Stéphane ATTAL

STOCHASTIC CALCULUS
(Cours de M2)

September 1, 2008

Springer
Berlin Heidelberg New York
Hong Kong London
Milan Paris Tokyo

Contents

1

STOCHASTIC PROCESSES

1.1 Background

A subset A of Ω is *negligible* if there exists $B \in \mathcal{F}$ such that $A \subset B$ and $P(B) = 0$. Note that this notion is relative to $\mathcal{F}$ and P but we shall not mention them. The probability space $(\Omega, \mathcal{F}, P)$ is called *complete* if $\mathcal{F}$ contains all the negligible sets.

An incomplete probability space $(\Omega, \mathcal{F}, P)$ can easily be completed in the following way. Let $\mathcal{N}$ denote the set of all negligible sets. Let $\overline{\mathcal{F}}$ be the σ-algebra generated by $\mathcal{F}$ and $\mathcal{N}$. It is easy to see that it coincides with the set of subsets of Ω which are of the form $B \cup N$ for some $B \in \mathcal{F}$ and $N \in \mathcal{N}$. One extends P to a probability measure $\overline{P}$ on $\overline{\mathcal{F}}$ by putting $\overline{P}(B \cup N) = P(B)$. The probability space $(\Omega, \overline{\mathcal{F}}, \overline{P})$ is then complete. It is called the *completion* of $(\Omega, \mathcal{F}, P)$. From now on, all our probability spaces are implicitely assumed to be complete.

1.2 Stochastic Processes

The main goal of the theory of stochastic processes is to study the behaviour of families $(X_t)_{t\in I}$ of random variables X_t indexed by a subset I of $\mathbb{R}$. In the sequel we shall mainly consider the cases where I is $\mathbb{R}^+$ or a sub-interval of $\mathbb{R}^+$; in the discrete time setup we are interested in $I = \mathbb{N}$. We shall give here a rigorous framework for the study of stochastic processes and their laws.

Let I be a subset of $\mathbb{R}$. A *process* (or *stochastic process*) *indexed by* I is a family $(X_t)_{t\in I}$ of random variables defined on a common probability space $(\Omega, \mathcal{F}, P)$. When I is $\mathbb{R}^+$ one simply says a *process* and $(X_t)_{t\in I}$ may be simply denoted by X.

For each finite subset $T = \{t_1 \dots t_n\}$ of I one denotes by $\mu_{X,T}$ the law of the n-tuple $(X_{t_1} \dots X_{t_n})$. If $(X_t)_{t\in I}$ and $(Y_t)_{t\in I}$ are two processes indexed by the same set I and if $\mu_{X,T} = \mu_{Y,T}$ for all finite subsets T of I, we say that $(X_t)_{t\in I}$ and $(Y_t)_{t\in I}$ are *versions* of each other.

Let I be a subset of $\mathbb{R}$. Suppose we are given for any finite subset T of I a probability measure μ^T on $\mathbb{R}^{|T|}$ (where $|T|$ denotes the cardinal of T). It is natural to wonder if one can construct a process $(X_t)_{t\in I}$ such that $\mu_{X,T} = \mu^T$ for all T. The answer is given by the famous Kolmogorov Consistency Theorem that we state without proof.

Let I be a set. Let $(E, \mathcal{E})$ be a measurable space. Suppose that for all finite subsets T of I we are given a probability measure μ^T on $(E^T, \mathcal{E}^T)$. The family $\{\mu^T\,;\, T \text{ finite subset of } I\}$ is called *consistent* if, for all finite subsets T_1, T_2 of I such that $T_1 \subset T_2$, the restriction of μ^{T_2} to the σ-algebra $\mathcal{E}^{T_1}$ coincides with μ^{T_1}.

On the product space $\Omega = E^I$, we consider the σ-algebra $\mathcal{E}^I$ generated by the finite cylinders

$$\{\omega \in \Omega; \omega(t_1) \in A_1, \dots, \omega(t_k) \in A_n\},$$

that is, the σ-algebra generated by the coordinate mappings

$$\begin{aligned} X_t : \Omega &\longrightarrow \quad \mathbb{R} \\ \omega &\longmapsto \omega(t)\,. \end{aligned}$$

Theorem 1.1 (Kolmogorov Consistency Theorem). *Let $(E, \mathcal{E})$ be a Polish space together with its Borel σ-algebra. Let I be a subset of $\mathbb{R}$. If for all finite subsets T of I there exists a probability measure μ^T on $(E^T, \mathcal{E}^T)$ such that the family $\{\mu^T\,;\, T \text{ finite subset of } I\}$ is consistent, then there exists a unique probability measure μ on $(E^I, \mathcal{E}^I)$ which extends all the probability measures μ^T.*

Consider a process $(X_t)_{t\in I}$ on $(\Omega, \mathcal{F}, P)$. Let μ_X be the unique probability measure on $(\mathbb{R}^I, \mathcal{B}(\mathbb{R})^I)$ which extends the consistent family of probability measures $\{\mu_{X,T};\, T \subset I,\, T \text{ finite}\}$. This measure μ_X on $(\mathbb{R}^I, \mathcal{B}(\mathbb{R})^I)$ is called the *law* of the process $(X_t)_{t\in I}$. With these definitions, the following result is straightforward.

Proposition 1.2. *Two processes $(X_t)_{t\in I}$ and $(Y_t)_{t\in I}$ are versions of each other if and only if they have the same law.*

Let $(X_t)_{t\in I}$ be a process on $(\Omega, \mathcal{F}, P)$. For all $t \in I$, define the function Y_t on $\mathbb{R}^I$ by $Y_t(x) = x(t)$, that is Y_t is the t-th coordinate mapping. This defines a process $(Y_t)_{t\in I}$ on $\left(\mathbb{R}^I, \mathcal{B}(\mathbb{R})^I, \mu_X\right)$ and we have

$$\begin{aligned} \mu_{Y,T}(A_{t_1} \times \cdots \times A_{t_n}) &= \mu_X\left(Y_{t_1}^{-1}(A_{t_1}) \cap \cdots \cap Y_{t_n}^{-1}(A_{t_n})\right) \\ &= \mu_X(\{x \in \mathbb{R}^I\,;\, x(t_1) \in A_{t_1} \dots x(t_n) \in A_{t_n}\}) \\ &= \mu_{X,T}(A_{t_1} \times \cdots \times A_{t_n}). \end{aligned}$$

Thus $(Y_t)_{t\in I}$ is a version of $(X_t)_{t\in I}$. This version is called the *canonical version* of $(X_t)_{t\in I}$. The probability space $(\mathbb{R}^I, \mathcal{B}(\mathbb{R}^I), \mu_X)$ is called the *canonical space* of $(X_t)_{t\in I}$. We shall say more about the notion of *canonical version* in next section.

From now on, unless otherwise stated, all the processes are indexed by $\mathbb{R}^+$.

Two processes X and Y defined on the same probability space are *modifications* of each other if, for all $t \in \mathbb{R}^+$, X_t equals Y_t almost surely. They are *indistinguishable* if, for almost all $\omega \in \Omega$, $X_t(\omega)$ equals $Y_t(\omega)$ for all t.

Note the subtle difference between the two definitions. "Modification" and "indistinguishable" both mean that $Y_t = X_t$ for all t and almost all ω, but in the first case the null set of ω may depend on t, whereas in the second defintion it does not depend on t.

A subset A of $\mathbb{R}^+ \times \Omega$ is *evanescent* if there exists a negligible subset B of Ω such that $A \subset \mathbb{R}^+ \times B$. From all these definitions one can check easily the following properties.

Proposition 1.3. *a) If X and Y are indistinguishable then they are modifications of each other.*

b) If X and Y are modifications of each other then they are versions of each other.

c) X and Y are indistinguishable if and only if the set of (ω, t) such that $X_t(\omega) \neq Y_t(\omega)$ is evanescent.

Let X be a process. For all $\omega \in \Omega$, the mapping $t \mapsto X_t(\omega)$ defines a function on $\mathbb{R}^+$. These functions are called the *paths* of the process X. One says that the process X has *continuous paths* (or simply, is *continuous*) if for almost all $\omega \in \Omega$ the path $X_\cdot(\omega)$ is continuous on $\mathbb{R}^+$. In the same way one defines *right-continuous processes*, *left-continuous processes* (we take the convention that any process is left-continuous at 0). A process is said to be *càdlàg* if its paths are right-continuous and admit left limits at all points.

Proposition 1.4. *Let X and Y be two right-continuous (resp. left-continuous) processes. If they are modifications of each other then they are indistinguishable.*

Proof. There exists a negligible set $\mathcal{N}$ such that for all $\omega \notin \mathcal{N}$ the paths $X_\cdot(\omega)$ and $Y_\cdot(\omega)$ are right-continuous and $X_t(\omega) = Y_t(\omega)$ for all t rational. Thus passing to the limit we have $X_t(\omega) = Y_t(\omega)$ for all t. □

A process X is *measurable* if it is measurable as a mapping from $\mathbb{R}^+ \times \Omega$ to $\mathbb{R}$, where $\mathbb{R}^+ \times \Omega$ is equipped with the σ-algebra $\mathcal{B}(\mathbb{R}^+) \otimes \mathcal{F}$.

1.3 Regularization

At that stage of this chapter we have to discuss a rather fine point about the definition of the canonical version of a stochastic process. This section can be avoided at first reading. We do not give all details.

There is quite an important problem attached to the way the canonical space of a process X has been constructed. Indeed, the canonical space furnished by Kolmogorov's Theorem 1.1 is not good enough, for the σ-field $\mathcal{F} = \mathcal{E}^I$ associated to this construction is too poor. Recall that this σ-algebra is the one generated by finite support cylinders in $\Omega = E^I$. One can show that the events of $\mathcal{E}$ are those which depend only on a countable set of coordinates of $\omega \in \Omega$. As a consequence, one cannot "ask" if the process we are interested in is continuous or not. Indeed, the set

$$\{\omega\,;\, t \mapsto X_t(\omega) \,\text{is continuous}\,\}$$

has no reason in general to be in $\mathcal{F}$. The same problems hold when one wants to study simple objects such as

$$\inf\{t; X_t(\omega) > 0\}\,,\; \lim_{s\to t} X_s(\omega)\,,\; \sup_{s\leq t} |X_s| \;... \tag{1.1}$$

This clearly implies strong restrictions for whom wants to study the properties of a given stochastic process!

The way this problem is solved in general follows the following path. For most of the processes we will be interested in (for example martingales, ...) we will be able to prove that the process in question X admits a modification Y which is càdlàg (or sometimes better: continuous). Let D denote the set of càdlàg functions from $\mathbb{R}^+$ to $\mathbb{R}$. This is a subset of $E = \mathbb{R}^{\mathbb{R}^+}$ which is not measurable for the cylinder σ-field $\mathcal{F} = \mathcal{B}(\mathbb{R})^{\mathbb{R}^+}$.

Lemma 1.5. *If a process X admits a càdlàg version then the measure P_X of any $\mathcal{F}$-measurable set F such that $F \supset D$ is equal to 1.*

In general measure theory, the property described above means that the *outer measure* of P is equal to 1 on D.

Proof. Assume that on some probability space $(\Omega', \mathcal{F}', P')$ there exists a process Y which is a version of X and which has almost all its paths càdlàg.

Let F be a $\mathcal{F}$-measurable set such that F contains D. The probability $P_X(F)$ is the probability that a certain countable set of constraints on X_t, $t \in \mathbb{R}^+$, is satisfied. It is then equal to the probability $P'(F')$ of the same set of constraints applied to Y_t, $t \in \mathbb{R}^+$. But the fact that F contains D implies that F' is the whole of Ω' (up to a null set maybe). Hence the probability is equal to 1. □

Let us give an example in order to make the proof above more clear. Let F be for example the set

$$F = \{\omega \in \Omega ; \lim_{\substack{s \to 0 \\ s \in \mathbb{Q}}} \omega(s) = \omega(0)\} .$$

This set is $\mathcal{F}$-measurable, for it can be written as

$$F = \bigcap_{\substack{\epsilon > 0 \\ \epsilon \in \mathbb{Q}}} \bigcup_{\substack{\eta > 0 \\ \eta \in \mathbb{Q}}} \bigcap_{\substack{s > 0 \\ s \in \mathbb{Q}}} \{\omega \in \Omega ; |\omega(s) - \omega(0)| \leq \epsilon\} .$$

Hence we have

$$\begin{aligned} P_X(F) &= P_X \left(\bigcap_{\substack{\varepsilon > 0 \\ \varepsilon \in \mathbb{Q}}} \bigcup_{\substack{\eta > 0 \\ \eta \in \mathbb{Q}}} \bigcap_{\substack{s > 0 \\ s \in \mathbb{Q}}} \{\omega \in \Omega ; |\omega(s) - \omega(0)| \leq \varepsilon\} \right) \\ &= P_X \left(\bigcap_{\substack{\varepsilon > 0 \\ \varepsilon \in \mathbb{Q}}} \bigcup_{\substack{\eta > 0 \\ \eta \in \mathbb{Q}}} \bigcap_{\substack{s > 0 \\ s \in \mathbb{Q}}} \{\omega \in \Omega ; |X_s(\omega) - X_0(\omega)| \leq \varepsilon\} \right) \\ &= P' \left(\bigcap_{\substack{\varepsilon > 0 \\ \varepsilon \in \mathbb{Q}}} \bigcup_{\substack{\eta > 0 \\ \eta \in \mathbb{Q}}} \bigcap_{\substack{s > 0 \\ s \in \mathbb{Q}}} \{\omega' \in \Omega' ; |Y_s(\omega') - Y_0(\omega')| \leq \varepsilon\} \right) \\ &= P'(\{\omega' \in \Omega' ; \lim_{\substack{s \to 0 \\ s \in \mathbb{Q}}} Y_s(\omega') = Y_0(\omega')\}) . \end{aligned}$$

But as the path of Y are almost all càdlàg we have $\lim_{\substack{s \to 0 \\ s \in \mathbb{Q}}} Y_s(\omega') = Y_0(\omega')$ to be satisfied by almost all ω'. Hence the above probability is equal to 1.

When the situation $P(F) = 1$ for all $F \in \mathcal{F}$, $F \supset D$, occurs we have the following result, whose easy proof is left to the reader.

Lemma 1.6. *On the space D define the coordinate mappings $Y_t(\omega) = \omega(t)$. Then the σ-algebra $\mathcal{F}_D$ generated by the Y_t, $t \in \mathbb{R}^+$ coincides with $\mathcal{F} \cap D$. Furthermore, putting, for all $F \in \mathcal{F}_D$*

$$Q(F) = P_X(\widehat{F})$$

for any $\widehat{F} \in \mathcal{F}$ such that $F = \widehat{F} \cap D$, defines a probability measure on $(D, \mathcal{F}_D)$.

Hence we have defined a version $(D, \mathcal{F}_D, Q, Y)$ of X whose paths are all càdlàg. On that probability space, elements such as in (1.1) are now all well-defined and measurable.

Actually the σ-algebra $\mathcal{F}_D$ is much richer than was $\mathcal{F}$. One can for example prove the following. The space D admits a natural topology, called the *Skorohod topology*, that we shall not describe here. Let C be the space of continuous functions from $\mathbb{R}^+$ to $\mathbb{R}$. The Skorohod topology coincides on C with the topology of uniform convergence on compact sets. Both D and C are Polish spaces with these topologies.

Theorem 1.7. *On D (resp. C) the σ-algebra $\mathcal{F}_D$ (resp. $\mathcal{F}_C$) coincide with the Borel σ-algebra $\mathcal{B}(D)$ (resp. $\mathcal{B}(C)$) associated to the Skorohod topology.*

1.4 Uniform Integrability

We concentrate here on a notion which is fundamental, in particular for the study of martingales. Let $U = \{X_i\,;\, i \in I\}$ be any family of integrable random variables on $(\Omega, \mathcal{F}, P)$, one can think of U as a subset of $L^1(\Omega, \mathcal{F}, P)$. The family U is *uniformly integrable* if

$$\sup_{X \in U} \mathbb{E}\left[|X|\ \mathbb{1}_{|X| \geq a}\right] \xrightarrow[a \to +\infty]{} 0. \tag{1.2}$$

Proposition 1.8. *Let U be a subset of $L^1(\Omega, \mathcal{F}, P)$. The following assertions are equivalent.*

i) U is uniformly integrable.

ii) One has

a) $\sup_{X \in U} \mathbb{E}\left[|X|\right] < \infty$

and

b) for all $\varepsilon > 0$, there exists a $\delta > 0$ such that $A \in \mathcal{F}$ and $P(A) \leq \delta$ imply $\mathbb{E}\left[|X|\ \mathbb{1}_A\right] \leq \varepsilon$, for all $X \in U$.

Proof. If U is uniformly integrable and $A \in \mathcal{F}$ then

$$\begin{aligned} \sup_{X \in U} \mathbb{E}\left[|X|\ \mathbb{1}_A\right] &= \sup_{X \in U} \mathbb{E}\left[|X|\ \mathbb{1}_{|X| \geq a} \mathbb{1}_A\right] + \sup_{X \in U} \mathbb{E}\left[|X|\ \mathbb{1}_{|X| < a} \mathbb{1}_A\right] \\ &\leq \sup_{X \in U} \mathbb{E}\left[|X|\ \mathbb{1}_{|X| \geq a}\right] + a \sup_{X \in U} \mathbb{E}\left[\mathbb{1}_{|X| < a} \mathbb{1}_A\right]. \end{aligned}$$

Let c be such that the first term above is smaller than $\varepsilon/2$ for all $a \geq c$. We then have $\sup_{X \in U} \mathbb{E}\left[|X|\ \mathbb{1}_A\right] \leq \varepsilon/2 + cP(A)$. Taking $A = \Omega$ gives *a)*. Taking $\delta = \varepsilon/2c$ gives b). We have proved that *i)* implies *ii)*.

Conversely, assume *ii)* is satisfied. Let $\varepsilon > 0$ and let $\delta > 0$ be as in b). Let $c = \sup_{X \in U} \mathbb{E}[|X|]/\delta < \infty$. Let A be the event $(|X| \geq c)$. We have $P(A) \leq \mathbb{E}[|X|]/c \leq \delta$. Thus $\mathbb{E}\left[|X|\ \mathbb{1}_A\right] \leq \varepsilon$ for all $X \in U$. □

One case of uniformly integrable familly which appears quite often is described by the following result.

Proposition 1.9. *Let X be an integrable random variable. The set of random variables $\mathbb{E}[X \,|\, \mathcal{B}]$, where $\mathcal{B}$ runs over all the sub σ-fields of $\mathcal{F}$, is a uniformly integrable family.*

Proof. Put $Y = \mathbb{E}[X \,|\, \mathcal{B}]$. We have

$$\mathbb{E}[|Y| \; \mathbb{1}_{|Y|\geq a}] \leq \mathbb{E}\left[\mathbb{E}[|X| \;|\, \mathcal{B}] \, \mathbb{1}_{|Y|\geq a}\right] = \mathbb{E}\left[|X| \; \mathbb{1}_{|Y|\geq a}\right] .$$

As we have

$$P(|Y| \geq a) \leq a^{-1}\mathbb{E}[|Y|] \leq a^{-1}\mathbb{E}[|X|] ,$$

one can choose a large enough such that $P(|Y| \geq a) \leq \delta$ and hence $\mathbb{E}\left[|X| \; \mathbb{1}_{|Y|\geq a}\right] \leq \varepsilon$. □

The main use of the notion of uniform integrability is the following result.

Theorem 1.10. *Let (X_n) be a sequence of random variables belonging to $L^1(\Omega, \mathcal{F}, P)$. Suppose that (X_n) converges almost surely to a random variable $X_\infty \in L^1(\Omega, \mathcal{F}, P)$. Then (X_n) converges to X_∞ in $L^1(\Omega, \mathcal{F}, P)$ if and only if (X_n) is uniformly integrable.*

Proof. If (X_n) converges to X_∞ in $L^1(\Omega, \mathcal{F}, P)$ we then have

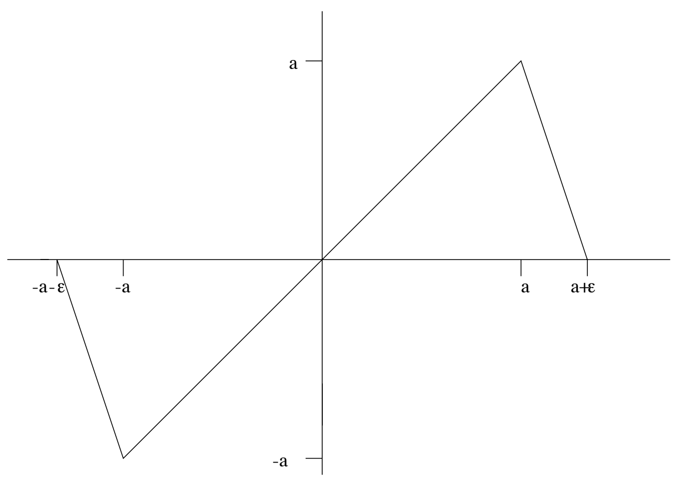
$$\sup_n \mathbb{E}[|X_n|] \leq \sup_n \mathbb{E}[|X_n - X_\infty|] + \mathbb{E}[|X_\infty|] < +\infty$$

and

$$\mathbb{E}[|X_n| \, \mathbb{1}_A] \leq \mathbb{E}[|X_\infty| \, \mathbb{1}_A] + \mathbb{E}[|X_n - X_\infty| \, \mathbb{1}_A] .$$

The second term of the right hand side is dominated by ε for n large enough (independently of A); as any finite sequence of random variables is always uniformly integrable, the conclusion follows.

Conversely, if (X_n) is uniformly integrable, let $\varepsilon > 0$ and a_ε be such that if $a \geq a_\varepsilon$ then $\sup_n \mathbb{E}[|X_n| \; \mathbb{1}_{|X_n|\geq a}] \leq \varepsilon/3$. Let ϕ_a be the following function.



By Fatou's lemma we have $\mathbb{E}[|X_\infty| \, \mathbb{1}_{|X_\infty|\geq a}] \leq \varepsilon/3$ and

$$\begin{aligned}\mathbb{E}\left[|X_n - X_\infty|\right] &\leq \mathbb{E}\left[|X_n - \phi_a(X_n)|\right] + \mathbb{E}\left[|\phi_a(X_n) - \phi_a(X_\infty)|\right] \\ &\qquad + \mathbb{E}\left[|\phi_a(X_\infty) - X_\infty|\right] \\ &\leq \mathbb{E}\left[|X_n|\, \mathbb{1}_{|X_n|\geq a}\right] + \mathbb{E}\left[|\phi_a(X_n) - \phi_a(X_\infty)|\right] \\ &\qquad + \mathbb{E}\left[|X_\infty|\, \mathbb{1}_{|X_\infty|\geq a}\right] .\end{aligned}$$

As $\mathbb{E}[|\phi_a(X_n) - \phi_a(X_\infty)|]$ tends to 0 by Lebesgue's theorem, this gives the result. □

1.5 Brownian Motion

In this section and in the next one, we shall consider two fundamental examples of stochastic processes: *Brownian motion* and the *Poisson process.*

Let us first study Brownian motion. We wish to construct a process $(W_t)_{t\in\mathbb{R}^+}$ with values in $\mathbb{R}$ such that

1) for all $s < t$, the random variable $W_t - W_s$ is independent of the random variables W_u, $u \leq s$,

2) for all $t \geq 0$, the random variable W_t follows the Gaussian $\mathcal{N}(0, t)$ probability law,

3) a.s. the paths $t \mapsto W_t$ are continuous.

Let us construct such a process. Recall the following well-known construction.

Theorem 1.11. *For any given probability measure μ on $\mathbb{R}$, there exists a probability space $(\Omega, \mathcal{F}, P)$ and a sequence (X_n) of random variables on Ω such that the random variables X_n are independent and each have the law μ.*

We can now prove the following.

Theorem 1.12. *Let $\mathcal{H}$ be a separable Hilbert space. There exists a probability space $(\Omega, \mathcal{F}, P)$ and a family $X(h)$, $h \in \mathcal{H}$, of random variables on Ω such that*

i) the mapping $h \longmapsto X(h)$ is linear,

ii) each random variable $X(h)$ follows the $\mathcal{N}(0, \|h\|)$ law.

Proof. Let (e_n) be an orthonormal basis of $\mathcal{H}$. By Theorem 1.11, there exists a probability space $(\Omega, \mathcal{F}, P)$ and a sequence (Z_n) of independent random variables on Ω, with individual law $\mathcal{N}(0, 1)$.

For any $h \in \mathcal{H}$, put $X(h) = \sum_n \langle e_n\,,\, h\rangle\, Z_n$. This series is convergent in $L^2(\Omega, \mathcal{F}, P)$ and defines a random variable $X(h)$ on Ω. The family $X(h)$, $h \in \mathcal{H}$, satisfies the assumptions of the theorem, as can be easily checked. □

Note that, by construction we have

$$\langle X(h') \, , \, X(h) \rangle_{L^2(\Omega,\mathcal{F},P)} = \langle h' \, , \, h \rangle \, .$$

Also note that the set $\{X(h),\, h \in \mathcal{H}\}$ is a *Gaussian subspace* of $L^2(\Omega, \mathcal{F}, P)$, that is, any linear combination of the $X(h)$ is a Gaussian random variable. By the fundamental characterization of Gaussian families, this means that every finite family $(X(h_1), \dots, X(h_n))$ is Gaussian.

Now the construction of a Brownian motion is easy. Take $\mathcal{H} = L^2(\mathbb{R}^+)$ and construct a family $X(h)$, $h \in \mathcal{H}$, as in Theorem 1.12. We claim that the process $W_t = X(\mathbb{1}_{[0,t]})$, $t \in \mathbb{R}^+$, is a Brownian motion. Indeed, the random variable W_t follows the law $\mathcal{N}(0,t)$ by ii) above and thus satisfies condition 2) in the definition of Brownian motion; furthermore, we have, for all $u \leq s \leq t$

$$\langle W_t - W_s \, , \, W_u \rangle_{L^2(\Omega,\mathcal{F},P)} = \langle \mathbb{1}_{[s,t]} \, , \, \mathbb{1}_{[0,u]} \rangle_{\mathcal{H}} = 0 \, .$$

But the pair $(W_t - W_s, W_u)$ is Gaussian, hence they are independent random variables. This gives the condition 1) in the definition of a Brownian motion.

We still need to prove that the process we obtained is continuous, or at least can be modified into a continuous process. Actually, we shall establish a stronger property for the paths: the Hölder continuity of order α for every $\alpha < 1/2$. This property is based on a general criterion due to Kolmogorov that we state here without proof.

Theorem 1.13. [Kolmogorov criterion] *Let X be a process such that there exists strictly positive constants γ, C and ε such that*

$$\mathbb{E}[|X_t - X_s|^\gamma] \leq C|t-s|^{1+\varepsilon} \, .$$

Then there exists a modification of X whose paths are Hölder continuous of order α for every $\alpha \in [0, \varepsilon/\gamma]$. □

Once this is admited, the fact that the increments of the Brownian motion are Gaussian gives

$$\mathbb{E}\left[(W_t - W_s)^{2p}\right] = C_p |t-s|^p$$

for every $p > 0$ (this fact is rather easy to check and left to the reader). This immediately yields the following.

Theorem 1.14. *Up to modification, the paths of Brownian motion are locally Hölder continuous of order α for every $\alpha < 1/2$. In particular they are continuous.* □

Another very important property of the Brownian motion is that it admits a non-trivial quadratic variation.

Theorem 1.15. *For any $t \in \mathbb{R}^+$, the quantity*

$$\sum_{i;t_i \le t} \left(W_{t_{i+1}} - W_{t_i}\right)^2$$

converges to t, in the L^2 sense when the diameter δ of the partition $\{t_i\,;\, i \in \mathbb{N}\}$ tends to 0.

Proof. Let us compute the L^2-norm of the difference (for simplicity, we assume that the partition ends at t):

$$\begin{aligned}\left\| \sum_{i;t_i \le t} (W_{t_{i+1}} - W_{t_i})^2 - t \right\|^2 &= \mathbb{E}\left[\left(\sum_{i;t_i \le t} (W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i) \right)^2 \right] \\ &= \sum_{i;t_i \le t} \mathbb{E}\left[\left((W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i) \right)^2 \right],\end{aligned}$$

where we used the independence of the increments and the fact that

$$\mathbb{E}[(W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i)] = 0\,.$$

Furthermore, for a Gaussian random variable Y with mean 0, we have $\mathbb{E}[Y^4] = 3\mathbb{E}[Y^2]^2$. This gives

$$\left\| \sum_{i;t_i \le t} (W_{t_{i+1}} - W_{t_i})^2 - t \right\|^2 = 2 \sum_{i;t_i \le t} (t_{i+1} - t_i)^2$$

which converges to 0 with the diameter of the partition. □

It is actually possible to prove the same result for almost sure convergence, but we do not give a proof of this fact here.

This quadratic variation property has many important consequences; one very important application of it is the following.

Theorem 1.16. *The paths of the Brownian motion are almost surely of infinite variation on any interval.*

The paths of the Brownian motion are almost surely nowhere locally Hölder continuous of order α for $\alpha > 1/2$.

In particular, the paths of the Brownian motion are almost surely nowhere differentiable.

Proof. Note that almost surely we have

$$\begin{aligned}&\sum_{i;t_i \le t} (W_{t_{i+1}}(\omega) - W_{t_i}(\omega))^2 \le \\ &\le \left(\sup_{i;t_i \le t} \left| W_{t_{i+1}}(\omega) - W_{t_i}(\omega) \right| \right) \left(\sum_{i;t_i \le t} \left| W_{t_{i+1}}(\omega) - W_{t_i}(\omega) \right| \right).\end{aligned}$$

The first term in the right hand side converges to 0 by the continuity of Brownian motion. The second term is dominated by the total variation of the Brownian path. As the left hand side converges to a finite quantity, this forces the total variation to be infinite.

The case of the non-Hölder property is treated by following the same idea: for all $\alpha > 1/2$ we have

$$\sum_{i;t_i\leq t} (W_{t_{i+1}}(\omega) - W_{t_i}(\omega))^2 \leq$$

$$\leq t \left(\sup_{i;t_i\leq t} |t_{i+1} - t_i|^{2\alpha-1} \right) \left(\sup_{i;t_i\leq t} \frac{|W_{t_{i+1}}(\omega) - W_{t_i}(\omega)|^2}{|t_{i+1} - t_i|^{2\alpha}} \right).$$

If the Brownian paths were Hölder of order α the last term above would be dominated independently of the partition. The rest of the right hand side converges to 0. This contradicts the fact that the left hand side converges to t. This proves the non Hölderian character of Brownian motion for $\alpha > 1/2$.

Non-differentiability is immediate now. □

We have not yet said if the Brownian paths are Hölder-continuous of order exactly 1/2 or not. It so happens that they are not, but this result needs further developments; we just mention it as a remark.

1.6 Poisson Processes

We now construct the *Poisson process*. Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let (T_n) be a strictly increasing sequence of positive random variables. The T_n's are thought as arrival times. A process X such that, for all $t \in \mathbb{R}^+$

$$X_t = \sum_n \mathbb{1}_{T_n \leq t}$$

is called a *counting process associated to* (T_n). It is valued in $\mathbb{N} \cup \{+\infty\}$. If $\sup_n T_n = \infty$ a.s. one says that X is a *non-exploding counting process*.

A *Poisson process* is a non-exploding counting process N whose increments are independent and stationary. That is,

1) $N_t - N_s$ is independent of all the random variables N_u, $u \leq s$
2) $N_t - N_s$ has the same law as $N_{t+h} - N_{s+h}$ for all $t \geq s \geq 0$ and $h \geq 0$.

Theorem 1.17. *Poisson processes exist and they are all of the following form: there exists* $\lambda \in \mathbb{R}^+$ *such that*

$$P(N_t = n) = \frac{(\lambda t)^n}{n!}\, e^{-\lambda t} \tag{1.3}$$

for all $n \in \mathbb{N}$, *all* $t \in \mathbb{R}^+$. *In other words, the associated sequence* (T_n) *consists of sums of independent times each of which follows an exponential distribution with parameter* λ.

Proof. One direction is easy. If (T_n) is the sequence given by $T_n = \sum_{m\leq n} S_m$ where the S_m's are independent, identically distributed random variables following the exponential law with parameter $\lambda \geq 0$, it is then easy to check that the associated counting process

$$N_t = \sum_{n=0}^{\infty} 1\!\!1_{T_n \leq t}$$

is a Poisson process and follows the Poisson law (1.3).

The converse is more difficult. From the hypothesis we have

$$P(N_t = 0) = P(N_s = 0)P(N_t - N_s = 0) = P(N_s = 0)P(N_{t-s} = 0)$$

and thus

$$P(N_t = 0) = e^{-\lambda t}$$

for some $\lambda \geq 0$ and for all $t \in \mathbb{R}^+$.

We now claim that $P(N_t \geq 2) = o(t)$. Indeed, divide $[0,1]$ into n intervals of the same length. Let S_n be the number of subintervals which contain at least two times of the sequence (T_m). Clearly S_n has a binomial distribution $B(n, P(N_{1/n} \geq 2))$. Therefore $\mathbb{E}[S_n] = nP(N_{1/n} \geq 2)$. For a fixed ω, for n sufficiently large there is no interval with more than one stopping time. Thus $\lim_{n\to+\infty} S_n(\omega) = 0$ a.s. We now wish to apply the dominated convergence theorem in order to conclude that $\lim_{n\to+\infty} \mathbb{E}(S_n) = 0$ and hence the announced estimate. As we clearly have $S_n \leq N_1$, we just need to prove that $\mathbb{E}[N_1] < \infty$. The intervals $T_{n+1} - T_n$ between the jumps are independent random variables, with the same law: the one of T_1. Hence

$$\mathbb{E}[e^{-T_n}] = \mathbb{E}[e^{-T_1}]^n = \alpha^n.$$

This proves that

$$P(|N_t| > n) \leq P(T_n < t) \leq \frac{\mathbb{E}[e^{-T_n}]}{e^{-t}} \leq e^t \alpha^n.$$

That is, N_t admits exponential moments.

Now, we have

$$P(N_t = 1) = 1 - P(N_t = 0) - P(N_t \geq 2)$$

and thus

$$\lim_{t\to 0} \frac{1}{t} P(N_t = 1) = \lim_{t\to 0} \frac{1 - e^{-\lambda t} + o(t)}{t} = \lambda\ .$$

Finally, for $\beta \in [0,1]$ put $f(t) = \mathbb{E}[\beta^{N_t}]$. Clearly $f(t+s) = f(t)f(s)$ and f is of the form $f(t) = e^{t\, g(\beta)}$. But

$$
\begin{aligned}
f(t) &= \sum_{n=0}^{\infty} \beta^n P(N_t = n) \\
&= P(N_t = 0) + \beta P(N_t = 1) + \sum_{n=2}^{\infty} \beta^n P(N_t = n)
\end{aligned}
$$

and $g(\beta) = f'(0)$. This gives

$$
\begin{aligned}
g(\alpha) &= \lim_{t \to 0} \frac{P(N_t = 0) - 1}{t} + \frac{\alpha P(N_t = 1)}{t} + \frac{1}{t} o(t) \\
&= -\lambda + \lambda\alpha,
\end{aligned}
$$

so

$$
f(t) = e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n \beta^n}{n!}
$$

and the required result follows. □

The parameter λ is called the *intensity* of N. In particular we have

$$
\begin{aligned}
\mathbb{E}[N_t] &= \lambda t \\
\mathrm{Var}[N_t] &= \lambda t \,.
\end{aligned}
$$

1.7 Filtrations, Conditional Expectations

From now on, unless otherwise stated, all our processes are indexed by $\mathbb{R}^+$.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. A *filtration* of the space $(\Omega, \mathcal{F}, P)$ is a family $(\mathcal{F}_t)_{t \in \mathbb{R}^+}$ of sub-σ-algebras of $\mathcal{F}$ such that $\mathcal{F}_s \subset \mathcal{F}_t$ for all $s \leq t$. We shall denote the filtration by $(\mathcal{F}_t)$ simply (we cannot simplify the notation to $\mathcal{F}$ as for processes, because there would be too much confusion with the σ-algebra $\mathcal{F}$).

The quadruple $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ is called a *filtered probability space.*

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ be a filtered probability space. For all $t \geq 0$ one defines

$$
\mathcal{F}_{t-} = \bigvee_{s<t} \mathcal{F}_s, \quad \mathcal{F}_{t+} = \bigcap_{s>t} \mathcal{F}_s
$$

with the convention $\mathcal{F}_{0-} = \mathcal{F}_0$. One also puts

$$
\mathcal{F}_\infty = \bigvee_{t \geq 0} \mathcal{F}_t.
$$

This way one defines two new filtrations of $(\Omega, \mathcal{F}, P)$, namely $(\mathcal{F}_{t-})_{t \in \mathbb{R}^+}$ and $(\mathcal{F}_{t+})_{t \in \mathbb{R}^+}$.

A filtration $\mathcal{F}$ is *right-continuous* if $\mathcal{F}_t = \mathcal{F}_{t+}$ for all $t \geq 0$. The filtration $(\mathcal{F}_{t+})_{t\in\mathbb{R}^+}$ is always right-continuous. In the same way one can speak of *left-continuous* filtrations or *continuous* filtrations.

A filtration $(\mathcal{F}_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, P)$ is called *complete* if $\mathcal{F}_0$ (and hence each $\mathcal{F}_t$, $t \in \mathbb{R}^+$) contains the negligible sets of $(\Omega, \mathcal{F}, P)$. If it is not the case we make it complete by adding the negligible sets to $\mathcal{F}_0$ in the same way as in Section 1.1.

From now on, all the filtered probability spaces are supposed to be complete and right-continuous (one replaces $(\mathcal{F}_t)$ by $(\mathcal{F}_{t+})_{t\in\mathbb{R}^+}$ if necessary).

A process X defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ is *adapted* if X_t is $\mathcal{F}_t$-measurable for all $t \in \mathbb{R}^+$.

The *natural filtration* of a process X is the filtration generated by X, that is the filtration $(\mathcal{F}_t)$ where $\mathcal{F}_t = \sigma\{X_s; s \leq t\}$ is the σ-algebra generated by all the random variables X_s, $s \leq t$. Once again, the natural filtration of a process is understood to be made complete and right-continuous. The natural filtration of a process is the smallest (complete and right-continuous) filtration that makes this process measurable and adapted (exercise).

A process X is *progressive* if for all $t \in \mathbb{R}^+$ the mapping $(s, \omega) \mapsto X_s(\omega)$ on $[0, t] \times \Omega$ is measurable for $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$. A subset A of $\mathbb{R}^+ \times \Omega$ is *progressive* if $(s, \omega) \mapsto \mathbb{1}_A(s, \omega)$ is a progressive process. The progressive subsets of $\mathbb{R}^+ \times \Omega$ form a σ-algebra of $\mathbb{R}^+ \times \Omega$. This σ-algebra is called the *progressive σ-algebra*. A process is progressive if and only if it is measurable with respect to this σ-algebra.

It is clear that a progressive process is measurable and adapted, but the converse is not true (cf. [**?**], p. 47), one needs a little more regularity.

Proposition 1.18. *An adapted process with right-continuous paths (or with left-continuous paths) is progressive.*

Proof. For all $n \in \mathbb{N}$ let

$$X^n_t = \sum_{k=0}^{\infty} X_{(k+1)2^{-n}} \mathbb{1}_{[k2^{-n},(k+1)2^{-n}[}(t).$$

As X is right-continuous, X^n_t converges to X_t for all t. But the process X^n is clearly progressive with respect to the filtration $(\mathcal{F}_{t+2^{-n}})_{t\in\mathbb{R}^+}$. Consequently X is progressive with respect to the filtration $(\mathcal{F}_{t+\varepsilon})_{t\in\mathbb{R}^+}$ for any $\varepsilon > 0$.

For all $s \leq t$ we have

$$X_s = \lim_{\varepsilon \to 0} X_s \mathbb{1}_{[0,t-\varepsilon]}(s) + X_t \mathbb{1}_{\{t\}}(s).$$

The term within the limit symbol is measurable with respect to $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$, thus so is X_s. This proves that X is progressive.

The case where X has left-continuous paths is treated in the same way. □

We now turn to the definitions and properties of conditional expectations.

Theorem 1.19. *Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-algebra of $\mathcal{F}$. Let X be an integrable random variable on $(\Omega, \mathcal{F}, P)$. Then there exists an integrable random variable Y which is $\mathcal{G}$-measurable and which satisfies*

$$\mathbb{E}[X \mathbb{1}_A] = \mathbb{E}[Y \mathbb{1}_A] \quad \textit{for all} \quad A \in \mathcal{G}. \tag{1.4}$$

Any other $\mathcal{G}$-measurable integrable random variable Y' satisfying (1.4) *is equal to Y almost surely.*

Proof. Assume first that X is square integrable that is, $X \in L^2(\Omega, \mathcal{F}, P)$. The space $L^2(\Omega, \mathcal{G}, P)$ is a closed subspace of $L^2(\Omega, \mathcal{F}, P)$. Let Y be the orthogonal projection of X on $L^2(\Omega, \mathcal{G}, P)$. Then Y belongs to $L^2(\Omega, \mathcal{G}, P)$ and satisfies

$$\mathbb{E}[XZ] = \mathbb{E}[YZ] \quad \text{for all} \quad Z \in L^2(\Omega, \mathcal{G}, P). \tag{1.5}$$

In particular $\mathbb{E}[X\, \mathbb{1}_A] = \mathbb{E}[Y \mathbb{1}_A]$ for all $A \in \mathcal{G}$. This also implies that if X is positive then so is Y, for Y has a positive integral on any set $A \in \mathcal{G}$.

Now, assume X is only integrable. Its positive part X^+ is also integrable. For all n, the random variable $X^+ \wedge n$ is square integrable. Let Y_n^+ be associated to $X^+ \wedge n$ in the same way as above. Then Y_n^+ is positive and the sequence (Y_n^+) is increasing. The random variable $Y^+ = \lim_n Y_n^+$ is integrable, for $\mathbb{E}[Y_n^+] \le \mathbb{E}[X^+]$. In the same way, associate Y^- to X^-. The random variable $Y = Y^+ - Y^-$ answers our statements.

Uniqueness follows immediately from our remark about positivity. □

The almost sure equivalence class of integrable, $\mathcal{G}$-measurable random variables Y such that (1.4) holds is called the *conditional expectation* of X with respect to $\mathcal{G}$. It is denoted by $\mathbb{E}[X \,|\, \mathcal{G}]$. We also denote by $\mathbb{E}[X \,|\, \mathcal{G}]$ a representative of the equivalence class $\mathbb{E}[X \,|\, \mathcal{G}]$.

Here are the main properties of conditional expectations.

Theorem 1.20. *Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$ and X be an integrable random variable on $(\Omega, \mathcal{F}, P)$.*

i) The mapping $X \mapsto \mathbb{E}[X \,|\, \mathcal{G}]$ is linear.

ii) If X is positive then so is $\mathbb{E}[X \,|\, \mathcal{G}]$.

iii) If (X_n) is an increasing sequence of integrable random variables a.s. *converging to an integrable random variable X then $\mathbb{E}[X_n \,|\, \mathcal{G}]$ converges* a.s. *to $\mathbb{E}[X \,|\, \mathcal{G}]$.*

iv) (Jensen inequality) *If f is a convex function on $\mathbb{R}$ and if $f(X)$ is integrable then*

$$f \circ \mathbb{E}[X \,|\, \mathcal{G}] \le \mathbb{E}[f(X) \,|\, \mathcal{G}] \quad \text{a.s.} \tag{1.6}$$

v) If X is $\mathcal{G}$-measurable then $\mathbb{E}[X \,|\, \mathcal{G}] = X$ a.s. *The converse is true if $\mathcal{G}$ is complete.*

vi) If $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{F}$ are σ-algebras then

$$\mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{G}_2\right] \mid \mathcal{G}_1\right] = \mathbb{E}[X \,|\, \mathcal{G}_1] \quad \text{a.s.} \tag{1.7}$$

In particular

$$\mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{G}\right]\right] = \mathbb{E}[X]. \tag{1.8}$$

vii) If Y is $\mathcal{G}$-measurable and XY is integrable then

$$\mathbb{E}[XY \,|\, \mathcal{G}] = Y\mathbb{E}[X \,|\, \mathcal{G}] \quad \text{a.s.} \tag{1.9}$$

viii) If X belongs to $L^p(\Omega, \mathcal{F}, P)$, for $1 \leq p \leq \infty$, then $\mathbb{E}[X \,|\, \mathcal{G}]$ belongs to $L^p(\Omega, \mathcal{G}, P)$ and $\|\,\mathbb{E}[X \,|\, \mathcal{G}]\,\|_p \leq \|X\|_p$.

Proof. *i)* is obvious. *ii)* has been already proved.
iii) Let $Y = \lim_n \mathbb{E}[X_n \,|\, \mathcal{G}]$. Then

$$\begin{aligned}|\mathbb{E}[Y \mathbb{1}_A] - \mathbb{E}[X\, \mathbb{1}_A]\,| &\leq |\mathbb{E}[Y \mathbb{1}_A] - \mathbb{E}\left[\mathbb{E}[X_n \,|\, \mathcal{G}]\, \mathbb{1}_A\right]\,| + |\mathbb{E}\left[\mathbb{E}[X_n \,|\, \mathcal{G}]\, \mathbb{1}_A\right] - \mathbb{E}[X\, \mathbb{1}_A]\,| \\ &\leq \mathbb{E}\left[\,|Y - \mathbb{E}[X_n \,|\, \mathcal{G}]|\, \mathbb{1}_A\right] + \mathbb{E}\left[\,|X_n - X|\, \mathbb{1}_A\right] \xrightarrow[n\to\infty]{} 0.\end{aligned}$$

iv) f is the superior convex envelop of a countable family of affine functions l_n and $l_n(X)$ is integrable, $n \in \mathbb{N}$. We have

$$l_n \circ \mathbb{E}[X \,|\, \mathcal{G}] = \mathbb{E}[l_n(X) \,|\, \mathcal{G}] \leq \mathbb{E}[f(X) \,|\, \mathcal{G}]. \tag{1.10}$$

Passing to the superior envelop gives *iv)*.
v) is easy and *vi)* is straightforward.
vii) If Y is simple then the statement is clear, for

$$\mathbb{E}[X\, \mathbb{1}_B \mathbb{1}_A] = \mathbb{E}[\mathbb{1}_B \mathbb{E}[X \,|\, \mathcal{G}]\, \mathbb{1}_A]$$

for all $A, B \in \mathcal{G}$. For a general Y one takes a monotone limit.
viii) This is Jensen inequality applied to the convex functions $f = \|\,\cdot\,\|_p$.
□

1.8 Stopping Times

We now enter into the definitions and properties of a notion which is very particular to Probability Theory: the notion of *stopping times*. It appears to be a rather unknown concept for most of the mathematicians from other areas (and from almost all the physicists). But it is a completely fundamental tool in the framework of stochastic processes.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ be a filtered probability space. A *stopping time* is a measurable map $T : \Omega \mapsto \mathbb{R}^+ \cup \{+\infty\}$ such that for all $t \in \mathbb{R}^+$ the event $(T \leq t)$ belongs to $\mathcal{F}_t$.

As the filtration $(\mathcal{F}_t)$ is right-continuous one may equivalently replace the condition "$(T \leq t) \in \mathcal{F}_t$ for all t" by the condition"$(T < t) \in \mathcal{F}_t$ for all t" (exercise).

The constant times: $T(\omega) = t$ for all ω, are stopping times. More generally, for any stopping time T and any $t \in \mathbb{R}^+$, then $T + t$ is a stopping time.

Let T be a stopping time. The set of $A \in \mathcal{F}$ such that

$$A \cap (T \leq t) \text{ belongs to } \mathcal{F}_t \text{ for all } t$$

is a σ-algebra (exercise). We denote it by $\mathcal{F}_T$ and call it the *σ-algebra of events anterior to T*. This σ-algebra coincides with the σ-algebra of $A \in \mathcal{F}$ such that

$$A \cap (T < t) \text{ belongs to } \mathcal{F}_t \text{ for all } t.$$

The terminology for $\mathcal{F}_T$ comes from the fact that its definition generalizes the idea that $\mathcal{F}_t$ is the σ-algebra of events occurring before the time t. Indeed, the constant stopping time $T(\omega) = t$ has its anterior σ-algebra $\mathcal{F}_T$ which coincides with $\mathcal{F}_t$.

One denotes by $\mathcal{F}_{T-}$ the σ-algebra generated by $\mathcal{F}_0$ and the events of the form

$$A \cap (T > t),\ t \geq 0,\ A \in \mathcal{F}_t.$$

The σ-algebra $\mathcal{F}_{T-}$ is called the *σ-algebra of events strictly anterior to T*. When $T(\omega) = t$, then clearly $\mathcal{F}_{T-}$ coincides with $\mathcal{F}_{t-}$.

A stopping time is *discrete* if the set of its values is (at most) countable.

Proposition 1.21. *Every stopping time is the limit of a decreasing sequence of discrete stopping times.*

Proof. Let T be a stopping time. For all $n \in \mathbb{N}$ put

$$S_n = +\infty\, \mathbb{1}_{T=+\infty} + \sum_{k \in \mathbb{N}} k2^{-n}\, \mathbb{1}_{(k-1)2^{-n} < T \leq k2^{-n}}.$$

Then the sequence (S_n) satisfies the statements. □

Here is a list of the main properties of stopping times and of their associated σ-algebras. Note that some of them generalize to stopping times some properties that were known for constant times.

Theorem 1.22. *a) If S, T are stopping times then so are $S \wedge T$ and $S \vee T$.*

b) Let (S_n) be a monotonic sequence of stopping times and put $S = \lim_{n\to\infty} S_n$, then S is also a stopping time.

c) For all stopping time T we have $\mathcal{F}_{T-} \subset \mathcal{F}_T$ and T is $\mathcal{F}_{T-}$-measurable.

d) If S, T are two stopping times and if $S \leq T$ then $\mathcal{F}_S \subset \mathcal{F}_T$ and $\mathcal{F}_{S-} \subset \mathcal{F}_{T-}$; if $S < T$ then $\mathcal{F}_S \subset \mathcal{F}_{T-}$.

e) For all stopping times S, T and all $A \in \mathcal{F}_S$ we have $A \cap (S \leq T) \in \mathcal{F}_T$ and $A \cap (S < T) \in \mathcal{F}_{T-}$. In particular $(S \leq T)$ belongs to $\mathcal{F}_S$ and $\mathcal{F}_T$, the event $(S = T)$ belongs to $\mathcal{F}_S$ and $\mathcal{F}_T$, finally the event $(S < T)$ belongs to $\mathcal{F}_S$ and $\mathcal{F}_{T-}$.

Proof. All the proofs are easy from the definitions and left to the reader. We just precise that for proving *d)* and *e)* it is useful to notice the following identities:

$$(S < T) = \bigcup_{r \in \mathbb{Q}^+} (S < r) \cap (r < T) \tag{1.11}$$

$$(S \leq T) = \bigcap_{r \in \mathbb{Q}^+} (S \leq r) \cup (r \leq T). \quad \square \tag{1.12}$$

Proposition 1.23. *For all $A \in \mathcal{F}_\infty$ and all stopping time T, the set $A \cap (T = \infty)$ belongs to $\mathcal{F}_{T-}$. In particular the events $(T = \infty)$, $(T < \infty)$ belong to $\mathcal{F}_{T-}$.*

Proof. As the set of $A \in \mathcal{F}_\infty$ such that $A \cap (T = \infty)$ belongs to $\mathcal{F}_{T-}$ is a σ-algebra, it is sufficient to prove the result for all $A \in \mathcal{F}_n$, $n \in \mathbb{N}$. But, in this case, $A \cap (T = \infty)$ is equal to $\cap_{m \geq n} \{A \cap (T > m)\}$ which clearly belongs to $\mathcal{F}_{T-}$. $\square$

Theorem 1.24. *Let (T_n) be a monotonic sequence of stopping times. Let $T = \lim_n T_n$.*
a) If (T_n) is decreasing then

$$\mathcal{F}_T = \bigcap_n \mathcal{F}_{T_n}.$$

b)If (T_n) is increasing then

$$\mathcal{F}_{T-} = \bigvee_n \mathcal{F}_{T_n-}.$$

Proof. *a)* Clearly $\mathcal{F}_T \subset \cap_n \mathcal{F}_{T_n}$. Now let $A \in \cap_n \mathcal{F}_{T_n}$. We have $A \cap (T < t) = \cup_n A \cap (T_n < t)$, which is an element of $\mathcal{F}_t$. Thus A belongs to $\mathcal{F}_T$.

b) Clearly $\mathcal{F}_{T-}$ contains $\vee_n \mathcal{F}_{T_n-}$. Now consider $A \cap (t < T)$, with $A \in \mathcal{F}_t$, a typical generator of $\mathcal{F}_{T-}$. This set also writes as $\cup_n A \cap (t < T_n)$, thus it belongs to $\vee_n \mathcal{F}_{T_n-}$. $\square$

Let S, T be two stopping times such that $S \leq T$. One denotes by $[\![S, T]\!]$ the following subset of $\mathbb{R} \times \Omega$:

$$\{(t, \omega) \text{ such that } t \in [S(\omega), T(\omega)]\}. \tag{1.13}$$

One defines in an analogous way the intervals $[\![S, T[\![$, $]\!]S, T]\!]$, $]\!]S, T[\![$. All these particular subsets of $\mathbb{R} \times \Omega$ are called *stochastic intervals*.

The stochastic interval $[\![S, S]\!]$ is denoted $[\![S]\!]$ and is called the *graph* of S; it corresponds to the set of (t, ω) in $\mathbb{R} \times \Omega$ such that $S(\omega) = t$.

Proposition 1.25. *Every stochastic interval is a progressive subset of* $\mathbb{R}^+ \times \Omega$.

Proof. The indicator function of $[\![S,T[\![$ is adapted and right-continuous, thus progressive. The indicator function of $]\!]S,T]\!]$ is adapted and left-continuous, thus progressive. Furthermore $]\!]S,T[\![= [\![S,T[\![\cap]\!]S,T]\!]$ and $[\![S,T]\!] = [\![0,S[\![^c \cap]\!]T,+\infty[\![^c$. Thus every stochastic interval is progressive. □

Proposition 1.26. *If X is a progressive process and T is a finite stopping time, then X_T is $\mathcal{F}_T$-measurable.*

Proof. We have to prove that for every Borel set A the set $(X_T \in A) \cap (T \leq t)$ belongs to $\mathcal{F}_t$. But this set is equal to $(X_{T\wedge t} \in A) \cap (T \leq t)$. Let us consider the stopping time $S = T \wedge t$, it is $\mathcal{F}_t$-measurable. As X is progressive then X_S is $\mathcal{F}_t$-measurable for it is the composition of the mappings $\omega \mapsto (S(\omega), \omega)$ and $(s,\omega) \mapsto X_s(\omega)$. □

2

MARTINGALES

The first cornerstone of Stochastic Integration Theory is a particular family of stochastic processes: the martingales. They provide a vaste class of non-trivial stochastic processes which constitute the archetypes of noise. This is a very much studied family and their properties are now well-known and established. In this chapter we explore in details the main ingredients of this theory.

2.1 First Properties

Let $(\Omega, \mathcal{F}, P)$ be a probability space together with a filtration $(\mathcal{F}_t)$. A process X is a *martingale for* $(\mathcal{F}_t)$ if

a) each random variable X_t is integrable,

b) for all $s \leq t$ one has

$$\mathbb{E}[X_t \,|\, \mathcal{F}_s] = X_s$$

almost surely.

A process X is a *martingale*, without any reference to a particular filtration, if it is a martingale for its natural filtration.

A process X is a *supermartingale* (*resp.* a *submartingale*) if the equality in *b)* is replaced by

$$\mathbb{E}[X_t \,|\, \mathcal{F}_s] \leq X_s\,,$$

resp.

$$\mathbb{E}[X_t \,|\, \mathcal{F}_s] \geq X_s\,.$$

The main example of a martingale is obtained as follows. Consider an integrable random variable H and put $X_t = \mathbb{E}[H \,|\, \mathcal{F}_t]$ for all t. The stochastic process X then defines a martingale for $(\mathcal{F}_t)$. Such a martingale is called a *closed martingale*, the random variable H is called a *closure* of the martingale X. Note that, in general, a closed martingale may admit several closures.

Note the following properties which are easily deduced from the definitions:

• X is a submartingale if and only if $-X$ is a supermartingale.

• If X and Y are supermartingales then $X+Y$ and $X\wedge Y$ are supermartingales also.

• If X is a martingale and f is a convex function such that $f(X_t)$ is integrable for all t, then $f(X)$ is a submartingale. In particular $|X|$ and X^2 are positive submartingales.

• If f is a convex increasing function then $f(X)$ is a submartingale once X is a submartingale, in particular X^+ is a submartingale for every submartingale X. Conversely, if f is a convex decreasing function then $f(X)$ is a submartingale once X is a supermartingale, in particular X^- is a submartingale for every supermartingale X.

Before coming to the main theorems concerning martingales (existence of a càdlàg version, Doob's Stopping Theorem,...) we first need to consider the theory of discrete time martingales. For a moment our processes, our filtrations are indexed by $\mathbb{N}$, that is, they are of the form $X=(X_n)$, $(\mathcal{F}_n)$... A *stopping time* relative to $(\mathcal{F}_n)$ is a random variable T valued in $\mathbb{N}\cup\{+\infty\}$ such that $(T=n)$ belongs to $\mathcal{F}_n$ for all n. A *martingale* is a process of integrable random variables (X_n), such that $\mathbb{E}[X_m\,|\,\mathcal{F}_n]=X_n$ a.s., for all $n\leq m$. All the other time-continuous definitions are adapted to this context in the same way.

For every adapted process X and every stopping time T we define the *stopped process* X^T by

$$X^T_n = X_{T\wedge n}.$$

Theorem 2.1. *If X is a martingale (resp. supermartingale, submartingale) and if T is a stopping time, then X^T is a martingale (resp. supermartingale, submartingale).*

Proof. We have

$$\begin{aligned}
\mathbb{E}[X^T_{n+1}-X^T_n\,|\,\mathcal{F}_n] &= \mathbb{E}[(X_{n+1}-X_n)\mathbb{1}_{T\geq n+1}\,|\,\mathcal{F}_n]\\
&= \mathbb{E}[(X_{n+1}-X_n)\,\mathbb{1}_{(T\leq n)^c}\,|\,\mathcal{F}_n]\\
&= \mathbb{E}[(X_{n+1}-X_n)\,|\,\mathcal{F}_n]\,\mathbb{1}_{(T<n)^c}\\
&= 0 \quad (\text{resp. } \leq 0\,,\ \geq 0)\,. \quad \square
\end{aligned}$$

Theorem 2.2 (Stopping Theorem). *Let X be a martingale (resp. supermartingale, submartingale). Let T_1 and T_2 be two bounded stopping times such that $T_1\leq T_2$. Then we have*

$$\mathbb{E}[X_{T_2}\,|\,\mathcal{F}_{T_1}] = X_{T_1} \quad (\textit{resp. } \leq\,,\ \geq)\,. \tag{2.1}$$

Proof. We give the proof in the martingale case only, the other cases are obtained in the same way. First consider the case where $T_2 = k$ is a constant time, with $k \geq T_1$. We have, for all $A \in \mathcal{F}_{T_1}$ and all $j \leq k$

$$\mathbb{E}[X_{T_1} \mathbb{1}_{A\cap(T_1=j)}] = \mathbb{E}[X_j \mathbb{1}_{A\cap(T_1=j)}] = \mathbb{E}[X_k \mathbb{1}_{A\cap(T_1=j)}].$$

Summing over j, we obtain

$$\mathbb{E}[X_{T_1} \mathbb{1}_A] = \sum_{j=0}^{k} \mathbb{E}[X_{T_1} \mathbb{1}_{A\cap(T_1=j)}] = \mathbb{E}[X_k \mathbb{1}_A] = \mathbb{E}[X_{T_2} \mathbb{1}_A].$$

This proves the result in that particular case. More generally, if $T_1 \leq T_2 \leq k$ we have

$$\mathbb{E}[X_{T_1} \mathbb{1}_A] = \mathbb{E}[X^{T_2}_{T_1} \mathbb{1}_A] = \mathbb{E}[X^{T_2}_k \mathbb{1}_A] = \mathbb{E}[X_{T_2} \mathbb{1}_A]\,. \quad \square$$

Theorem 2.3. *An adapted process X is a martingale (resp. supermartingale, submartingale) if and only if it satisfies*

$$\mathbb{E}[X_T] = \mathbb{E}[X_0] \quad (\textit{resp.} \ \leq 0, \geq 0) \tag{2.2}$$

for every bounded stopping time T.

Proof. If X is a martingale (*resp.* supermartingale, submartingale) and T is a bounded stopping time, then Equation (2.2) is a direct application of Theorem 2.2, with $T_2 = T$ and $T_1 = 0$ and taking the expectation on the identity (2.1).

In the converse direction, assume that the equality holds true in (2.2) for every bounded stopping time T. Take $n \in \mathbb{N}$ and $A \in \mathcal{F}_n$. Consider the (bounded) stopping time

$$T_A = \begin{cases} n & \text{on } A \\ n+1 & \text{on } A^c\,. \end{cases}$$

Then, by hypothesis, we have

$$\mathbb{E}[X_0] = \mathbb{E}[X_{T_A}] = \mathbb{E}[X_n\, \mathbb{1}_A] + \mathbb{E}[X_{n+1}\, \mathbb{1}_{A^c}]$$

but also

$$\mathbb{E}[X_0] = \mathbb{E}[X_{n+1}] = \mathbb{E}[X_{n+1}\, \mathbb{1}_A] + \mathbb{E}[X_{n+1}\, \mathbb{1}_{A^c}]\,.$$

Alltogether this gives

$$\mathbb{E}[X_{n+1}\, \mathbb{1}_A] = \mathbb{E}[X_n\, \mathbb{1}_A]\,,$$

that is,

$$\mathbb{E}[X_{n+1}\,|\,\mathcal{F}_n] = X_n\,.$$

We have proved that X is a martingale. The cases of supermartingales and submartingales are obtained in the same way. $\square$

A process A is *predictable* if A_n is $\mathcal{F}_{n-1}$–measurable for all $n \geq 1$ and if $A_0 = 0$.

Theorem 2.4 (Doob Decomposition). *Every submartingale X can be decomposed in a unique way as*

$$X = M + A$$

where M is a martingale and A is a predictable, integrable, increasing process.

Proof. Put $A_0 = 0$ and $A_{n+1} = A_n + \mathbb{E}[X_{n+1} - X_n \,|\, \mathcal{F}_n]$. Clearly the process A is increasing, predictable and integrable.

If we put $M_n = X_n - A_n$ for all n then we check easily that $\mathbb{E}[M_{n+1} - M_n \,|\, \mathcal{F}_n] = 0$. This proves the existence of the decomposition.

Let us now prove uniqueness. If $X = M' + A' = M + A$ are two such decompositions of X then

$$A'_{n+1} - A'_n = X_{n+1} - X_n - (M'_{n+1} - M'_n).$$

Conditioning with respect to $\mathcal{F}_n$ gives

$$A'_{n+1} - A'_n = \mathbb{E}[X_{n+1} - X_n \,|\, \mathcal{F}_n] = A_{n+1} - A_n.$$

That is, $A = A'$ and $M = M'$. □

This unique process A is called the *compensator* of X. There is a special case which is very often considered is the following. Let X be a martingale, consider the submartingale M^2, the compensator of M^2 is denoted by $\langle M \rangle$.

2.2 Inequalities

In this section we prove fundamental inequalities concerning martingales, submartingales, etc ... They will be essential for the convergence theorems and for the regularization theorems.

In the following, for any process X we make use of the following notation:

$$X^*_n = \sup_{k \leq n} X_k$$

for any $n \in \mathbb{N}$.

Theorem 2.5 (Doob's Maximal Inequality). *Let X be a positive submartingale. Then for all $\lambda \in \mathbb{R}^+$, all $n \in \mathbb{N}$, we have*

$$\lambda\, P(X^*_n \geq \lambda) \leq \mathbb{E}[X_n \, \mathbb{1}_{X^*_n \geq \lambda}] \leq \mathbb{E}[X_n]\,.$$

Proof. Let $T = \inf\{k \le n; X_k \ge \lambda\}$. Then clearly T is a stopping time. Furthermore, the event $(X_n^* \ge \lambda)$ belongs to $\mathcal{F}_T$ (exercise). We have

$$\lambda\, P(X_n^* \ge \lambda) \le \mathbb{E}[X_T\, \mathbb{1}_{X_n^* \ge \lambda}]$$

for on the event $(X_n^* \ge \lambda)$ we have $X_T \ge \lambda$. But we also have

$$\begin{aligned}
\mathbb{E}[X_T\, \mathbb{1}_{X_n^* \ge \lambda}] &= \mathbb{E}[X_T\, \mathbb{1}_{X_n^* \ge \lambda}\, \mathbb{1}_{T \le n}] \\
&= \sum_{i=0}^{n} \mathbb{E}[X_i\, \mathbb{1}_{X_n^* \ge \lambda}\, \mathbb{1}_{T=i}] \\
&\le \sum_{i=0}^{n} \mathbb{E}[X_n\, \mathbb{1}_{X_n^* \ge \lambda}\, \mathbb{1}_{T=i}] \\
&= \mathbb{E}[X_n\, \mathbb{1}_{X_n^* \ge \lambda}]\,.
\end{aligned}$$

This gives the result. □

Note that the first inequality in the above theorem does not make use of the assumption "X is positive". This theorem immediately gives the following corollary.

Corollary 2.6. *Let X be a martingale. Then for all $\lambda \in \mathbb{R}^+$, all $n \in \mathbb{N}$, we have*

$$\lambda\, P(\sup_{k \le n} |X_k| \ge \lambda) \le \mathbb{E}[|X_n|\, \mathbb{1}_{\sup_{k \le n} |X_k| \ge \lambda}] \le \mathbb{E}[|X_n|]\,. \quad \square$$

For a process X and for any $p \in [1, +\infty]$ we define

$$\|X\|_p = \sup_n \|X_n\|_p$$

where the last norm is the usual L^p-norm.

Theorem 2.7 (Doob's Inequality). *Let X be a positive submartingale and fix $p > 1$. Then we have the inequality*

$$\|X^*\|_p \le \frac{p}{p-1} \|X\|_p\,.$$

Proof. For a moment put $x = X_n$ and $y = X_n^* \wedge n$. Let ϕ be a right-continuous, increasing function on $\mathbb{R}^+$ such that $\phi(0) = 0$. By Fubini's theorem we have

$$\begin{aligned}
\mathbb{E}\left[\phi(y)\right] &= \mathbb{E}\left[\int_0^y d\phi(\lambda)\right] = \int_0^\infty P(y \ge \lambda)\, d\phi(\lambda) \\
&\le \int_0^\infty d\phi(\lambda)\, \frac{1}{\lambda} \int_{\{y \ge \lambda\}} x\, dP \quad \text{(by Doob's Maximal Inequality)} \\
&= \int x \left(\int_0^y \frac{1}{\lambda}\, d\phi(\lambda) \right) dP\,.
\end{aligned}$$

Taking $\phi(\lambda) = \lambda^p$ gives

$$\int_0^y \frac{1}{\lambda}\, d\phi(\lambda) = \int_0^y p\, \lambda^{p-2}\, d\lambda = \frac{p}{p-1}\, y^{p-1}\,.$$

By Hölder's Inequality we have

$$\mathbb{E}\left[y^p\right] \leq \frac{p}{p-1}\, \mathbb{E}\left[x\, y^{p-1}\right] \leq \frac{p}{p-1}\, \|X\|_p\, \|y^{p-1}\|_{\frac{p}{p-1}} = \frac{p}{p-1}\, \|X\|_p\, \|y\|_p^{p-1}\,.$$

This gives the required inequality after letting n tend to $+\infty$. □

Again, like in Corollary 2.6, note that the theorem above is valid for $X = |Y|$ where Y is a martingale.

The last important inequality is the one estimating the "down-crossings" of a submartingale.

Let (x_n) be a sequence of real numbers and $[a, b]$ an interval of $\mathbb{R}$. We put $A_0 = B_0 = 0$ and

$$\begin{aligned} B_n &= \inf\{p > A_{n-1}; x_p > b\} \\ A_n &= \inf\{p > B_n; x_p < a\} \end{aligned}$$

with the convention $\inf \emptyset = +\infty$. We put

$$D^{a,b} = \sup\{n; A_n < +\infty\}.$$

This corresponds to the number of *down-crossings* of the sequence (x_n) through the interval $[a, b]$.

Lemma 2.8. *Let (x_n) be a sequence of real numbers and $x^* = \sup_n |x_n|$. The sequence (x_n) is convergent if and only if x^* is finite and $D^{a,b}$ is finite for all rational numbers $a < b$.*

Proof. If x^* is finite, the sequence (x_n) is bounded. If $D^{a,b}$ is finite for all rational numbers $a < b$, put $m = \liminf_n x_n$ and $M = \limsup_n x_n$. If $m < M$ we choose $a, b \in \mathbb{Q}$ such that $m < a < b < M$. We then clearly have $D^{a,b} = +\infty$. This gives a contradiction and hence $m = M$.

Conversely, if the sequence (x_n) converges then it is a bounded sequence and x^* is finite. If there exists $a < b$ such that $D^{a,b} = +\infty$ then $\liminf_n x_n < a$ and $\limsup_n x_n > b$. This is impossible. □

In the sequel we put

$$D_n^{a,b} = \sup\{p; A_p < n\}.$$

Theorem 2.9. *If X is a submartingale then for all $a < b$ we have*

$$\mathbb{E}[D_n^{a,b}] \leq \frac{1}{(b-a)}\mathbb{E}[(X_n - b)^+].$$

Proof. For every fixed n we stop at time n the submartingale X, which becomes the submartingale X^n (Theorem 2.1). Put $A_p^n = A_p \wedge n$ and $B_p^n = B_p \wedge n$, which is smaller than A_p^n. They all are bounded stopping times. Hence by Theorem 2.2 we have

$$\mathbb{E}[X_{A_p^n} - X_{B_p^n}] \geq 0\,.$$

This gives

$$\mathbb{E}[(X_{B_p^n} - X_{A_p^n})\mathbb{1}_{A_p<n}] \leq \mathbb{E}[(X_{A_p^n} - X_{B_p^n})\mathbb{1}_{A_p\geq n}]\,.$$

On the set $(A_p < n)$ we have

$$X_{B_p^n} - X_{A_p^n} = X_{B_p} - X_{A_p} \geq b - a\,.$$

On the set $(A_p \geq n)$ we have

$$X_{A_p^n} - X_{B_p^n} = X_n - X_{B_p^n} = \mathbb{1}_{B_p<n}(X_n - X_{B_p^n}) \leq \mathbb{1}_{B_p<n}(X_n - b)\,.$$

Alltogether we obtain

$$(b-a)\mathbb{E}[\mathbb{1}_{A_p<n}] \leq \mathbb{E}[(X_n - b)^+\mathbb{1}_{B_p<n\leq A_p}].$$

Summing over p, we find the announced inequality. □

2.3 Convergence Theorems

We are now coming to the main theorems concerning martingales: the convergence theorems. They give very nice and simple conditions for having a martingale (or submartingale, supermartingale) to converge at $+\infty$.

We start with the simplest case: the *square integrable* martingales. Let M be a martingale such that $\mathbb{E}[M_n^2] < \infty$ for all n. As M^2 is a submartingale, the sequence $(\mathbb{E}[M_n^2])$ is increasing.

For $p \geq 1$ we have $\mathbb{E}[M_n M_{n+p} \,|\, \mathcal{F}_n] = M_n^2$, whence

$$\mathbb{E}[(M_{n+p} - M_n)^2 \,|\, \mathcal{F}_n] = \mathbb{E}[M_{n+p}^2 - M_n^2 \,|\, \mathcal{F}_n]\,,$$

and

$$\mathbb{E}[(M_{n+p} - M_n)^2] = \mathbb{E}[M_{n+p}^2 - M_n^2]\,. \tag{2.3}$$

Theorem 2.10. *If M is a martingale such that $\sup_n \mathbb{E}[M_n^2] < \infty$, then M converges almost surely and in L^2.*

Proof. Let $m^* = \sup_n \mathbb{E}[M_n^2]$. We have seen that the sequence $(\mathbb{E}[M_n^2])$ is increasing, it converges to m^*. In particular by (2.3) we have

$$\mathbb{E}[(M_{n+p} - M_n)^2] \leq m^* - \mathbb{E}[M_n^2].$$

This shows that

$$\sup_p \mathbb{E}[(M_{n+p} - M_n)^2] \longrightarrow_n 0$$

and (M_n) is a Cauchy sequence in L^2, hence it converges in L^2.

Now, let us show the a.s. convergence. Let $V_n = \sup_{i,j\geq n} |M_i - M_j|$. Clearly the sequence (V_n) is positive and decreasing. Let V be its limit. If we show that that $V = 0$ a.s., then we are done. Let us apply Doob's inequality to the martingale $(M_i - M_n)_{i\geq n}$ (where n is fixed and i is varying). For all $\rho > 0$ we have

$$\begin{aligned} P(V_n > \rho) &= P(\sup_{i,j\geq n} |M_i - M_j| > \rho) \\ &\leq P(\sup_{i\geq n} |M_i - M_n| > \frac{\rho}{2}) \\ &\leq \frac{16}{\rho^2} \sup_{i\geq n} \mathbb{E}[(M_i - M_n)^2] \end{aligned}$$

for

$$\begin{aligned} P(\sup_{i\geq n} |M_i - M_n| > \frac{\rho}{2}) &= \mathbb{E}[\mathbb{1}_{\sup_{i\geq n} |M_i - M_n|^2 > \frac{\rho^2}{4}}] \\ &\leq \frac{4}{\rho^2} \mathbb{E}\left[\sup_{i\geq n} |M_i - M_n|^2 \, \mathbb{1}_{\sup_{i\geq n} |M_i - M_n|^2 > \frac{\rho^2}{4}}\right] \\ &\leq \frac{4}{\rho^2} \mathbb{E}[\sup_{i\geq n} |M_i - M_n|^2] \\ &\leq \frac{16}{\rho^2} \sup_{i\geq n} \mathbb{E}[|M_i - M_n|^2] \, . \end{aligned}$$

This last quantity converges to 0 and by consequence $P(V > 0) = 0$. □

The next convergence theorem concern submartingales which are bounded in L^1.

Theorem 2.11. *If X is a martingale (resp. submartingale, supermartingale) which is bounded in L^1 then it converges almost surely.*

Proof. Let us first consider the case where X is a positive submartingale. Let $M = \sup_n \mathbb{E}[|X_n|]$. For all $a < b$, we have by Theorem 2.9

$$\mathbb{E}[D_n^{a,b}] \leq \frac{1}{b-a}(M + |b|).$$

Taking the limit $n \to +\infty$ we see that $D^{a,b}$ is integrable, hence it is almost surely finite. Let $N_{a,b} = (D^{a,b} = +\infty)$. The set $N = \cup_{a<b\in\mathbb{Q}} N_{a,b}$ is negligeable.

By Doob's inequality we have

$$\lambda\, P(X_n^* > \lambda) \leq \sup_n \mathbb{E}[|X_n|] = M < +\infty\,.$$

Letting n tend to $+\infty$ we obtain

$$\lambda\, P(X^* > \lambda) \leq M$$

and thus $P(X^* > \lambda)$ tends to 0 when λ tends to $+\infty$. This proves that X^* is a.s. finite.

We have proved that all the hypothesis of Lemma 2.8 are almost surely satisfied. We have proved the almost sure convergence in that particular case.

Now, if X is a martingale bounded in L^1 the processes X^+ and X^- are positive submartingales which are bounded in L^1, hence they converge almost surely. So does X.

If X is a general submartingale, bounded in L^1, it can be decomposed as $X = M + A$ where M is a martingale and A is an increasing sequence, vanishing at 0 (Theorem 2.4). We have

$$\mathbb{E}[|A_n|] = \mathbb{E}[A_n] = \mathbb{E}[X_n - M_n] \leq \mathbb{E}[|X_n|] + |\mathbb{E}[M_n]| \leq 2K\,.$$

Hence A is bounded in L^1 and as a consequence M is bounded in L^1 too. Hence the martingale M converge a.s. and A too (by monotone convergence). This proves that X converges a.s.

The case of supermartingales is easily deduced from the one of submartingales. □

Corollary 2.12. *Every supermartingale greater than a constant converges almost surely. Every submartingale smaller than a constant converges almost surely.*

Proof. A supermartingale greater than a constant is always bounded in L^1 (exercise). The same holds for a submartingale smaller than a constant. □

Note that the definition of supermartingale or submartingale is in the large sense: the martingales are supermartingales and submartingales. Hence the above theorems have very strong consequences such as: every positive martingale converges a.s., etc.

With the help of the notion of uniform integrability, we can pass from a.s. convergence to L^1 convergence for martingales.

Theorem 2.13. *Let X be a martingale bounded in L^1 and let X_∞ be its a.s. limit (Theorem 2.11). The following assertions are equivalent.*

i) X converges to X_∞ in L^1.

ii) X is uniformly integrable.

iii) $X_n = \mathbb{E}[X_\infty \,|\, \mathcal{F}_n]$.

iv) There exists an integrable random variable Y such that $X_n = \mathbb{E}[Y \,|\, \mathcal{F}_n]$ for all n. In that case, $X_\infty = \mathbb{E}[Y \,|\, \mathcal{F}_\infty]$.

When any of these assertions is satisfied, we have, for every stopping times S and T such that $S \leq T$

$$\mathbb{E}[X_T \,|\, \mathcal{F}_S] = X_S \,.$$

Proof. We already know that i) is equivalent to ii) (Theorem 1.10). Proposition 1.9 shows that iii) implies ii).

Let us show that i) implies iii). We have, for all $p \geq n$

$$X_n = \mathbb{E}[X_p \,|\, \mathcal{F}_n].$$

The conditional expectation is continuous in L^1 and X_p converges to X_∞ in L^1. This gives iii). We have proved the equivalence of the first 3 assertions.

Let us now show that if $X_n = \mathbb{E}[Y \,|\, \mathcal{F}_n]$ then $X_\infty = \mathbb{E}[Y \,|\, \mathcal{F}_\infty]$. First of all, it is clear that X_∞ is $\mathcal{F}_\infty$-measurable. Hence we need to show that for all $A \in \mathcal{F}_\infty$ we have $\mathbb{E}[Y \, \mathbb{1}_A] = \mathbb{E}[X_\infty \, \mathbb{1}_A]$. This relation is valid for $A \in \mathcal{F}_n$. Hence it is valid for all $A \in \cup_n \mathcal{F}_n$. It remains valid for the σ-field $\mathcal{F}_\infty$ generated by $\cup_n \mathcal{F}_n$, by a monotone class argument. All the equivalences are proved.

We now prove the Stopping Theorem-like part. Let T be a stopping time and $A \in \mathcal{F}_T$. We have

$$\begin{aligned}\mathbb{E}[X_\infty \mathbb{1}_A] &= \sum_n \mathbb{E}[X_\infty \mathbb{1}_{A \cap (T=n)}] \\ &= \sum_n \mathbb{E}[X_n \mathbb{1}_{A \cap (T=n)}] \\ &= \sum_n \mathbb{E}[X_T \mathbb{1}_{A \cap (T=n)}] \\ &= \mathbb{E}[X_T \mathbb{1}_A].\end{aligned}$$

This shows that $\mathbb{E}[X_\infty \,|\, \mathcal{F}_T] = X_T$. Now, if $S \leq T$ is another stopping time, we have

$$\mathbb{E}[X_T \,|\, \mathcal{F}_S] = \mathbb{E}[\mathbb{E}[X_\infty \,|\, \mathcal{F}_T] \,|\, \mathcal{F}_S] = \mathbb{E}[X_\infty \,|\, \mathcal{F}_S] = X_S \,. \quad \square$$

Our last convergence theorem to be proved is one which concerns "inverse submartingales".

For a moment we consider a filtration $(\mathcal{F}_{-n})_{n \in \mathbb{N}}$ indexed by $-\mathbb{N}$ instead of $\mathbb{N}$. This means that $\cdots \subset \mathcal{F}_{-n-1} \subset \mathcal{F}_{-n} \subset \cdots \subset \mathcal{F}_0$. An *inverse submartingale* is an integrable, adapted process $(X_{-n})_{n \in \mathbb{N}}$ such that

$X_{-n-1} \leq \mathbb{E}[X_{-n} \mid \mathcal{F}_{-n-1}]$. When the equality holds for all n, we speak of an *inverse martingale*. Note that an inverse martingale is always uniformly integrable for $X_{-n} = \mathbb{E}[X_0 \mid \mathcal{F}_{-n}]$ (by Proposition 1.9).

Theorem 2.14 (Inverse Submartingale Convergence Theorem). *Every inverse submartingale X which is uniformly integrable converges almost surely and in L^1 to a random variable $X_{-\infty}$ which satisfies*

$$X_{-\infty} \leq \mathbb{E}[X_0 \,|\, \mathcal{F}_{-\infty}]\,.$$

If X is inverse martingale then the inequality above is an equality.

Proof. In the case of inverse submartingales the basic inequalities (Theorems 2.5 and 2.9) are stated in a somehow different way but they are proved in the same way as for usual submartingales:

– Let X be a positive inverse submartingale. Then for all $\lambda \in \mathbb{R}^+$, all $n \in \mathbb{N}$, we have

$$\lambda\, P(X^*_{-n} \geq \lambda) \leq \mathbb{E}[X_0\, \mathbb{1}_{X^*_{-n} \geq \lambda}] \leq \mathbb{E}[X_{-n}]\,.$$

– If X is an inverse submartingale then for all $a < b$ we have

$$\mathbb{E}[D^{a,b}_{-n}] \leq \frac{1}{(b-a)} \mathbb{E}[(X_0 - b)^+]\,.$$

The existence of an almost sure limit $X_{-\infty}$ is then proved in the same way as before. The uniform integrability gives easily the L^1 convergence and all the other properties, in the same way as for usual submartingales. □

Proposition 2.15.

i) Every inverse martingale is uniformly integrable.

ii) Every inverse submartingale which is bounded in L^1 is uniformly integrable.

iii) Every inverse submartingale X such that $\inf_n \mathbb{E}[X_{-n}] > -\infty$ *is uniformly integrable.*

Proof.

i) If M is an inverse martingale then $M_{-n} = \mathbb{E}[M_0 \,|\, \mathcal{F}_{-n}]$ and M is uniformly integrable by Proposition 1.9.

ii) For inverse submartingales we also have a Doob decomposition: $X = M + A$ where and A is predictable, decreasing, positive and bounded in L^1. This means that A is uniformly integrable too. As M is also uniformly integrable, we conclude that X is uniformly integrable.

iii) We have $|X| = -X + 2X^+$ and X^+ is a submartingale (exercise). As a consequence we have

$$\mathbb{E}[|X_{-n}|] = -\mathbb{E}[X_{-n}] + 2\mathbb{E}[X^+_{-n}] \leq -\inf_n \mathbb{E}[X_{-n}] + 2\mathbb{E}[X^+_0]\,.$$

That is, X is bounded in L^1 and hence uniformly integrable by ii). □

2.4 Regularization Theorems

We are now back to the usual setup, that is, filtrations and processes indexed by $\mathbb{R}^+$. We shall prove a fundamental theorem on the path regularity of submartingales. We have already discussed in Section 1.3 the importance of such a regularity.

We first start with an easy extension of the Downcrossing Theorem 2.9. For a given submartingale X and a given interval $[a,b] \subset \mathbb{R}$, for any finite set $F \subset \mathbb{R}^+$ we put $D(X, F, [a,b])$ to be the number of downcrossings over the interval [a,b] for the process X restricted to the indexes $t \in F$. For any countable set $Q \subset \mathbb{R}^+$ we put

$$D(X, Q, [a,b]) = \sup\{D(X, F, [a,b])\,;\ F \subset Q, F \text{ finite}\}\,.$$

Theorem 2.9 immediately extends to the following result.

Theorem 2.16. *Let X be a submartingale, let $I = [r,s]$ be an interval of $\mathbb{R}^+$ and $a < b \in \mathbb{R}$. We have*

$$(b-a)\mathbb{E}\left[M(X, I \cap \mathbb{Q}, [a,b])\right] \leq \mathbb{E}\left[(X_s - b)^+\right]\,.$$

This extension has the following important consequence.

Theorem 2.17. *Let X be a submartingale. There exists a set $\Omega_0 \in \mathcal{F}_\infty$ with $P(\Omega_0) = 1$ and such that for all $\omega \in \Omega_0$, all $t \in \mathbb{R}^+$*

$$\lim_{\substack{s \to t \\ s > t \\ s \in \mathbb{Q}}} X_s(\omega) \quad \textit{exists}$$

and for all $t \in {\mathbb{R}^+}^*$

$$\lim_{\substack{s \to t \\ s < t \\ s \in \mathbb{Q}}} X_s(\omega) \quad \textit{exists.}$$

Proof. Let $t_0 \in \mathbb{R}^+$ be such that $\lim_{s \to t_0, s > t_0, s \in \mathbb{Q}} X_s(\omega)$ does not exist. Then there exists $a, b \in \mathbb{Q}$, $a < b$ such that

$$\liminf_{\substack{s \to t \\ s > t \\ s \in \mathbb{Q}}} X_s(\omega) < a < b < \limsup_{\substack{s \to t \\ s > t \\ s \in \mathbb{Q}}} X_s(\omega).$$

Then, if $t_0 \in I = [r,s]$, we have $M(X, I \cap \mathbb{Q}, [a,b]) = +\infty$. But by Theorem 2.16 we know that

$$\mathbb{E}\left[M(X, I \cap \mathbb{Q}, [a,b])\right] \leq \frac{\mathbb{E}\left[(X_s - b)^+\right]}{b-a} < +\infty.$$

This in particular implies that there exists a set $\Omega_0 \in \mathcal{F}_\infty$, with $P(\Omega_0) = 1$, such that $M(X(\omega), I \cap \mathbb{Q}, [a,b]) < \infty$ for all $\omega \in \Omega_0$. This contradicts the hypothesis above.

The case of the left-limit is obtained in a similar way. □

The limits obtained in the theorem above are respectively denoted by X_{t+} and X_{t-}.

Proposition 2.18. *Let X be a submartingale such that $t \mapsto \mathbb{E}[X_t]$ is right-continuous. Then all the random variables X_{t+} belong to $L^1(\Omega, \mathcal{F}, P)$ and we have*

$$\mathbb{E}\left[X_{t+} \mid \mathcal{F}_t\right] = X_t \, .$$

Furthermore the process $(X_{t+})_{t\geq 0}$ is a $(\mathcal{F}_{t+})_{t\geq 0}$-submartingale. If X is a martingale then $(X_{t+})_{t\geq 0}$ is $(\mathcal{F}_{t+})_{t\geq 0}$-martingale.

Proof. Let $t \in \mathbb{R}^+$ and (t_n) be a decreasing sequence converging to t with $t_n > t$. Then $(X_{t_n})_{n\in\mathbb{N}}$ is a $(\mathcal{F}_{t_n})_{n\in\mathbb{N}}$-submartingale and it converges to X_{t+}. As $t \mapsto \mathbb{E}[X_t]$ is right-continuous, the sequence $(\mathbb{E}[X_{t_n}])$ converges to $\mathbb{E}[X_t]$, hence this sequence is bounded. We can apply the Inverse Submartingale Convergence Theorem 2.14 so that X_{t_n} converges almost surely and in L^1 to X_{t+}. As $X_t = \mathbb{E}\left[X_{t_n} \mid \mathcal{F}_t\right]$, we get, passing to the limit, $\mathbb{E}\left[X_{t+} \mid \mathcal{F}_t\right] = X_t$. Now take $s < t$ and $A \in \mathcal{F}_{s+}$. Let (s_n) be a decreasing sequence converging to s with $s_n > s$. We have $A \in \mathcal{F}_{s_n} \subset \mathcal{F}_t$ (for n large enough) and

$$\mathbb{E}\left[X_{t+}\, \mathbb{1}_A\right] = \mathbb{E}\left[X_t\, \mathbb{1}_A\right] \geq \mathbb{E}\left[X_{s_n}\, \mathbb{1}_A\right] \to_n \mathbb{E}\left[X_{s+}\, \mathbb{1}_A\right] .$$

This proves that $\mathbb{E}[X_{t+} \,|\, \mathcal{F}_{s+}] \geq X_{s+}$. □

Now, using the fact that our filtration are always right-continuous we get the following important theorem, which was the aim of this section.

Theorem 2.19. *Every submartingale admits a càdlàg modification.*

Proof. Put $\widetilde{X_t} = X_{t+}$ on Ω_0 and $\widetilde{X_t} = 0$ on Ω_0^c. Then $\widetilde{X}$ is a $(\mathcal{F}_t)$-submartingale (by the right-continuity of the filtration) and it is càdlàg. □

Here we are! All the martingales, submartingales and supermartingales admit a càdlàg modification. Hence, by the discussion of Section 1.3, we can always consider our martingales (submartingales, supermartingales) to be living on the canonical space $(D, \mathcal{B}(D), P_X)$ of càdlàg functions.

Let us end up this section with the generalization to continuous time of the main convergence theorems for martingales. We have already seen the extension of the Down-crossing Theorem (Theorem 2.16). Now comes the extension to continuous time of Doob's L^p inequality (Theorem 2.7).

Theorem 2.20 (Doob's inequality). *Let X be a positive submartingale. Let $M_t^* = \sup_{s\leq t} X_s$. Let $1 < p < \infty$ and $q = p/(p-1)$. Then we have the inequality*

$$\|X^*\|_p \leq q\|X\|_p \, .$$

Proof. Let Q be a countable dense subset of $\mathbb{R}^+$. As X is càdlàg we have $\sup_{s\leq t} X_s = \sup s \leq t, s \in QX_s$. The result follows no easily from the discrete time corresponding inequality (Theorem 2.7). □

As all the convergence theorems rely on these two inequalities we get the corresponding extensions of the main convergence theorems.

Theorem 2.21. *If M is a martingale such that $\sup_t \mathbb{E}[M_t^2] < \infty$, then M converges a.s. and in L^2.*

Theorem 2.22. *Every martingale (resp. supermartingale, submartingale), bounded in L^1, converges a.s.*

Corollary 2.23. *Every supermartingale larger than a constant converges a.s. Every submartingale smaller than a constant converges a.s.*

Theorem 2.24. *Let X be a martingale bounded in L^1 and let X_∞ be its a.s. limit. The following assertions are then equivalent.*

i) X_t converges to X_∞ in L^1.

ii) X is uniformly integrable.

iii) $X_t = \mathbb{E}[X_\infty \,|\, \mathcal{F}_t]$.

When any of these assertions is satisfied, we have for all stopping times $S \leq T$

$$\mathbb{E}[X_T \,|\, \mathcal{F}_S] = X_S.$$

2.5 Brownian Motion and Poisson Processes

As basic ewamples of continuous time martingales we recover our two favorite stochastic processes: Brownian motion and Poisson processes (cf Sections 1.5 and 1.6).

Theorem 2.25.

1) Let W be a Brownian motion defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$. Then the following processes are martingales (with respect to $(\mathcal{F}_t)$):

– $W_t,\ t \in \mathbb{R}^+$,

– $W_t^2 - t,\ t \in \mathbb{R}^+$,

– $\exp(\alpha W_t - \alpha^2 t/2),\ t \in \mathbb{R}^+$.

2) Let N be a Poisson process with intensity λ defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$. Then the following processes are martingales (with respect to $(\mathcal{F}_t)$):

– $N_t - \lambda t,\ t \in \mathbb{R}^+$,

– $(N_t - \lambda t)^2 - \lambda t,\ t \in \mathbb{R}^+$,

– $\exp(-\alpha\lambda t)(1+\alpha)^{N_t},\ t \in \mathbb{R}^+$.

Proof. Let us first show that the Brownian motion itself is a martingale. For all $s \le t$ we have, using the independence of the increments of W and the fact that the law of $W_t - W_s$ is $\mathcal{N}(0, \sqrt{t-s})$

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

In the same way we have

$$\begin{aligned}\mathbb{E}[W_t^2 - t \,|\, \mathcal{F}_s] &= \mathbb{E}[(W_t - W_s)^2 \,|\, \mathcal{F}_s] + 2\mathbb{E}[W_t W_s \,|\, \mathcal{F}_s] - \mathbb{E}[W_s^2 \,|\, \mathcal{F}_s] - t \\ &= \mathbb{E}[(W_t - W_s)^2] + 2W_s^2 - W_s^2 - t \\ &= t - s + W_s^2 - t = W_s^2 - s\,.\end{aligned}$$

Finally, noting that for a random variable with law $\mathcal{N}(m, \sigma)$ we have (exercise)

$$\mathbb{E}[\exp(\alpha X)] = \exp(\alpha^2\sigma^2/2 - \alpha m)\,,$$

we have

$$\begin{aligned}\mathbb{E}[\exp(\alpha W_t - \alpha^2 t/2) \,|\, \mathcal{F}_s] &= \mathbb{E}[\exp(\alpha(W_t - W_s)) \exp(\alpha W_s) \,|\, \mathcal{F}_s] \;\exp(-\alpha^2 t/2) \\ &= \mathbb{E}[\exp(\alpha(W_t - W_s)) \,|\, \mathcal{F}_s] \;\exp(\alpha W_s)\; \exp(-\alpha^2 t/2) \\ &= \mathbb{E}[\exp(\alpha(W_t - W_s))] \;\exp(\alpha W_s)\; \exp(-\alpha^2 t/2) \\ &= \exp(\alpha^2(t-s)/2)\; \exp(\alpha W_s)\; \exp(-\alpha^2 t/2) \\ &= \exp(\alpha W_s - \alpha^2 s/2)\,.\end{aligned}$$

Let us now prove the corresponding results for the Poisson processes. Recall that it has independent increments and that the law of $N_t - N_s$ is $\mathcal{P}(\lambda(t-s))$. Hence we have

$$\begin{aligned}\mathbb{E}[N_t - \lambda t \,|\, \mathcal{F}_s] &= \mathbb{E}[N_t - N_s \,|\, \mathcal{F}_s] + \mathbb{E}[N_s \,|\, \mathcal{F}_s] - \lambda t \\ &= \mathbb{E}[N_t - N_s] + N_s - \lambda t \\ &= \lambda(t-s) + N_s - \lambda t = N_s - \lambda s\,.\end{aligned}$$

In the same way we have

$$\begin{aligned}\mathbb{E}\left[(N_t - \lambda t)^2 - \lambda t \,|\, \mathcal{F}_s\right] &= \mathbb{E}\left[(N_t - N_s - \lambda(t-s))^2 \,|\, \mathcal{F}_s\right] + \\ &\quad + 2\mathbb{E}\left[(N_t - \lambda t)(N_s - \lambda s) \,|\, \mathcal{F}_s\right] - \\ &\quad - \mathbb{E}\left[(N_s - \lambda s)^2 \,|\, \mathcal{F}_s\right] - \lambda t \\ &= \mathbb{E}\left[(N_t - N_s - \lambda(t-s))^2\right] + (N_s - \lambda s)^2 - \lambda t \\ &= \lambda(t-s) + (N_s - \lambda s)^2 - \lambda t = (N_s - \lambda s)^2 - \lambda s\,.\end{aligned}$$

Finally, noting that for a random variable X with law $\mathcal{P}(\tau)$ we have (exercise)

$$\mathbb{E}[\exp(\alpha X)] = \exp(\tau(e^\alpha - 1))\,,$$

we have

$$\begin{aligned}\mathbb{E}[\exp(\alpha N_t) \,|\, \mathcal{F}_s] &= \mathbb{E}[\exp(\alpha(N_t - N_s)) \,|\, \mathcal{F}_s] \,\exp(\alpha N_s)\\ &= \mathbb{E}[\exp(\alpha(N_t - N_s))] \,\exp(\alpha N_s)\\ &= \exp\left(\lambda(t-s)(e^\alpha - 1)\right) \exp(\alpha N_s)\,.\end{aligned}$$

In particular the process

$$\exp\left(\alpha N_t - \lambda t(e^\alpha - 1)\right), \quad t \in \mathbb{R}^+ \tag{2.4}$$

is a martingale. As $(1+\alpha)^{N_t} = \exp\left(\ln(1+\alpha)N_t\right)$ we obtain the requiered result, replacing α by $\ln(1+\alpha)$ in (2.4). □

2.6 Quadratic Variation

In the previous section we have seen that both the Brownian motion and the Poisson proocesses are martingales. These two examples are very different. We have seen in Section 1.5 that the Brownian motion is a continuous process whose paths are of infinite variation over any interval. The Poisson processes have of course non-continuous paths, but they are of finite variation, for they are increasing.

The following result illustrates this difference: there is no non-trivial continuous martingale with finite variations. It is going to be a very usefull result in the sequel.

Theorem 2.26. *Every finite variation continuous martingale is constant.*

Proof. Let X be a martingale with continuous and finite variation paths. We have by the martingale property, for any subdivision $0 = t_0 < t_1 < \ldots < t_n = t$ of the interval $[0, t]$

$$\begin{aligned}\mathbb{E}\left[\sum_{i=0}^{n-1}(M_{t_{i+1}} - M_{t_i})^2\right] &= \mathbb{E}\left[\sum_{i=0}^{n-1}(M^2_{t_{i+1}} - M^2_{t_i})\right]\\ &= \mathbb{E}[M_t^2]\,.\end{aligned}$$

Let V_t be the total variation of X on the interval $[0, t]$. By stopping this martingale (for example at $T = \inf\{t\,;\ |V_t| > n\}$) we may assume that our martingale has its total variation bounded (Theorem 2.1). We get

$$\mathbb{E}[M_t^2] \le \mathbb{E}\left[V_t \sup_i \left|X_{t_{i+1}} - X_{t_i}\right|\right] \le K\,\mathbb{E}\left[\sup_i \left|X_{t_{i+1}} - X_{t_i}\right|\right]\,.$$

The last quantity goes to 0 when the diameter of the partition converges to 0 for X is continuous. Hence we have proved $\mathbb{E}[M_t^2] = 0$ and as a consequence $M = 0$ a.s. □

In the same way as for the Brownian motion (cf Theorem 1.15), every continuous martingale has a non-trivial quadratic variation.

Theorem 2.27. *Let M be a continuous martingale, locally square integrable. Then there exists a unique increasing adapted continuous process $[M,M]$ vanishing at 0 and such that $M^2-[M,M]$ is a uniformly integrable martingale. This process $[M,M]$ is the quadratic variation of M, that is,*

$$[M,M]_t=\lim\sum_{i=0}^{n-1}(M_{t_{i+1}}-M_{t_i})^2\,,$$

where the limit is almost sure and taken over a sequence of partitions $\{0=t_0<t_1<\ldots<t_n=t\}$ of the interval $[0,t]$ whose diameter tends to 0.

Proof. We consider that M is bounded, up to restricting ourselves to a compact interval of time if necessary. For every $n\in\mathbb{N}$, we define the stopping times

$$T_0^n=0\,,\ T_{k+1}^n=\inf\{t>T_k^n;\left|M_t-M_{T_k^n}\right|\geq 2^{-n}\}\,,\ k\in\mathbb{N}\,.$$

We choose to denote $t\wedge T_k^n$ by t_k^n simply. We then have

$$M_t^2=2\sum_k M_{t_{k-1}^n}(M_{t_k^n}-M_{t_{k-1}^n})+\sum_k(M_{t_k^n}-M_{t_{k-1}^n})^2\,. \qquad (2.5)$$

Put

$$H_t^n=\sum_k M_{T_{k-1}^n}\,\mathbb{1}_{]T_{k-1}^n,T_k^n]}(t)\,.$$

Let $H^n\bullet M$ be the first term of the right hand side of (2.5) and let A^n be the second term of the same right hand side. Note that $\sup_t\left|H_t^n-H_t^{n+1}\right|\leq 2^{-n-1}$ and $\sup_t\left|H_t^n-M_t\right|\leq 2^{-n}$. Also note that $(T_k^n)_{k\in\mathbb{N}}$ is a subsequence of $(T_k^{n+1})_{k\in\mathbb{N}}$. We consider the expression

$$\mathbb{E}\left[(H^n\bullet M-H^{n+1}\bullet M)_\infty^2\right]\,.$$

It can be written as

$$\mathbb{E}\left[\left(\sum_k Z_k\Delta_k\right)^2\right]$$

where $\Delta_k=M_{T_k^{n+1}}-M_{T_{k-1}^{n+1}}$ and Z_k is a random variable measurable for $\mathcal{F}_{T_{k-1}^{n+1}}$ satisfying $|Z_k|\leq 2^{-n-1}$. The above expression can be written as

$$\mathbb{E}\left[\sum_k|Z_k|^2\,|\Delta_k|^2\right]$$

for when $k<k'$ we have

$$\mathbb{E}\left[Z_kZ_{k'}\Delta_k\Delta_{k'}\right]=\mathbb{E}\left[Z_kZ_{k'}\Delta_k\,\mathbb{E}\left[\Delta_{k'}\mid\mathcal{F}_{T_{k'-1}^{n+1}}\right]\right]=0$$

for M is a martingale. Finally we obtain

$$\mathbb{E}\left[(H^n \bullet M - H^{n+1} \bullet M)^2_\infty\right] \leq 4^{-n-1} \|M_\infty\|_2^2 \,.$$

It is easy to check that $H^n \bullet M$ is a continuous martingale. It converges uniformly, almost surely, to a martingale N. Hence the process A^n converges uniformly, almost surely, to a process A. We have

$$M_t^2 = 2N_t + A_t \,.$$

It is easy to see that A is an increasing process, at least on the set $\{T_k^n; n, k \in \mathbb{N}\}$. But if I is an open interval in the complementray set of this set, then no T_k^n lies in I and M is constant on I. Hence so is A.

We have constructed a continuous increasing process A, vanishing at 0, such that $M^2 - A = N$ is a uniformly integrable martingale.

Let us show uniqueness. If A' is another such process then $A - A'$ is a uniformly integrable, continuous martingale. But this martingale is also of finite variations, hence $A - A' = 0$ by Theorem 2.26.

We did not quite prove the theorem in that we did not prove that the process $[M, M]$ is the quadratic variation of M. We have proved that it is the limit of the quadratic variations of M on a particular family of subdivisions made of *stopping times*. We need to prove that the result holds true for any fixed sequence of *deterministic* subdivisions. We shall not give the proof of that extension and admit it (see [?] Theorem 1.3 for a complete proof). □

3

STOCHASTIC INTEGRALS

Stochastic processes such as martingales are typically used as noises, perturbating ordinary differential equations with stochastic terms (cf Chapter 4). But as we have seen with Theorem 2.26, even the simplest martingales (the continuous ones, the Brownian motion) have trajectories with infinite variation over any interval. This means that integrating with respect to such a process naively (that is, path by path) is impossible. Another theory of integration has to be developed for such stochastic processes, a theory which makes use of the martingales globally and not trajectory by trajectory. This is the aim of Stochastic Integration Theory which we develop here.

3.1 Predictable Processes

In Stochastic Integration Theory we need a particular notion of measurability for stochastic processes which is stronger than the notion of adaptedness or of progressivity, we shall consider *predictable processes.*

We are given a fixed complete and right-continuous filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ for the whole of this chapter.

The *predictable* σ-algebra on $\mathbb{R}^+ \times \Omega$ is the σ-algebra $\mathcal{P}$ generated by adapted and left-continuous processes. A process X is *predictable* if it is measurable with respect to the predictable σ-algebra $\mathcal{P}$.

Proposition 3.1. *The following σ-algebras all coincide with the predictable σ-algebra.*

i) The σ-algebra generated by adapted, continuous processes.

ii) The σ-algebra generated by the stochastic intervals $[\![0_A]\!]$, $A \in \mathcal{F}_0$ and $]\!]S, T]\!]$, where S, T runs over all stopping times.

iii) The σ-algebra generated by the sets $A \times \{0\}$ with $A \in \mathcal{F}_0$ and the sets $A \times]s, t]$ with $A \in \mathcal{F}_s$ and $s < t$ rationals.

Proof. Let us call $\mathcal{P}_1, \mathcal{P}_2$ and $\mathcal{P}_3$, respectively, the 3 σ-algebras introduced above. Clearly $\mathcal{P}_1 \subset \mathcal{P}$. We also have $\mathcal{P}_3 \subset \mathcal{P}$ for the indicator function of $A \times]s,t]$ is adapted and left-continuous.

Let X be an adapted and left-continuous process. Let

$$X_t^n = X_0 \mathbb{1}_{[0]}(t) + \sum_k X_{k/n} \mathbb{1}_{]k/n,(k+1)/n]}(t) . \tag{3.1}$$

The process X^n converges to X. This proves that $\mathcal{P}$ is included in $\mathcal{P}_3$ and hence they are equal.

The functions $\mathbb{1}_{]u,v]}$ can be approached by continuous functions f_n with compact support in $]u, v+1/n]$. For any H which is $\mathcal{F}_u$-measurable, the process Hf_n is continuous and adapted. This shows that $\mathcal{P}_3$ is included in $\mathcal{P}_1$, hence $\mathcal{P}_1 = \mathcal{P}$.

The indicator functions of stochastic intervals $]\!]S,T]\!]$ are left-continuous, hence $\mathcal{P}_2 \subset \mathcal{P}$. Let $A \in \mathcal{F}_s$, by putting $S = s$ on A and $S = t$ on A^c, by putting $T = t$, we define two stopping times such that $S \leq T$ and $]\!]S,T]\!] = A \times]s,t]$. Hence $\mathcal{P}_3 \subset \mathcal{P}_2$ and hence $\mathcal{P}_2 = \mathcal{P}$. □

Theorem 3.2. *If X is a predictable process, then $X_T \mathbb{1}_{T<\infty}$ is $\mathcal{F}_{T-}$-measurable for every stopping time T.*

Proof. Let $\mathcal{H}$ be the set of predictable processes X such that $X_T \mathbb{1}_{T<\infty}$ is $\mathcal{F}_{T-}$-measurable for every stopping time T. By a monotone class argument one sees easily that it is sufficient to prove that $\mathcal{H}$ contains the indicators of stochastic intervals of the form $[\![0_A]\!]$, $A \in \mathcal{F}_0$ and $]\!]U,V]\!]$, with $U \leq V$ being stopping times (cf proposition above). Let T be a stopping time. In the first case $X_T \mathbb{1}_{T<\infty}$ equals $\mathbb{1}_{A \cap (T=0)}$ which is $\mathcal{F}_{T-}$-measurable. In the second case we have

$$X_T \mathbb{1}_{T<\infty} = (\mathbb{1}_{U<T} - \mathbb{1}_{V<T}) \mathbb{1}_{T<\infty}$$

which is also clearly $\mathcal{F}_{T-}$-measurable. □

3.2 Stochastic Stieltjes Integrals

The very first step in Stochastic Integration Theory is to integrate with respect to finite variation stochastic processes.

Let us first recall few basic facts about finite variation functions. We are interested here only into functions f from $\mathbb{R}^+$ to $\mathbb{R}$. A function f is a *finite variation* function if it is right-continuous and if its total variation

$$V_t(f) = \lim \sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)| ,$$

is finite on every interval $[0,t]$, where the limit above is taken over any sequence of partitions $\{0 = t_0 < t_1 < \ldots < t_n = t\}$ of the interval $[0,t]$ whose diameters tend to 0.

Recall that a finite variation function is always difference of two increasing functions and always admits left-limits in all points (hence they are càdlàg). By convention we always put

$$f(0_-) = f(0)\,.$$

For all $s > 0$ we put

$$\Delta f(s) = f(s) - f(s_-)\,.$$

In particular, with our convention in 0 we have

$$\Delta f(0) = 0\,.$$

Recall that the cardinal of jumps Δf of f is at most countable.

To any finite variation function f is associated a measure μ_f on $[0, +\infty[$ defined by $\mu_f(]a, b]) = f(b) - f(a)$. It is easy to show that this is enough to define a (signed, real) measure, called the *Stieltjes measure* associated to f, which satisfies

$$\begin{aligned}\mu_f(]a, b[) &= f(b_-) - f(a)\,,\\ \mu_f([a, b]) &= f(b) - f(a_-)\,,\\ \mu_f([a, b[) &= f(b_-) - f(a_-)\,.\end{aligned}$$

In particular note that

$$\mu_f(\{a\}) = \Delta f(a)\,.$$

Let g be any function on $[0, +\infty[$, we say that g is *integrable with respect to μ_f* if

$$\int_0^\infty |g(s)|\, |d\mu_f|\,(s) < \infty\,.$$

In that case the integral of g with respect to μ_f is well-defined and we shall denote it by

$$\int_0^\infty g(s)\, df(s)\,.$$

Note that, when dealing with general Stieltjes measures, one has to be careful with the domain of integration: integrating over $[s, t]$, or $]s, t]$ etc... may give different results. But, with our convention in 0, the integrals $\int_{[0,t]}$ and $\int_{]0,t]}$ coincide. They are naturally denoted by $\int_0^t$ as there is no ambiguity anymore.

In the following theorem we list basic properties of these integrals, without proof.

Theorem 3.3.

i) The function $t \mapsto h(t) = \int_0^t g(s)\, df(s)$ is again a finite variation function on $\mathbb{R}^+$, with associated Stieltjes measure

$$\mu_h(ds) = g(s)\, d\mu_f(s)\,.$$

ii) $\Delta h(t) = g(t)\, \Delta f(t)$, *for all* $t \in \mathbb{R}^+$.

Finite variation functions also admit non-trivial quadratic variations.

Proposition 3.4. *Let* f *be a finite variation function on* $\mathbb{R}^+$. *The quadratic variation of* f *on the interval* $[0, t]$ *is equal to*

$$\sum_{0 \leq s \leq t} \Delta f(s)^2\,.$$

Proof. We can decompose f as $f^c + f^d$ where

$$f^d(t) = \sum_{0 \leq s \leq t} \Delta f(s)$$

and $f^c(t) = f(t) - f^d(t)$. In particular f^c is continuous. We have

$$\begin{aligned}\sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)|^2 = &\sum_{i=0}^{n-1} |f^c(t_{i+1}) - f^c(t_i)|^2 + \\ &+ \sum_{i=0}^{n-1} |f^c(t_{i+1}) - f^c(t_i)| \left|f^d(t_{i+1}) - f^d(t_i)\right| + \\ &+ \sum_{i=0}^{n-1} \left|f^d(t_{i+1}) - f^d(t_i)\right|^2\,.\end{aligned}$$

The two first terms are dominated by

$$\begin{aligned}\sup_{|s-t| \leq \delta} |f^c(t) - f^c(s)| &\left[\sum_{i=0}^{n-1} |f^c(t_{i+1}) - f^c(t_i)| + \sum_{i=0}^{n-1} \left|f^d(t_{i+1}) - f^d(t_i)\right|\right] \\ &\leq 2V_t(f) \sup_{|s-t| \leq \delta} |f^c(t) - f^c(s)|\end{aligned}$$

and hence they tend to 0 as the diameter δ of the partition tends to 0. Let us show that the last term $\sum_{i=0}^{n-1} \left|f^d(t_{i+1}) - f^d(t_i)\right|^2$ converges to $\sum_{0 \leq s \leq t} \Delta f(s)^2$.

The sum $\sum_{0 \leq s \leq t} \Delta f(s)^2$ can be actually written as a convergent series $\sum_n \Delta f(s_n)^2$, where the points s_n are the jumping points of f. Define $\delta(s_n)$ to be the length of the largest interval around s_n which meets no other s_m. We can define a new sequence (u_n) which is the sequence (s_n) reordered in decreasing order with respect to the values $\delta(s_n)$. The sum $\sum_n \Delta f(s_n)^2$ is equal to $\sum_n \Delta f(u_n)^2$. In particular, the sum

$$\sum_{n; \delta(s_n) \leq \delta} \Delta f(s_n)^2$$

converges to 0 as δ tends to 0. Now, if δ is the diameter of the partition $\{0 = t_0 < t_1 < \ldots < t_n = t\}$, we can divide the sum $\sum_{i=0}^{n-1} \left|f^d(t_{i+1}) - f^d(t_i)\right|^2$ into the sum over the t_i's such that $[t_i, t_{i+1}[$ contains only one s_n and the sum over those intervals $[t_i, t_{i+1}[$ which contains more than one s_n (the terms with no s_n in the interval $[t_i, t_{i+1}[$ give no contribution). The first sum contains $\sum_{n;\delta(s_n)\geq\delta} \Delta f(s_n)^2$ plus a term dominated by $\sum_{n;\delta(s_n)\leq\delta} \Delta f(s_n)^2$. The second sum is also dominated by $\sum_{n;\delta(s_n)\leq\delta} \Delta f(s_n)^2$. This gives the result. □

In the following we shall take the following notation:

$$[f, f](t) = \sum_{0\leq s\leq t} \Delta f(s)^2 \,.$$

Finally, for any two finite variations functions f and g, we denote by $[f, g]$ the function

$$[f, g](t) = \sum_{0<s\leq t} \Delta f(s)\Delta g(s) \,.$$

With all these notations, we have the following Integration by Part Formula for finite variation functions.

Proposition 3.5. *If f and g are finite variation functions then*

$$f(t)g(t) = f(0)g(0) + \int_0^t f(s_-)\, dg(s) + \int_0^t g(s_-)\, df(s) + [f, g](t) \,. \tag{3.2}$$

Proof. Consider the square $[0, t] \times [0, t]$. It can be decomposed into 3 disjoint pieces

$$\begin{aligned}[0, t] \times [0, t] =& \{(u, v); 0 \leq u < v \leq t\} \cup \{(u, v); 0 \leq v < u \leq t\} \cup \\ &\cup \{(u, v); 0 \leq u = v \leq t\} \,.\end{aligned}$$

Applying the measure $\mu_f \times \mu_g$ to this decomposition gives (exercise)

$$\begin{aligned}(f(t) - f(0))(g(t) - g(0)) = &\int_0^t (f(s_-) - f(0))\, dg(s) + \\ &+ \int_0^t (g(s_-) - g(0))\, df(s) + [f, g](t) \,.\end{aligned}$$

Rearranging the terms, we obtain (3.2). □

This Integration by Part Formula extends into a formula which we shall call the Ito Formula for Finite Variation Functions.

Theorem 3.6 (Ito's Formula for Finite Variation Functions). *Let h be a C^1 function on $\mathbb{R}$ and let f be a finite variation function on $\mathbb{R}^+$. Then we have*

$$h \circ f(t) = h \circ f(0) + \int_0^t h' \circ f(s_-)\, df(s) + \\ + \sum_{0<s\le t} \big(\Delta h \circ f(s) - h' \circ f(s_-)\, \Delta f(s)\big)\,. \quad (3.3)$$

Proof. We consider that the finite variation function f is fixed, together with a fixed compact interval of time $[0,T]$ on which we want to prove the relation (3.3). Let $\mathcal{A}$ be the set of functions h for which (3.3) holds true. Then clearly $\mathcal{A}$ is a vector space. Let us prove it is an algebra, that is, it is stable under products.

Assume that h_1 and h_2 belong to $\mathcal{A}$. Put $F_1(t) = h_1 \circ f(t)$ and $F_2(t) = h_2 \circ f(t)$. Both F_1 and F_2 are finite variation functions satisfying (3.3) which we write (with obvious notations)

$$\begin{aligned} dF_1(t) &= h_1' \circ f(t_-)\, df(t) + \big(\Delta F_1(t) - h_1' \circ f(t_-)\, \Delta f(s)\big) \\ dF_2(t) &= h_2' \circ f(t_-)\, df(t) + \big(\Delta F_2(t) - h_2' \circ f(t_-)\, \Delta f(s)\big)\,. \end{aligned}$$

Put $H(t) = F_1(t)F_2(t)$, applying Proposition 3.5 we get

$$dH(t) = F_1(s_-)\, dF_2(s) + F_2(s_-)\, dF_1(s) + d[F_1, F_2](t)\,.$$

In particular

$$\Delta H(t) = F_1(s_-)\, \Delta F_2(s) + F_2(s_-)\, \Delta F_1(s) + \Delta F_1(t)\, \Delta F_2(t)\,.$$

Alltogether we have, putting $h = h_1 h_2$

$$\begin{aligned} dH(t) = &\big(h_1 \circ f(t_-)\, h_2' \circ f(t_-) + h_2 \circ f(t_-)\, h_1' \circ f(t_-)\big)\, df(t) + \\ &+ \big(F_1(t_-)\, \Delta F_2(t) + F_2(t_-)\, \Delta F_1(t) + \Delta F_1(t)\, \Delta F_2(t) + \\ &+ (h_1 h_2' + h_1' h_2) \circ f(t_-)\, \Delta f(t)\big) \\ &= h' \circ f(t_-)\, df(t) + \Delta H(t) - h'(t_-)\, \Delta f(t)\,. \end{aligned}$$

This is to say that $h = h_1 h_2$ belongs to $\mathcal{A}$. We have proved that $\mathcal{A}$ is an algebra. As $\mathcal{A}$ obviously contains the functions $\mathbb{1}$ and $f(t) = t$, it contains all the polynomials. We conclude by a usual approximation argument. □

We finally write without proof, the analogue of Theorem 3.6 for multidimensional functions.

Theorem 3.7. *Let $f = (f_1, \dots, f_n)$ be a finite variation function from $\mathbb{R}^+$ to $\mathbb{R}^n$. For any function $h \in C^1(\mathbb{R}^n)$ we have*

$$h \circ f(t) = h \circ f(0) + \sum_{i=1}^n \int_0^t D_i h \circ f(s_-)\, df_i(s) + \\ + \sum_{0\le s\le t} \big(\Delta h \circ f(s) - D_i h \circ f(s_-)\, \Delta f_i(s)\big)\,. \quad (3.4)$$

Coming back to stochastic processes, we shall say that a stochastic process X is a *finite variation process* if the trajectory $t \mapsto X_t(\omega)$ is a finite variation function on $\mathbb{R}^+$, for almost all $\omega \in \Omega$.

The Stieltjes integral of a stochastic process H with respect to a finite variation process X is defined if

$$\int_0^\infty |H_t(\omega)| \ |dX_t(\omega)| < \infty$$

holds for almost all ω. In that case, we define the stochastic integral

$$\int_0^\infty H_s \, dX_s$$

by

$$\left[\int_0^\infty H_s \, dX_s\right](\omega) = \int_0^\infty H_s(\omega) \, dX_s(\omega) \, .$$

The role of predictable processes already appears in these simple stochastic integrals.

Theorem 3.8. *If X is an finite variation martingale and if H is a predictable process such that*

$$\int_0^\infty |H_t| \, |dX_s| < \infty \quad a.s.$$

then the stochastic process $\int_0^t H_s \, dX_s$, $t \in \mathbb{R}^+$ is a finite variation martingale too.

Proof. Let us start with a simple predictable process H of the form

$$H_t(\omega) = \mathbb{1}_A(\omega) \mathbb{1}_{]u,v]}(t)$$

with $u < v$ and $A \in \mathcal{F}_u$. We then have

$$\int_0^t H_s \, dX_s = \mathbb{1}_A \left(X_{v \wedge t} - X_{u \wedge t}\right) .$$

Note that the above vanishes when $t < u$. Hence, if $s \le u \le t$ we have

$$\begin{aligned}
\mathbb{E}\left[\int_0^t H_\tau \, dX_\tau \,\middle|\, \mathcal{F}_s\right] &= \mathbb{E}\left[\, \mathbb{1}_A \left(X_{v \wedge t} - X_u\right) \mid \mathcal{F}_s\right] \\
&= \mathbb{E}\left[\, \mathbb{1}_A \, \mathbb{E}\left[\left(X_{v \wedge t} - X_u\right) \mid \mathcal{F}_u\right] \mid \mathcal{F}_s\right] \\
&= 0 \\
&= \int_0^s H_\tau \, dX_\tau \, .
\end{aligned}$$

When $u < s < t$ we get

$$\begin{aligned}\mathbb{E}\left[\int_0^t H_\tau\, dX_\tau \,\middle|\, \mathcal{F}_s\right] &= \mathbb{E}\left[\mathbb{1}_A\left(X_{v\wedge t} - X_u\right) \mid \mathcal{F}_s\right] \\ &= \mathbb{1}_A\, \mathbb{E}\left[\left(X_{v\wedge t} - X_u\right) \mid \mathcal{F}_s\right] \\ &= \mathbb{1}_A\left(X_{v\wedge s} - X_u\right) \\ &= \int_0^s H_\tau\, dX_\tau\,.\end{aligned}$$

We conclude by a monotone class argument for general predictable processes H (cf Proposition 3.1). □

The quadratic variation of a finite variation process X is denoted by $[X,X]_t$, $t \in \mathbb{R}^+$. The Ito Formula for finite variation processes takes the following form (direct application of Proposition 3.5, Theorems 3.6 and 3.7)

Theorem 3.9 (Ito Formula for Finite Variation Processes).
1) Let X and Y be two finite variation processes. Then we have

$$X_t Y_t = X_0 Y_0 + \int_0^t X_{s-}\, dY_s + \int_0^t Y_{s-}\, dX_s + [X,Y]_t\,. \tag{3.5}$$

2) Let X be a finite variation process. For any function $h \in C^1(\mathbb{R})$ we have

$$h(X_t) = h(X_0) + \int_0^t h'(X_{s-})\, dX_s + \sum_{0\le s\le t}\big(\Delta h(X_s) - h'(X_{s-})\,\Delta X_s\big)\,. \tag{3.6}$$

3) Let $X = (X^1,\dots,X^n)$ be a finite variation process valued in $\mathbb{R}^n$. For any function $h \in C^1(\mathbb{R}^n)$ we have

$$\begin{aligned}h(X_t) = h(X_0) + \sum_{i=1}^n \int_0^t D_i h(X_{s-})\, dX_s^i + &\\ + \sum_{0\le s\le t}\big(\Delta h(X_s) - D_i h(X_{s-})\,\Delta X_s^i\big)\,.&\end{aligned} \tag{3.7}$$

A usefull application of this Ito formula is the one which leads to the notion of *stochastic exponential.*

Theorem 3.10 (Stochastic Exponentials, the Finite Variation Case).
Let X be a finite variation process such that $X_0 = 0$. Then the process

$$Z_t = e^{X_t} \prod_{0\le s\le t}(1+\Delta X_s)\, e^{-\Delta X_s} = e^{X_t^c} \prod_{0\le s\le t}(1+\Delta X_s)\,, \tag{3.8}$$

is a well-defined finite variation process, it is the unique solution of the equation

$$Z_t = 1 + \int_0^t Z_{s-}\, dX_s\,. \tag{3.9}$$

If X is a martingale then Z is a martingale too.

Proof. Let us first check that Z is well-defined. The problem is to check the convergence of the infinite product

$$V_t = \prod_{0\le s\le t} (1+\Delta X_s)\, e^{-\Delta X_s} \,.$$

But we have that the series

$$\widetilde{V}_t = \sum_{0\le s\le t} (\ln(1+\Delta X_s) - \Delta X_s)$$

is absolutely convergent for the series $\sum_{0\le s\le t} \Delta X_s^2$ converges. It defines a finite variation process $\widetilde{V}$. Hence the infinite product V above is convergent, it defines a finite variation process as it is the exponential of $\widetilde{V}$ (Theorem 3.9). Note that

$$\Delta V_t = V_{t-}\left((1+\Delta X_t)\, e^{-\Delta X_t} - 1\right) \,.$$

We put $K_t = e^{X_t}$ and $Z_t = K_t V_t$. By the Ito formula (Theorem 3.9) we have

$$K_t = 1 + \int_0^t K_{s-}\, dX_s + \sum_{0\le s\le t} K_{s-}\left(e^{\Delta X_s} - 1 - \Delta X_s\right)$$

and

$$\Delta K_t = K_{t-}\left(e^{\Delta X_t} - 1\right) \,.$$

This gives

$$\begin{aligned}
Z_t &= 1 + \int_0^t K_{s-}\, dV_s + \int_0^t V_{s-}\, dK_s + [K,V]_t \\
&= 1 + \sum_{0\le s\le t} K_{s-}V_{s-}\left((1+\Delta X_s)\, e^{-\Delta X_s} - 1\right) + \\
&\quad + \int_0^t V_{s-}K_{s-}\, dX_s + \sum_{0\le s\le t} V_{s-}K_{s-}\left(e^{\Delta X_s} - 1 - \Delta X_s\right) + \\
&\quad + \sum_{0\le s\le t} V_{s-}K_{s-}\left((1+\Delta X_s)\, e^{-\Delta X_s} - 1\right)\left(e^{\Delta X_s} - 1\right) \\
&= 1 + \int_0^t Z_{s-}\, dX_s \,.
\end{aligned}$$

This proves that Z is indeed a solution of Equation (3.9).

Let us prove uniqueness. If Z' is another solution of (3.9) then the process $Y = Z - Z'$ is a solution to

$$Y_t = \int_0^t Y_{s-}\, dX_s \,.$$

In particular, Y is a finite variation process. Put $B_t = \int_0^t |dX_s|$ and $K = \sup_{s\le t} |Y_s|$. We have

$$|Y_t| \le \int_0^t |Y_{s-}| \, dB_s \le K \, B_t$$

$$\begin{aligned}|Y_t| &\le \int_0^t |Y_{s-}| \, dB_s \le \int_0^t K B_{s-} \, dB_s \\ &\le K \, \frac{B_t^2}{2} \quad \text{(by the Ito formula)}\end{aligned}$$

and so on ... we get $|Y_t| \le K B_t^n / n!$ and finally $Y_t = 0$.

The martingale property comes from Theorem 3.8. □

We denote by $\mathcal{E}(X)$ the solution of (3.9). In the case where X is a martingale, the process e^X is not in general a martingale. whereas the process $\mathcal{E}(X)$ is a martingale, called the *exponential martingale* associated to X.

On the other hand $\mathcal{E}(X)$ does not satisfy anymore the exponential property. Anyway, we have the following very nice multiplication formula.

Theorem 3.11. *Let X and Y be two finite variation processes. Then*

$$\mathcal{E}(X) \, \mathcal{E}(Y) = \mathcal{E}(X + Y + [X, Y]) \, .$$

Proof. This is just again an application of the Ito formula (Theorem 3.9). If $Z_t = \mathcal{E}(X)_t$ and $W_t = \mathcal{E}(Y)_t$ then

$$\begin{aligned}Z_t W_t &= 1 + \int_0^t Z_{s-} \, dW_s + \int_0^t W_{s-} \, dZ_s + \sum_{0 \le s \le t} \Delta Z_s \, \Delta W_s \\ &= 1 + \int_0^t Z_{s-} W_{s-} \, dY_s + \int_0^t Z_{s-} W_{s-} \, dX_s + \sum_{0 \le s \le t} Z_{s-} W_{s-} \Delta X_s \, \Delta Y_s \\ &= 1 + \int_0^t Z_{s-} W_{s-} \, d(Y_s + X_s + [X, Y]_s) \, .\end{aligned}$$

That is, ZW satisfies the characteristic equation of $\mathcal{E}(X + Y + [X, Y])$. □

3.3 Continuous Semimartingales

We shall now forget the finite variation processes for a while and enter into a completely different family of processes, the continuous martingales, local martingales and semimartingales.

With Theorem 2.26 we have seen that continuous martingales are really a different object compared to finite variation martingales. In Theorem 2.27 we have introduced the notion of quadratic variation for continuous martingales which are locally square integrable. This has led to the definition of the *square bracket* of such a martingale M, that is, the quadratic variation process $[M, M]$

which is the uniquue increasing, adapted, continuous process, vanishing at 0, such that $M^2 - [M, M]$ is a martingale.

By polarization it easy to check the following.

Theorem 3.12. *Let M and N be two continuous martingales, locally square integrable. Then there exists a unique finite variation, adapted, continuous process $[M, N]$ vanishing at 0 and such that $MN - [M, N]$ is a uniformly integrable martingale. This process $[M, N]$ is also given by*

$$[M, N]_t = \lim \sum_{i=0}^{n-1} (M_{t_{i+1}} - M_{t_i})\,(N_{t_{i+1}} - N_{t_i})\,,$$

where the limit is almost sure and taken over a sequence of partitions $0 = t_0 < t_1 < \ldots < t_n = t$ of the interval $[0, t]$ whose diameter tends to 0.

By the Stopping Theorem 2.1 we know that stopped martingales remain martingales. The square bracket behaves in a good way relatively to the action of stopping.

Proposition 3.13. *Let M and N be two continuous, locally square integrable, martingales and let T be a stopping time. Then we have*

$$[M^T, N^T] = [M^T, N] = [M, N]^T\,.$$

Proof. We have, by stopping, that $(M^2)^T - [M, M]^T$ is a martingale. But $(M^2)^T = (M^T)^2$, so by uniqueness in the Theorem 2.27 we have that $[M, M]^T = [M^T, M^T]$.

The general case is easily obtained by polarization. □

This "locality" property for martingales and their bracket is going to be of great importance for a generalization of the notion of martingale which appears to be necessary for the theory of stochastic integration.

An adapted, right-continuous stochastic process X is *local martingale* if there exists a sequence (T_n) of stopping times such that:

i) the sequence (T_n) is increasing and $\lim_n T_n = +\infty$ almost surely,

ii) for every n, the process X^{T_n} is a uniformly integrable martingale.

Of course, a martingale is a local martingale. The converse is not true, but we shall wait next section (and the introduction to stochastic calculus) before exhibiting a counter-example.

Note that, up to changing the sequence of stopping times (T_n) into $(T_n \wedge S_n)$, where $S_n = \inf\{t\,;\ |X_t| > n\}$, we can always assume that the martingales associated to a local martingale are bounded (hence square integrable).

The square bracket extends to continuous local martingales.

Proposition 3.14. *Let M be a continuous local martingale. There exists a unique continuous, adapted, increasing process $[M,M]$, vanishing at 0, such that $M^2-[M,M]$ is a local martingale. This process $[M,M]$ is such that for all $\varepsilon>0$*

$$P\left(\sup_{s\leq t}\left|[M,M]_t-\sum_{i=0}^{n-1}(M_{t_{i+1}}-M_{t_i})^2\right|\geq\varepsilon\right)$$

converges to 0 when the diameter of the partition $\{0=t_0<t_1<\ldots<t_n=t\}$ converges to 0.

Proof. Let (T_n) be a sequence of stopping time such that T_n increases to $+\infty$ and M^{T_n} is a continuous martingale. By Theorem 2.27 there exists an increasing, continuous, adapted process A^n, vanishing at 0, such that $(M^{T_n})^2-A^n$ is a martingale. But $\left((M^{T_{n+1}})^2-A^{n+1}\right)^{T_n}$ is equal to $(M^{T_n})^2-(A^{n+1})^{T_n}$. By uniqueness of the square bracket, this means that $(A^{n+1})^{T_n}=A^n$. As a consequence we can define unambiguously a continuous, adapted, increasing process A such that $A^{T_n}=A^n$ for all n. The process $Y=M^2-A$ is such that Y^{T_n} is a martingale for all n, hence Y is a local martingale. The process A is the square bracket we were looking for. Uniqueness is obvious from the uniqueness property for martingales.

Let us prove the last property. For all stopping time S such that M^S is a bounded martingale we have

$$P\left(\sup_{s\leq t}\left|[M,M]_t-\sum_{i=0}^{n-1}(M_{t_{i+1}}-M_{t_i})^2\right|\geq\varepsilon\right)\leq P(S\leq t)+ \\ +P\left(\sup_{s\leq t}\left|[M^S,M^S]_t-\sum_{i=0}^{n-1}(M^S_{t_{i+1}}-M^S_{t_i})^2\right|\right)\,. \quad (3.10)$$

The last term converges to 0 by Theorem 2.27 and S can be choosen such that $P(S\leq t)$ is arbitrarily small. This gives the result. □

Again, these result can be extended by polarization in order to define the square brackets $[M,N]$ for local martingales. We do not develop here this obvious extension. In particular, the property

$$[M,N]^T=[M^T,N^T]=[M^T,N]$$

remains also true in the context of local martingales.

Finally, the result of Theorem 2.26 extends obviously to continuous local martingales: *a continuous local martingale which is of finite variations is constant.*

We now prove a fundamental inequality concerning square brackets.

Theorem 3.15 (Kunita-Watanabe's inequality). *If M and N are two local martingales, if H and K are two measurable processes, then one has almost surely*

$$\int_0^\infty |H_s|\,|K_s|\;|d[M,N]_s| \le \left(\int_0^\infty H_s^2\,d[M,M]_s\right)^{1/2}\left(\int_0^\infty K_s^2\,d[N,N]_s\right)^{1/2}.$$

Proof. Let $s < t$ be two rational numbers. The fact that for all $\lambda \in \mathbb{Q}$ we have

$$[M+\lambda N, M+\lambda N]_t - [M+\lambda N, M+\lambda N]_s \ge 0$$

gives

$$\left|[M,N]_s^t\right| \le [M,M]_s^{t\;1/2}[N,N]_s^{t\;1/2}$$

(where X_s^t means $X_t - X_s$).

Let $\{0 = t_0 < t_1 < \cdots < t_n < t_{n+1} = +\infty\}$ be a subdivision of $\mathbb{R}^+$. Take $H_0, K_0, H_{t_i}, K_{t_i}$ to be bounded random variables, take $0 \le i \le n$ and define

$$H_t = H_0 \mathbb{1}_{t=0} + \sum_i H_{t_i} \mathbb{1}_{]t_i,t_{i+1}]}(t)$$

and the same definition for K_t. If one takes the previous inequalities for $s = t_i$ and $t = t_{i+1}$, multiply them by $|H_{t_i}K_{t_i}|$, sum them and apply Schwarz inequality one gets

$$\begin{aligned}\left|\int H_sK_s\,d[M,N]_s\right| &\le \sum_i |H_{t_i}K_{t_i}|\;|[M,N]_{t_i}^{t_{i+1}}| \\ &\le \left(\sum_i H_{t_i}^2\,[M,M]_{t_i}^{t_{i+1}}\right)^{1/2}\left(\sum_i K_{t_i}^2\,[N,N]_{t_i}^{t_{i+1}}\right)^{1/2}.\end{aligned}$$

That is the required inequality with the absolute value symbol outside the integrals.

The simple processes H and K as above form an algebra which generates the product σ-algebra on $\mathbb{R}^+ \times \Omega$. A monotone class arguments allows to extend the inequalities (with the absolute value symbol outside the integrals) to general bounded measurable processes H and K.

Let $J.$ be the $\{-1,1\}$-valued process such that $|d\langle M,N\rangle_s| = J_s\;d\langle M,N\rangle_s$. If one applies the previous results to H and JK we get the inequality with the absolute value symbol inside the integral. The inequality is brought to the positive case and one extends to the non-bounded case by truncation and monotone convergence. □

We can now introduce the class of processes which is central in the theory of stochastic integration (at least in this first approach): the continuous semimartingales. A process X is a *continuous semimartingale* if it can be decomposed as $X = M + A$ where M is a continuous local martingale and A is a continuous, adapted, finite variation process.

Because of Theorem 2.26 (and its extension to local martingales) the decomposition of X as $M + A$ is unique.

Quadratic variations also extend to continuous semimartingales.

Theorem 3.16. *A continuous semimartingale $X = M + A$ has a finite quadratic variation and $[X, X] = [M, M]$.*

Proof. This is easily deduced from observing that

$$\left| \sum_i \left(M_{t_{i+1}} - M_{t_i}\right) \left(A_{t_{i+1}} - A_{t_i}\right) \right| \leq \left(\sup_i \left| M_{t_{i+1}} - M_{t_i} \right| \right) V_t(A)$$

where $V_t(A)$ is the total variation of A on the interval $[0, t]$. $\square$

3.4 Stochastic Integral

We are now able to enter into the definition of the stochastic integrals. The first step consists in integrating with respect to continuous square integrable martingales.

Let $\mathcal{P}_e$ denote the space of bounded simple predictable processes on $\mathbb{R}^+$, that is, processes H such that there exists a subdivision $\{0 = t_0 < t_1 < \cdots < t_n < t_{n+1} = +\infty\}$ of $\mathbb{R}^+$, random variables H_{t_i} which are $\mathcal{F}_{t_i}$-measurable and bounded, with

$$H_t = H_0 \mathbb{1}_{t=0} + \sum_i H_{t_i} \mathbb{1}_{]t_i, t_{i+1}]}(t) \,.$$

Note that H_0 is necessarily a constant.

We denote by $\mathcal{M}^c$ the space of continuous square integrable martingales. Let $M \in \mathcal{M}^c$ be fixed. Define the *stochastic integral*

$$\int_0^t H_s \, dM_s = H_0 M_0 + \sum_i H_{t_i} (M_{t_{i+1} \wedge t} - M_{t_i \wedge t}). \tag{3.11}$$

Lemma 3.17. *The process $\left(\int_0^\cdot H_s \, dM_s\right)$ is an element of $\mathcal{M}^c$ and*

$$\mathbb{E}\left[\left(\int_0^\infty H_s \, dM_s \right)^2 \right] = \mathbb{E}\left[\int_0^\infty H_s^2 \, d[M, M]_s \right] .$$

Proof. Put $Y_t = \int_0^t H_s \, dM_s$. From (3.11) it is clear that Y is a continuous process. The martingale property is also immediate. Let us compute $\mathbb{E}[Y_t^2]$. We get, denoting by N the martingale $M^2 - [M, M]$,

$$
\begin{aligned}
\mathbb{E}[Y_t^2] &= \mathbb{E}\left[\left(H_0 M_0 + \sum_i H_{t_i}(M_{t_{i+1}\wedge t} - M_{t_i\wedge t})\right)^2\right] \\
&= H_0^2 \mathbb{E}[M_0^2] + 2\sum_i \mathbb{E}\left[H_0 M_0\, H_{t_i}(M_{t_{i+1}\wedge t} - M_{t_i\wedge t})\right] + \\
&\quad + \sum_{i,j} \mathbb{E}\left[H_{t_i}(M_{t_{i+1}\wedge t} - M_{t_i\wedge t})\, H_{t_j}(M_{t_{j+1}\wedge t} - M_{t_j\wedge t})\right] \\
&= H_0^2 \mathbb{E}[M_0^2] + 2\sum_i \mathbb{E}\left[H_0 M_0\, H_{t_i}\mathbb{E}\left[(M_{t_{i+1}\wedge t} - M_{t_i\wedge t})\,|\,\mathcal{F}_{t_i}\right]\right] + \\
&\quad + 2\sum_{i<j} \mathbb{E}\left[H_{t_i}(M_{t_{i+1}\wedge t} - M_{t_i\wedge t})\, H_{t_j}\mathbb{E}\left[(M_{t_{j+1}\wedge t} - M_{t_j\wedge t})\,|\,\mathcal{F}_{t_j}\right]\right] \\
&\quad + \sum_i \mathbb{E}\left[H_{t_i}^2(M_{t_{i+1}\wedge t} - M_{t_i\wedge t})^2\right] \\
&= H_0^2 \mathbb{E}[M_0^2] + \sum_i \mathbb{E}\left[H_{t_i}^2(M_{t_{i+1}\wedge t}^2 - M_{t_i\wedge t}^2)\right] \\
&= H_0^2 \mathbb{E}[M_0^2] + \sum_i \mathbb{E}\left[H_{t_i}^2([M,M]_{t_{i+1}\wedge t}^2 - [M,M]_{t_i\wedge t}^2)\right] + \\
&\quad + \sum_i \mathbb{E}\left[H_{t_i}^2(N_{t_{i+1}\wedge t} - N_{t_i\wedge t})\right] \\
&= \mathbb{E}\left[\int_0^\infty H_s^2\, d[M,M]_s\right]. \quad \square
\end{aligned}
$$

Let M be a fixed continuous square integrable martingale. Define $L^2(M)$ to be the space of predictable processes H such that

$$\|H\|_{L^2(M)} = \left(\mathbb{E}\left[\int_0^\infty H_s^2\, d[M,M]_s\right]\right)^{1/2} < \infty\,.$$

Theorem 3.18 (Itô's theorem). *The linear mapping $H \mapsto \int_0^\infty H_s\, dM_s$ from $\mathcal{P}_e$ to $\mathcal{M}^c$ extends in a unique way to an isometry from $L^2(M)$ to $\mathcal{M}^c$, still denoted $H \mapsto \int_0^\infty H_s\, dM_s$.*

Proof. The space $L^2(M)$ is the L^2-space of some measure on the predictable σ-algebra. Thus the space $\mathcal{P}_e$ of simple processes is dense in $L^2(M)$. The previous lemma shows that the mapping $H \mapsto \int_0^\infty H_s\, M_s$ is an isometry on $\mathcal{P}_e$, thus it extends uniquely into an isometry on $L^2(M)$. $\square$

We now have an important result due to Kunita-watanabe, which characterizes the stochastic integral $\int H\, dM$ as a process and not only as an operator on $L^2(M)$.

Theorem 3.19. *Let $H \in L^2(M)$. Then for all $N \in \mathcal{M}^c$ we have*

$$\mathbb{E}\left[\int_0^\infty |H_s|\; |d[M,N]_s|\right] < \infty.$$

The stochastic integral $Y_t = \int_0^t H_s \, dM_s$ is the only element of $\mathcal{M}^c$ such that, for all $N \in \mathcal{M}^c$

$$\mathbb{E}\left[L_\infty N_\infty\right] = \mathbb{E}\left[\int_0^\infty H_s \, d[M,N]_s\right] . \tag{3.12}$$

Furthermore we have

$$[L,N]_t = \int_0^t H_s \, d[M,N]_s \, .$$

Proof. The first inequality is a consequence of Kunita-Watanabe's Inequality (Theorem 3.15).

The Equality (3.12) come from the fact that the mapping

$$H \mapsto \mathbb{E}\left[\left(\int_0^\infty H_s \, dM_s\right) N_\infty - \int_0^\infty H_s \, d[M,N]_s\right]$$

is continuous on $L^2(M)$ (again Kunita-Watanabe's inequality) and the fact that this mapping is null on $\mathcal{P}_e$ (immediate). Thus it is null on all $L^2(M)$.

In order to establishing the last identity, define the process $J_t = L_t N_t - \int_0^t H_s \, d[M,N]_s$. Applying the previous equality to N^T we get $\mathbb{E}\left[J_T\right] = 0$. Thus J is a martingale (Theorem 2.3) and $[L,N] - \int_0^\cdot H_s \, d[M,N]_s$ is a finite variation continuous martingale, null at 0, thus it vanishes (Theorem 2.26). This gives the requiered identity.

We just have to prove uniqueness. The second member of (3.12) is a continuous linear form in N known when one is given H and M. The first member is a linear form $N \mapsto \langle M \, , \, N\rangle = \mathbb{E}[M_\infty N_\infty]$ which is a scalar product on $\mathcal{M}^c$. Hence N is completely determined by this identity. □

An easy consequence of the previous theorem is the following.

Theorem 3.20. *Let $H \in L^2(M)$ and put $Y = \int H \, dM$. If K is any bounded predictable process, then*

$$\int_0^t K_s \, dY_s = \int_0^t H_s K_s \, dM_s .$$

Proof. By Theorem 3.19 we have, for all local martingale N

$$[Y,N] = \left[\int_0^\cdot H_s \, dM_s, N\right] = \int_0^\cdot H_s \, d[M,N]_s \, .$$

This gives

$$\begin{aligned} \left[\int_0^\cdot K_s \, dY_s, N\right] &= \int_0^\cdot K_s \, d[Y,N]_s \\ &= \int_0^\cdot K_s H_s \, d[M,N]_s \\ &= \left[\int_0^\cdot K_s H_s \, dM_s, N\right] . \end{aligned}$$

We conclude by the uniqueness property of Theorem 3.19. □

Stochastic integrals also behave naturally with respect to stopping.

Proposition 3.21. *Let $M \in \mathcal{M}^c$ and $H \in L^2(M)$ be fixed. Then for any stopping time T we have*

$$\left(\int_0^\infty H_s \, dM_s\right)^T = \int_0^\infty H_s \mathbb{1}_{[0,T]}(s) \, dM_s = \int_0^\infty H_s \, dM_s^T .$$

Proof. Note that

$$M^T = \int_0^\infty \mathbb{1}_{[0,T]}(s) \, dM_s$$

for

$$\begin{aligned}[M^T, N] &= [M^T, N^T] = [M, N]^T \\ &= \int_0^\cdot \mathbb{1}_{[0,T]}(s) \, d[M,N]_s = \left[\int_0^\cdot \mathbb{1}_{[0,T]}(s) \, dM_s, N\right]\end{aligned}$$

holds true for all $N \in \mathcal{M}^c$ and we apply Theorem 3.19.

Hence, by Theorem 3.20 we have

$$\int_0^\infty H_s \, dM_s^T = \int_0^\infty H_s \mathbb{1}_{[0,T]}(s) \, dM_s$$

and on the other hand we have

$$\left(\int_0^\infty H_s \, dM_s\right)^T = \int_0^\infty \mathbb{1}_{[0,T]}(s) \, d\left(\int_0^\cdot H_s \, dM_s\right) = \int_0^\infty \mathbb{1}_{[0,T]}(s) H_s \, dM_s .$$

We have proved the requiered equality. □

In the following we make use of the following notation, for any stopping time T

$$\int_0^T H_s \, dM_s = \int_0^\infty H_s \mathbb{1}_{[0,T]} H_s \, dM_s .$$

The theorem above allows to extend stochastic integrals to continuous local martingales.

Let M be a continuous local martingale. We denote by $L^2_{\mathrm{loc}}(M)$ the space of predictable processes H such that there exists a sequence (T_n) of stopping times, increasing to $+\infty$ and such that

$$\mathbb{E}\left[\int_0^{T_n} H_s^2 \, d[M,M]_s\right] < \infty .$$

Theorem 3.22. *For any continuous local martingale M and any $K \in L^2_{\mathrm{loc}}(M)$, there exists a unique continuous local martingale, vanishing at 0, denoted $\int_0^\cdot H_s \, dM_s$ such that, for all continuous local martingale N we have*

$$\left[\int_0^\cdot H_s \, dM_s, N\right] = \int_0^\cdot H_s \, d[M,N]_s .$$

Proof. We can choose a sequence of stopping times (T_n) increasing to $+\infty$ such that M^{T_n} is an element of $\mathcal{M}^c$ and H^{T_n} belongs to $L^2(M^{T_n})$. Thus we can define the stochastic integral $X^n = \int_0^\infty H_s^{T_n}\, dM_s^{T_n}$. But the result of Proposition 3.21 shows that $(X^{n+1})^{T_n} = X^n$ on $[0, T_n]$. Hence we can define a process X such that $X^{T_n} = X^n$. It is now easy to check that X is the process announced in the theorem. □

The last step is to integrate with respect to continuous semimartingales.

We denote by $\mathcal{P}_{\mathrm{lb}}$ the space of locally bounded predictable processes, that is, predictable processes H such that there exists a sequence (T_n) of stopping times increasing to $+\infty$ with H^{T_n} being a bounded process. Note that $\mathcal{P}_{\mathrm{lb}}$ is included into $L^2_{\mathrm{loc}}(M)$ for all continuous local martingale M.

For any continuous semimartingale $X = M + A$ and any process $H \in \mathcal{P}_{\mathrm{lb}}$ we define the stochastic integral

$$\int_0^\infty H_s\, dX_s = \int_0^\infty H_s\, dM_s + \int_0^\infty H_s\, dA_s\,,$$

that is, a sum of a stochastic integral with respect to a local martingale and a Stieltjes stochastic integral.

3.5 Itô's Formula

We are now going to prove the most important result of Stochastic Integration Theory: Itô's Change of Variable Formula.

Theorem 3.23 (Itô's Formula). *Let X be a continuous semimartingale. Let f be a function on $\mathbb{R}$ of class C^2. Then $f(X)$ is a continuous semimartingale with*

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\, dX_s + \frac{1}{2}\int_0^t f''(X_s)\, d[X, X]_s\,. \tag{3.13}$$

Proof. Writing

$$\sum_i (X_{t_{i+1}} - X_{t_i})^2 = X_t^2 - X_0^2 - 2\sum_i X_{t_i}(X_{t_{i+1}} - X_{t_i})$$

and passing to the limit, when the subdivision $\{t_i\,;\ i \in \mathbb{N}\}$ gets finer and finer, gives the result for $f(x) = x^2$.

In the same way one gets easily convinced that if Formula (3.13) holds true for a function f, it will also hold true for the function $xf(x)$. Thus the formula is true for polynomial functions.

By stopping, one can reduce the problem to the case when X takes values in a compact set $K \subset \mathbb{R}$. But on K any $C^2(\mathbb{R})$ function f is the limit in $C^2(K)$ of polynomials. One concludes easily. □

We state the multi-valued version of Theorem 3.23 without proof.

Theorem 3.24 (Itô's Formula in $\mathbb{R}^n$). *Let $X = (X^1, \ldots, X^n)$ be a continuous semimartingale valued in $\mathbb{R}^n$. Let f be a functionfrom $\mathbb{R}^n$ to $\mathbb{R}$ of class C^2. Then $f(X)$ is a continuous semimartingale on $\mathbb{R}$ with*

$$f(X_t) = f(X_0) + \sum_{i=1}^n \int_0^t D_i f(X_s)\, dX_s^i + \\ + \frac{1}{2} \sum_{i,j=1}^n \int_0^t D^2_{i,j} f(X_s)\, d[X^i, X^j]_s \,. \quad (3.14)$$

We isolate a very important special case: Itô's Integration by Part Formula.

Theorem 3.25 (Itô's Integration by Part Formula). *Let X and Y be a continuous semimartingales. Then XY is a continuous semimartingale with*

$$X_t\, Y_t = X_0\, Y_0 + \int_0^t X_s\, dY_s + \int_0^t Y_s\, dX_s + [X, Y]_t \,. \quad (3.15)$$

In particular

$$X_t^2 = X_0^2 + 2 \int_0^t X_s\, dX_s + [X, X]_t \,. \quad (3.16)$$

We are now going to meet a class of random variables which will be of great use for us: the stochastic exponentials. They are very famous in classical stochastic calculus for they are the solutions of the simplest stochastic differential equation: the exponential one. They will be of great use in quantum probability for they play the role of coherent vectors in the Fock space.

Theorem 3.26. *Let X be a continuous semimartingale such that $X_0 = 0$. Then there exists one and only one continuous semimartingale Z such that, for all $t \in \mathbb{R}^+$*

$$Z_t = 1 + \int_0^t Z_s\; dX_s \,. \quad (3.17)$$

This solution is given by the formula

$$Z_t = \exp\left(X_t - \frac{1}{2} [X, X]_t \right) \quad (3.18)$$

for all $t \in \mathbb{R}^+$.

Proof. Put $Y_t = X_t - \frac{1}{2}[X, X]_t$ and $Z_t = \exp(Y_t)$. By the Ito Formula we have

$$\begin{aligned} Z_t &= 1 + \int_0^t \exp(Y_t)\, dY_t + \frac{1}{2} \int_0^t \exp(Y_t)\, d[Y, Y]_t \\ &= 1 + \int_0^t \exp(Y_t)\, dX_t - \frac{1}{2} \int_0^t \exp(Y_t)\, d[X, X]_t + \frac{1}{2} \int_0^t \exp(Y_t)\, d[X, X]_t \\ &= 1 + \int_0^t Z_t\, dX_t \,. \end{aligned}$$

This shows that Z is indeed a solution of (3.17). Let us prove uniqueness. If Z' is another solution to (3.18), then the process $W = Z - Z'$ is a solution of

$$W_t = \int_0^t W_s \, dX_s \, .$$

In particular, put

$$Q_t = \exp(-X_t) = 1 - \int_0^t \exp(-X_s) \, dX_s + \frac{1}{2} \int_0^t \exp(-X_s) \, d[X,X]_s \, .$$

By Itô's formula we have

$$\begin{aligned} W_t Q_t &= \int_0^t W_s \, dQ_s + \int_0^t Q_s \, dW_s + [W,Q]_t \\ &= -\int_0^t W_s Q_s \, dX_s + \frac{1}{2} \int_0^t W_s Q_s \, d[X,X]_s + \int_0^t Q_s W_s \, dX_s - \\ &\quad - \int_0^t W_s Q_s \, d[X,X]_s \\ &= -\frac{1}{2} \int_0^t W_s Q_s \, d[X,X]_s \, . \end{aligned}$$

We conclude in the same way as for the proof of uniqueness in Theorem 3.10. □

The solution of Equation (3.17) is called the *stochastic exponential* of the semimartingale X and is denoted by $\mathcal{E}(X)$.

We have seen that, due to the correction term in the Itô formula, the solution of the stochastic exponential equation is not a straight exponential. The next result shows how much the stochastic exponential fails to be a true exponential.

Theorem 3.27. *If X and Y are semimartingales then*

$$\mathcal{E}(X) \, \mathcal{E}(Y) = \mathcal{E}(X + Y + [X,Y]) \, .$$

Proof. Let $U = \mathcal{E}(X)$ and $V = \mathcal{E}(Y)$, then

$$U_t V_t = 1 + \int_0^t U_s V_s \; dY_s + \int_0^t U_s V_s \; dX_s + \int_0^t U_s V_s \; d[X,Y]_s \, .$$

If we put $W_t = U_t V_t$, then we have proved that

$$W_t = 1 + \int_0^t W_s \; d\big(X_s + Y_s + [X,Y]_s\big).$$

This proves our theorem. □

3.6 Brownian Motion and Poisson Processes

We have seen how to define the stochastic integral with respect to any finite variation process, but also with respect to any continuous semimartingale. This means that we are now able to integrate with respect to our two prefered stochastic processes: the Poisson processes and the Brownian motion.

Let us first consider the case of the Poisson process N with intensity $\lambda > 0$. It is clearly a finite variation process, with no continuous part and all jumps equal to 1.

Following the results of Sections 2.5 and 3.2, we see that the process $X_t = N_t - \lambda t$ is a (finite variation) martingale. Clearly we have

$$[N, N]_t = \sum_{0 \leq s \leq t} \Delta N_s = N_t \tag{3.19}$$

and by consequence

$$[X, X]_t = [N, N]_t = N_t = X_t + \lambda t \,. \tag{3.20}$$

The Ito formula (for finite variation processes) takes the following form when specialized to the process N:

$$f(N_t) = f(0) + \int_0^t f'(N_{s-}) \, dN_s + \sum_{0 \leq s \leq t} \big(\Delta f(N_s) - f'(N_{s-}) \big) \,. \tag{3.21}$$

In particular

$$N_t^2 = 2 \int_0^t N_{s-} \, dN_s + N_t \tag{3.22}$$

and

$$X_t^2 = 2 \int_0^t X_{s-} \, dX_s + X_t + \lambda t \,. \tag{3.23}$$

The formula for the stochastic exponential of N takes the following form:

$$\begin{aligned} \mathcal{E}(N)_t &= e^{N_t} \prod_{0 \leq s \leq t} (1 + \Delta N_s) e^{\Delta N_s} \\ &= e^{N_t - \sum_{0 \leq s \leq t} \Delta N_s} \prod_{0 \leq s \leq t} (1 + 1) \\ \mathcal{E}(N)_t &= 2^{N_t} \,. \end{aligned} \tag{3.24}$$

In the same way, we get

$$\begin{aligned} \mathcal{E}(X)_t &= e^{X_t} \prod_{0 \leq s \leq t} (1 + \Delta X_s) e^{\Delta X_s} \\ &= e^{X_t - \sum_{0 \leq s \leq t} \Delta X_s} \prod_{0 \leq s \leq t} (1 + 1) \\ \mathcal{E}(X)_t &= e^{-\lambda t} \, 2^{N_t} \,. \end{aligned} \tag{3.25}$$

We will also be interested in the following. Let $h \in L^2(\mathbb{R};\mathbb{C})$ and put $Y_t = \int_0^t h(s)\,dX_s$. Then

$$\mathcal{E}(Y)_t = \exp\left(\int_0^t h(s)\,dX_s\right) \prod_{\substack{0\leq s\leq t\\ s \text{ jump of } X}} (1+h(s))\exp(-h(s)\Delta X_s)$$

$$\mathcal{E}(Y)_t = \exp\left(-\lambda\int_0^t h(s)\,ds\right) \prod_{\substack{0\leq s\leq t\\ s \text{ jump of } X}} (1+h(s))\,. \tag{3.26}$$

Let us come now to the case of the Brownian motion W. It is a continuous martingale. By Theorem 2.25 we have

$$[W,W]_t = t\,. \tag{3.27}$$

Hence the Ito formula takes the following form:

$$f(W_t) = f(0) + \int_0^t f'(W_s)\,dW_s + \frac{1}{2}\int_0^t f"(W_s)\,ds\,. \tag{3.28}$$

In particular

$$W_t^2 = 2\int_0^t W_s\,dW_s + t\,. \tag{3.29}$$

Let h be a function in $L^2(\mathbb{R}^+;\mathbb{C})$. Consider the martingale

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

Then the stochastic exponential of M is given by

$$\mathcal{E}(M)_t = \exp\left(\int_0^t h(s)\ dW_s - \frac{1}{2}\int_0^t h(s)^2\ ds\right)\,. \tag{3.30}$$

Stochastic exponentials also allow a very nice characterization of the Brownian motion.

Theorem 3.28 (Levy's Theorem). *A process X is a Brownian motion if and only if it is a continuous martinale with $[X,X]_t = t$ for all $t \in \mathbb{R}^+$.*

Proof. We have already proved the "only if" part. Let us prove the converse.

By the Ito formula we have

$$e^{iuX_t} = 1 + iu\int_0^t e^{iuX_s}\ dX_s - \frac{u^2}{2}\int_0^t e^{iuX_s}\ ds$$

for $[X,X]_s = s$. In particular

$$e^{iu(X_t-X_s)} = 1 + iu \int_s^t e^{i(X_\tau - X_s)}\, dX_\tau - \frac{u^2}{2} \int_s^t e^{iu(X_\tau - X_s)}\, d\tau$$

and

$$\mathbb{E}\left[e^{iu(X_t-X_s)}\right] \mathcal{F}_s = 1 - \frac{u^2}{2} \int_s^t \mathbb{E}\left[e^{iu(X_\tau-X_s)}\right] \mathcal{F}_s\, d\tau\,.$$

The above equation is easily solved and we obtain

$$\mathbb{E}\left[e^{iu(X_t-X_s)}\right] \mathcal{F}_s = e^{-(t-s)u^2/2}\,.$$

That is, the conditional law of $X_t - X_s$, given $\mathcal{F}_s$, is gaussian $\mathcal{N}(0, \sqrt{t-s})$. This gives the result. □

4

STOCHASTIC DIFFERENTIAL EQUATIONS

4.1 Markov Processes

We first start with a basic introduction to Markov processes, their semigroup and their generator.

Intuitively speaking, a process X, with state space $(E, \mathcal{E})$ is a *Markov process* if, to make a prediction at time s on what is going to happen in the future, it is useless to know anything more about the whole past up to time s than the present state X_s. The minimal past of X at time s is the σ-algebra $\mathcal{F}_s = \sigma\{X_u; u \leq s\}$. Let us think of the conditionnal probability

$$P[X_t \in A \,|\, \mathcal{F}_s]$$

where $A \in \mathcal{E}$, $s < t$. If X is Markov in the intuitive sense described above, this should be a function of X_s, that is, $g(X_s)$ where g is a $\mathcal{E}$-measurable, $[0, 1]$-valued function. As a function of A the above ought to be a probability measure, clearly. Thus the above is better written

$$g_{s,t}(X_s, A).$$

This motivates the following definition.

Let $(E, \mathcal{E})$ be a measurable space. A *transition probability* on E is a map Π from $E \times \mathcal{E}$ to $[0, 1]$ such that

i) for all $x \in E$, the map $A \longmapsto \Pi(x, A)$ is a probability measure on $\mathcal{E}$,

ii) for all $A \in \mathcal{E}$, the map $x \longmapsto \Pi(x, A)$ is $\mathcal{E}$-measurable.

Transition probabilities act on bounded functions by

$$\Pi f(x) = \int_E f(y)\, \Pi(x, dy)$$

for all $f \in \mathcal{L}^\infty(E, \mathcal{E})$.

Transition probabilities act on probability measures by

$$\mu\Pi(A) = \int_E \Pi(x, A)\, d\mu(x)\,.$$

for all probability measure μ on E.

Transition probabilities can be composed by

$$\Pi_1\Pi_2(x, A) = \int_E \Pi_2(y, A)\, \Pi_1(x, dy)\,,$$

defining this way a new transition probability.

Transition probabilities describe how an initial position x for X_0 is distributed later on for X_T, where $[0, T]$ is a fixed interval of time. In order to consider Markov processes we have to make this definition time-dependent.

A *transition function* on $(E, \mathcal{E})$ is a familly $P_{s,t}$, $0 \leq s < t$ of transition probabilities on $(E, \mathcal{E})$ such that for every $s < t < u$ we have

$$P_{s,u}(x, A) = \int_E P_{t,u}(y, A)\, P_{s,t}(x, dy)$$

for all $x \in E$, $A \in \mathcal{E}$, that is, $P_{s,t}P_{t,u} = P_{s,u}$. Think of $P_{s,t}$ as the transition probability describing what happens to the Markov process in between time s and time t. The transition function is *homogeneous* if $P_{s,t}$ depends on $t - s$ only. In that case we write P_t for $P_{0,t}$ and the above relation reduces to

$$P_{t+s}(x, A) = \int_E P_t(y, A)\, P_s(x, dy)\,,$$

that is, $P_sP_t = P_{t+s}$. In other words, (P_t) is a semigroup (cf Chapter **??**).

A *Markov process* on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$, with transition function P, is an adapted process X such that, for any $f \in \mathcal{L}^\infty(E, \mathcal{E})$, any $s < t$

$$\mathbb{E}\left[f(X_t) \mid \mathcal{F}_s\right] = P_{s,t}f(X_s)\,.$$

In the homogeneous case this gives more simply

$$\mathbb{E}\left[f(X_t) \mid \mathcal{F}_s\right] = P_{t-s}f(X_s)\,.$$

The measure $\nu = X_0(P)$, that is the law of X_0, plays an important role and is called the *initial measure*.Initial measure

We have a first existence theorem for Markov processes.

Theorem 4.1. *Given a transition function P on $(E, \mathcal{E})$ and a probability measure ν on $(E, \mathcal{E})$, there exists a unique probability measure* Prob_ν *on $(E^{\mathbb{R}^+}, \mathcal{E}^{\mathbb{R}^+})$ such that the coordinate mapping X (given by $X_t(\omega) = \omega(t)$) is a Markov process with respect to its natural filtration, with transition function P and initial measure ν.*

Proof. Consider the measures

$$P_\nu^{t_1,\ldots,t_n}(A_0 \times \ldots \times A_n) = \\ = \int_{A_n} \cdots \int_{A_0} P_{t_{n-1},t_n}(x_{n-1}, dx_n) \ldots P_{0,t_1}(x_0, dx_1)\, \nu(dx_0)\,.$$

They constitute a consistent familly of probability measure and then we apply Kolmogorov's Theorem 1.1. □

Of course, the discussion we had in Section 1.3 is still valid in this context. This means that the above existence theorem is not enough to work decently with a Markov process. We will need to know some regularity of the paths of the processes we are interested in. We do not discuss this point for the moment.

In the sequel we concentrate only on the homogeneous case. Note that, in that context we have

$$P_\nu\left[X_0 \in A_0, X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\right] = \\ = \int_{A_n} \cdots \int_{A_0} P_{t_n - t_{n-1}}(x_{n-1}, dx_n) \ldots P_{t_1}(x_0, dx_1)\, \nu(dx_0)\,.$$

Also note that for all $x \in E$, if δ_x denotes the Dirac measure on x, then we put P_x for P_{δ_x} and we have

$$P_x[X_t \in A] = P_t(x, A)\,.$$

This is all for the general theory on Markov processes and their transition functions. We shall concentrate now on a particular class of homogeneous Markov processes which is very workable, for we ask a little more regularity.

A *Feller semigroup* on $C_0(E)$ is a familly T_t, $t \geq 0$, of positive linear operators on $C_0(E)$ such that:

i) $T_0 = I$ and $\|T_t\| \leq 1$, for all t,

ii) $T_{t+s} = T_t \circ T_s$ for all $s, t \in \mathbb{R}^+$,

iii) $\lim_{t \to 0} \|T_t f - f\| = 0$ for all $f \in C_0(E)$.

Note that (T_t) is then a strongly continuous semigroup on $C_0(E)$ (cf Chapter **??**).

Proposition 4.2. *To each Feller semigroup T on E is associated a unique homogeneous transition function P on $(E, \mathcal{E})$ such that*

$$T_t f(x) = P_t f(x)$$

for all $f \in C_0(E)$, all $x \in E$.

Proof. The mapping $f \mapsto T_t f(x)$ is positive linear and continuous on $C_0(E)$ and thus, by Riesz's theorem, there exists a measure $P_t(x,\cdot)$ on $\mathcal{E}$ such that

$$T_t f(x) = \int_E f(y)\, P_t(x,dy)\,.$$

The mapping $x \mapsto \int_E f(y)\, P_t(x,dy)$ is C_0 hence Borel. In particular $x \mapsto P_t(x,A)$ is measurable for all $A \in \mathcal{E}$. The rest of the proof is easy and is left to the reader. □

A transition function P which is associated to a Feller semigroup T is called a *Feller transition function.*

The following result gives a weaker characterization of Feller transition semigroups.

Proposition 4.3. *A (homogeneous) transition function P is Feller if and only if*

i) $P_t C_0(E) \subset C_0(E)$ for all t,

ii) For all $f \in C_0(E)$, all $x \in E$, we have $\lim_{t\to 0} P_t f(x) = f(x)$.

Proof. Of course, only one direction remains to be proved.

As $P_t f$ belong to $C_0(E)$ then, by hypothesis, $\lim_{s\to 0} P_{t+s} f(x) = P_t f(x)$. The function $(t,x) \mapsto P_t f(x)$ is thus right-continuous in t and therefore measurable on $\mathbb{R}^+ \times E$. The function

$$x \mapsto U_p f(x) = \int_0^\infty e^{-pt} P_f f(x)\, dt,$$

for $p > 0$, is measurable and by ii)

$$\lim_{p\to+\infty} p\, U_p f(x) = f(x).$$

One checks easily that $U_p f$ belongs to $C_0(E)$. Note that one also gets easily

$$U_p f - U_q f = (q-p) U_p U_q f = (q-p) U_q U_p f.$$

As a result the image $\mathcal{D} = U_p(C_0(E))$ does not depend on $p > 0$. Finally observe that $\|p\, U_p f\| \leq \|f\|$.

Let us show that $\mathcal{D}$ is dense in $C_0(E)$. Indeed if μ is a bounded measure vanishing on $\mathcal{D}$ then, for any $f \in C_0(E)$

$$\int f\, d\mu = \lim_{p\to+\infty} \int p U_p f\, d\mu = 0$$

and thus $\mu = 0$. Finally, we have

$$P_t U_p f(x) = e^{pt} \int_t^{+\infty} e^{-ps} P_s f(x)\, ds$$

hence

$$\|P_t U_p f - U_p f\| \leq (e^{pt} - 1)\|U_p f\| + t\|f\|.$$

It follows that $\lim_{t\to 0} \|P_t f - f\| = 0$ for $f \in \mathcal{D}$. One concludes easily. □

A Markov process whose transition function is Feller is called a *Feller process*.

The following theorem shows that Feller processes have a sufficient regularity in order to apply the regularization procedure developed in Section 1.3. We do not prove this result (see [**?**], Theorem 2.7).

Theorem 4.4. *Every Feller process admits a cadlad modification.*

We now refer to Chapter **??** and in particular to the results of Section **??**.

Let X be a Feller process and P be its semigroup. In particular P admit a generator A, it is called the *infinitesimal generator* of X.

Note that in particular we have

$$\mathbb{E}\left[f(X_{t+h}) - f(X_t) \mid \mathcal{F}_t\right] = hAf(X_t) + o(h)\,.$$

In other words, the operator A describes the infinitesimal motion of X.

We have a property which is very usefull for the study of Markov processes in general.

Proposition 4.5. *If f belongs to* $\operatorname{Dom} A$ *then the process*

$$M_t^f = f(X_t) - f(X_0) - \int_0^t Af(X_s)\, ds$$

is a martingale. In particular, if $Af = 0$ then $f(X_t)$ is a martingale.

Proof. Since f and Af are bounded function, the random variable M_t^f is integrable and

$$\mathbb{E}\left[M_t^f \mid \mathcal{F}_s\right] = M_s^f + \mathbb{E}\left[f(X_t) - f(X_s) - \int_s^t Af(X_u)\, du \mid \mathcal{F}_s\right].$$

By the Markov property, the last term is equal to

$$\mathbb{E}_{X_s}\left[f(X_{t-s}) - f(X_0) - \int_0^{t-s} Af(X_u)\, du\right].$$

But we have, for any $y \in E$

$$\mathbb{E}_y\left[f(X_{t-s}) - f(X_0) - \int_0^{t-s} Af(X_u)\, du\right] = \\ = P_{t-s}f(y) - f(y) - \int_0^{t-s} P_u Af(y)\, du = 0\,.$$

This gives the result. □

The following theorem will never be used here in its full generality but we give it for information and for comparison with the Lindblad generators that we will obtain later on.

Theorem 4.6. *If $P.$ is a Feller semigroup on $\mathbb{R}^d$ and $C_K^\infty \subset \mathrm{Dom}\, A$ then*
i) $C_K^2 \subset \mathrm{Dom}\, A$,
ii) for every relatively compact open set U, there exist functions a_{ij}, b_i, c on U and a kernel N such that for $f \in C_K^2$ and $x \in U$

$$\begin{aligned} Af(x) = c(x)f(x) &+ \sum_i b_i(x)\frac{df}{dx_i}(x) + \sum_{i,j} a_{ij}(x)\frac{d^2 f}{dx_i \partial x_j}(x) \\ &+ \int_{\mathbb{R}^d \setminus \{x\}} \left[f(y) - f(x) - \mathbb{1}_U(y)\sum_i (y_i - x_i)\frac{df}{dx_i}(x)\right] N(x, dy) \end{aligned}$$

where $N(x,\cdot)$ is a Radon measure on $\mathbb{R}^d \setminus \{x\}$, the matrix $a(x) = \|a_{ij}(x)\|$ is positive, c is negative.

4.2 Brownian motion and Poisson processes

Our two prefered stochastic processes are basic examples of Markov processes. Let us start with the Brownian motion.

Let $C_0^2(\mathbb{R})$ be the space of C^2 functions f on $\mathbb{R}$ such that f, f' and f'' tend to 0 at $\pm\infty$.

Theorem 4.7. *The Brownian motion is a Feller process, with semigroup*

$$P_t f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-y^2/2} f(x + y\sqrt{t})\, dy\,.$$

Its generator is $\frac{1}{2}\Delta$, with domain C_0^2.

Proof. To prove the Markov character of the Brownian motion let us first establish a usefull lemma.

Lemma 4.8. *Let Y and Z be two random variables and $\mathcal{G}$ be a σ-algebra such that Y is independent of $\mathcal{G}$ and Z is $\mathcal{G}$-measurable. Then for every bounded measurable function f on $\mathbb{R}^2$ we have*

$$\mathbb{E}[f(Y,Z)\,|\,\mathcal{G}] = g(Z)$$

where g is given by

$$g(z) = \mathbb{E}[F(Y,z)]\,.$$

Proof (of the lemma). The result is immediate when f is of the form $f(y,z) = h(y)k(z)$. One concludes easily by an approximation argument. □

We are now back to the proof of Theorem 4.7. Let W be a Brownian motion on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$. For any bounded measurable function f on $\mathbb{R}$ and all $s \leq t$ we have

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

where $g(x) = \mathbb{E}[f(W_{t+s} - W_s + x)]$ (by Lemma 4.8). More precisely we have

$$\begin{aligned} g(x) &= \mathbb{E}[[f(W_t + x)] \\ &= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{+\infty} e^{-y^2/2t}\, f(y+x)\, dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-y^2/2}\, f(y\sqrt{t}+x)\, dy\,. \end{aligned}$$

This proves the Markovian character of W and the explicit form of the semigroup. Checking that this semigroup is Feller is easy and left to the reader.

Let us compute the generator of W with its exact domain. If $f \in C_0^2$ then by Taylor's formula we have

$$P_t f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-y^2/2} \left[f(x) + y\sqrt{t} f'(x) + \frac{y^2 t}{2} f''(\theta) \right] dy$$

for some $\theta \in [x, x + y\sqrt{t}]$. This gives

$$P_t f(x) = f(x) + \frac{1}{\sqrt{2}} f''(x) + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-y^2/2}\, \frac{y^2 t}{2} (f''(\theta) - f''(x))\, dy\,.$$

Let us denote by $R(t,x)$ the last term of the right hand side. Let M be a finite bound for f'' (it exists by hypothesis). Let $R > 0$ be fixed, we have

$$\begin{aligned} |R(t,x)| &\leq Ct \int_{|y|\leq R} y^2 e^{-y^2/2}\, |f''(\theta) - f''(x)|\; dy + CMt \int_{|y|>R} y^2 e^{-y^2/2}\, dy \\ &\leq Ct \sup_{|u-v|\leq Rt} |f''(u) - f''(v)| \int_{|y|\leq R} y^2 e^{-y^2/2}\, dy + CMt \int_{|y|>R} y^2 e^{-y^2/2}\, dy\,. \end{aligned}$$

The first term of the right hand side tends to 0 with t by the uniform continuity of f'' on compact sets. The second term tends to 0 as R tends to $+\infty$. Hence $|R(t,x)|$ can be made arbitrarily small. We have proved that for all $f \in C_0^2$ we have

$$\lim_{t\to 0} \frac{P_t f(x) - f(x)}{t} = \frac{1}{2} \Delta f(x)\,.$$

This proves that the generator A is the announced operator, with domain containing C_0^2. We need to prove that this domain coincides with C_0^2. By Proposition **??** we know that $\mathrm{Dom}\, A = R_p(A)(C_0(\mathbb{R}))$ where

$$R_p(A) = \int_0^{\infty} e^{-pt}\, P_t f(x)\, dt\,.$$

A direct computation shows that

$$R_p(A)f(x) = \int_{-\infty}^{+\infty} f(y) \, \frac{1}{\sqrt{2p}} \, e^{-\sqrt{2p}\,|x-y|} \, dy \,.$$

It is now an exercise to check that for all $f \in C_0(\mathbb{R})$ we have $f \in C_0^2(\mathbb{R})$ and $pR_p(A)f - f = \frac{1}{2}(R_p(A)f)''$. □

Here is for the Brownian motion. Let us consider the Poisson processes.

Theorem 4.9. *Let N be a Poisson process with intensity $\lambda > 0$. Then N is a Feller process with semigroup*

$$P_t f(x) = \sum_{n=0}^{\infty} f(n+x) \, \frac{(\lambda t)^n}{n!} \, e^{-\lambda t}$$

and generator

$$Af(x) = \lambda(f(x+1) - f(x))$$

defined on all $C_0(\mathbb{R})$.

Proof. The proof is very similar to the one of the previous theorem. We have, using again Lemma 4.8

$$\mathbb{E}[f(N_{t+s}) \,|\, \mathcal{F}_s] = \mathbb{E}[f(N_{t+s} - N_t + N_t) \,|\, \mathcal{F}_s] = g(N_s)$$

where

$$g(x) = \mathbb{E}[f(N_{t+s} - N_t + x)] = \mathbb{E}[f(N_t + x)] \,.$$

A direct computation with the Poisson law of N_t gives the explicit form of the semigroup as above.

Deriving with respect to t at time $t = 0$ gives trivially the generator. □

4.3 Stochastic Differential Equations

Let us first recall well-known results concerning ordinary differential equations, they are all very classical and easy to prove (we leave the proofs to the reader).

Theorem 4.10. *Let f be a Lipschitz function on $\mathbb{R}$ and let $x \in \mathbb{R}$ be fixed. Consider the equation*

$$X_t^x = x + \int_0^t f(X_s^x) \, ds \,, \quad t \in \mathbb{R}^+. \tag{4.1}$$

Then Equation (4.1) admits a unique solution (X_t^x). The family of operators P_t on $C_0(\mathbb{R})$ defined by

$$P_t h(x) = h(X_t^x)$$

forms a strongly continuous semigroup, with generator

$$A = f(x)\frac{d}{dx}$$

defined on $C^1(\mathbb{R})$*. This semigroup is invertible, with inverse* P_{-t} *being given by*

$$P_{-t}h(x) = h(Y_t^x)$$

where Y^x *is the unique solution of*

$$Y_t^x = x - \int_0^t f(X_s^x)\, ds\,, \quad t \in \mathbb{R}^+.$$

With stochastic differential equations we are considering perturbations of equations of the form (4.1) by addition of a supplementary term, a noise. That is, the general form of the equation becomes

$$X_t^x = x + \int_0^t f(X_s^x)\, ds + \int_0^t g(X_{s-}^x)\, dY_s\,, \tag{4.2}$$

where Y is a fixed semimartingale. Of course, more general equations may be considered. For example, the coefficients inside the integrals may depend on the whole trajectory: $f(X_u; u < s)$, $g(X_u; u < s)$; they may also depend on the time itself: $f(s, X_u; u < s)$, $g(s, X_u; u < s)$ etc. We do not aim to develop a serious theory of stochastic differential equations here, but just consider few examples which illustrate the connection with the Markov property, in order to make a parallel with the quantum version of these equations. The two main examples we consider here are equations of the form (4.2) where Y is a Brownian motion or a Poisson process.

Theorem 4.11. *Let* Y *be a Brownian motion or a compensated Poisson process. If* b *and* σ *are bounded Lipschitz functions on* $\mathbb{R}$ *then the equation*

$$X_t = x + \int_0^t \sigma(X_s)\, dY_s + \int_0^t b(X_s)\, ds \tag{4.3}$$

in the Brownian case,

$$X_t = x + \int_0^t \sigma(X_{s-})\, dY_s + \int_0^t b(X_{s-})\, ds \tag{4.4}$$

in the Poisson case, admits a unique solution.

Proof. We proceed, for proving the existence, by a usual Picard iteration method. Let us describe it in the Brownian case. Let $X_t^0 = x$, for all t. For all $n \geq 1$ put

$$X_t^{n+1} = x + \int_0^t \sigma(X_s^n)\, dY_s + \int_0^t b(X_s^n)\, ds.$$

Then

$$\left\|X_t^{n+1} - X_t^n\right\|^2 \leq 2\int_0^t \left\|\sigma(X_s^n) - \sigma(X_s^{n-1})\right\|^2 ds + 2t\int_0^t \left\|b(X_s^n) - b(X_s^{n-1}\right\|^2 ds. \tag{4.5}$$

By the Lipschitz assumptions the above is bounded by

$$C\int_0^t \left\|X_s^n - X_s^{n-1}\right\|^2 ds.$$

Iterating we get

$$\begin{aligned}\left\|X_t^{n+1} - X_t^n\right\|^2 &\leq C^n \int_{0<s_1<\ldots<s_n<t} \left\|X_s^1 - X_s^0\right\|^2 ds_1 \,\ldots\, ds_n \\ &\leq KC^n \frac{t^n}{n!}.\end{aligned}$$

The series $\sum_n X_t^{n+1} - X_t^n$ is thus normaly convergent to a limit X_t. The arguments are afterwards rather standard. By Doob's maximal inequality, one proves the uniformly almost sure convergence to X_t. One then prove, by an easy approximation argument, that this X is a solution of the stochastic differential equation.

The uniqueness is obtained in the same way as for the exponential equation: using the isometry formula and Gronwall's lemma, the difference between two solutions can only be 0.

The main ingredient we used above is the fact that, for the Brownian motion, we have $[Y,Y]_t = t$. Fact which gives the norm estimate (4.5). In the Poisson case we have $[Y,Y]_t = \lambda t + Y_t$, hence

$$\mathbb{E}\left[\int_0^t H_s\, dY_s\right] = \mathbb{E}\left[\int_0^t H_s^2\,(dY_s + \lambda ds)\right] = \lambda\mathbb{E}\left[\int_0^t H_s^2\, ds\right].$$

The arguments then go on in the same way. □

4.4 Markov Property

Theorem 4.12. *The solution of the equation* (4.3) *or* (4.4) *is a homogeneous Markov process.*

Proof. Let us prove the Brownian case. Let us denote, for a moment $X_t^{x,s}$, the unique solution of the equation

$$X_t = x + \int_s^t \sigma(X_u)\, dY_u + \int_s^t b(X_u)\, du, \quad t \geq s.$$

As
$$X_t = X_s + \int_s^t \sigma(X_u)\,dY_u + \int_s^t b(X_u)\,du$$
we have
$$X_t = X_t^{X_s,s}.$$
In other words, if we define
$$F(x,s,t,\omega) = X_t^{x,s}(\omega)$$
we have
$$X_t(\omega) = F(X_s,s,t,\omega)$$
for all $t \geq s$. Note that $\omega \mapsto F(x,s,t,\omega)$ is independent of $\mathcal{F}_s$.

We aim to prove that, almost surely
$$\mathbb{E}_x\left[f(X_{t+h}) \mid \mathcal{F}_t\right](\omega) = \mathbb{E}_{X_t(\omega)}\left[f(X_h)\right](\omega)$$
which we may rewrite as
$$\mathbb{E}_x\left[f(F(X_t,t,t+h,\cdot)) \mid \mathcal{F}_t\right](\omega) = \mathbb{E}\left[f(F(x,0,h,\cdot))\right]_{x=X_t(\omega)}.$$
Put $g(x,\omega) = f \circ F(x,t,t+h,\omega)$ and approximate g pointwise by functions of the form
$$\sum_k \phi_k(x)\psi_k(\omega).$$
We get
$$\begin{aligned}\mathbb{E}\left[g(X_t,\cdot) \mid \mathcal{F}_t\right] &= \mathbb{E}\left[\lim \sum_k \phi_k(X_t)\psi_k(\cdot) \mid \mathcal{F}_t\right]\\ &= \lim \sum_k \phi_k(X_t)\mathbb{E}\left[\psi_k(\cdot) \mid \mathcal{F}_t\right]\\ &= \lim \sum_k \mathbb{E}\left[\phi_k(y)\psi_k(\cdot) \mid \mathcal{F}_t\right]_{y=X_t}\\ &= \mathbb{E}\left[g(y,\cdot) \mid \mathcal{F}_t\right]_{y=X_t} = \mathbb{E}\left[g(y,\cdot)\right]_{y=X_t}.\end{aligned}$$
But one easily checks that
$$\begin{aligned}\mathbb{E}\left[f(F(X_t,t,t+h,\cdot)) \mid \mathcal{F}_t\right] &= \mathbb{E}\left[f(F(y,t,t+h,\cdot))\right]_{y=X_t}\\ &= \mathbb{E}\left[f(F(y,0,h,\cdot))\right]_{y=X_t}.\end{aligned}$$
Which is the requiered identity.

These arguments are clearly working in the same way for the Poisson case.
□

Theorem 4.13. *In the case where Y is a Brownian motion, the infinitesimal generator of the solution of* (4.3) *is*

$$A = b(x)\frac{d}{dx} + \frac{1}{2}\sigma(x)^2\frac{d^2}{dx^2}\,.$$

In the case where Y is a compensated Poisson process with parameter λ, the infinitesimal generator of the solution of (4.4) *is*

$$Af(x) = (b(x) - \lambda\sigma(x))\frac{d}{dx}f(x) + \lambda(f(x + \sigma(x)) - f(x))\,.$$

Proof. Let us first consider the Brownian case. By the Ito formula we have

$$\begin{aligned} f(X_t) &= f(X_0) + \int_0^t f'(X_s)\,dX_s + \frac{1}{2}\int_0^t f''(X_s)\,d[X,X]_s \\ &= f(x) + \int_0^t f'(X_s)\sigma(X_s)\,dY_s + \int_0^t f'(X_s)b(X_s)\,ds + \\ &\quad + \frac{1}{2}\int_0^t f''(X_s)\sigma(x)^2\,ds\,. \end{aligned}$$

In particular

$$\mathbb{E}\left[f(X_t)\right] = f(x) + \int_0^t \mathbb{E}\left[Af(X_s)\right]\,ds.$$

This gives the result, recording that

$$Af(x) = \frac{d}{dx}\mathbb{E}[f(X_t)]_{|t=0}\,.$$

In the Poisson case, the Ito formula gives

$$\begin{aligned} f(X_t) &= f(X_0) + \int_0^t f'(X_{s-})\,dX_s + \sum_{0\le s\le t} \Delta f(X_s) - f'(X_{s-})\,\Delta X_s \\ &= f(X_0) + \int_0^t f'(X_{s-})\,\sigma(X_{s-})\,dY_s + \int_0^t f'(X_{s-})\,b(X_{s-})\,ds + \\ &\quad + \sum_{0\le s\le t}\left(f(X_{s-} + \sigma(X_{s-})) - f(X_{s-}) - f'(X_{s-})\sigma(X_{s-})\right)\,. \end{aligned}$$

For a general continuous function k, let us compute a quantity like

$$\frac{d}{dt}\mathbb{E}\left[\sum_{0\le s\le t} k(X_{s-})\right]_{|t=0}\,.$$

Let $T_1, \ldots, T_n, \ldots$ be the successive jumping times of X (that is, those of Y). We have

$$\mathbb{E}\left[\sum_{0\le s\le t} k(X_{s-})\right] =$$
$$= \sum_{n\in\mathbb{N}} \mathbb{E}\left[\mathbb{1}_{T_1+\ldots+T_n\le t<T_1+\ldots+T_{n+1}}(k(X_{T_1-}) + \ldots k(X_{(T_1+\ldots T_n)-}))\right] .$$

The random variable $k(X_{T_1-}) + \ldots + k(X_{(T_1+\ldots T_n)-})$ is a certain function j of x and $T_1, T_2, \ldots, T_n$ only. This gives

$$\mathbb{E}\left[\sum_{0\le s\le t} k(X_{s-})\right] =$$
$$= \int_{0\le t_1+\ldots t_n\le t\le t_1+\ldots t_{n+1}} j(x, t_1, \ldots, t_n)\lambda^{n+1}e^{\lambda(t_1+\ldots+t_{n+1})}\, dt_1 \,\ldots\, dt_{n+1} .$$

Changing the variables into $u_1 = t_1$, $u_2 = t_1 + t_2$, $\ldots$, $u_{n+1} = t_1 + \ldots + t_{n+1}$, we get

$$\int_{0\le u_1\le\ldots\le u_n\le t\le u_{n+1}} \widetilde{j}(x, u_1, \ldots, u_n)\lambda^{n+1}e^{\lambda u_{n+1}}\, du_1 \,\ldots\, du_{n+1} .$$

When deriving with respect to t at $t = 0$, only the term with $n = 1$ survives and the derivative in 0 is equal to

$$\int_0^\infty \widetilde{j}(x, 0)\lambda^2 e^{\lambda u_2}\, du_2 = \lambda k(x) .$$

The conclusion is now easy. □

4.5 Dynamical Systems

In this section we connect the stochastic differential equations and their Markov generator to the notion of dynamical system. This section is to be put in perspective with Section **??**. It also prepares to what we are going to develop in the quantum context.

Let us consider the Brownian motion W on its canonical space $(\Omega, \mathcal{F}, P)$ where $\Omega = C_0(\mathbb{R}^+)$. We define, for all $s \in \mathbb{R}^+$, on $C_0(\mathbb{R}^+)$, the shift θ_s as an application from Ω to Ω by

$$\theta_s(\omega)(t) = \omega(t + s) - \omega(s) .$$

From this, we define the shift operators Θ_s as follows. For any random variable $F \in L^2(\omega, \mathcal{F}, P)$ we put

$$\Theta_s(F)(\omega) = F(\theta_s(\omega)) .$$

In particular we have $\Theta_s(W_t) = W_{t+s} - W_s$.

As the process $Y_t = W_{t+s} - W_s$, $t \in \mathbb{R}^+$, is again a Brownian motion, this implies that the mapping θ_s preserves the measure P. As a consequence Θ_s is an isometry of $L^2(\Omega, \mathcal{F}, P)$.

Lemma 4.14. *If H is a predictable process, then, for all fixed $s \in \mathbb{R}^+$, the process $K_t = \Theta_s(H_{t-s})$, $t \geq s$ is also predictable.*

Proof. The process K as a mapping from $\Omega \times \mathbb{R}^+$ to $\mathbb{R}$ is the composition of H with the mapping $\phi(\omega, t) = (\theta_s(\omega), t - s)$ from $\Omega \times [s, +\infty[$ to $\Omega \times \mathbb{R}^+$. We just need to check that ϕ is measurable for the predictable σ-algebra $\mathcal{P}$.

If $A \times]u, v]$, with $u < v$ and $A \in \mathcal{F}_u$ is a basic predictable set, then

$$\phi^{-1}(A \times]u, v]) = \theta_s^{-1}(A) \times]u + s, v + s] .$$

We need to check that $\theta_s^{-1}(\mathcal{F}_u) \subset \mathcal{F}_{u+s}$. The σ-algebra $\mathcal{F}_u$ is generated by events of the form $(W_t \in [a, b])$, $fort \leq u$. The set $\theta_s^{-1}(W_t \in [a, b])$ is equal to $(W_{t+s} - W_s \in [a, b])$, hence it belongs to $\mathcal{F}_{u+s}$. □

Lemma 4.15. *Let H be a predictable process such that $\int_0^{t+s} \mathbb{E}[H_u^2]\, du < \infty$. Then we have*

$$\Theta_s \left(\int_0^t H_u \, dW_u \right) = \int_s^{t+s} \Theta_s(H_{u-s}) \, dW_u . \qquad (4.6)$$

Proof. If H is an elementary predictable process then the identity (4.6) is obvious (from the fact that $\Theta_s(FG) = \Theta_s(F)\,\Theta_s(G)$). As a general stochastic integral $\int_0^t H - s\, dW_s$ is obtained as a limit in the norm

$$\| \int_0^t H_s \, dW_s \|^2 = \int_0^t \|H_s\|^2 \, ds ,$$

and as Θ_s is an isometry, it is clear that Equation (4.6) holds true for any stochastic integral. □

Theorem 4.16. *Let f and g be two bounded Lipschitz functions. Denote by X^x the unique process solution of the stochastic differential equation*

$$X_t^x = x + \int_0^t f(X_u^x)\, du + \int_0^t g(X_u^x)\, dW_u .$$

Then, for all $s, t \in \mathbb{R}^+$, we have

$$\Theta_s \left(X_t^{X_s^x} \right) = X_{s+t}^x ,$$

in the sense that, for almost all $\omega \in \Omega$, we have

$$X_t^{X_s^x(\omega)}(\theta_s(\omega)) = X_{s+t}^x(\omega) .$$

Proof. Define

$$Y_u^x(\omega) = \begin{cases} X_u^x(\omega) & \text{if } u \le s\,, \\ X_{u-s}^{X_s^x(\omega)}(\theta_s(\omega)) & \text{if } u > s\,. \end{cases}$$

Then Y_{s+t}^x is solution of

$$\begin{aligned}
Y_{s+t}^x &= \Theta\left(X_t^{X_s^x}\right) \\
&= X_s^x + \Theta_s\left[\int_0^t f(X_u^{X_s^x})\,du\right] + \Theta_s\left[\int_0^t g(X_u^{X_s^x})\,dW_u\right] \\
&= x + \int_0^s f(X_u^x)\,du + \int_0^s g(X_u^x)\,dW_u + \int_0^t \Theta_s\left(f(X_u^{X_s^x})\right)\,du+ \\
&\quad + \int_s^{s+t} \Theta_s\left(f(X_{u-s}^{X_s^x}\right)\,dW_u \\
&= x + \int_0^s f(Y_u^x)\,du + \int_0^s g(Y_u^x)\,dW_u + \int_0^t f(Y_{u+s}^x)\,du+ \\
&\quad + \int_s^{s+t} f(Y_u^x\,dW_u \\
&= x + \int_0^s f(Y_u^x)\,du + \int_0^s g(Y_u^x)\,dW_u + \int_s^{s+t} f(Y_u^x)\,du+ \\
&\quad + \int_s^{s+t} f(Y_u^x\,dW_u \\
&= x + \int_0^{s+t} f(Y_u^x)\,du + \int_0^{s+t} g(Y_u^x)\,dW_u\,.
\end{aligned}$$

This shows that Y^x is solution of the same stochastic differential equation as X^x. We conclude easily by uniqueness of the solution. □

We are now ready to establish a parallel between stochastic differential equations and dynamical systems. Recall that we defined discrete time dynamical systems in Section **??**. In continuous time the definition extends in the following way. A *continuous-time dynamical system* on a set E is a one-parameter family of applications T_t, $t \in \mathbb{R}^+$, on E such that $T_s \circ T_t = T_{s+t}$ for all s, t. That is, T is a semigroup of applications on E. Each mapping T_t can be lifted into an operator on $\mathcal{L}^\infty(E)$, denoted by $\widehat{T}_t$ and defined by

$$\widehat{T}_t f(x) = f(T_t x)\,.$$

The following result is now a direct application of Theorem 4.16.

Corollary 4.17. *Let $(\Omega, \mathcal{F}, P)$ be the canonical space of the Brownian motion W. Consider the stochastic differential equation*

$$X_t^x = x + \int_0^t f(X_u^x)\,du + \int_0^t g(X_u^x)\,dW_u\,.$$

Then the mappings T_t on $\mathbb{R} \times \Omega$ defined by

$$T_t(x,\omega) = (X_t^x(\omega), \theta_t(\omega))$$

define a continuous time dynamical system on $\mathbb{R}$.

This is to say that a stochastic differential equation is nothing more than a deterministic dynamical system on a product set $\mathbb{R} \times \Omega$, that is, it is a semigroup of point transformations of this product set. We then have a result analogous to the one of Theorem **??** when this dynamical system is restricted to the $\mathbb{R}$-component.

Before establishing this result, we need few technical lemmas. In the following $\Omega_{t]}$ denotes the space of continuous functions on $[0,t]$. For all $\omega \in \Omega$ we denote by $\omega_{t]}$ the restriction of ω to $[0,t]$. Finally $P_{t]}$ denotes the restriction of the measure P to $(\Omega_{t]}, \mathcal{F}_t)$.

Lemma 4.18. *The image of the measure P under the mapping*

$$\begin{aligned} \Omega &\to \quad \Omega_{t]} \times \Omega \\ \omega &\mapsto (\omega_{t]}, \theta_t(\omega)) \end{aligned}$$

is the measure $P_{t]} \otimes P$.

Proof. Recall that $\omega(s) = W_s(\omega)$ and $\theta_t(\omega)(s) = W_{t+s}(\omega) - W_t(\omega)$. If A is a finite cylinder of $\Omega_{t]}$ and B a finite cylinder of Ω, then the set

$$\{\omega \in \Omega\,;\ (\omega_{t]}, \theta_t(\omega)) \in A \times B\}$$

is of the form

$$\begin{aligned} \{\omega \in \Omega\,;\ W_{t_1}(\omega) \in A_1, \dots, W_{t_n}(\omega) \in A_n,& \\ (W_{s_1} - W_t)(\omega)\ B_1, \dots, (W_{s_k} - W_t)(\omega) \in B_k\}& \end{aligned}$$

for some $t_1, \dots, t_n \le t$ and some $s_1, \dots, s_k > t$. By the independence of the Brownian motion increments, the probability of the above event is equal to

$$\begin{aligned} P(\{\omega \in \Omega_{t]}\,;\ W_{t_1}(\omega) \in A_1, \dots, W_{t_n}(\omega) \in A_n\})& \\ P(\{\omega \in \Omega\,;\ (W_{s_1} - W_t)(\omega)\ B_1, \dots, (W_{s_k} - W_t)(\omega) \in B_k\})\,.& \end{aligned}$$

This is to say that

$$\begin{aligned} P(\{\omega \in \Omega\,;\ (\omega_{t]}, \theta_t(\omega)) \in A \times B\}) =& \\ P(\{\omega \in \Omega\,;\ \omega_{t]} \in A)\, P(\{\omega \in \Omega\,;\ \theta_t(\omega) \in B\})\,.& \end{aligned}$$

This is exactly the claim of the lemma for the cylinder sets. As the measures P and $P \otimes P$ are determined by their values on the cylinder sets, we conclude easily. □

Lemma 4.19. *Let g be a bounded measurable function on $\Omega_{t]} \times \Omega$. Then we have*

$$\int_{\Omega_{t]}} \int_{\Omega} g(\omega,\omega')\, dP_{t]}(\omega)\, dP(\omega') = \int_{\Omega} g(\omega_{t]},\theta_t(\omega))\, dP(\omega)\,.$$

Proof. First consider g of the form $\mathbb{1}_{A\otimes B}$. We have

$$\begin{aligned} P_{t]}\otimes P(A\times B) &= P_{t]}(A)P(B) \\ &= P_{t]}(A)P(\theta_t^{-1}(B)) \quad \text{(for } \theta_t \text{ preserves } P) \\ &= P(\{\omega\in\Omega\,;\ \omega_{t]}\in A, \theta_t(\omega)\in B\}) \quad \text{(by Lemma 4.18)} \\ &= P\otimes P(\{(\omega,\omega')\in\Omega\times\Omega\,;\ \omega_{t]}\in A, \theta_t(\omega)\in B\} \\ &= P\otimes P(\phi^{-1}(A\times B))\,. \end{aligned}$$

This proves the lemma for such functions g. It now easy to extend to general functions g by a usual monotone class argument. □

Theorem 4.20. *Let $(\Omega,\mathcal{F},P)$ be the canonical space of the Brownian motion W. Consider the stochastic differential equation*

$$X_t^x = x + \int_0^t b(X_u^x)\, du + \int_0^t \sigma(X_u^x)\, dW_u$$

and the associated dynamical system

$$T_t(x,\omega) = (X_t^x(\omega),\theta_t(\omega))\ .$$

For any bounded function h on $\mathbb{R}$ consider the mapping

$$P_t h(x) = \mathbb{E}\left[\widehat{T}_t(h\otimes \mathbb{1})(x,\,\cdot\,)\right]\ .$$

Then P is a Markov semigroup on $\mathbb{R}$ with generator

$$A = b(x)\frac{d}{dx} + \frac{1}{2}\sigma(x)^2\frac{d^2}{dx^2}\ .$$

Proof. The fact that each P_t is a Markov operator is a consequence of Theorem **??**. Let us check that they form a semigroup. By definition we have

$$\begin{aligned} P_t(P_s h)(x) &= \mathbb{E}\left[\widehat{T}_t(P_s h\otimes \mathbb{1})(x,\,\cdot\,)\right] \\ &= \int_{\Omega} P_s h(X_t^x(\omega))\, dP(\omega) \\ &= \int_{\Omega}\int_{\Omega} h(X_s^{X_t^x(\omega)}(\omega')\, dP(\omega')\, dP(\omega) \\ &= \int_{\Omega} h(X_s^{X_t^x(\omega_{t]})}(\theta_t(\omega))\, dP(\omega) \quad \text{(by Lemma 4.19)} \\ &= \int_{\Omega} h(X_{s+t}^x(\omega)\, dP(\omega) \quad \text{(by Theorem 4.16)} \\ &= P_{s+t}h(x)\,. \end{aligned}$$

We have proved the semigroup property. □

We have proved the continuous time analog of Theorem **??**. Every Markov semigroup, with a generator of the form above, can be dilated on a larger set (a product set) into a dynamical system. This means again an open system point of view on Markov processes: Markov processes are obtained by restriction of certain types of dynamical systems on a product space, when one is averaging over one inaccessible component.

In this section we have developed the Brownian case only. But it is clear that all this discussion extends exactly in the same way to the case of the Poisson process. Indeed, the arguments developed above are only based on the independent increment property.

Index