{t\geq 0}$ is adapted and continuous.** For all $n \geq 1$, $\int_0^t u_s 1_{[0,T_n]}(s)\, dB_s$ is $\mathcal{F}_t$ -measurable as it is the integral of a random process in $\mathrm{L}^2(Prog)$, and we deduce that it is the same for $\int_0^t u_s\, dB_s$ that is the a.s. limit of the sequence. Moreover, for all $0 \leq v \leq t$, almost surely as soon as $T_n \geq t$, $\int_0^v u_s\, dB_s$ coincides with $\int_0^v u_s 1_{[0,T_n]}(s)\, dB_s$ which gives the continuity of our random process on $[0, t]$.

3. The random process $(\int_0^t u_s\, dB_s)_{t\geq 0}$ is not necessarily a bounded martingale in L^2 nor even necessarily a martingale. However, it is a **local martingale**:

   Definition 21. *An adapted and continuous process $(M_t)_{t\geq 0}$ is called a local martingale if there exists an non-decreasing sequence of stopping times $(T_n)_{n\geq 1}$ such that*

   (a) almost surely $T_n \nearrow +\infty$ $(n \to +\infty)$,

   (b) for all $n \geq 1$, $(M_{t\wedge T_n})_{t\geq 0}$ is a uniformly integrable martingale.

   Such a stopping time sequence is said to be reducing $(M_t)_{t\geq 0}$.

   It should be noted that this definition does not imply that $(M_t)_{t\geq 0}$ is integrable.

   Thus if $u \in \mathrm{L}^2_{loc}$, by using the $(T_n)_{n\geq 1}$ defined in (7), for all $n \geq 1$ and $t \geq 0$,
$$\int_0^{t\wedge T_n} u_s\, dB_s = \int_0^t 1_{[0,T_n]} u_s\, dB_s$$
   is a bounded martingale in L^2 and thus uniformly integrable. We obtain that $(\int_0^t u_s\, dB_s)_{t\geq 0}$ is indeed a local martingale.

4. The integral defined on L^2_{loc} is no longer an isometry and we only have the inequality
$$\mathbb{E}\left[\left(\int_0^t u_s\, dB_s\right)^2\right] \leq \mathbb{E}\left(\int_0^t u_s^2\, ds\right). \tag{8}$$

Indeed, if $\mathbb{E}\left(\int_0^t u_s^2\, ds\right) < +\infty$, then $u1_{[0,t]} \in \mathrm{L}^2(Prog)$ and we have the equality. And if $\mathbb{E}\left(\int_0^t u_s^2\, ds\right) = +\infty$, there is nothing to prove!

Now that we have defined the stochastic integral we can define an important class of random processes:

Definition 22 (Itô Process)**.** *A random process $(X_t)_{t\geq 0}$ is an Itô process if it can be written as*

$$\text{a.s. for all } t \geq 0 \quad X_t = X_0 + \int_0^t v_s\, ds + \int_0^t u_s\, dB_s, \tag{9}$$

where

1. *X_0 is $\mathcal{F}_0$-measurable;*
2. *the random process v is progressively measurable and satisfies $\int_0^t |v_s|\, ds < +\infty$ for all $t \geq 0$,*
3. *the random process u belongs to L^2_{loc}.*

For an Itô process u, we will often use the notation

$$dX_t = v_t\, dt + u_t\, dB_t,$$

which means nothing but (9).

5.3 Itô's Formula

Let's start with a somewhat informal calculation to understand the necessity of Itô's formula. Consider a function ϕ of class $\mathcal{C}^2$ on $\mathbb{R}$ and x of class $\mathcal{C}^1$ on $[0, t]$. We then obtain the following expansion:

$$\begin{aligned}
\phi(x(t)) &= \phi(x(0)) + \sum_{i=1}^{n} \left(\phi(x(\frac{it}{n})) - \phi(x(\frac{(i-1)t}{n})) \right) \\
&= \phi(x(0)) + \sum_{i=1}^{n} \phi' \left(x(\frac{(i-1)t}{n}) \right) \left(x(\frac{it}{n}) - x(\frac{(i-1)t}{n}) \right) \\
&\quad + \sum_{i=1}^{n} \frac{1}{2}\phi'' \left(x(\frac{(i-1)t}{n}) \right) \left(x(\frac{it}{n}) - x(\frac{(i-1)t}{n}) \right)^2 \\
&\quad + \text{Remainder.}
\end{aligned}$$

As n tends to infinity, the first sum converges to $\int_0^t \phi'(x(s))x'(s)ds = \int_0^t \phi'(x(s))dx(s)$. The second sum converges to 0 because the second derivative is bounded and the quadratic variation of x is null. Therefore, we finally obtain the classical formula:

$$\phi(x(t)) = \phi(x(0)) + \int_0^t \phi'(x(s))dx(s).$$

This calculation becomes incorrect if we take Brownian motion for x because the quadratic variation is no longer null and the second term cannot be neglected. Thus, Itô's formula proposes a corrected version that takes this term into account:

Theorem 13 (Itô's Formula)**.** *Let $\phi \in \mathcal{C}^2(\mathbb{R})$ such that ϕ, ϕ', and ϕ'' are bounded. Then almost surely, for all $t \geq 0$,*

$$\phi(B_t) = \phi(B_0) + \int_0^t \phi'(B_s)dB_s + \frac{1}{2}\int_0^t \phi''(B_s)ds.$$

Proof. (This presentation comes from [5]) Let $t_i = \frac{ti}{n}$ for $i = 0, \cdots, n$ and write the Taylor expansion to order 2 at each time step:

$$\phi(B_t) = \phi(B_0) + \sum_{i=1}^n \phi'(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) + \frac{1}{2}\sum_{i=1}^n \phi''(B_{\theta_i})(B_{t_i} - B_{t_{i-1}})^2,$$

where for each i, θ_i is an appropriate point in $[t_{i-1}, t_i]$ (we used the intermediate value theorem).

First step. Interpret $\sum_{i=1}^n \phi'(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}})$ as $\int_0^t X_s^n \, dB_s$ where $X_s^n = \sum_{i=1}^n \phi'(B_{t_{i-1}})1_{[t_{i-1},t_i[}(s)$ for $s \in [0,t]$. Show that $(X_s^n)s \geq 0$ converges in $\mathbb{L}^2(Prog)$ to $(\phi'(B_s))s \geq 0$. This follows from the continuity of ϕ' and convergence $\mathbb{P}$-a.s. and a.s. for X_s^n to $\phi'(B_s)$. Also, since ϕ' is bounded, we deduce from the dominated convergence theorem

$$\mathbb{E}\int \left[X_s^n - \phi'(B_s)\right]^2 \, ds \to 0.$$

Using the fact that the Itô integral is an isometry, we obtain

$$\sum_{i=1}^n \phi'(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) \xrightarrow{\mathbb{L}^2} \int \phi'(B_s) \, dB_s.$$

Second step. To study the second sum

$$U_n = \sum_{i=1}^n \phi''(B_{\theta_i})(B_{t_i} - B_{t_{i-1}})^2,$$

make two substitutions: replace θ_i with t_{i-1} to define

$$V_n = \sum_{i=1}^{n} \phi''(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}})^2,$$

then replace $(B_{t_i} - B_{t_{i-1}})^2$ with $t_i - t_{i-1}$ and define

$$W_n = \sum_{i=1}^{n} \phi''(B_{t_{i-1}})(t_i - t_{i-1}).$$

Firstly, control the $\mathbb{L}^1$ distance between U_n and V_n:

$$\begin{aligned} \mathrm{E}|U_n - V_n| &\leq \mathrm{E}\left(\sup_i |\phi''(B_{\theta_i}) - \phi''(B_{t_{i-1}})| \sum_{i=1}^{n} (B_{t_i} - B_{t_{i-1}})^2\right) \\ &\overset{C.S.}{\leq} \mathrm{E}[\sup_i |\phi''(B_{\theta_i}) - \phi''(B_{t_{i-1}})|^2]^{1/2} \mathrm{E}\left((\sum_{i=1}^{n} (B_{t_i} - B_{t_{i-1}})^2)^2\right)^{1/2} \end{aligned}$$

For the first term in the product: it tends to 0 a.s. by uniform continuity and is dominated by 2 times the supremum of ϕ''. Thus, this term tends to 0 by the dominated convergence theorem. For the second term of the product, note that $\sum_{i=1}^{n} (B_{t_i} - B_{t_{i-1}})^2$ tends to t in $\mathbb{L}^2$, and it follows that this second term converges to t. Finally, it is proved that

$$U_n - V_n \xrightarrow{\mathbb{L}^1} 0.$$

Third step. Now control the $\mathbb{L}^2$ distance between V_n and W_n:

$$\begin{aligned} \mathrm{E}(|V_n - W_n|^2) &= \mathrm{E}\left(\left|\sum_{i=1}^{n} \phi''(B_{t_{i-1}})[(B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})]\right|^2\right) \\ &= \sum_{i=1}^{n} \mathrm{E}\left(\left|\phi''(B_{t_{i-1}})\right|^2 [(B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})]^2\right) \\ &\leq ||\phi''||_\infty^2 \sum_{i=1}^{n} \mathrm{E}\left([(B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})]^2\right) \end{aligned}$$

where the second equality follows from the fact that the expectations of the crossed terms cancel out. For $i = 1, \cdots, n$, using X as a random variable with distribution $\mathcal{N}(0, t_i - t_{i-1})$,

$$\mathbb{E}\left([(B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})]^2\right) = \mathbb{E}((X^2 - \mathbb{E}(X^2))^2) \quad = \mathrm{Var}(X^2) = (t_i - t_{i-1})^2 \mathrm{Var}(U^2),$$

where U is a standard normal random variable. It follows that $V_n - W_n$ converges to 0 in $\mathbb{L}^2$.

Fourth step. Finally, study the convergence of W_n: it is a Riemann sum, thus a.s.

$$W_n \to \int_0^t \phi''(B_s)\ ds.$$

Furthermore, since the sequence (W_n) is bounded by $||\phi''||_\infty t$, convergence in $\mathbb{L}^p$ for all $p \geq 1$ follows from the dominated convergence theorem. This concludes the proof for fixed t. To exchange almost sure convergence and $t \geq 0$, it suffices to restrict to rationals and use continuity. □

To conclude this section, we provide other more general versions of Itô's formula. We will not prove them this year, but it is important to know them as they are very useful in exercises.

Theorem 14 (Itô's Formula for an Itô Process)**.** *Let*

$$X_t = X_0 + \int_0^t u_s\ dB_s + \int_0^t v_s\ ds$$

be an Itô process with u in $\mathbb{L}^2_{loc}$ and v progressively measurable such that for all $t \geq 0$ a.s., $\int_0^t |v_s|\ ds < +\infty$. Let ϕ be a $\mathcal{C}^2$ function. Then almost surely, for all $t \geq 0$,

$$\begin{aligned}\phi(X_t) &= \phi(X_0) + \int_0^t \phi'(X_s)\ dX_s + \frac{1}{2}\int_0^t \phi''(X_s) d < X >_s \\ &= \phi(X_0) + \int_0^t \phi'(X_s) u_s\ dB_s + \int_0^t \phi'(X_s) v_s\ ds + \frac{1}{2}\int_0^t \phi''(X_s) u_s^2 ds.\end{aligned}$$

Theorem 15 (Function of Time and an Itô Process)**.** *Consider an Itô process $(X_t)t \geq 0$ as in the previous theorem and $\phi : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ a $\mathcal{C}^2$ function. Then almost surely, for all $t \geq 0$,*

$$\begin{aligned}\phi(t, X_t) &= \phi(0, X_0) + \int_0^t \partial_s \phi(X_s)\ ds + \int_0^t \partial_x \phi(X_s)\ dX_s + \frac{1}{2}\int_0^t \partial_{xx}\phi(X_s) d < X >_s \\ &= \phi(0, X_0) + \int_0^t \partial_x \phi(X_s) u_s\ dB_s + \int_0^t \partial_s \phi(X_s) + \partial_x \phi(X_s) v_s + \frac{1}{2}\partial_{xx}\phi(X_s) u_s^2\ ds.\end{aligned}$$

6 Stochastic Differential Equations

See [7] for a reference used in this section.

The aim of this section is to provide meaning and solve the following *stochastic differential equation* (SDE):

$$\begin{aligned} dX_t &= b(X_t)dt + \sigma(X_t)dB_t \\ X_0 &= Z \end{aligned} \tag{10}$$

where:

- Z is a random variable called the *initial condition*,
- b is a locally bounded measurable function from $\mathbb{R}$ to $\mathbb{R}$ called the *drift*,
- σ is a locally bounded measurable function from $\mathbb{R}$ to $\mathbb{R}$ called the *diffusion coefficient*,
- $(B_t)_{t\geq 0}$ is a Brownian motion.

Let $(\mathcal{F}_t)$ denote the filtration generated by $(B_t)_{t\geq 0}$ (which we assume completed).

6.1 Existence and Uniqueness of Strong Solutions

Definition 23 (Strong Solution). *A strong solution of our stochastic differential equation is a process $(X_t)_{t\geq 0}$ with continuous trajectories such that:*

1. $(X_t)_{t\geq 0}$ *is adapted to* $(\mathcal{F}_t)_{t\geq 0}$,
2. *Almost surely, for all* $t \geq 0$,

$$X_t = Z + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dB_s. \tag{11}$$

The solution is called *strong* because it depends on the initial condition and the Brownian motion, which are given data of the problem.

Theorem 16. *Suppose $Z \in \mathbb{L}^2$ and b and σ are **Lipschitz**. Then the stochastic differential equation*

$$\begin{aligned} dX_t &= b(X_t)dt + \sigma(X_t)dB_t \\ X_0 &= Z \end{aligned} \tag{12}$$

*admits a **unique strong solution**.*

Proof. Uniqueness. Let $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ be two strong solutions. Fix $M>0$ and let τ be the first time the distance M is reached by either $(X_t)_{t\geq 0}$ or $(Y_t)_{t\geq 0}$:

$$\tau = \inf\{t\geq 0 \mid |X_t|\geq M \text{ or } |Y_t|\geq M\}. \tag{13}$$

Fix $T>0$ and consider $0\leq t\leq T$. We have

$$\begin{aligned}
\mathrm{E}(|X_{t\wedge\tau}-Y_{t\wedge\tau}|^2) &= \mathrm{E}\left(|\int_0^{t\wedge\tau} b(X_s)-b(Y_s)ds + \int_0^{t\wedge\tau}\sigma(X_s)-\sigma(Y_s)dB_s|^2\right)\\
&\leq 2\left\{\mathrm{E}\left(|\int_0^{t\wedge\tau} b(X_s)-b(Y_s)ds|^2\right)+\mathrm{E}\left(|\int_0^{t\wedge\tau}\sigma(X_s)-\sigma(Y_s)dB_s|^2\right)\right\}\\
&\overset{C.S.+(8)}{\leq} 2\left\{\mathrm{E}\left(t\int_0^{t\wedge\tau}(b(X_s)-b(Y_s))^2ds\right)+\mathrm{E}\left(\int_0^{t\wedge\tau}(\sigma(X_s)-\sigma(Y_s))^2ds\right)\right\}\\
&\leq 2K^2(T+1)\int_0^t \mathrm{E}(|X_{s\wedge\tau}-Y_{s\wedge\tau}|^2)ds.
\end{aligned} \tag{14}$$

By defining $u(s)=\mathrm{E}(|X_{s\wedge\tau}-Y_{s\wedge\tau}|^2)$ for $0\leq s\leq T$, we obtain $u(0)=0$ and for all $0\leq t\leq T$,

$$u(t)\leq 2K^2(T+1)\int_0^t u(s)ds. \tag{15}$$

We can use the

Lemma 7 (Gronwall's Lemma)**.** *Let u be a positive function locally bounded on $\mathbb{R}^+$ such that for all $t\geq 0$,*

$$u(t)\leq a+b\int_0^t u(s)ds, \tag{16}$$

where a and b are positive constants. Then for all $t\geq 0$,

$$u(t)\leq ae^{bt}. \tag{17}$$

Proof of Gronwall's Lemma. By induction for all $n\geq 1$,

$$u(t)\leq a+ab+\cdots+a\frac{(bt)^n}{n!}+b^{n+1}\int_0^t\int_0^{t_1}\cdots\int_0^{t_n}u(t_{n+1})dt_{n+1}\cdots dt_1. \tag{18}$$

Since u is bounded on $[0,t]$ (say by a constant C), the last term is bounded by $C\frac{(bt)^{n+1}}{(n+1)!}$ and tends to 0. Hence the result. □

In our case, $a=0$ and u is bounded by $4M^2$ by definition of τ. We conclude that $u(t)$ is zero for all $t\geq 0$, almost surely, and therefore $X_{t\wedge\tau}=Y_{t\wedge\tau}$. We let M tend to infinity to conclude that X and Y are modifications of each other. Since these two processes are continuous, we also conclude that they are indistinguishable.

Existence. We establish existence using a Picard iteration method (i.e., a fixed-point argument). We define the process

$$X_t^{(0)} = Z \quad \text{for } t \geq 0, \tag{19}$$

and for all $n \geq 0$,

$$X_t^{(n+1)} = Z + \int_0^t b(X_s^{(n)})ds + \int_0^t \sigma(X_s^{(n)})dB_s \quad \text{for } t \geq 0. \tag{20}$$

Fix $T \geq 0$. We want to use Doob's inequality for $(\int_0^t \sigma(X_s^{(n)})dB_s)_{0\leq t\leq T}$ and for that we have to check before that it is a true martingale. For this, we show by induction (see [7]) using a calculation similar to the one we just did for uniqueness, that for all $n \geq 0$,

$$\sup_{0\leq t\leq T} \mathrm{E}((X_t^{(n)})^2) < +\infty. \tag{21}$$

We then deduce that $(X_t^{(n)})_{0\leq t\leq T}$ is in $\mathbb{L}^2(Prog, [0,T])$, and, using the Lipschitz property of σ, that $(\sigma(X_t^{(n)}))_{0\leq t\leq T}$ is also. For all $n \geq 1$, the process $(\int_0^t \sigma(X_s^{(n)})dB_s)_{0\leq t\leq T}$ is therefore a true martingale bounded in $\mathbb{L}^2$, and we can use Doob's inequality. We obtain for all $t \leq T$,

$$\begin{aligned}
\mathrm{E}(\sup_{0\leq s\leq t} |X_s^{(n+1)} - X_s^{(n)}|^2) &\leq 2\mathrm{E}\left(\sup_{0\leq s\leq t}\left|\int_0^s b(X_u^{(n)}) - b(X_u^{(n-1)})du\right|^2\right) \\
&\quad + 2\mathrm{E}\left(\sup_{0\leq s\leq t}\left|\int_0^s \sigma(X_u^{(n)}) - \sigma(X_u^{(n-1)})dB_u\right|^2\right) \\
&\overset{Doob}{\leq} 2\mathrm{E}\left(\left(\int_0^t |b(X_u^{(n)}) - b(X_u^{(n-1)}|du\right)^2\right) \\
&\quad + 8\mathrm{E}\left(\left|\int_0^t \sigma(X_u^{(n)}) - \sigma(X_u^{(n-1)})dB_u\right|^2\right) \\
&\overset{C.S.+(8)}{\leq} 2(4+T)K^2\,\mathrm{E}\left(\int_0^t |X_u^{(n)} - X_u^{(n-1)}|^2 du\right) \\
&\leq C\int_0^t \mathrm{E}(\sup_{0\leq v\leq u} |X_v^{(n)} - X_v^{(n-1)}|^2)du,
\end{aligned} \tag{22}$$

where $C = 2(4+T)K^2$. For $n=1$, we obtain

$$\mathrm{E}(\sup_{0\le s\le t} |X_s^{(2)} - X_s^{(1)}|^2) \le C \int_0^t \mathrm{E}(\sup_{0\le v\le u} |X_v^{(1)} - X_v^{(0)}|^2) du \le Ca\ t, \tag{23}$$

where $a = \mathrm{E}(\sup_{t\le T} |X_t^{(1)} - X_t^{(0)}|^2)$ (which is finite since $Z \in \mathbb{L}^2$). By iteration, we deduce that for all $n \ge 1$, and all $0 \le t \le T$,

$$\left\| \sup_{0\le t\le T} |X_t^{(n+1)} - X_t^{(n)}| \right\|_2^2 = \mathrm{E}\left(\sup_{0\le s\le t} |X_s^{(n+1)} - X_s^{(n)}|^2 \right) \le aC^n \frac{t^n}{n!}, \tag{24}$$

which implies

$$\sum_{n\ge 0} \left\| \sup_{0\le t\le T} |X_t^{(n+1)} - X_t^{(n)}| \right\|_2 < +\infty. \tag{25}$$

We conclude that

$$\sum_{n\ge 0} \sup_{0\le t\le T} |X_t^{(n+1)} - X_t^{(n)}| \tag{26}$$

is the almost sure limit of an increasing sequence of positive variables, and also converges in $\mathbb{L}^2$. In particular, the limit is almost surely finite:

$$\sum_{n\ge 0} \sup_{0\le t\le T} |X_t^{(n+1)} - X_t^{(n)}| < +\infty \quad \text{almost surely.} \tag{27}$$

This implies that almost surely, the sequence $(X_t^{(n)})_{0\le t\le T}$ is Cauchy in the space of continuous functions on $[0,T]$ endowed with the uniform norm. Hence, almost surely, this sequence converges uniformly to a limit that we denote $(X_t)_{0\le t\le T}$. The trajectories of $(X_t)_{0\le t\le T}$ are continuous as uniform limits of continuous functions. Moreover, $(X_t)_{0\le t\le T}$ is adapted to the Brownian filtration as the limit of $(X_t^{(n)})_{0\le t\le T}$, which are adapted to the same filtration (we even have better than almost sure convergence of $X_t^{(n)}$ to X_t).

The next step is to demonstrate that this process is a solution to our SDE. For this, let's return to the recurrence equation (20). First, we notice, by a calculation similar to (22), that

$$\mathrm{E}(\sup_{0\le t\le T} |X_t^{(n)} - X_t|^2) \underset{n\to+\infty}{\to} 0, \tag{28}$$

which implies that for all $0 \le t \le T$, $X_t^{(n)}$ converges in $\mathbb{L}^2$ to X_t. Using the fact that σ is Lipschitz, we obtain

$$\mathrm{E}\int_0^T (\sigma(X_s^{(n)}) - \sigma(X_s))^2 ds \le K^2\ \mathrm{E}\int_0^T (X_s^{(n)} - X_s)^2 ds,$$

from which, by using (28), we deduce that for all $0 \le t \le T$, $(\int_0^t \sigma(X_s^{(n)})dB_s)_{n\ge0}$ converges in $\mathbb{L}^2$ to $\int_0^t \sigma(X_s)dB_s$. Similarly,

$$\mathrm{E}\left|\int_0^T b(X_s^{(n)})ds - \int_0^T b(X_s)ds\right|^2 \le K^2 T\ \mathrm{E}\left(\int_0^T \left|X_s^{(n)} - X_s\right|^2 ds\right),$$

which also tends to 0. Therefore, we have $\mathbb{L}^2$ convergence of $(\int_0^T b(X_s^{(n)})ds)_{n\ge0}$ to $\int_0^T b(X_s)ds$. Consequently, we can pass to the $\mathbb{L}^2$ limit in (20), and we indeed obtain that $(X_t)_{0\le t\le T}$ is a strong solution to our SDE. Since T is arbitrary and there is uniqueness, this concludes the proof. □

In fact, we can prove with the same tools the following stronger theorem, which also deals with the non-homogeneous case in time.

Theorem 17. *Let $Z \in \mathbb{L}^2$. We assume that σ and b are continuous on $\mathbb{R}^+ \times \mathbb{R}$ and Lipschitz in the spatial variable (i.e., for any fixed $t \ge 0$, $b(t,\cdot)$ and $\sigma(t,\cdot)$ are Lipschitz). Then there exists a unique strong solution to the equation*

$$X_t = Z + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s.$$

6.2 Langevin Equation and Ornstein-Uhlenbeck Process

[**Refer to [5] for a detailed exposition in this section**]. In 1908, Paul Langevin proposed an equation to account for the displacement of a particle suspended in a fluid based on the laws of Newtonian mechanics. It is assumed that the particle is subject to a frictional force in the fluid (proportional to its velocity) and to collisions due to the thermal agitation of the liquid. If $(x(t))_{t\ge0}$ denotes the displacement of a particle with mass 1, then for all $t \ge 0$,

$$x''(t) = -bx'(t) + F(t)$$

where b is the coefficient of friction and F is the external force acting on the particle. The velocity $v = x'$ of the particle thus satisfies

$$v'(t) = -bv(t) + F(t).$$

Considering that the force F represents independent collisions, it is reasonable to assume that $\int_a^b F(t)dt = \sigma(B(t) - B(a))$ where $(B_t)_{t\ge0}$ is a Brownian motion and σ is a real number measuring the agitation of the medium. This leads to the following stochastic differential equation (SDE) for the velocity:

$$dX_t = -bX_t dt + \sigma dB_t$$

with initial condition $Z \in \mathbb{L}^2$ independent of $(B_t)_{t\geq 0}$. To determine the solution of this equation, we apply Itô's formula to $(e^{bt}X_t)_{t\geq 0}$:

$$\begin{aligned} e^{bt}X_t &= Z + \int_0^t be^{bs}X_s ds + \int_0^t e^{bs}\ dX_s \\ &= Z + \sigma \int_0^t e^{bs}\ dB_s. \end{aligned}$$

The process

$$X_t = Ze^{-bt} + \sigma \int_0^t e^{-b(t-s)}\ dB_s, \qquad t \geq 0$$

is thus a solution of the Langevin equation. This process is called the Ornstein-Uhlenbeck process. The integral part is a Wiener process (see TD 5) since the integrated process is deterministic, making it a Gaussian process. For all $0 \leq r \leq t$,

$$\mathrm{E}(X_t) = \mathrm{E}(Z)e^{-bt}$$

and

$$\begin{aligned} \mathrm{Cov}(X_r, X_t) &= e^{-b(t+r)}\mathrm{Var}(Z) + \sigma^2 \int_0^r e^{-b(t-s)}e^{-b(r-s)}\ ds \\ &= e^{-b(t+r)}\mathrm{Var}(Z) + \frac{\sigma^2}{2b}(e^{-b(t-r)} - e^{-b(t+r)}). \end{aligned}$$

Thus, we obtain the following asymptotics: $\mathrm{E}(X_t) \to 0$ as t tends to infinity, and for all $h \geq 0$,

$$\lim_{r\to+\infty} \mathrm{Cov}(X_r, X_{r+h}) = \frac{\sigma^2}{2b}e^{-bh}.$$

In particular, $\mathrm{E}(X_t^2)$ tends to $\frac{\sigma^2}{2b}$, meaning that the average kinetic energy converges. The process $(X_t)_{t\geq 0}$ converges in law to a Gaussian $\mathcal{N}(0, \frac{\sigma^2}{2b})$ since a limit of Gaussians is Gaussian. This probability is invariant in the sense that if Z follows the law $\mathcal{N}(0, \frac{\sigma^2}{2b})$, then for all $t \geq 0$, X_t follows the same law: indeed, X_t is Gaussian, and its mean (zero) and variance $(\frac{\sigma^2}{2b})$ can be explicitly calculated. More generally, in this case, the process $(X_t)_{t\geq 0}$ is a centered Gaussian process with covariance

$$K_X(r,t) = \frac{\sigma^2}{2b}e^{-b(t-r)}, \qquad 0 \leq r \leq t.$$

This process is stationary: for all $t_0 \geq 0$, the process $(X_{t_0+t})_{t\geq 0}$ has the same law as $(X_t)_{t\geq 0}$.

Finally, considering the position $(Y_t)_{t\geq 0}$ of the particle itself rather than its velocity, in the case where the velocity is given by the Gaussian process above, and where the particle starts from a deterministic position $x \in \mathbb{R}$, we obtain for all $t \geq 0$,

$$Y_t = x + \int_0^t X_s \, ds.$$

The process $(Y_t)_{t\geq 0}$ is Gaussian (since X is Gaussian and we can view the integral as a sequence of Riemann sums), centered at x, and with covariance defined for all $0 \leq r \leq t$ by

$$\begin{aligned} K_Y(r,t) &= \int_0^r \int_0^t \mathrm{Cov}(X_{s_1}, X_{s_2}) ds_1 ds_2 \\ &= \frac{\sigma^2}{b^2} r - \frac{\sigma^2}{2b^3}(2 - 2e^{-bt} - 2e^{-br} + e^{-b(t-r)} + e^{-b(t+r)}). \end{aligned}$$

If we let the friction and agitation coefficients b and σ tend to infinity simultaneously in such a way that the ratio σ/b converges to a constant $\kappa \in]0, +\infty[$, then

$$K_Y(r,t) \to \kappa^2 r.$$

In this limit, our process converges in law to $(x + \kappa B_t)_{t\geq 0}$ where B is a Brownian motion since in the Gaussian case it suffices to verify the convergence of means and variance functions. Thus, in this limit, we recover the process proposed by Einstein to describe a particle suspended in a liquid.

References

[1] F. Baudoin. *Diffusion processes and stochastic calculus.* EMS Textbooks in Mathematics. European Mathematical Society (EMS), Zürich, 2014.

[2] P. Billingsley. *Convergence of probability measures.* Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999.

[3] P. Bougerol. *https://www.lpsm.paris/pageperso/bougerol/CalcSto15-16.pdf.*

[4] D. Chafaï. *https://djalil.chafai.net/docs/M2/m2-stochastic-calculus-course-2020-2021.pdf.*

[5] F. Comets and T. Meyre. *Calcul stochastique et modèles de diffusions : Cours et exercices corrigés.* Dunod, 2006. Coll. Sciences sup.

[6] N. Guillotin. *http://math.univ-lyon1.fr/ guillotin/livre.pdf.*

[7] J.-F. Le Gall. *Mouvement brownien, martingales et calcul stochastique*, volume 71 of *Mathématiques & Applications (Berlin) [Mathematics & Applications].* Springer, Heidelberg, 2013.

[8] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, third edition, 1999.