Stochastic calculus lecture

Serge Cohen [1]

November 12, 2024

[1]Institut de Mathématiques de Toulouse; UMR 5219, Université de Toulouse; CNRS, UT3 F-31062 Toulouse Cedex 9, France. First-Name.Name@math.univ-toulouse.fr

Chapter 1

Construction of Brownian Motion and first properties

1 Introduction

The aim of this lecture is to construct the stochastic integral. The primary motivation is to develop an integral and differential calculus capable of handling computations with random "noises." Historically, the concept of random "noises" originated from experimental sciences. In probability theory, the most classical example of "noise" is Brownian motion. Brownian motion was introduced by Robert Brown in 1828 to study the movement of pollen particles in water. Later, in 1905, Einstein used Brownian motion to model the trajectories of gas molecules. Additionally, Bachelier applied Brownian motion to model stock option prices.

Let us give a first definition.

Definition 1.1. *Brownian motion* $(B_t)_{t\geq 0}$ *is a* $\mathbb{R}^d$ *"process"* $(d \geq 1)$ *(i.e. a family of random variables in short r.v.) such that*

1. $\forall n \in \mathbb{N}^*$ *and* $t_0 < t_1 < \ldots < t_n$ *the r.v.'s* $B_{t_0},\ B_{t_1} - B_{t_0}, \ldots, B_{t_{n+1}} - B_{t_n}$ *are independent (BM is a process with independent increments PII in short.)*
2. *If* $s < t,\ B_t - B_s$ *is a centered Gaussian random vector with covariance matrix* $(t-s)Id$.

$(B_t)_{t\geq 0}$ *starts from* $0 \in \mathbb{R}^d$ *if* $B_0 = 0,\ \mathbb{P}a.s.$

Remark 1.1.

- *If $d=1$, $Var(B_t - B_s) = t - s$.*
- *Random vectors from $(\Omega, \mathcal{A}, P)$ to $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)$, Independence, Gaussian random vectors are supposed to be known.*
- *"Sample paths" : Heuristically we fix $\omega \in \Omega$, and we are interested in $t \mapsto B_t(\omega)$ Wiener (1923, 1924), Paley-Zygmund. P almost surely the sample paths are contnuous and nowhere differentiable.*
- *If $B(t) = (B^1(t), \dots, B^d(t))$, then $\forall i = 1$ to d $B^i(t)$ are real valued Brownian motions and if $i \neq j$, B^i is independent of B^j.*

The aim of stochastic calculus is to give a rigorous meaning to (stochastic) differential equations of the type

$$y_t = \int_0^t f(y_s)\dot{B}_s ds$$

which have many applications. We will show that the Brownian motion is a continuous martingale, and that integrals can be defined in this framework. Another goal is to have a chain rule associated to these integrals. It is called the Itô formula which claims that $\forall f \in C^2(\mathbb{R}^d, \mathbb{R})$,

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

Actually one can show if $(X_t)_{t\geq 0}$ is real valued process with independent increments and if the distribution of $X_t - X_s$ does depend only of $t-s$, (stationary increments) with continuous sample paths then $X_t = X_0 + \sigma B_t$. More generally one can study PIIS process with stationary independent increments not with continuous sample paths. Another example of PIIS is the Poisson process.

Références :

- I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus.* Springer Verlag, 1988.
- Damien Lamberton and Bernard Lapeyre. *Introduction to stochastic calculus applied to finance.* Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL, second edition, 2008.

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

2 Definition and continuity of Brownian motion

2.1 Distribution of a process

Definition 1.2. *Let T be a set and $(E, \mathcal{E})$ a measurable set. A E valued stochastic process indexed by T is a family $(X_t,\ t \in T)$ of r.v.'s X_t : $(\Omega, \mathcal{A}, \mathbb{P}) \mapsto (E, \mathcal{E})$.*

Reminder : If $Y : (\Omega, \mathcal{A}, \mathbb{P}) \mapsto (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ is real valued random variable, the distribution $\mathbb{P}_Y$ of Y is a probability measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$. It is the push forward of the probability $\mathbb{P}$ by the measurable function Y. It is defined by $\mathbb{P}_Y(A) = \mathbb{P}(Y \in A)$, $\forall A \in \mathscr{B}(\mathbb{R})$.

For a process a sample path is associated to every $\omega \in \Omega$, $t \mapsto X_t(\omega)$. Hence the distribution of a process is a probability on a set of functions from $T \mapsto E$ which is denoted by E^T, and endowed with the cylindrical sigma field. A Cylinder is indexed by $t_0 < t_1 < \ldots < t_n$, it is a subset of E^T,

$$C_{t_0,\, t_1,\, \ldots,\, t_n} = \{f \in E^T,\ (f(t_0),,\ \ldots,\ f(t_n)) \in A_0 \times \ldots \times A_n,\ \text{with } A_i \in \mathcal{E}\},$$

where $C_{t_0,\, t_1,\, \ldots,\, t_n}$ actually depends also on A_i's. Then the distribution of the process is a probability on E^T endowed with the smallest sigma field that contains all cylinders. This sigma field is called the cylindrical sigma field denoted by $\mathcal{E}^{\otimes T}$.

Definition 1.3. *The distribution of a E-valued process is a probability on E^T endowed with the smallest sigma field $\mathcal{E}^{\otimes T}$ that contains all cylinders. It is uniquely defined by*

$$\mathbb{P}(X(t_0) \in A_0,\ \ldots, X(t_n) \in A_n),$$

$\forall n \in \mathbb{N}^*$ and $t_0 < t_1 < \ldots < t_n$.

Example the Brownian motion that starts from 0 :

If $t_0 < t_1 < \ldots < t_n$, and f is bounded Borel function $(\mathbb{R}^d)^{n+1} \mapsto \mathbb{R}$, let

$$p_t(x) = \frac{1}{(2\pi t)^{d/2}} \exp\left(\frac{-\|x\|^2}{2t}\right), \tag{1.1}$$

where $\|.\|$ is the Euclidean norm on $\mathbb{R}^d$. Then $p_{t-s}(x)dx$ is the distribution of $B_t - B_s$, so

$$\begin{aligned}\mathbb{E}f(B_{t_0}, \ldots,\ B_{t_n}) &= \int f(x_0, \ldots,\ x_n) d\mathbb{P}_{(B_{t_0}, \ldots,\ B_{t_n})}(x_0, \ldots,\ x_n)\\ &= \int f(x_0, \ldots,\ x_n) d\mathbb{P}_{(B_{t_0}, B_{t_1}-B_{t_0}, \ldots,\ B_{t_n}-B_{t_{n-1}})}(x_0,\ x_1 - x_0, \ldots,\ x_n - x_{n-1})\\ &= \int f(x_0, \ldots,\ x_n) p_{t_0}(x_0) p_{t_1-t_0}(x_1 - x_0) \ldots p_{t_n - t_{n-1}}(x_n - x_{n-1}) dx_0 \ldots dx_n.\end{aligned}$$

Hence

$$\begin{aligned}&\mathbb{P}(B(t_0) \in A_0,\ \ldots, B(t_n) \in A_n)\\ &\quad = \int_{A_0 \times \ldots \times A_n} p_{t_0}(x_0) p_{t_1-t_0}(x_1 - x_0) \ldots p_{t_n - t_{n-1}}(x_n - x_{n-1}) dx_0 \ldots dx_n\end{aligned} \tag{1.2}$$

and

$$\mathbb{P}(B(t_0) \in A_0,\ \ldots, B(t_n) \in A_n) = \mathbb{P}(X_0 \in A_0,\ \ldots,,\ X_0 + \ldots + X_n \in A_n),$$

where $X_0 \overset{(d)}{=} \mathcal{N}(0, t_0)$, $X_i \overset{(d)}{=} \mathcal{N}(0, t_i - t_{i-1})$ for all $i = 1$ to n and X_i's are independent.

Remark 1.2. *This remark is also an exercise. If $A \in \mathscr{B}(\mathbb{R}^d)^{\otimes \mathbb{R}_+}$, then $\exists (t_n)_{n\in\mathbb{N}}$ and $\exists B \in \mathscr{B}(\mathbb{R}^d)^{\otimes \mathbb{N}}$ such that*

$$A = \{f \in (\mathbb{R}^d)^{\mathbb{R}_+},\ (f(t_n))_{n\in\mathbb{N}} \in B\}.$$

Does exist on $(\mathbb{R}^d)^{\mathbb{R}_+}$ a probability such that

$$\mu(C_{t_0,\ t_1,\ \ldots,\ t_n}) = \mathbb{P}(B(t_0) \in A_0,\ \ldots, B(t_n) \in A_n)?$$

The answer"yes" is given by a Kolmogorov theorem. Let us first introduce definitions and a necessary condition.

Definition 1.4. *Let $\mathcal{T}$ be the set of finite increasing sequences $t = (t_0, \ldots, t_n)$ of numbers, where the length $n+1$ of these sequences ranges over the set of positive integers. Suppose that for each $\underline{t}$ of length $n+1$, we have a probability measure $Q_{\underline{t}}$ on $\left(\mathbb{R}^{n+1}, \mathscr{B}\left(\mathbb{R}^{n+1}\right)\ \right)$. Then the collection $\{Q_t\}_{t\in\mathcal{T}}$ is called a family of finite-dimensional distributions. This family is said to be consistent provided that the following condition is satisfied: If $\underline{t} = (t_0, t_1, \ldots, t_n)$ with $n \geq 1, \underline{t}^i = (t_0, t_1, \ldots, t_n)$, where t_i is missing then $\forall i \leq n$*

$$Q_{\underline{t}^i}(A_0 \times \cdots \times A_{i-1} \times A_{i+1} \times \cdots \times A_n) = Q_{\underline{t}}(A_0 \times \cdots \times \mathbb{R} \times \cdots \times A_n).$$

If we have a probability measure μ on $\left(\mathbb{R}^{[0,\infty)}, \mathcal{B}(\mathbb{R})^{\otimes[0,+\infty)}\right)$, then we can define a family of finite-dimensional distributions by

$$Q_{\underline{t}}(A) = \mu\left[\omega \in \mathbb{R}^{[0,\infty)}; (\omega\left(t_1\right), \ldots, \omega\left(t_n\right)) \in A\right], \tag{1.3}$$

where $A \in \mathscr{B}(\mathbb{R}^n)$ and $t = (t_1, \ldots, t_n) \in T$. This family is easily seen to be consistent. We are interested in the converse of this fact, because it will enable us to construct a probability measure $\mathbb{P}$ from the finite-dimensional distributions of Brownian motion.

Theorem 1.1 ((Daniell (1918), Kolmogorov (1933))**.** *Let $\{Q_t\}$ be a consistent family of finite-dimensional distributions. Then there is a probability measure $\mathbb{P}$ on $\left(\mathbb{R}^{[0,\infty)}, \mathcal{B}(\mathbb{R}^d)^{\otimes[0,+\infty)}\right)$, such that (1.3) holds for every $t \in T$.*

Proof. A proof can be read in p 50 of Karatzas and Shreeve. One can use for instance Carathéodory theorem, that may also be used to construct Lebesgue measure. □

To verify that we can apply the Theorem to the construction of the Brownian motion, it is enough to show consistency in this case. Coming back to (1.2), we are left to check

$$\int_{\mathbb{R}^d} p_{t_i-t_{i-1}}(x_i - x_{i-1})p_{t_{i+1}-t_i}(x_{i+1} - x_i)dx_i = p_{t_{i+1}-t_{i-1}}(x_{i+1} - x_{i-1}).$$

But it is the same as $p_{t_i-t_{i-1}} * p_{t_{i+1}-t_i} = p_{t_{i+1}-t_{i-1}}$, or $\mathcal{N}(0, t_i - t_{i-1}) + \mathcal{N}(0, t_{i+1} - t_i) = \mathcal{N}(0, t_{i+1} - t_{i-1})$ where independence is assumed on the left hand side.

2.2 Regularity of BM sample paths

In this section we want to convince ourselves that BM sample paths are almost surely continuous. We hope that $\mathcal{C}$ the set of continuous functions is of probability 1 under the distribution of BM.

Remark 1.3. *Unfortunately $\mathcal{C}$ is not measurable in $\mathscr{B}(\mathbb{R}^d)^{\otimes \mathbb{R}_+}$. Prove this fact as an exercise, that you use the previous Remark/exercise.*

To circumvent this problem we will show that there exists a process with almost sure continuous sample paths that has the same distribution as BM.

To clean a bit the situation some definitions are introduced.

Definition 1.5. *1. Processes X and Y have the same finite-dimensional distributions if, for any integer $n \geq 1$, real numbers $0 \leq t_1 < t_2 < \cdots < t_n < \infty$, and $A \in \mathscr{B}\left(\mathbb{R}^{nd}\right)$, we have:*

$$\mathbb{P}\left[(X_{t_1}, \ldots, X_{t_n}) \in A\right] = \mathbb{P}\left[(Y_{t_1}, \ldots, Y_{t_n}) \in A\right].$$

2. Y is a modification of X if, for every $t \geq 0$, we have $\mathbb{P}\left[X_t = Y_t\right] = 1$.

3. X and Y are called indistinguishable if almost all their sample paths agree:

$$\mathbb{P}\left[X_t = Y_t; \forall 0 \leq t < \infty\right] = 1.$$

Exercise 1.1. *If X is a modification of X' then the distribution of X and X' are the same.*

If X and X' are indistinguishable then there are modifications of each other. The converse is false.

Example. *Let $\Omega = [0,1]$, $\mathcal{A} = \mathscr{B}([0,1])$, $\mathbb{P} = dx$ the Lebesgue measure. Let us take $X_t(\omega) = \mathbf{1}(\omega \neq t)$, and $Y_t = 1$. Then*

$$\mathbb{P}(X_t = Y_t) = \mathbb{P}(\omega \neq t) = 1.$$

Hence X is a modification of Y. But $\mathbb{P}(\forall t \in [0,1],\ X_t = Y_t) = 0$. They are not indistinguishable.

Exercise 1.2. *Let Y be a modification of X, and suppose that both processes have a.s. right-continuous sample paths. Then X and Y are indistinguishable.*

The next theorem will show that there exits a modification of the BM with almost sure continuous sample paths.

Theorem 1.2 (Kolmogorov, Čentsov (1956))**.**

Suppose that a process $\{X_t; 0 \le t \le T\}$ on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ satisfies the condition

$$\mathbb{E}\,|X_t - X_s|^\alpha \le C|t-s|^{1+\beta}, \quad 0 \le s,t \le T,$$

for some positive constants α, β, and C. Then there exists a continuous modification $\tilde{X} = \left\{\tilde{X}_t; 0 \le t \le T\right\}$ of X, which is locally Hölder-continuous with exponent γ for every $\gamma \in (0, \beta/\alpha)$, i.e.,

$$\mathbb{P}\left[\omega; \sup_{\substack{0<t-s<h(\omega)\\ s,t\in[0,T]}} \frac{\left|\tilde{X}_t(\omega) - \tilde{X}_s(\omega)\right|}{|t-s|^\gamma} \le \delta\right] = 1, \tag{1.4}$$

where $h(\omega)$ is an a.s. positive random variable and $\delta > 0$ is an appropriate constant.

Proof. For simplicity, we take $T = 1$. Much of what follows is a consequence of the Cebyšev inequality. First, for any $\varepsilon > 0$, we have

$$\mathbb{P}\left[|X_t - X_s| \ge \varepsilon\right] \le \frac{\mathbb{E}\,|X_t - X_s|^\alpha}{\varepsilon^\alpha} \le C\varepsilon^{-\alpha}|t-s|^{1+\beta},$$

and so $X_s \to X_t$ in probability as $s \to t$. Second, setting $t = k/2^n, s = (k-1)/2^n$, and $\varepsilon = 2^{-\gamma n}$ (where $0 < \gamma < \beta/\alpha$) in the preceding inequality, we obtain

$$\mathbb{P}\left[\left|X_{k/2^n} - X_{(k-1)/2^n}\right| \ge 2^{-\gamma n}\right] \le C2^{-n(1+\beta-\alpha\gamma)},$$

and consequently,

$$\begin{aligned}
&\mathbb{P}\left[\max_{1\le k\le 2^n} \left|X_{k/2^n} - X_{(k-1)/2^n}\right| \ge 2^{-\gamma n}\right] \\
&\quad = \mathbb{P}\left[\bigcup_{k=1}^{2^n} \left|X_{k/2^n} - X_{(k-1)/2^n}\right| \ge 2^{-\gamma n}\right] \\
&\quad \le C2^{-n(\beta-\alpha\gamma)}.
\end{aligned}$$

The last expression is the general term of a convergent series; by the Borel-Cantelli lemma, there is a set $\Omega^* \in \mathscr{F}$ with $\mathbb{P}(\Omega^*) = 1$ such that for each $\omega \in \Omega^*$,

$$\max_{1 \leq k \leq 2^n} \left| X_{k/2^2}(\omega) - X_{(k-1)/2^n}(\omega) \right| < 2^{-\gamma n}, \quad \forall n \geq n^*(\omega), \tag{1.5}$$

where $n^*(\omega)$ is a positive, integer-valued random variable. For each integer $n \geq 1$, let us consider the partition $D_n = \{(k/2^n)\,; k = 0,\ 1, \ldots, 2^n\}$ of $[0,1]$, and let $D = \bigcup_{n=1}^{\infty} D_n$ be the set of dyadic rationals in $[0,1]$. We shall fix $\omega \in \Omega^*, n \geq n^*(\omega)$, and show that for every $m > n$, we have

$$|X_t(\omega) - X_s(\omega)| \leq 2 \sum_{j=n+1}^{m} 2^{-\gamma j}; \quad \forall t, s \in D_m, 0 < t - s < 2^{-n}. \tag{1.6}$$

For $m = n+1$, we can only have $t = (k/2^m)\,, s = ((k-1)/2^m)$, and (1.6) follows from (1.5). Suppose (1.6) is valid for $m = n+1, \ldots, M-1$. Take $s < t, s, t \in D_M$, consider the numbers $t^1 = \max\{u \in D_{M-1}; u \leq t\}$ and $s^1 = \min\{u \in D_{M-1}; u \geq s\}$, and notice the relationships $s \leq s^1 \leq t^1 \leq t, s^1 - s \leq 2^{-M}, t - t^1 \leq 2^{-M}$. From (1.5) we have $|X_{s^1}(\omega) - X_s(\omega)| \leq 2^{-\gamma M}, |\ X_t(\omega) - X_{t^1}(\omega)\ | \leq 2^{-\gamma M}$, and from (1.6) with $m = M-1$,

$$|X_{t^1}(\omega) - X_{s^1}(\omega)| \leq 2 \sum_{j=n+1}^{M-1} 2^{-\gamma j}.$$

We obtain (1.6) for $m = M$.

We can show now that $\{X_t(\omega); t \in D\}$ is uniformly continuous in t for every $\omega \in \Omega^*$. For any numbers $s, t \in D$ with $0 < t - s < h(\omega) \triangleq 2^{-n^*(\omega)}$, we select $n \geq n^*(\omega)$ such that $2^{-(n+1)} \leq t - s < 2^{-n}$. We have from (1.6)

$$|X_t(\omega) - X_s(\omega)| \leq 2 \sum_{j=n+1}^{\infty} 2^{-\gamma j} \leq \delta |t-s|^{\gamma}, \quad 0 < t - s < h(\omega), \tag{1.7}$$

where $\delta = 2/\left(1 - 2^{-\gamma}\right)$. This proves the desired uniform continuity. We define $\tilde{X}$ as follows. For $\omega \notin \Omega^*$, set $\tilde{X}_t(\omega) = 0, 0 \leq t \leq 1$. For $\omega \in \Omega^*$ and $t \in D$, set $\tilde{X}_t(\omega) = X_t(\omega)$. For $\omega \in \Omega^*$ and $t \in [0,1] \cap D^c$, choose a sequence $\{s_n\}_{n=1}^{\infty} \subseteq D$ with $s_n \to t$; uniform continuity and the Cauchy criterion imply that $\{X_{s_n}(\omega)\}_{n=1}^{\infty}$ has a limit which depends on t but not on the particular sequence $\{s_n\}_{n=1}^{\infty} \subseteq D$ chosen to converge to t, and we set

$\widetilde{X}_t(\omega) = \lim_{s_n \to t} X_{s_n}(\omega)$. The resulting process $\widetilde{X}$ is thereby continuous; indeed, $\widetilde{X}$ satisfies (1.7), so (1.4) is established.

To see that $\tilde{X}$ is a modification of X, observe that $\tilde{X}_t = X_t$ a.s. for $t \in D$; for $t \in [0,1] \cap D^c$ and $\{s_n\}_{n=1}^{\infty} \subseteq D$ with $s_n \to t$, we have $X_{s_n} \to X_t$ in probability and $X_{s_n} \to \tilde{X}_t$ a.s., so $\widetilde{X}_t = X_t$ a.s. □

This long proof is an example of the "chaining" argument, which is used in many other proofs.

To show that BM has a continuous modification we still have to show that the bound on the expectations of the increments of BM that is an assumption of the Theorem is satisfied

Proposition 1.1. *If B is a real valued BM $d = 1$,*

$$\forall n \in \mathbb{N}, \forall t, \in [0, +\infty),\ \mathbb{E}(B_t - B_s)^{2n} = C_n |t-s|^n \tag{1.8}$$

There exists a modification of the BM with locally hölder continuous paths for every exponent $0 < \gamma < \frac{1}{2}$.

Proof.

$$B_t - B_s \stackrel{(d)}{=} \mathcal{N}(0, |t-s|) \stackrel{(d)}{=} \sqrt{|t-s|}\mathcal{N}(0,1)$$

Hence $\mathbb{E}(B_t - B_s)^2 = |t-s|$ and $\mathbb{E}(B_t - B_s)^{2n} = C_n|t-s|^n$ where $C_n = \mathbb{E}X^{2n}$, $X \stackrel{(d)}{=} \mathcal{N}(0,1)$. Applying (1.8) for n fixed we get Hölder continuity for $\gamma < \frac{n-1}{2n}$. □

Remark 1.4. *Further on we always take continuous modifications of BM.*

3 Quadratic variations of Brownian motions

We may wonder if the previous result is optimal. For instance could it be that sample paths of Brownian motions are locally Liptschitz continuous ? Actually elementary definitions of integrals of the type $\int H_s dB_s$ are possible if the sample paths have almost surely finite variations. Let us first recall some facts concerning functions with finite variations.

3.1 Reminder of functions with finite variations

If $f : (0,+\infty) \mapsto \mathbb{R}$ is a non decreasing function, right continuous, we can associate a measure μ on $(0,+\infty)$ with $\forall 0 < s < t$

$$\mu((s,t]) = f(t) - f(s)$$

and if g is a bounded Borel function one can define

$$\int_0^t g(s)df(s) \stackrel{def}{=} \int \mathbf{1}_{[0,t]}(s)g(s)d\mu(s).$$

If f is C^1 it is equal to $\int_0^t g(s)f'(s)ds$.

Definition 1.6. *For $t > 0$, Let $\mathcal{P}_t$ be the set of finite subdivisions Δ of $[0,t]$: $\Delta = (t_i)_{i=1,\dots,n} \in \mathcal{P}_t$, if $0 \leq t_1 \leq \dots t_n \leq t$. The mesh of Δ is denoted by $|\Delta| = \sup_{i=1,\dots,n-1}(t_{i+1} - t_i)$. For $f : (0,+\infty) \mapsto \mathbb{R}$, the variation of f on $[0,t]$ is denoted by*

$$V_t(f) \stackrel{def}{=} \sup_{\Delta \in \mathcal{P}_t} \sum_{i=1}^{n} |f(t_{i+1}) - f(t_i)|$$

and is said to have finite variations if $\forall t > 0$, $V_t(f) < \infty$.

Example.

- *If f is monotone or a difference of non decreasing functions, f has finite variations.*
- *If f is locally Lipschitz, f has finite variations.*

We will admit two facts for Riemmann-Stieljes integral

- Every function f with finite variations is a a difference of non decreasing functions f_1, f_2, one can write

$$\int_0^t g(s)df(s) \stackrel{def}{=} \int_0^t g(s)df_1(s) - \int_0^t g(s)df_2(s).$$

- If g is continuous and $(\Delta^n)_{n\in\mathbb{N}}$ a sequence of subdivisions with meshes $|\Delta^n| \to 0$,

$$\int_0^t g(s)df(s) = \lim_{n\to\infty} \sum_{t_i^n \in \Delta^n} g(t_i^n)(f(t_{i+1}^n) - f(t_i^n)).$$

3.2 Quadratic variations

Actually quadratic variations of BM sample paths are almost surely : positive finite and BM sample paths do not have finite variations.

Definition 1.7. *A real valued process X has finite quadratic variations denoted by $\langle X, X\rangle$ or $(\langle X\rangle)$ if $\forall(\Delta^n)_{n\in\mathbb{N}}$ sequence of subdivisions of $\mathcal{P}_t$ such that $|\Delta^n| \to 0$ and $\Delta^n = (t_i^n)_{1\le i\le N(n)}$ with $t_1^n = 0$ and $t_{N(n)}^n = t$ and*

$$T_{[0,t]}^{\Delta^n} \overset{def}{=} \sum_{t_i^n\in\Delta^n} (X_{t_{i+1}^n} - X_{t_i^n})^2 \overset{(P)}{\to} \langle X\rangle_t.$$

Proposition 1.2. *If B is a Brownian motion $\langle B\rangle_t = t$ a.s.*

Proof. We prove $T_{[0,t]}^{\Delta^n} \to t$ in $L^2(\Omega)$. If $\Delta^n = (t_i^n)_{1\le i\le N(n)}$

$B_{t_{i+1}^n} - B_{t_i^n} \overset{(d)}{=} \mathcal{N}(0, t_{i+1}^n - t_i^n) \overset{(d)}{=} \sqrt{t_{i+1}^n - t_i^n}\mathcal{N}(0,1)$. Then $\mathbb{E}((T_{[0,t]}^{\Delta^n}) = \sum_{t_i^n\in\Delta^n} \mathbb{E}(B_{t_{i+1}^n} - B_{t_i^n})^2 = \sum_{t_i^n\in\Delta^n} t_{i+1}^n - t_i^n = t$. Hence

$$\begin{aligned}
\mathbb{E}((T_{[0,t]}^{\Delta^n} - t)^2) &= Var(T_{[0,t]}^{\Delta^n} - t)\\
&= Var(\sum_{t_i^n\in\Delta^n} (B_{t_{i+1}^n} - B_{t_i^n})^2 - (t_{i+1}^n - t_i^n))\\
&= \sum_{t_i^n\in\Delta^n} Var((B_{t_{i+1}^n} - B_{t_i^n})^2 - (t_{i+1}^n - t_i^n))\\
&= \sum_{t_i^n\in\Delta^n} (t_{i+1}^n - t_i^n)^2 Var(N^2 - 1)\\
&\le C|\Delta^n| \sum_{t_i^n\in\Delta^n} (t_{i+1}^n - t_i^n) \to 0,
\end{aligned}$$

where $N = \mathcal{N}(0,1)$. This implies convergence in probability. □

To get almost sure convergence some additional assumptions are needed for Δ^n...

Proposition 1.3. *Almost surely the sample paths of Brownian motion have infinite variations on every intervals $[0,t]$ for $t > 0$.*

Proof. If ω is such that $V_{[0,t]}(B(\omega)) < +\infty$ then

$$\sum_{t_i^n\in\Delta^n} (B_{t_{i+1}^n} - B_{t_i^n})^2 \le \sup|(B_{t_{i+1}^n} - B_{t_i^n}| \sum_{t_i^n\in\Delta^n} |(B_{t_{i+1}^n} - B_{t_i^n}|.$$

Since $s \mapsto B_s(\omega)$ is uniformly continuous on $[0,t]$ $\sup|B_{t^n_{i+1}} - B_{t^n_i}| \to 0$. Hence the quadratic variations of the sample paths should vanish, which is true only on a negligible set. □

4 Brownian motion as a Gaussian process

4.1 Elementary properties

Definition 1.8. *A real valued process $(X_t, t \in T)$ is a Gaussian process if $\forall n \in \mathbb{N},\ t_1, \ldots, t_n \in T,\ (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n,\ \sum_{i=1}^n \alpha_i X_{t_i}$ is a Gaussian random variable. The process X is centered if $\forall t \in T,\ \mathbb{E}X_t = 0$ and $\Gamma(s,t) \stackrel{def}{=} Cov(X_s, X_t)$ is the covariance function.*

Remark 1.5.

- *If $\forall i = 1$ to n, $\alpha_i = 0$ $\sum_{i=1}^n \alpha_i X_{t_i} = 0$. It means that we consider $\mathcal{N}(0,0)$ as a generalized degenerated Gaussian random variable with variance 0. Gaussian processes are generalization of random Gaussian vectors (where T is a finite set). (See for instance N. Bouleau Processus stochastique et applications 1988.)*
- *If (X_1, X_2) is a Gaussian vector and $Cov(X_1, X_2) = 0$ then X_1 and X_2 are independent.*

Proposition 1.4. *The Brownian motion which starts from 0 is the unique centered Gaussian process with covariance $\Gamma(s,t) = \min(s,t)$.*

Proof. The proof relies on the fact that the covariance always characterizes the distribution of a Gaussian centered process. If X is a Gaussian centered process and $t_1, \ldots, t_n \in T$, then $(X_{t_1}, \ldots,\ X_{t_n})$ Gaussian vector implies the characteristic function

$$\mathbb{E}\exp(i\sum_{i=1}^n \alpha_i X_{t_i}) = \exp(-\frac{1}{2}\langle C\alpha, \alpha\rangle)$$

where $\alpha = (\alpha_1, \ldots,\ \alpha_n)$, $\langle ., . \rangle$ is The Euclidean scalar product in $\mathbb{R}^n$, and $C_{i,j} = \mathbb{E}(X_{t_i} X_{t_j}) = \Gamma(t_i, t_j)$. Hence the matrix $(\Gamma(t_i, t_j))$ characterizes the distribution of the finite dimensional margins $(X_{t_1}, \ldots,\ X_{t_n})$ $\forall t_1, \ldots, t_n \in T$, and henceforth the distribution of the process X.

Let us compute $\Gamma(s,t)$ for Brownian motion. Let $s \le t$

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

Hence $\Gamma(s,t) = \min(s,t)$.

□

Proposition 1.5. *Let $(B_t,\ t \geq 0)$ be a Brownian motion.*

1. $(B_{t+s} - B_s, \quad t \geq 0)$ *is a BM independent of* $\mathcal{F}_s^B = \sigma\left(B_u, \quad u \leq s\right)$.
2. $(-B_t, \quad t \geq 0)$ *is a Brownian motion,*
3. *(Self-similarity) for all* $\lambda > 0$, $\left(B_t^{(\lambda)}, \quad t \geq 0\right)$ *where* $B_t^{(\lambda)} := \frac{1}{\sqrt{\lambda}} B_{\lambda t}, \quad t \geq 0$ *is a Brownian motion.*

Proof. $W_t = B_{t+s} - B_s$ is a centered Gaussian process such that $W_0 = 0$ a.s. Its covariance

$\mathbb{E}(W_t W_{t'}) = \mathbb{E}((B_{t+s} - B_s)(B_{t'+s} - B_s)) = \min(t+s, t'+s) - s = \min(t,t')$.

$\forall u_1 \leq \ldots \leq u_n \leq s \leq t_1 \leq \ldots \leq t_n$ $(B_{u_1}, \ldots,\ B_{u_n},\ B_{t_1+s} - B_s, \ldots,\ B_{t_n+s} - B_s)$ is a Gaussian vector and $(B_{u_1}, \ldots,\ B_{u_n})$ is independent of $(B_{t_1+s} - B_s, \ldots,\ B_{t_n+s} - B_s)$ since $\forall i,\ j$

$$\mathbb{E}(B_{u_i}(B_{t_j+s} - B_s)) = 0.$$

Actually for Gaussian random vectors, a vanishing covariance yields independence. Then independence of sigma field is a consequence of independence of the random variable that generate them. To prove other parts of the Proposition, compute covariances. □

4.2 Brownian bridge

We may condition $(B_t,\ 0 \leq t \leq 1)$ to the event $B_1 = 0$. In this case we obtain a Brownian bridge.

Definition 1.9. *The process $X_t^0 = B_t - tB_1$ is called a Brownian bridge.*

Proposition 1.6. *X^0 is independent of B_1.*

Proof. $\forall 0 \leq t \leq 1,\ \mathbb{E}(X_t^0 - tB_1) = 0$. And it is a centered Gaussian process. Furthermore $\forall 0 \leq t \leq 1,\ \mathbb{E}(X_t^0 B_1) = \mathbb{E}((B_t - tB_1)B_1) = 0$. □

Proposition 1.7. *Let $X_t^b = B_t - tB_1 + tb$. The distribution of X^b is a regular version of the conditional distribution of $(B_t,\ t \leq 1)$ given $B_1 = b$.*

Proof. Reminder : $\mathbb{P}(dx, y)$ is a regular version of the conditional distribution of X given $Y = y$ if and only if (in short iff) $\forall \varphi$ bounded and measurable

$$\mathbb{E}(\varphi(X)|Y) = \int \varphi(x)\mathbb{P}(dx, Y)$$

$\mathbb{P}$ almost surely. It can be characterized by $\forall \varphi,\ g$ bounded and measurable

$$\mathbb{E}(\varphi(X)g(Y)) = \int (\int \varphi(x)\mathbb{P}(dx, y))g(y)d\mathbb{P}_Y(y).$$

In our case we have to show

$$\mathbb{E}(\varphi(B_s,\ s \leq 1)g(B_1)) = \int \mathbb{E}(\varphi(X_s^b,\ s \leq 1)g(b)\frac{e^{-\frac{b^2}{2}}db}{\sqrt{2\pi}}. \tag{1.9}$$

Let $\psi((f(s),\ s \leq 1), b) = \varphi((f(s) + sb, s \leq 1))$. Since $B_s = X_s^0 + sB_1$,

$$\begin{aligned}\mathbb{E}(\varphi(B_s,\ s \leq 1)g(B_1)) &= \mathbb{E}(\psi((X_s^0,\ s \leq 1), B_1)g(B_1)) \\ &= \int \mathbb{E}(\psi((X_s^0,\ s \leq 1), b)g(b))\frac{e^{-\frac{b^2}{2}}db}{\sqrt{2\pi}} \\ &= \int \mathbb{E}(\varphi((X_s^b,\ s \leq 1))g(b))\frac{e^{-\frac{b^2}{2}}db}{\sqrt{2\pi}}.\end{aligned}$$

□

4.3 Wiener integral

In this part the integral $\int f(s)dB_s$ is defined for a deterministic function using the Gaussianity of the Brownian motion. (Later the stochastic integral is defined for f a stochastic process.) If f is a simple function

$$f = \sum_{i=1}^{n} \alpha_i \mathbf{1}_(a_i, b_i]$$

for $a_1 < b_1 \leq a_2 < b_2 \leq \ldots$ Then

$$\int f(s)dB_s \stackrel{def}{=} \sum_{i=1}^{n} \alpha_i(B(b_i) - B(a_i)), \tag{1.10}$$

this random variable is denoted by $I(f)$. It is a centered Gaussian random variable with variance

$$\mathbb{E}(I(f)^2) = \sum_{i=1}^{n} \alpha_i^2(b_i - a_i) = \|f\|^2_{L^2(0,\infty)}.$$

Theorem 1.3. *There exists a unique map I from $L^2(0,\infty)$ to $\mathcal{H}$ the set that is the closure in $L^2(\Omega,\mathcal{A},\mathbb{P})$ of the linear combinations $\sum_{i=1}^n \alpha_i(B(b_i)-B(a_i))$ such that*

1. $I(\mathbf{1}_(a_i,b_i]) = B(b_i) - B(a_i)$.
2. *I is a linear map*
3. *I is an isometry i.e. $\forall f \in L^2(0,\infty)$ $\|f\|_{L^2(0,\infty)} = \|I(f)\|_{L^2(\Omega,\mathcal{A},\mathbb{P})}$.*

Proof. If $f \in L^2(0,\infty)$ $\exists (f_n)_{n\in\mathbb{N}}$ with f_n simple functions and $\lim_{n\to\infty} f_n = f$ in $L^2(0,\infty)$. Then let $I(f) \stackrel{def}{=} \lim_{n\to\infty} I(f_n)$ in $L^2(0,\infty)$. Please remark that $I(f)$ does not depend on the sequence $(f_n)_{n\in\mathbb{N}}$ since if $\lim_{n\to\infty} g_n = f$ then $\lim_{n\to\infty} I(g_n) - I(f_n) = 0$ because of the isometry property. Moreover we get the uniqueness of I by density of simple functions in L^2. □

Remark 1.6. 1. *Since I is an isometry $\forall f,\ g \in L^2(0,\infty)$,*

$$\langle f,g\rangle_{L^2(0,\infty)} = \langle I(f), I(g)\rangle_{L^2(\Omega,\mathcal{A},\mathbb{P})}. \tag{1.11}$$

If we denote by $\int_0^\infty f(s)dB_s = I(f)$, the so-called Wiener integral this can be rewritten :

$$\int_0^\infty f(s)g(s)ds = \mathbb{E}(\int_0^\infty f(s)dB_s \int_0^\infty g(s)dB_s).$$

2. *Conversely if $J\ :\ L^2(0,\infty) \mapsto L^2(\Omega,\mathcal{A},\mathbb{P})$ is such that $J(f)$ is a centered Gaussian random variable and*

$$\int_0^\infty f(s)g(s)ds = \mathbb{E}(J(f)J(g))$$

then J is a linear map. Moreover $J(\mathbf{1}_{(0,t]})$ is a real Brownian motion. (Since $(J(\mathbf{1}_{(0,t]}),\ t \le 0)$ is a centered Gaussian process and $\mathbb{E}(J(\mathbf{1}_{(0,s]})J(\mathbf{1}_{(0,t]}) = \min(s,t)$.) If we denote by $B_t = J(\mathbf{1}_{(0,t]})$, then J is the isometry I associated to the BM B.

The same construction can be generalized to all intervals I' and $L^2(I')$. For $I' = \mathbb{R}$, the process $X_t = I(\mathbf{1}_{(0,t]}), \forall t \in \mathbb{R}$ can be obtained from two independent real valued Brownian motion $X_t = B_t^1, \forall t \le 0$, and $X_t = B_t^2, \forall t \ge 0$. One can easily check that

$$\mathbb{E}(X_t - X_s)^2 = |t-s|,\ \forall t < 0 < s.$$

4.4 Second construction of Brownian motion

If $(e_n)_{n\in\mathbb{N}}$ is an orthonormal basis (ONB) of $L^2(0,1)$ then $(I(e_n))_{n\in\mathbb{N}}$ is a sequence of Gaussian independent random variables with distribution $\mathcal{N}(0,1)$. Actually $(1,\ \sqrt{2}\cos(2\pi ks),\ \sqrt{2}\sin(2\pi ks))_{k\in\mathbb{N}^*}$ is an ONB of $L^2(0,1)$, $\forall t\in(0,1)$,

$$\mathbf{1}_{(0,t]}(s) \overset{L^2(0,1)}{=} a_0(t)+\sum_{k=1}^{\infty} a_k(t)\sqrt{2}\cos(2\pi ks)+b_k(t)\sqrt{2}\sin(2\pi ks)$$

$$I(\mathbf{1}_{(0,t]}) \overset{L^2(\Omega,\mathcal{A},\mathbb{P})}{=} a_0(t)\xi_0+\sum_{k=1}^{\infty} a_k(t)\xi_k+\sum_{k=1}^{\infty} b_k(t)\eta_k,$$

with $(\xi_0,\ \xi_k,\ \eta_k)$ i.i.d. Gaussian random variables with distribution $\mathcal{N}(0,1)$. Moreover $a_0(t)=\int_0^1 \mathbf{1}_{(0,t]}(s)ds=t$, $\forall k\geq 1$

$$a_k(t)=\sqrt{2}\int_0^t \cos(2\pi ks)ds=\frac{\sin(2\pi kt)}{\sqrt{2}\pi k}$$

$$b_k(t)=\sqrt{2}\int_0^t \sin(2\pi ks)ds=\frac{(1-\cos(2\pi kt))}{\sqrt{2}\pi k}.$$

Hence we get a series expansion of BM, a priori in L^2 sense...

$$I(\mathbf{1}_{(0,t]})=t\xi_0+\sum_{k=1}^{\infty}\xi_k\frac{\sin(2\pi ks)}{\sqrt{2}\pi k}+\sum_{k=1}^{\infty}\eta_k\frac{(1-\cos(2\pi ks))}{\sqrt{2}\pi k}. \tag{1.12}$$

Since $I(\mathbf{1}_{(0,1]})=\xi_0$, hence (1.12) can be viewed as tB_1 plus the expansion of a Brownian bridge.

Theorem 1.4. *If $(\xi_0,\ \xi_k,\ \eta_k)$ are i.i.d. Gaussian random variables with distribution $\mathcal{N}(0,1)$*

$$t\xi_0+\sum_{k=1}^{\infty}\xi_k\frac{\sin(2\pi kt)}{\sqrt{2}\pi k}+\sum_{k=1}^{\infty}\eta_k\frac{(1-\cos(2\pi kt))}{\sqrt{2}\pi k}$$

almost surely converges to a process $(B_t)_{t\in(0,1)}$ with the distribution of a BM.

Proof. We may refer to criteria for convergences of random Fourier series in Kahane Some random series of functions Theorem 2 p 236 second edition, we get almost surely the uniform (but not normal) convergence of the series. Since $I(\mathbf{1}_{(0,t]})$ is a BM we get the distribution of the limit of the series. □

Remark 1.7. *With this construction almost sure continuity of the sample paths is for free !*

Chapter 2

Reminder for martingales indexed by $\mathbb{N}$

To integrate processes $H_s(\omega)$ against BM "$dB_s(\omega)$" we will assume that $H : [0, +\infty) \mapsto \mathbb{R}$, depends is "previsible". Roughly it means that $H(t, .)$ is measurable with respect to the sigma-field $\sigma(X_s,\ s < t)$ of the past. Then the time dependence of $t \mapsto \int_0^t H_s dB_s$, will be achieved so that $\int_0^t H_s dB_s$ is a martingale. First we recall results for martingales indexed by $\mathbb{N}$ especially convergence results. Then we will extend these results to martingales indexed by $[0, +\infty)$. The main issue in this case is that $[0, +\infty)$ is not denumerable.

1 Definitions and first examples

Definition 2.1. *A filtration on* $(\Omega, \mathcal{F}, \mathbb{P})$ *is a non-decreasing sequence of sub-*σ*-fields of* $\mathcal{F}$*:*

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_n \subset \cdots \subset \mathcal{F}.$$

One says that $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbb{P})$ *is a filtered probability space.*

Example 2.1. Let $(\Omega, \mathcal{F}, \mathbb{P}) = ([0, 1[, \mathcal{B}([0, 1[), \lambda)$, where λ is Lebesgue measure. The filtration $(\mathcal{F}_n)_{n\geq 0}$ defined by

$$\mathcal{F}_n = \sigma\left(\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right[, i = 0, \ldots, 2^n - 1\right), \quad n \geq 0$$

is called the dyadic filtration.

If the parameter n denotes time, then $\mathcal{F}_n$ is interpreted as available information up to time n.

Example 2.2. For a stochastic process $(X_n)_{n\geq 0}$, we define its natural filtration $\mathcal{F}^X = (\mathcal{F}_n^X)_{n\geq 0}$ by: for all $n \geq 0$,

$$\mathcal{F}_n^X = \sigma(X_0, X_1, \cdots, X_n),$$

which is the smallest σ-field such that $X_0, \ldots, X_n$ are measurable.

Definition 2.2. *We say that a stochastic process $X = (X_n)_{n\geq 0}$ is adapted to the filtration $(\mathcal{F}_n)_{n\geq 0}$, if for all $n \geq 0$, X_n is $\mathcal{F}_n$-measurable. We say that a stochastic process $(X_n)_{n\geq 0}$ is adapted if it is adapted to some filtration.*

A stochastic process is obviously adapted to its natural filtration.

Remark 2.1. If $(\mathcal{F}_n)_{n\geq 0}$ and $(\mathcal{G}_n)_{n\geq 0}$ are two filtrations such that $\mathcal{G}_n \subset \mathcal{F}_n$ for all $n \geq 0$, and if $(X_n)_{n\geq 0}$ is adapted to $(\mathcal{G}_n)_{n\geq 0}$, then $(X_n)_{n\geq 0}$ is adapted to $(\mathcal{F}_n)_{n\geq 0}$.

Definition 2.3. *A filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ is a non-decreasing sequence of sub-σ-fields of $\mathcal{F}$:*

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_n \subset \cdots \subset \mathcal{F}.$$

One says that $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbb{P})$ is a filtered probability space.

Example 2.3. For a stochastic process $(X_n)_{n\geq 0}$, we define its natural filtration $\mathcal{F}^X = (\mathcal{F}_n^X)_{n\geq 0}$ by: for all $n \geq 0$,

$$\mathcal{F}_n^X = \sigma(X_0, X_1, \cdots, X_n),$$

which is the smallest σ-field such that $X_0, \ldots, X_n$ are measurable.

Definition 2.4. *We say that a stochastic process $X = (X_n)_{n\geq 0}$ is adapted to the filtration $(\mathcal{F}_n)_{n\geq 0}$, if for all $n \geq 0$, X_n is $\mathcal{F}_n$-measurable. We say that a stochastic process $(X_n)_{n\geq 0}$ is adapted if it is adapted to some filtration.*

A stochastic process is obviously adapted to its natural filtration.

Remark 2.2. If $(\mathcal{F}_n)_{n\geq 0}$ and $(\mathcal{G}_n)_{n\geq 0}$ are two filtrations such that $\mathcal{G}_n \subset \mathcal{F}_n$ for all $n \geq 0$, and if $(X_n)_{n\geq 0}$ is adapted to $(\mathcal{G}_n)_{n\geq 0}$, then $(X_n)_{n\geq 0}$ is adapted to $(\mathcal{F}_n)_{n\geq 0}$.

Definition 2.5. *Let $X = (X_n)_n$ be an adapted process on filtered probability space $(\Omega, \mathcal{A}, (\mathcal{F}_n, n \in \mathbb{N}), \mathbb{P})$ such that for all n, X_n is integrable.*

The process X is a **martingale** *if for all n,*

$$\mathbb{E}[X_{n+1}/\mathcal{F}_n] = X_n, \quad \textit{almost surely.}$$

The process $X$is a **sub-martingale** *if for all integer n n,*

$$\mathbb{E}[X_{n+1}/\mathcal{F}_n] \geq X_n, \quad \textit{almost surely.}$$

The process X is a **super martingale** *if for all integer n,*

$$\mathbb{E}[X_{n+1}/\mathcal{F}_n] \leq X_n, \quad \textit{almost surely.}$$

Examples

(See exercises at the end of the chapter for some proofs of the following properties are left to the reader.)

1. If $X \in L^1(\Omega, \mathcal{A})$, $X_n = \mathbb{E}[X/\mathcal{F}_n]$ is a martingale.This process is also uniformly integrable.
2. (**Fundamental example.**) Let $(Z_n, n \in \mathbb{N}^*)$ be a sequence of independent and integrable random variables and X_0 be an integrable random variable independent of the sequence (Z_n). (Most of the time, X_0 is constant.) Let $X_n := X_0 + \sum_{i=1}^n Z_i$. Then the filtrations $\mathcal{F}_n^X$ and $\mathcal{F}_n = \sigma(X_0, Z_1, \ldots, Z_n)$ are equal and for this filtration :
   (a) if for all integer n, $\mathbb{E}(Z_n) = 0$, X is a martingale;
   (b) if for all integer n, $\mathbb{E}(Z_n) \geq 0$, X is a sub martingale;
   (c) if for all integer n $\mathbb{E}(Z_n) \leq 0$, X is an super martingale;
   (d) if all r.v. Z_i have same expectation m, $X_n - nm$ is a martingale.
3. A special case of the example 2 comes from the game theory. In this case the distribution of the r.v. Z_n is the Bernoulli distribution with parameter p : $\mathbb{P}(Z_i = 1) = p$, $\mathbb{P}(Z_i = -1) = 1 - p$. with values $+1$ et -1. In this case X_n is the fortune of the player after n bets, when its initial fortune is X_0. The process $(M_n)_n$ where $M_n = X_n - n(2p - 1)$ is a martingale for its natural filtration $\mathcal{F}^X$.

4. In the example 2, if we assume that $\mathbb{E}[\exp(aZ_n)] := \exp(r_n)$ exists and is finite, let $R_n = r_1 + \cdots + r_n$. (Here $R_0 = 0$.)

   Then $M_n = \exp(aX_n - R_n)$ is a martingale for the natural filtration $\mathcal{F}^X$.

A process X can be a martingale (resp. super, resp sub) with respect to several filtrations.

Proposition 2.1. *If X is a martingale (resp. a super-martingale, a sub-martingale) with respect to a filtration $(\mathcal{F}_n)$ and the process X is adapted to an other filtration $(\mathcal{G}_n)$ smaller than$(\mathcal{F}_n)$ (that means for all n, $\mathcal{G}_n \subset \mathcal{F}_n$), Then X is a martingale (resp. a super-martingale, a sub-martingale)with respect to the filtration $\mathcal{G}_n$. A martingale (resp. a super-martingale, a sub–martingale) is a martingale (resp. a super-martingale, a sub–martingale)with respect to its natural filtration .*

Proof. Use successive conditioning. . □

We can also increase filtrations by adding to each σ fields $\mathcal{F}_n$ an independent σ field:

Proposition 2.2. *Let (X_n) be a martingale (resp. a sub-martingale, an super-martingale), with respect to a filtration $\mathcal{F}_n$. Let $\mathcal{B}$ be a σ field independent of $\mathcal{F}_\infty$, and let $\mathcal{G}_n = \mathcal{F}_n \vee \mathcal{B}$. Then (X_n) is a martingale (resp. a sub-martingale, an super-martingale) with respect to the filtration $\mathcal{G}_n$.*

Proof. Left to the reader. □

Notation 2.1. *In the sequel*

$$(\Delta X)_n := X_n - X_{n-1} \tag{2.1}$$

is the increments process of (X_n).

Proposition 2.3. *Let X be a $\mathcal{F}$-martingale. Then*

1. $\forall n \geq 0$, $\forall k \geq 0$, $\mathbb{E}[X_{n+k}/\mathcal{F}_n] = X_n$; $\mathbb{E}[X_n] = \mathbb{E}[X_0]$.

2. *If the martingale is square integrable the increments $(\Delta X)_n$ of X are orthogonal :*

$$n \neq m \implies \mathbb{E}[(\Delta X)_n(\Delta X)_m] = 0.$$

3. *If X is a super-martingale, $-X$ is a sub-martingale.*

4. *The set of martingales with respect to a given filtration is a linear space.*

5. *If X is a martingale and ϕ is a convex application such that $Y_n = \phi(X_n)$ is integrable then , Y_n is a sub-martingale.*

6. *If X is a sub-martingale, and if ϕ is increasing and convex, $\phi(X)$ is a sub-martingale if $\phi(X_n)$ is integrable.*

Proof. The proof is left to the reader.

The point 1 relies on successive conditioning and induction.

The point 2 is obtained by conditioning by $\mathcal{F}_{m-1}$ for $n < m$.

The points 3 et 4 are immediate.

The points 5 et 6 rely on Jensen conditionnal inequality.

□

For square integrable martingale, we have

Proposition 2.4. *If M_n is a square integrable martingale*

$$\forall n \leq p,\ \mathbb{E}[(M_p - M_n)^2] = \sum_{k=n+1}^{p} \mathbb{E}[(\Delta M)_k^2].$$

Proof. Apply the property of orthogonal increments

2 of Proposition 2.3. □

Corollary 2.1. *A martingale bounded in L^2 converges in L^2.*

Proof. By definition, since the martingale is bounded in L^2 there exists a constant C such that for all n,

$$\mathbb{E}(X_n^2) \leq C^2.$$

Then,

$$\mathbb{E}(X_n - X_0)^2 \leq 4C^2,$$

and Proposition 2.4 allows to prove that the series

$$\sum_k \mathbb{E}[(\Delta M)_k^2]$$

converges. As a consequence,

$$\lim_{n\to\infty} \sup_{p\geq q\geq n} \sum_q^p \mathbb{E}[(\Delta M)_k^2] = 0.$$

Using the previous proposition again

$$\lim_n \sup_{p,q\geq n} \mathbb{E}(M_p - M_q)^2 = 0,$$

and the sequence is Cauchy in L^2 and converges.

□

2 Doob's decomposition

Definition 2.6. *Let $(A_n)_{n\geq 0}$ be a process indexed by $\mathbb{N}$, A is predictable with respect to the sigma field $\mathcal{F}_n$ if $\forall n$ A_n is $\mathcal{F}_{n-1}$measurable.*

Theorem 2.1. D DOOB 'S DECOMPOSITION *: Let X be a sub-martingale ; there exists a martingale M and a predictable increasing process A, null at 0, unique, such that for all integer n, $X_n = M_n + A_n$.*

The process A is called "compensator" of X.

Proof. Let $A_0 = 0$ and $M_0 = X_0$. For $n \geq 1$, define A_n in the following way : let $\Delta_n = \mathbb{E}(X_n/\mathcal{F}_{n-1}) - X_{n-1}$, and

$$A_n = \Delta_1 + \cdots + \Delta_n.$$

Moreover$M_n = X_n - A_n$. By construction A_n is predictable, and since X_n is a sub-martingale, $\Delta_n \geq 0$, and A_n is increasing. Moreover,

$$\mathbb{E}(M_{n+1}/\mathcal{F}_n) = \mathbb{E}(X_{n+1}/\mathcal{F}_n) - A_{n+1} = X_n + \Delta_n - A_{n+1} = M_n.$$

and M_n is a martingale.

Uniqueness comes from the fact that if such a decomposition exists then

$$\mathbb{E}(X_{n+1} - X_n/\mathcal{F}_n) = A_{n+1} - A_n,$$

This characterize A_n if $A_0 = 0$. □

In the particular case of square integrable martingale we obtain the following.

Proposition 2.5. *Let M_n be a square integrable martingale. Recall (notation 2.1) and $(\Delta M)_n = M_n - M_{n-1}$ and let*

$$U_n = \mathbb{E}[(\Delta M)_n^2/\mathcal{F}_{n-1}).$$

Then $M_n^2 - \sum_{k=1}^n U_k$ is a martingale.

Proof. It is the Doob's decomposition applying to the sub-martingale M_n^2, since

$$\mathbb{E}[(\Delta M)_n^2/\mathcal{F}_{n-1}) = \mathbb{E}(M_n^2/\mathcal{F}_{n-1}] - M_{n-1}^2.$$

□

3 Stopping times

3.1 Definition

Definition 2.7. *A random variable $T : \Omega \to \mathbb{N} \cup \{+\infty\}$ is called a stopping time (with respect to the filtration $(\mathcal{F}_n)_{n\geq 0}$) if for all $n \geq 0$,*

$$\{T \leq n\} \in \mathcal{F}_n.$$

Remark 2.3. Since $\{T = n\} = \{T \leq n\} \setminus \{T \leq n-1\}$, T is a stopping time if and only if for all $n \geq 0$,

$$\{T = n\} \in \mathcal{F}_n.$$

Remark 2.4. A stopping time is thus a random time, which can be interpreted as a stopping rule for deciding whether to continue or stop a process on the basis of the present information and past events, for instance playing until you go broke or you break the bank, etc. . .

Example 2.4. 1. If $T = n$ a.s., then clearly T is a stopping time.

2. Let $(X_n)_{n\geq 0}$ be an adapted stochastic process, and consider the first time X_n reaches the borel set A:

$$T_A = \inf\{n \geq 0 \,|\, X_n \in A\},$$

with the convention that $\inf \varnothing = +\infty$. It is called the hitting time of A. Then T_A is a stopping time. Indeed,

$$\begin{aligned}\{T_A = n\} &= \{X_0 \notin A, X_1 \notin A, \dots, X_{n-1} \notin A, X_n \in A\} \\ &= \bigcap_{k=0}^{n-1} \{X_k \notin A\} \cap \{X_n \in A\} \in \mathcal{F}_n.\end{aligned}$$

3. Show that $\tau_A = \sup\{n \geq 1 \,|\, X_n \in A\}$ the last passage time in A is not a stopping time in general.

Recall the notations: $x \wedge y = \inf(x, y)$ and $x \vee y = \max(x, y)$.

Proposition 2.6. *If S and T are two stopping times, then $S \wedge T$, $S \vee T$ and $S + T$ are also stopping times.*

Proof. Writing

$$\{S \wedge T \leq n\} = \{S \leq n\} \cup \{T \leq n\}$$

and

$$\{S \vee T \leq n\} = \{S \leq n\} \cap \{T \leq n\}$$

gives the result for $S \wedge T$ and $S \vee T$. For $S + T$, we write:

$$\{S + T \leq n\} = \bigcup_{k\leq n} \{S = k\} \cap \{T \leq n - k\} \in \mathcal{F}_n,$$

since $\mathcal{F}_k \subset \mathcal{F}_n$ for all $k \leq n$. □

Remark 2.5. In particular, if T is a stopping time, then for all $n \geq 0$, $T \wedge n$ is a bounded stopping time.

Proposition 2.7. *If $(T_k)_k$ is a sequence of stopping times, then $\inf_k T_k$, $\sup_k T_k$, $\liminf_k T_k$ and $\limsup_k T_k$ are also stopping times.*

Proof. Exercise. □

Proposition 2.8. *Let T be a stopping time. Then,*

$$\mathcal{F}_T = \{A \in \mathcal{F} \,|\, \forall n \geq 0, A \cap \{T = n\} \in \mathcal{F}_n\}$$

is a σ-field, called the σ-field of T-past.

Remark 2.6. Obviously, T is $\mathcal{F}_T$-measurable.

Proof. It is obvious that $\Omega \in \mathcal{F}_T$. If $A \in \mathcal{F}_T$, then for all n,

$$A^c \cap \{T = n\} = \{T = n\} \setminus A = \{T = n\} \setminus (A \cap \{T = n\}) \in \mathcal{F}_n,$$

hence $A^c \in \mathcal{F}_T$. If $(A_k)_k$ is countable collection of $\mathcal{F}_T$-mesurable set, then

$$\left(\bigcup_k A_k\right) \cap \{T = n\} = \bigcup_k (A_k \cap \{T = n\}) \in \mathcal{F}_n,$$

hence $\bigcup_k A_k \in \mathcal{F}_T$. □

Proposition 2.9. *Let S and T be two stopping times such that $S \leq T$. Then, $\mathcal{F}_S \subset \mathcal{F}_T$.*

Proof. Let $A \in \mathcal{F}_S$. Then, for all $n \geq 0$,

$$A \cap \{T = n\} = A \cap \{S \leq n\} \cap \{T = n\} = \bigcup_{k=0}^{n} A \cap \{S = k\} \cap \{T = n\} \in \mathcal{F}_n. \ \square$$

Definition 2.8. *Let $(X_n)_{n\geq 0}$ be an adapted stochastic process and T a stopping time. If $T < \infty$ a.s., we define the random variable X_T by*

$$X_T(\omega) = X_{T(\omega)}(\omega) = X_n(\omega) \quad \textit{if } T(\omega) = n.$$

Note that X_T is $\mathcal{F}_T$-measurable, since

$$\{X_T \in B\} \cap \{T = n\} = \{X_n \in B\} \cap \{T = n\} \in \mathcal{F}_n,$$

for any Borel set B.

4 Martingales transformations

Proposition 2.10. *Let (X_n) be an adapted process and (H_n) be a predictable process such that for all n, the r.v. $H_n(X_n - X_{n-1})$ is integrable. Let $(H.X)$ be the process defined by*

$$(H.X)_n = H_0 X_0 + \sum_{k=1}^{n} H_k(X_k - X_{k-1}).$$

Then, if X is a martingale, $(H.X)$ is a martingale. If X is a super- (resp. sub-) martingale, and if H is positive, then $(H.X)$ is a(n) super- (resp. sub-) martingale.

Proof. Using the notation 2.1 , the process $(H.X)$ satisfies

$$(\Delta(H.X))_n = H_n(\Delta X)_n.$$

The proof is then left to the reader. □

In a casino for example, the process H corresponds to a player's strategy : according to all observations he has at time n, he bets at time $n+1$ an H_{n+1}, to earn a gain $H_{n+1}(X_{n+1} - X_n)$.

An important particular case of Proposition 2.10 is the following

Corollary 2.2. *Let (X_n) be a martingale (resp. a sub-, an super-martingale), and let be a T stopping time. Then the process X^T defined by $X_n^T = X_{T\wedge n}$ is a martingale (resp. a sub-, an super-martingale).*

Proof. It is enough to consider the predictable (right ?) process $H = \mathbf{1}_{[0,T]}$. In this case the process $(H.X)$ is nothing but X^T :

$$(H.X)_n = X_0 + \sum_{k=1}^{T\wedge n}(X_k - X_{k-1}) = X_0 + \sum_{k=1}^{n}(X_k - X_{k-1})\mathbf{1}_{k\leq T}.$$

□

Note that the process $T\wedge n$ is adapted to the filtration $\mathcal{G}_n = \mathcal{F}_{T\wedge n}$ smaller than $\mathcal{F}_n$.

Using the predictable process $H = \mathbf{1}_A \mathbf{1}_{[T,\infty[}$, for $A \in \mathcal{F}_T$, we obtain

Corollary 2.3. *If T is a stopping time, then $\mathbf{1}_A(X_{T\vee n} - X_{T)})$ is a martingale (resp. a sub-, an super-martingale).*

5 Stopping theorem :bounded stopping time's case.

Theorem 2.2. *(***Stopping theorem***.)*

Let $(X_n, n \in \mathbb{N})$ be a martingale and S and T be two bounded stopping times (that means there exists an integer n such that $S \vee T \leq n$, almost surely). Then,

$$\mathbb{E}(X_T/\mathcal{F}_S) = X_{S\wedge T}. \tag{2.2}$$

If X is a sub (resp. an super-)martingale,

$$X_{S\wedge T} \leq \text{ (resp } \geq \text{) } \mathbb{E}(X_T/\mathcal{F}_S). \tag{2.3}$$

In particular if $(X_n, n \in \mathbb{N})$ is a sub-martingale and S and T are two bounded stopping time, then

$$\mathbb{E}(X_S \mathbf{1}_{S\leq T}) \leq \mathbb{E}(X_T \mathbf{1}_{S\leq T}). \tag{2.4}$$

We have the inverse inequality for an super-martingale.

Proof. We give only the proof for the martingale case.

First, we study the case where

$T = n$ and $S \leq n$. The equality (2.2) to obtain can be written as

$$X_S = \mathbb{E}(X_n/\mathcal{F}_S).$$

By definition,

$$X_S = \sum_{k=0}^{n} X_k \mathbf{1}_{S=k}.$$

We know that X_S is $\mathcal{F}_S$ measurable; and also integrable as finite linear combination of integrable variables.

It is enough to prove that for all $A \in \mathcal{F}_S$,

$$\mathbb{E}(X_S \mathbf{1}_A) = \mathbb{E}(X_n \mathbf{1}_A).$$

This can be written as

$$\sum_{k=0}^{n} \mathbb{E}(X_k \mathbf{1}_{A\cap\{S=k\}}) = \sum_{k=0}^{n} \mathbb{E}(X_n \mathbf{1}_{A\cap\{S=k\}}).$$

But since $A \in \mathcal{F}_S$, then $A \cap \{S = k\} \in \mathcal{F}_k$, and using the martingale property we obtain, for all $k \leq N$,

$$\mathbb{E}(X_k \mathbf{1}_{A \cap \{S=k\}}) = \mathbb{E}(X_n \mathbf{1}_{A \cap \{S=k\}}).$$

We now study the general case. Let an integer n such that $S \vee T \leq n$.

Using the previous case for the stopped martingale X^T, and the stopping time S. We have $X^T_n = X_T$ since $T \leq n$, $X^T_S = X_{S \wedge T}$. We have

$$\mathbb{E}(X_T/\mathcal{F}_S) = X_{S \wedge T}.$$

Note that the variable $X_{S \wedge T}$ is $\mathcal{F}_{S \wedge T}$ measurable, and as a consequence

$$X_{S \wedge T} = \mathbb{E}(X_T/\mathcal{F}_{S \wedge T}).$$

To obtain inequality (2.4), it is enough to note that inequality (2.3) means that for all $A \in \mathcal{F}_S$

$$\mathbb{E}(X_{S \wedge T} \mathbf{1}_A) \leq \mathbb{E}(X_T \mathbf{1}_A).$$

We apply it to the set $A = \{S \leq T\}$.

□

Corollary 2.4. *Let (T_n) be an increasing sequence of bounded stopping time, and X be a martingale (resp. a sub-martingale, an super-martingale) ; then $(X_{T_n}, n \in \mathbb{N})$ is a martingale (resp. a sub-martingale, an super-martingale) for the filtration $(\mathcal{F}_{T_n}, n \in \mathbb{N})$.*

Proof. (on exercise) □

Corollary 2.5. *Let X be an integrable r.v., and let (X_n) be the martingale $\mathbb{E}(X/\mathcal{F}_n)$. If T is a bounded stopping time then*

$$\mathbb{E}(X/\mathcal{F}_T) = X_T.$$

If S et T are two bounded stopping time;

$$\mathbb{E}(X/\mathcal{F}_S/\mathcal{F}_T) = \mathbb{E}(X/\mathcal{F}_T/\mathcal{F}_S) = \mathbb{E}(X/\mathcal{F}_{S \wedge T}) = X_{S \wedge T}.$$

Proof. Let N such that $T \vee S \leq N$. Using successive conditioning for the martingale $X_n = \mathbb{E}(X/\mathcal{F}_n)$ $\mathbb{E}(X/\mathcal{F}_T) = \mathbb{E}(X_N/\mathcal{F}_T)$. The stopping theorem yields $\mathbb{E}(X_N/\mathcal{F}_T) = X_T$.

If S and T are two bounded stopping,

$$\mathbb{E}(X_T/\mathcal{F}_S) = X_{S\wedge T} = \mathbb{E}(X/\mathcal{F}_{S\wedge T}).$$

□

6 Finite stopping times

In this section we extend the stopping theorem to the case of finite stopping times. Its requires some additional integrability conditions on martingales (resp. sub-martingales, super-martingales).

Proposition 2.11. *Let $(X_n, n \in \mathbb{N})$ be a martingale (resp. a sub-martingale) and T be S two almost surely finite stopping times.*

If the sequences $(X_{T\wedge n})$ are $(X_{S\wedge n})$ uniformly integrable, then

$$X_{S\wedge T} = E[X_T/\mathcal{F}_S]. \quad \textit{(resp. } X_{S\wedge T} \leq E[X_T/\mathcal{F}_S]).$$

This is the case when there exists a r.v. $Y \in L^1$ such that for all n, $|X_{T\wedge n}| \leq Y$),or when (X_n) is uniformly integrable

In particular for the martingale, we have $\mathbb{E}(X_T) = \mathbb{E}(X_0)$ for all finite stopping time which satisfies this assumption.

Proof. We only study the martingale case.

First note that the r.v. X_T is integrable, as almost sure limit of uniformly integrable sequence $(X_{T\wedge n})$. (Note that T is a.s. finite.) It is the same for X_S. We have to prove that for all bounded variable Z $\mathcal{F}_S$-measurable , we have

$$\mathbb{E}(X_S Z) = \mathbb{E}(X_T Z).$$

We can use the monotone class theorem monotones and restrict ourself to the case where Z is $\mathcal{F}_{S\wedge n}$ measurable using the fact that $\mathcal{F}_S = \vee_n \mathcal{F}_{S\wedge n}$.

Let such an n. Using the Stopping Theorem 2.2 for the stopping time $S \wedge p$ and $T \wedge p$, and $p \geq n$ we obtain

$$\mathbb{E}(Z X_{S\wedge p}) = \mathbb{E}(Z X_{T\wedge p}).$$

Using uniform integrability we can let n going to infinity.

For the last point, if (X_n) is uniformly integrable, it is enough to note that $X_{T\wedge n} = \mathbb{E}(X_n/\mathcal{F}_T)$. The desired conclusion follows from the fact that a family of conditional expectation of uniformly integrable family is uniformly integrable □

7 Inequalities and convergence

7.1 Inequalities

Theorem 2.3. *(*Doob*'s maximal inequality.) Let $(X_n, n \in \mathbb{N})$ be a positive sub martingale and $\lambda \geq 0$. let $X_n^* = \sup_{k=0}^n X_k$. Then*

$$\forall n \in \mathbb{N},\ \lambda \mathbb{P}\{X_n^* \geq \lambda\} \leq \mathbb{E}[X_n \mathbf{1}_{\{X_n^* \geq \lambda\}}] \leq E[X_n].$$

Proof. Let $T = \inf\{k \in \mathbb{N}, X_k \geq \lambda\}$ a stopping time. Then,

$$\{T \leq n\} = \{X_n^* \geq \lambda\}.$$

Take $S = T \wedge (n+1)$,which is a bounded stopping time . We have

$$A = \{S \leq n\} = \{T \leq n\} \in \mathcal{F}_S.$$

Using the Stopping Theorem for this sub-martingale, between n and $S \wedge n$.

$$\mathbb{E}(X_{S\wedge n} \mathbf{1}_A) \leq \mathbb{E}(X_n \mathbf{1}_A).$$

This can be written

$$\mathbb{E}(X_T \mathbf{1}_{T\leq n}) \leq \mathbb{E}(X_n \mathbf{1}_{T\leq n}). \tag{2.5}$$

On the set $\{T \leq n\}$, $X_T \geq \lambda$, hence $\lambda\mathbb{P}(T \leq n) \leq \mathbb{E}(X_T \mathbf{1}_{T\leq n})$. Then $\lambda\mathbb{P}(T \leq n) \leq\leq \mathbb{E}(X_n \mathbf{1}_{T\leq n})$ is the desired inequality. □

Corollary 2.6. *If $(X_n, n \in \mathbb{N})$ is a martingale, $(|X_n|, n \in \mathbb{N})$ is a positive sub-martingale and*

$$\forall n \in \mathbb{N},\ \lambda \mathbb{P}\{\max_{k\leq n} |X_k| \geq \lambda\} \leq E[|X_n| \mathbf{1}_{\max_{k\leq n} |X_k| \geq \lambda}] \leq E[|X_n|].$$

Theorem 2.4. *Let $(X_n, n \in \mathbb{N})$ be a positive sub-martingale and $p > 1$. Then, if $X_n \in L^p$,*

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

Proof. If $X_n \in L^p$ then variables $X_{k\in L^p}$ for $k \leq n$.

Let U be a positive r.v. in L^p,

$$E(U^p) = p\int_0^\infty t^{p-1}\mathbb{P}(U \geq t)dt.$$

Then

$$\begin{aligned} E[(X_n^*)^p] &= p\int_0^\infty t^{p-1}\mathbb{P}(X_n^* \geq t)dt \\ &\leq p\int_0^\infty t^{p-2}E[X_n\mathbf{1}_{\{X_n^*\geq t\}}]dt \\ &= pE[X_n\int_0^\infty t^{p-2}\mathbf{1}_{\{X_n^*\geq t\}}dt] \\ &= \frac{p}{p-1}E[X_n(X_n^*)^{p-1}]. \end{aligned}$$

Using Hölder inequality

$$E[X_n(X_n^*)^{p-1}] \leq \|X_n\|_p\|X_n^*\|_p^{p-1}.$$

Since X_n^* is bounded by $\sum_0^n X_k$, it belongs to L^p. The desired result is obtained by cancellation □

In particular

Corollary 2.7. *Let (X_n) be a positive sub-martingale bounded in L^1. Then the variable $X^* = \sup_n X_n$ is finite almost surely. If (X_n) is bounded in L^p $(p > 1)$, then X^* belongs to L^p. (This last result is false for $p = 1$.) The same conclusions hold for martingales (not necessary positive).*

Proof. The increasing sequence X_n^* converges towards X^*. It is enough to apply Doob's inequality and

$$\lambda\mathbb{P}(X_n^* > \lambda) \leq \sup_n \mathbb{E}(|X_n|) = K < \infty.$$

Letting n going to infinity

$$\lambda\mathbb{P}(X^* > \lambda) \leq K,$$

$\mathbb{P}(X^* > \lambda) \to 0$ $(\lambda \to \infty)$. The r.v. X^* is finite.

For the second part use the Theorem 2.4.

The case of martingales is obtained by applying the previous result to the positive sub-martingale $|X_n|$.

□

7.2 Convergences

Results

The following results are given without any proof.

Proposition 2.12. *Let X_n be a martingale, or a sub-martingale, or an super-martingale, bounded in L^1. Then X_n converges almost surely towards a variable X_∞.*

Using Fatou's lemma, M_∞ the limit of a bounded in L^1 martingale M_n is integrable. In general $M_n \neq \mathbb{E}(M_\infty/\mathcal{F}_n)$.

It is the case for uniformly integrable martingales .

Proposition 2.13. *let M_n be a bounded martingale in L^1, and let M_∞ the limit of M_n when $n \to \infty$. The following statements are equivalent*

1. *M_n converges in L^1 towards M_∞.*
2. *$M_n$is uniformly integrable.*
3. *$M_n = \mathbb{E}[M_\infty/\mathcal{F}_n]$.*
4. *There exists an integrable r.v. M such that $M_n = \mathbb{E}[M/\mathcal{F}_n]$. Moreover in this case , $M_\infty = \mathbb{E}[M/\mathcal{F}_\infty]$.*

(Here $\mathcal{F}_\infty = \vee_n \mathcal{F}_n$.)

Proof. For a sequence of r.v. which converges almost surely, it is equivalent to converge in L^1 or to be uniformly integrable. For all integrable r.v. M, the set of the r.v. $\mathbb{E}(M/\mathcal{B})$, where $\mathcal{B}$ is running in all sub σ fields of $\mathcal{A}$ is an uniformly integrable family. It is enough to prove the following points:

1. If (M_n) is uniformly integrable, then $M_n = \mathbb{E}(M_\infty/\mathcal{F}_n)$;
2. If M is an integrable r.v., the martingale $M_n = \mathbb{E}(M/\mathcal{F}_n)$ converges towards $\mathbb{E}(M/\mathcal{F}_\infty)$.

For the first point note that for $p \geq n$ $M_n = \mathbb{E}(M_p/\mathcal{F}_n)$,letting p going to infinity using the fact that the expectation is continuous in L^1, and that M_p converges towards M_∞ in L^1 by assumption. We get the desired result.

For the second point, note that M_∞ is $\mathcal{F}_\infty$ measurable by construction. It is enough to show that, for a $A \in \mathcal{F}_\infty$, we have $\mathbb{E}(M_\infty \mathbf{1}_A) = \mathbb{E}(M\mathbf{1}_A)$. This is true when A belongs to sub σ fields of $\mathcal{F}_n$,since

$$\mathbb{E}(M\mathbf{1}_A) = \mathbb{E}(M_n\mathbf{1}_A) = \mathbb{E}(M_\infty\mathbf{1}_A).$$

The desired identity is then, true for all element of $\cup_n \mathcal{F}_n$, and for all element σ -field generated by $\cup_n \mathcal{F}_n$ using a monotone class theorem argument. The desired inequality is true for $\mathcal{F}_\infty$. □

Remarks

1. A similar statement as in Proposition 2.13 is true for sub-and super-martingales; the proof is left to the reader.
2. A bounded martingale L^p for $p > 1$, is dominated by an L^p variable and converges in L^p.

We now are in position to enunciate the Stopping theorem for general stopping-times.

Theorem 2.5. *(***Stopping Theorem.***) Let M_n be a uniformly integrable martingale and let T be a stopping time (not necessarily finite). Then for $M_T = M_\infty$ on $\{T = \infty\}$,we have*

1. *$M_T = \mathbb{E}(M_\infty/\mathcal{F}_T)$.*
2. *The set (M_T), where T is a stopping time is uniformly integrable.*
3. *If S and T are two stopping time, we have*

$$\mathbb{E}(M_T/\mathcal{F}_S) = M_{S\wedge T}.$$

4. *Let M be a $\mathcal{A}$-measurable integrable r.v. and $M_n = \mathbb{E}(M/\mathcal{F}_n)$, then $M_T = \mathbb{E}(M/\mathcal{F}_T)$.*

Proof. For the first point, it is enough to write the proof of Stopping theorem in this case. If A belongs to $\mathcal{F}_T$, then

$$\begin{aligned}\mathbb{E}(M_T \mathbf{1}_A) &= \sum_{k\in\mathbb{N}\bigcup\infty} \mathbb{E}(M_k \mathbf{1}_{A\bigcap\{T=k\}}) \\ &= \sum_{k\in\mathbb{N}\bigcup\infty} \mathbb{E}(M_\infty \mathbf{1}_{A\bigcap\{T=k\}}) = \mathbb{E}(M_\infty \mathbf{1}_A).\end{aligned}$$

The family M_T is contained in the family $\mathbb{E}(M_\infty/\mathcal{B})$, where $\mathcal{B}$ is running in the subσ fields of $\mathcal{A}$. This last family is uniformly integrable.

The stopping martingale M^T is uniformly integrable. Using the Stopping theorem at time S, we obtain

$$\mathbb{E}(M_T/\mathcal{F}_S) = M_{S\wedge T}.$$

It is enough to write

$$\mathbb{E}(M/\mathcal{F}_T) = \mathbb{E}(\mathbb{E}(M/\mathcal{F}_\infty)/\mathcal{F}_T) = \mathbb{E}(M_\infty/\mathcal{F}_T) = M_T.$$

□

8 Exercises

1. Prove the claim 2 of examples 1 in section 1.
2. Recall a definition of uniform integrability (U.I.) that claims that X_i is U.I. if $\sup_i \mathbb{E}|X_i| < \infty$ and if a property sometimes called equiintegrability (to be recalled) is fulfilled.
3. Prove the claim 1 of examples 1 in section 1.
4. Show that M_n in the claim 3 of examples 1 in section 1 is square integrable. What is the Doob decomposition of M_n^2 ?

Chapter 3

Martingales in continuous time

1 Filtrations in continuous time

The definition of filtrations in continuous time is given under the usual conditions. Those conditions are assumed to avoid nasty technical problems related to the fact $(0, +\infty)$ is not denumerable and therefore there exists obstructions to measurabilty.

Definition 3.1. *A family of sigma fields $(\mathcal{F}_t)_{0 \le t \le +\infty}$ on a sigma field $\mathcal{A}$ associated to a probability space $(\Omega,\ \mathcal{A},\ \mathbb{P})$ is a filtration if $\forall s \le t,\ \mathcal{F}_s \subset \mathcal{F}_t$. It satisfies the usual conditions if*

1. *$\mathcal{F}_0$ contains all negligible sets of $\mathcal{A}$ (it is complete)*
2. *$\mathcal{F}_t = \cap_{s>t} \mathcal{F}_s, \forall t \le 0,$ (it is right continuous.)*

In this lecture all filtrations satisfy the usual condition.

We introduce a measurability assumption for processes that states that the process depends only on the past of the filtration.

Definition 3.2. *The stochastic process X is adapted to the filtration $\{\mathscr{F}_t\}$ if, for each $t \ge 0, X_t$ is an $\mathscr{F}_t$-measurable random variable.*

Obviously, every process X is adapted to $\left\{\mathscr{F}_t^X\right\} \stackrel{def}{=} \sigma(X_s, 0 \le s \le t)$. Moreover, if X is adapted to $\{\mathscr{F}_t\}$ and Y is a modification of X, then Y is also adapted to $\{\mathscr{F}_t\}$ provided that $\mathscr{F}_0$ contains all the P-negligible sets in $\mathscr{F}$.

Definition 3.3. *The stochastic process X is called progressively measurable with respect to the filtration $\{\mathscr{F}_t\}$ if, for each $t \geq 0$ and $A \in \mathscr{B}\left(\mathbb{R}^d\right)$, the set $\{(s,\omega); 0 \leq s \leq t, \omega \in \Omega, X_s(\omega) \in A\}$ belongs to the product σ-field $\mathscr{B}([0,t]) \otimes \mathscr{F}_t$; in other words, if the mapping*

$$(s,\omega) \mapsto X_s(\omega) : ([0,t] \times \Omega, \mathscr{B}([0,t]) \otimes \mathscr{F}_t) \rightarrow \left(\mathbb{R}^d, \mathscr{B}\left(\mathbb{R}^d\right)\right)$$

is measurable, for each $t \geq 0$.

Proposition 3.1. *If the stochastic process X is adapted to the filtration $\{\mathscr{F}_t\}$ and every sample path is right-continuous or else every sample path is left continuous, then X is also progressively measurable with respect to $\{\mathscr{F}_t\}$.*

Proof. We treat the case of right-continuity. With $t > 0, n \geq 1, k = 0, 1, \ldots, 2^n - 1$, and $0 \leq s \leq t$, we define:

$$X_s^{(n)}(\omega) = X_{(k+1)t/2^n}(\omega) \text{ for } \quad \frac{kt}{2^n} < s \leq \frac{k+1}{2^n} t,$$

as well as $X_0^{(n)}(\omega) = X_0(\omega)$. The so-constructed map $(s,\omega) \mapsto X_s^{(n)}(\omega)$ from $[0,t] \times \Omega$ into $\mathbb{R}^d$ is demonstrably $\mathscr{B}([0,t]) \otimes \mathscr{F}_t$-measurable. Besides, by right-continuity we have: $\lim_{n\to\infty} X_s^{(n)}(\omega) = X_s(\omega), \forall (s,\omega) \in [0,t] \times \Omega$. Therefore, the (limit) map $(s,\omega) \mapsto X_s(\omega)$ is also $\mathscr{B}([0,t]) \otimes \mathscr{F}_t$-measurable. □

2 Stopping times in continuous time

Definition 3.4. *Let us consider a measurable space $(\Omega, \mathscr{F})$ equipped with a filtration $\{\mathscr{F}_t\}$. A random variable $T : \Omega \mapsto [0, +\infty]$ is a stopping time of the filtration, if the event $\{T \leq t\}$ belongs to the σ-field $\mathscr{F}_t$, for every $t \geq 0$.*

Proposition 3.2. *Show that $\forall t \geq 0$, $\{T < t\}$ belongs to the σ-field $\mathscr{F}_t$ is equivalent to T is a stopping time.*

Proof. The proof is based on the observation $\{T < t\} = \bigcup_{n=1}^{\infty} \{T \leq t - (1/n)\} \in \mathscr{F}_t$, because if T is a stopping time, then $\{T \leq t - (1/n)\} \in \mathscr{F}_{t-(1/n)} \subseteq \mathscr{F}_t$ for $n \geq 1$. For the converse, suppose that $\forall t \geq 0$, $\{T <$

$t\} \in \mathscr{F}_t$ of the right-continuous filtration $\{\mathscr{F}_t\}$. Since for every positive integer m, we have $\{T \leq t\} = \bigcap_{n=m}^{\infty}\{T < t + (1/n)\}$, we deduce that $\{T \leq t\} \in \mathscr{F}_{t+(1/m)}$; whence $\{T \leq t\} \in \cap_{s>t}\mathcal{F}_s = \mathscr{F}_t$. □

Consider a subset $A \in \mathscr{B}\left(\mathbb{R}^d\right)$ of the state space of the process, and define the hitting time

$$D_A(\omega) = \inf\{t \geq 0; X_t(\omega) \in A\}$$

and

$$T_A(\omega) = \inf\{t > 0; X_t(\omega) \in A\}.$$

Remark 3.1. *By convention we set* $\inf(\emptyset) = +\infty$.

If $X_0 \notin A,\ D_A = T_A$.

Proposition 3.3. *1. If X is right continuous and adapted, if A is open then T_A is a stopping time.*

2. If the process X is continuous and adapted, if A is closed then D_A is a stopping time.

Proof. Let us prove the second claim.

$$\{D_A \leq t\} = \cap_{n\in\mathbb{N}^*} \cup_{s\in\mathbb{Q},\, 0<s\leq t} \{d(X_s, A) < \frac{1}{n}\}.$$

Let us first show that the set on the left is included in the set on the right. Since X is continuous and A closed $X_{D_A} \in A$. Moreover there exists a sequence $s_n \in \mathbb{Q}$ such that it is increasing to $D_A(\omega)$ and $d(X_{s_n}, X_{D_A}) < \frac{1}{n}$. Then $d(X_{s_n}, A) \leq d(X_{s_n}, X_{D_A})$. The inclusion is proved, let us prove the inclusion the other way around.

$\forall n \in \mathbb{N}^*\ \exists s_n \leq t$ and $d(X_{s_n}, A) < \frac{1}{n}$, We consider a susbsequence $s_{n_k} \to t' \leq t$, then $d(X_{t'}, A) = 0$ by continuity of X and of $x \mapsto d(x, A)$. This implies $D_A \leq t'$. Since $\{d(X_s, A) < \frac{1}{n}\} \in \mathcal{F}_s$ the proof is finished.

Let us prove the first claim. Because of the Proposition 3.2 it is enough to show that $\{T_A < t\} \in \mathcal{F}_t$.

$$\begin{aligned}\{T_A < t\} &= \{\omega,\ \exists s < t,\ X_s(\omega) \in A\} \\ &= \{\omega,\ \exists s < t, s \in \mathbb{Q}\ X_s(\omega) \in A\}\end{aligned}$$

since A is open and X right continuous. Then

$$\{T_A < t\} = \cup_{s<t,s\in\mathbb{Q}}\{\omega, X_s(\omega) \in A\}$$

and each of the set in the last union belongs to $\mathcal{F}_t$. □

Exercise 3.1. *If S and T are two stopping times, then $S \wedge T$, $S \vee T$, $T + S$ are also stopping times. (The last one is more difficult).*

Proposition 3.4. *Every stopping time is the limit of a nonincreasing sequence of stopping times that take only a finite number of values.*

Proof. Let us denote by T the stopping time and let us define for $n \in \mathbb{N}^*$ and $1 \leq k \leq 2^n$,

$$T_n(\omega) = \frac{k}{n}$$

if $\frac{k-1}{n} < T \leq \frac{k}{n}$ and $T_n = +\infty$ if $T > \frac{2^n}{n}$. On can check that T_n is a a nonincreasing sequence converging to T. Moreover

$$\{T_n = \frac{k}{n}\} = \{T \leq \frac{k}{n}\} \setminus \{T \leq \frac{k-1}{n}\} \in \mathcal{F}_{\frac{k}{n}}.$$

If $t \geq 0$,

$$\{T_n \leq t\} = \cup_{k/n \leq t} \{T_n = \frac{k}{n}\}$$

then T_n is a stopping time. □

Definition 3.5. *Let T be a stopping time of the filtration $\{\mathscr{F}_t\}$. The σ-field $\mathscr{F}_T$ of events determined prior to the stopping time T consists of those events $A \in \mathscr{F}$ for which $A \cap \{T \leq t\} \in \mathscr{F}_t$ for every $t \geq 0$.*

Exercise 3.2. *Verify that $\mathscr{F}_T$ is actually a σ-field and T is $\mathscr{F}_T$-measurable. Show that if $T(\omega) = t$ for some constant $t \geq 0$ and every $\omega \in \Omega$, then $\mathscr{F}_T = \mathscr{F}_t$.*

Proposition 3.5. *If X is progressively measurable valued in $(E, \mathcal{E})$ and T is a stopping time then $X_T \mathbf{1}(T < \infty)$ is $\mathcal{F}_T$-measurable.*

We set $X_T \mathbf{1}(T < \infty) = 0$ if $T = \infty$.

Proof. Let us suppose $T < +\infty$. For a fixed $t \geq 0$, let us set

$$\tilde{\Omega} = \{\omega | T(\omega) \leq t\}$$

endowed with the sigma field $\mathcal{F}_t$ restricted to $\tilde{\Omega}$ that is $\mathcal{F}_t^{\tilde{\Omega}} \stackrel{def}{=} \mathcal{F}_t \cap \{T(\leq t\}$. The map

$$\tilde{\Omega} \mapsto [0, t] \tag{3.1}$$

$$\omega \mapsto T(\omega) \tag{3.2}$$

is measurable from $\mathcal{F}_t^{\tilde{\Omega}}$ on $\mathcal{B}([0,t])$ because for $s \le t$ $\{T \le s\} \in \mathcal{F}_s$. Then Φ defined by

$$\tilde{\Omega} \mapsto [0,t] \times \Omega \quad (3.3)$$

$$\omega \mapsto (T(\omega), \omega) \quad (3.4)$$

is measurable if we endow $[0,t] \times \Omega$ with the sigma field $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$. Since X is progressively measurable $X_T = X \circ \Phi(\omega)$ is $\mathcal{F}_t^{\tilde{\Omega}}$ measurable i.e. $\forall A \in \mathcal{E}$, $\{\omega | X_{T(\omega)} \in A\} \cap \{\omega | T(\omega) \le t\} \in \mathcal{F}_t$, $\{\omega | T(\omega) < +\infty\}$ is $\mathcal{F}_T$ measurable hence $X_T \mathbf{1}(T < \infty)$ also.

Exercise 3.3. 1. *Show that $\mathcal{F}_T = \sigma(X_T,\ X \text{progressively measurable})$. The previous Proposition yields one inclusion out of two...*

2. *Show that if S, T stopping times and $S \le T$ a.s. $\mathcal{F}_S \subset \mathcal{F}_T$.*

□

3 Martingale in continuous time

In this section we shall consider exclusively real-valued processes $X = \{X_t; 0 \le t < \infty\}$ on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, adapted to a given filtration $\{\mathscr{F}_t\}$ and such that $\mathbb{E}\,|X_t| < \infty$ holds for every $t \ge 0$.

Definition 3.6. *The process $\{X_t, \mathscr{F}_t; 0 \le t < \infty\}$ is said to be a submartingale (respectively, a supermartingale) if, for every $0 \le s < t < \infty$, we have, a.s. $\mathbb{P}$; $\mathbb{E}\,(X_t \mid \mathscr{F}_s) \ge X_s$ (respectively, $\mathbb{E}\,(X_t \mid \mathscr{F}_s) \le X_s$).*

We shall say that $\{X_t, \mathscr{F}_t; 0 \le t < \infty\}$ is a martingale if it is both a submartingale and a supermartingale.

Example. 1. *If $(B_t)_{t \ge 0}$ is a Brownian Motion (BM) we consider the natural filtration $\mathcal{F}_t^B = \sigma(B_s,\ s \le t)$ and $\mathcal{F}_t = \sigma(\mathcal{F}_t^B, \mathcal{N})$, where $\mathcal{N}$ is the set of negligible sets. We take for granted that $\mathcal{F}_t$ is right continuous and hence satisfies the usual conditions. B_t* **is a $\mathcal{F}_t$ martingale** *since for $0 \le s \le t$*

$$\mathbb{E}(B_t | \mathcal{F}_s) = \mathbb{E}(B_t - B_s + B_s | \mathcal{F}_s) = \mathbb{E}(B_t - B_s) + B_s.$$

2. *$(B_t^2 - t)_{t \ge 0}$ is a $\mathcal{F}_t$ martingale. Actually $0 \le s \le t$ we have to show that $\mathbb{E}(B_t^2 - t | \mathcal{F}_s) = B_s^2 - s$, which is equivalent to $\mathbb{E}(B_t^2 - B_s^2 | \mathcal{F}_s) = t - s$. Observe that $\mathbb{E}((B_s + B_t - B_s)^2 - B_s^2 | \mathcal{F}_s) = \mathbb{E}(2B_s(B_t - B_s) + (B_t - B_s)^2 | \mathcal{F}_s) = t - s$.*

3. *$\forall \lambda \in \mathbb{C}$, $M_\lambda(t) = e^{\lambda B_t - \frac{\lambda^2}{2}t}$ is a $\mathcal{F}_t$ martingale. (It is important to allow λ to be complex valued since we will take $\lambda = iu$ where $u \in \mathbb{R}$ which is related to the characteristic function $\mathbb{E}e^{iuB_t}$; the definition of complex valued martingales just amounts to say that the real part and the imaginary part of the process are martingales.) To ensure integrability of $M_\lambda(t)$ we recall that the Laplace transform of Gaussian random variables is always finite. Let us show that $\mathbb{E}(\frac{M_\lambda(t)}{M_\lambda(s)}|\mathcal{F}_s) = 1$.*

$$\begin{aligned}\mathbb{E}(e^{\lambda(B_t - B_s) - \frac{\lambda^2}{2}(t-s)}|\mathcal{F}_s) &= \mathbb{E}(e^{\lambda(B_t - B_s)})e^{-\frac{\lambda^2}{2}(t-s)} \\ &= 1.\end{aligned}$$

Let $X = \{X_t; 0 \leq t < \infty\}$ be a real-valued stochastic process. Consider two numbers $\alpha < \beta$ and a finite subset F of $[0, \infty)$. We define the number of upcrossings $U_F(\alpha, \beta; X(\omega))$ of the interval $[\alpha, \beta]$ by the restricted sample path $\{X_i; t \in F\}$ as follows. Set

$$\tau_1(\omega) = \min\{t \in F; X_t(\omega) \leq \alpha\},$$

and define recursively for $j = 1, 2, \ldots$

$$\begin{aligned}\sigma_j(\omega) &= \min\{t \in F; t \geq \tau_j(\omega), X_t(\omega) > \beta\} \\ \tau_{j+1}(\omega) &= \min\{t \in F; t \geq \sigma_j(\omega), X_t(\omega) < \alpha\}\end{aligned}$$

The convention here is that the minimum of empty set is $+\infty$, and we denote by $U_F(\alpha, \beta; X(\omega))$ the largest integer j for which $\sigma_j(\omega) < \infty$. If $I \subset [0, \infty)$ is not necessarily finite, we define

$$U_I(\alpha, \beta; X(\omega)) = \sup\{U_F(\alpha, \beta; X(\omega)); F \subseteq I, F \text{ is finite }\}$$

The number of downcrossings $D_I(\alpha, \beta; X(\omega))$ is defined similarly.

The following theorem extends to the continuous-time case results of discrete martingales.

Theorem 3.1. *Let $\{X_t, \mathscr{F}_t; 0 \leq t < \infty\}$ be a submartingale whose every path is right-continuous, let $[\sigma, \tau]$ be a subinterval of $[0, \infty)$, and let $\alpha < \beta, \lambda > 0$ be real numbers. We have the following results:*

1. *Submartingale inequality:*

$$\lambda \cdot \mathbb{P}\left[\sup_{\sigma \le t \le \tau} X_t \ge \lambda\right] \le \mathbb{E}\left(X_\tau^+\right).$$

2. *Upcrossings and downcrossings inequalities:*

$$\mathbb{E}U_{[\sigma,\tau]}(\alpha,\beta;X(\omega)) \le \frac{\mathbb{E}\,(X_\tau^+) + |\alpha|}{\beta-\alpha}, \quad \mathbb{E}D_{[\sigma,\tau]}(\alpha,\beta;X(\omega)) \le \frac{\mathbb{E}\,(X_\tau - \alpha)^+}{\beta-\alpha}.$$

3. *Doob's maximal inequality:*

$$\mathbb{E}\left(\sup_{\sigma \le t \le \tau} X_t\right)^p \le \left(\frac{p}{p-1}\right)^p \mathbb{E}\,(X_\tau^p), \quad p > 1$$

provided $X_t \ge 0$ *a.s.* $\mathbb{P}$ *for every* $t \ge 0$*, and* $\mathbb{E}\,(X_\tau^p) < \infty$*.*

There exist modifications of martingales, which are right continuous.

Theorem 3.2. *If* X_t *is a* $\mathcal{F}_t$ *(with the usual conditions) sub-martingale there exists a modification of* X_t *with right continuous paths iff* $t \mapsto \mathbb{E}X_t$ *is right continuous. Moreover this modification has paths with left limits and it is a sub-martingale.*

Corollary 3.1. *If* X_t *is a* $\mathcal{F}_t$ *(with the usual conditions) martingale there exists a modification of* X_t *with r.c.l.l. right continuous left limits paths and it is a martingale*

The proof of the Theorem 3.2 can be found in Karatzas and Shreeve. It uses the upcrossings inequality and the backward martingales that I did not recalled.

We have convergence results for martingales in continuous time similar to those in discrete time. We always use the right continuous modification.

Theorem 3.3. *1. If* X_t *is a* $\mathcal{F}_t$ *sub-martingale and* $\sup \mathbb{E}(X_t^+) < \infty$ *then* $\lim_{t\to+\infty} X_t = X_\infty$ *exists a.s. and* $\mathbb{E}|X_\infty| < \infty$*.*

2. *If* X_t *is a* $\mathcal{F}_t$ *martingale the following properties are equivalent.*

   (a) X_t *converges in* L^1 *towards* X_∞*.*

   (b) X_t *is uniformly integrable.*

   (c) $\exists X_\infty \in L^1$ *and* $X_t = \mathbb{E}[X_\infty/\mathcal{F}_t]$*.*

Under any of this hypothesis we have also a.s. convergence of X_t *when* $t \to \infty$.

Please remark the slightly different assumption for the first part of the theorem, when it is compared to the result for martingales indexed by $\mathbb{N}$. This assumption works also for martingales indexed by $\mathbb{N}$.

4 Stopping time theorems

Let us generalize the stopping time theorems already obtained for martingales indexed by $\mathbb{N}$ to martingales in continuous time. Here we consider the case when X_t is uniformly integrable.

Definition 3.7. *If* X_t *is a martingale uniformly integrable and* T *a stopping time, let us define* $X_T(\omega) = X_\infty(\omega)$ *if* $T(\omega) = \infty$ *and* $X_T(\omega) = X_{T(\omega)}(\omega)$ *if* $T(\omega) < \infty$.

Theorem 3.4. *If* X_t *is a martingale uniformly integrable then the family of* (X_S) *where* S *is any stopping time is also uniformly integrable. If* $S \leq T$ *are two stopping times* $X_S = \mathbb{E}(X_T|\mathcal{F}_S) = \mathbb{E}(X_\infty|\mathcal{F}_S)$.

Proof. Let us first show
$$X_S = \mathbb{E}(X_\infty|\mathcal{F}_S). \tag{3.5}$$

For the first step I assume that S only takes a finite number of values $\{s_1 < \ldots < s_n\}$, and add by convention $s_0 = -\infty$. Let $Y_k = X_{s_k}$ for $k \leq n$ and $Y_{n+1} = X_\infty$. Then Y is a martingale for $\mathcal{F}_{s_k}$ and $Y_k = \mathbb{E}(X_\infty|\mathcal{F}_{s_k})$. Let $A \in \mathcal{F}_S$, i.e. $A \cap \{S \leq t\} \in \mathcal{F}_t$

$$\begin{aligned} X_S \mathbf{1}_A &= \sum_{k=1}^{n+1} Y_k \mathbf{1}_{\{S=s_k\}\cap A} \\ &= \sum_{k=1}^{n} Y_k (\mathbf{1}_{\{S\leq s_k\}\cap A} - \mathbf{1}_{\{S\leq s_{k-1}\}\cap A}) + Y_{n+1}\mathbf{1}_{\{S>s_n\}\cap A}. \end{aligned}$$

Hence

$$\begin{aligned} \mathbb{E}(X_S\mathbf{1}_A) &= \sum_{k=1}^{n+1} \mathbb{E}(X_\infty(\mathbf{1}_{\{S\leq s_k\}\cap A} - \mathbf{1}_{\{S\leq s_{k-1}\}\cap A}) + \mathbb{E}(X_\infty\mathbf{1}_{\{S>s_n\}\cap A}) \\ &= \mathbb{E}(X_\infty\mathbf{1}_A). \end{aligned}$$

In the general case S is the limit of a non increasing sequence of stopping times S_n that take a finite number of values and $\forall n \in \mathbb{N},\ X_{S_n} = \mathbb{E}(X_\infty|\mathcal{F}_{S_n})$. Since X is right continuous, $X_S = \lim_{n\to\infty} X_{S_n}$. Moreover $\mathcal{F}_{S_{n+1}} \subset \mathcal{F}_{S_n}$ hence X_{S_n} is a backward martingale and we admit that every backward martingale converges in L^1. So

$$X_S = \lim_{N\to\infty} \mathbb{E}(X_\infty| \cap_{n=1}^N \mathcal{F}_{S_n})$$

in L^1. Since $\mathcal{F}_S \subset \cap_{n=1}^N \mathcal{F}_{S_n}$,

$$\begin{aligned} X_S &= \mathbb{E}(\lim_{N\to\infty} \mathbb{E}(X_\infty| \cap_{n=1}^N \mathcal{F}_{S_n})|\mathcal{F}_S) \\ &= \lim_{N\to\infty} \mathbb{E}(\mathbb{E}(X_\infty| \cap_{n=1}^N \mathcal{F}_{S_n})|\mathcal{F}_S) \quad = \lim_{N\to\infty} \mathbb{E}(X_\infty|\mathcal{F}_S) = \mathbb{E}(X_\infty|\mathcal{F}_S). \end{aligned}$$

□

Since $\mathbb{E}(X_\infty|\mathcal{F}_S)$ is uniformly integrable so the family of (X_S) where S is any stopping time is also uniformly integrable. Moreover, if $S \le T$,

$$\mathbb{E}(X_T|\mathcal{F}_S) = \mathbb{E}(\mathbb{E}(X_\infty|\mathcal{F}_T)|\mathcal{F}_S) = \mathbb{E}(X_\infty|\mathcal{F}_S) = X_S$$

because $\mathcal{F}_S \subset \mathcal{F}_T$.

Corollary 3.2. *If X is a martingale and $S \le T$ bounded stopping times then $\mathbb{E}(X_T|\mathcal{F}_S)) = X_S$.*

Proof. Let $M > 0$ such that $0 \le S(\omega) \le T(\omega) \le M$ a.s. Let us define $Y_t = X_{t\wedge M}, \forall t \ge 0$. Then $Y_t = \mathbb{E}(Y_M|\mathcal{F}_t)$, it is uniformly integrable and we apply the previous theorem. Then $\mathbb{E}(Y_T|\mathcal{F}_S) = Y_S$ which is also $\mathbb{E}(X_{T\wedge M}|\mathcal{F}_S) = X_{S\wedge M}$. Hence $\mathbb{E}(X_T|\mathcal{F}_S) = X_S$ since $0 \le S(\omega) \le T(\omega) \le M$. □

Remark 3.2. *But, if X is not uniformly integrable, one cannot only suppose that $S \le T$ finite a.s.*

Exercise 3.4. *Let $B_t,\ t \ge 0$ be a BM starting from 0 and for $a > 0$ $T_a = \inf\{t > 0,\ B_T = a\}$. Here we assume that $T_a < \infty$ a.s. Compute $\mathbb{E}(e^{-\mu T_a})$ for $\mu \ge 0$. Deduce that $\mathbb{E}T_A = +\infty$.*

Solution :

Let $M^\lambda(t) = e^{\lambda B_t - \frac{\lambda^2}{2}t}$ be the exponential martingale for $\lambda \ge 0$. The stopping time $t \wedge T_a$ is bounded. Since $s \wedge T_a \le t \wedge T_a$, then

$$\mathbb{E}(M^\lambda_{t\wedge T_a}|\mathcal{F}_{s\wedge T_a}) = M^\lambda_{s\wedge T_a}.$$

Hence $M^\lambda_{t\wedge T_a}$ is a martingale. So $\mathbb{E}M^\lambda_{t\wedge T_a} = \mathbb{E}M^\lambda_0 = 1$. In

$$\mathbb{E}(e^{\lambda B_{t\wedge T_a} - \frac{\lambda^2}{2}(t\wedge T_a)}) = 1,$$

we let $t \to +\infty$, use $T_a < \infty$. a.s. and a dominated convergence argument, so that $\mathbb{E}(e^{\lambda B_{T_a} - \frac{\lambda^2}{2}T_a}) = 1$. Then we remark that $B_{T_a} = a$ and we get

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

Let us take $\mu = \frac{\lambda^2}{2}$, then $\mathbb{E}(e^{-\mu T_a}) = e^{-\sqrt{2\mu}a}$. We have the Laplace transform of the positive random variable T_a. Classically $\mathbb{E}T_A$ is obtained as the derivative of this Laplace transform for $\mu = 0$ which is here $+\infty$. Hence $\mathbb{E}T_A = +\infty$.

Theorem 3.5. *If X_t is a non-negative supermartingale and $S \leq T$ are two stopping times $X_S \geq \mathbb{E}(X_T|\mathcal{F}_S)$.*

Remark 3.3. *We know that the $\lim_{t\to+\infty} = X_\infty$ exists a.s. so we don't need to have the stopping time a.s. finite in this case.*

We admit the proof.

Chapter 4

Stochastic integral

The results for martingales are used to build $X_t = \int_0^t H_s dMs$ where H is progressively measurable and M is a martingale. Because of the martingale transform in the discrete case we expect X_t to be a martingale. But the real life is more complicated for integrability reasons... Hence we are forced to define local martingale associated to a sequence of stopping times i.e. we assume that there is a non decreasing sequence of stopping times T_n such that $X_{t \wedge T_n}$ is a martingale uniformly integrable. Our basic tool will be quadratic variations that we will generalize from Brownian motion to continuous martingales. In this part we assume the processes are a.s. continuous.

1 Quadratic variations

Using Riemmann Stieljes integral we know how to integrate with respect to processes A with finite variations. We will show that we can define an integral with respect to local martingales M. Once we have done that we will be able to define integrals with respect to processes of the form $M + A$... That is the goal of the chapter. Let us come back to a technical question "Do the martingales have finite variations ?". First we show that a process cannot be a continuous martingale and with finite variations unless it is trivial.

Proposition 4.1. *Every continuous martingale with finite variations is constant.*

Proof. Let Δ be a subdivision of $[0, t]$ and

$$V_t(f) = \sup_{\Delta \in \mathcal{P}_t} \sum_{i=1}^{n} |M_{t_{i+1}} - M_{t_i}|$$

where M is a continuous martingale with finite variations. Let $T_n = \inf\{t \geq 0,\ V_t \geq n \text{ } or \text{ } |M_t| \geq n\}$. As a hitting time T_n is a stopping time and we can show by contradiction that T_n almost surely converges to $+\infty$. We denote by $M_t^{T_n} = M_{t \wedge T_n}$ the stopped martingale : it is a bounded by n continuous process. It is also a martingale because of the stopping time theorem for bounded stopping time. If $s \leq t,\ \ s \wedge T_n \leq t \wedge T_n$

$$\mathbb{E}(M_{t \wedge T_n} | \mathcal{F}_{s \wedge T_n}) = M_{s \wedge T_n}.$$

Hence we can assume without a loss of generality that M is a bounded continuous martingale with bounded variations. If $\Delta \in \mathcal{P}_t$, by orthogonality of the increments of L^2 martingales

$$\begin{aligned}\mathbb{E}(M_t - M_s)^2 &= \sum_{t_i \in \Delta} \mathbb{E}(M_{t_{i+1}} - M_{t_i})^2 \\ &= \mathbb{E}(V_t \sup_{t_i \in \Delta} |M_{t_{i+1}} - M_{t_i}|).\end{aligned}$$

Since M is uniformly continuous on $[0, t]$ $M_t = M_0$ a.s. □

We have checked that $B_t^2 - t$ is martingale, we will show that $M^2 - \langle M, M \rangle$ is a martingale. In this chapter we consider Δ with an infinite number of points t_i such that $\lim_{i \to \infty} t_i = +\infty$ and rewrite $T_{[0,t]}^{\Delta} \stackrel{def}{=} \sum_{t_i \in \Delta} (X_{t_{i+1} \wedge t} - X_{t_i \wedge t})^2$.

Theorem 4.1. *A continuous and bounded martingale M is of finite quadratic variation and $\langle M, M \rangle$ is the unique continuous non decreasing adapted process vanishing at zero such that $M^2 - \langle M, M \rangle$ is a martingale.*

Proof. Uniqueness is an easy consequence of Proposition 4.1, since if there were two such processes A and B, then $A - B$ would be a continuous martingale of vanishing at zero with finite variations.

To prove the existence of $\langle M, M \rangle$, we first observe that since for $t_i < s < t_{i+1}$,

$$E\left[\left(M_{t_{i+1}} - M_{t_i}\right)^2 \mid \mathcal{F}_s\right] = E\left[\left(M_{t_{i+1}} - M_s\right)^2 \mid \mathscr{F}_s\right] + (M_s - M_{t_i})^2$$

it is easily proved that

$$E\left[T_t^{\Delta}(M) - T_s^{\Delta}(M) \mid \mathscr{F}_s\right] = E\left[(M_t - M_s)^2 \mid \mathscr{F}_s\right]$$
$$= E\left[M_t^2 - M_s^2 \mid \mathscr{F}_s\right] \tag{4.1}$$

As a result, $M_t^2 - T_t^{\Delta}(M)$ is a continuous martingale. In the sequel, we write T_t^{Δ} instead of $T_t^{\Delta}(M)$.

We now fix $a > 0$ and we are going to prove that if $\{\Delta_n\}$ is a sequence of subdivisions of $[0, a]$ such that $|\Delta_n|$ goes to zero, then $\left\{T_a^{\Delta_n}\right\}$ converges in L^2.

If Δ and Δ' are two subdivisions we call $\Delta\Delta'$ the subdivision obtained by taking all the points of Δ and Δ'. By (4.1) the process $X = T^{\Delta} - T^{\Delta'}$ is a martingale and, by (4.1) again, applied to X instead of M, we have

$$E\left[X_a^2\right] = E\left[\left(T_a^{\Delta} - T_a^{\Delta'}\right)^2\right] = E\left[T_a^{\Delta\Delta'}(X)\right]$$

Because $(x+y)^2 \leq 2\left(x^2 + y^2\right)$ for any pair (x, y) of real numbers,

$$T_a^{\Delta\Delta'}(X) \leq 2\left\{T_a^{\Delta\Delta'}\left(T^{\Delta}\right) + T_a^{\Delta\Delta'}\left(T^{\Delta'}\right)\right\}$$

and to prove our claim, it is enough to show that $E\left[T_a^{\Delta\Delta'}\left(T^{\Delta}\right)\right]$ converges to 0 as $|\Delta| + |\Delta'|$ goes to zero.

Let then s_k be in $\Delta\Delta'$ and t_l be the rightmost point of Δ such that $t_l \leq s_k < s_{k+1} \leq t_{l+1}$; we have

$$T_{s_{k+1}}^{\Delta} - T_{s_k}^{\Delta} = \left(M_{s_{k+1}} - M_{t_l}\right)^2 - \left(M_{s_k} - M_{t_l}\right)^2$$
$$= \left(M_{s_{k+1}} - M_{s_k}\right)\left(M_{s_{k+1}} + M_{s_k} - 2M_{t_l}\right)$$

and consequently,

$$T_a^{\Delta\Delta'}\left(T^{\Delta}\right) \leq \left(\sup_k \left|M_{s_{k+1}} + M_{s_k} - 2M_{t_t}\right|^2\right) T_a^{\Delta\Delta'}$$

By Schwarz's inequality,

$$E\left[T_a^{\Delta\Delta'}\left(T^{\Delta}\right)\right] \leq E\left[\sup_k \left|M_{s_{k+1}} + M_{s_k} - 2M_{t_l}\right|^4\right]^{1/2} E\left[\left(T_a^{\Delta\Delta'}\right)^2\right]^{1/2}$$

Whenever $|\Delta| + |\Delta'|$ tends to zero, the first factor goes to zero because M is continuous; it is therefore enough to prove that the second factor is bounded by a constant independent of Δ and Δ'. To this end, we write with $a = t_n$,

$$\begin{aligned}\left(T_a^{\Delta}\right)^2 &= \left(\sum_{k=1}^{n}\left(M_{t_k} - M_{t_{k-1}}\right)^2\right)^2 \\ &= 2\sum_{k=1}^{n}\left(T_a^{\Delta} - T_{t_k}^{\Delta}\right)\left(T_{t_k}^{\Delta} - T_{t_{k-1}}^{\Delta}\right) + \sum_{k=1}^{n}\left(M_{t_k} - M_{t_{k-1}}\right)^4\end{aligned}$$

Because of (4.1), we have $E\left[T_a^{\Delta} - T_{t_k}^{\Delta} \mid \mathscr{F}_{t_k}\right] = E\left[(M_a - M_{t_k})^2 \mid \cdot\mathscr{F}_{t_k}\right]$ and consequently

$$\begin{aligned}E\left[\left(T_a^{\Delta}\right)^2\right] =& 2\sum_{k=1}^{n} E\left[(M_a - M_{t_k})^2\left(T_{t_k}^{\Delta} - T_{t_{k-1}}^{\Delta}\right)\right] \\ &+ \sum_{k=1}^{n} E\left[\left(M_{t_k} - M_{t_{k-1}}\right)^4\right] \\ \leq& E\left[\left(2\sup_k |M_a - M_{t_k}|^2 + \sup_k \left|M_{t_k} - M_{t_{k-1}}\right|^2\right) T_a^{\Delta}\right]\end{aligned}$$

Let C be a constant such that $|M| \leq C$; by (4.1), it is easily seen that $E\left[T_a^{\Delta}\right] \leq 4C^2$ and therefore

$$E\left[\left(T_a^{\Delta}\right)^2\right] \leq 12C^2 E\left[T_a^{\Delta}\right] \leq 48C^4.$$

We have thus proved that for any sequence $\{\Delta_n\}$ such that $|\Delta_n| \to 0$, the sequence $\left\{T_a^{\Delta_n}\right\}$ has a limit $\langle M, M\rangle_a$ in L^2 hence in probability. It remains to prove that $\langle M, M\rangle_a$ may be chosen within its equivalence class in such a way that the resulting process $\langle M, M\rangle$ has the required properties.

Let $\{\Delta_n\}$ be as above; by Doob's inequality applied to the martingale $T^{\Delta_n} - T^{\Delta_m}$,

$$E\left[\sup_{t\le a}\left|T_t^{\Delta_n} - T_t^{\Delta_m}\right|^2\right] \le 4E\left[\left(T_a^{\Delta_n} - T_a^{\Delta_m}\right)^2\right].$$

Since, from a sequence converging in L^2, one can extract a subsequence converging a.s., there is a subsequence $\{\Delta_{n_k}\}$ such that $T_t^{\Delta_{e_k}}$ converges a.s. uniformly on $[0,a]$ to a limit $\langle M,M\rangle_t$ which perforce is a.s. continuous. Moreover, the original sequence might have been chosen such that Δ_{n+1} be a refinement of Δ_n and $\bigcup_n \Delta_n$ be dense in $[0,a]$. For any pair (s,t) in $\bigcup_n \Delta_n$ such that $s<t$, there is an n_0 such that s and t belong to Δ_n for any $n \ge n_0$. We then have $T_s^{\Delta_n} \le T_t^{\Delta_n}$ and as a result $\langle M,M\rangle$ is non decreasing on $\bigcup_n \Delta_n$; as it is continuous, it is increasing everywhere (although the T^{Δ_n} are not necessarily non decreasing).

Finally, that $M^2 - \langle M,M\rangle$ is a martingale follows upon passing to the limit in (4.1). The proof is thus complete. □

To enlarge the scope of the above result we will need the

Proposition 4.2. *For every stopping time T,*

$$\left\langle M^T, M^T\right\rangle = \langle M,M\rangle^T$$

Much as it is interesting, Theorem 4.1 is not sufficient for our purposes; it does not cover, for instance, the case of the Brownian motion B which is not a bounded martingale. Nonetheless, we have seen that B has a "quadratic variation", namely t, and that $B_t^2 - t$ is a martingale exactly as in Theorem 4.1. We now show how to subsume the case of BM and the case of bounded martingales in a single result by using the fecund idea of localization.

Definition 4.1. *An adapted, right-continuous process X is an $(\mathscr{F}_t,\mathbb{P})$-local martingale if there exist stopping times $T_n, n\ge 1$, such that i) the sequence $\{T_n\}$ is increasing and $\lim_n T_n = +\infty$ a.s.; ii) for every n, the process $X^{T_n}1_{[T_n>0]}$ is a uniformly integrable $(\mathscr{F}_t,\mathbb{P})$-martingale.*

We will drop $(\mathscr{F}_t,\mathbb{P})$ when there is no risk of ambiguity. In condition ii) we can drop the uniform integrability and ask only that $X^{T_n}1_{[T_n>0]}$ be a martingale; indeed, one can always replace T_n by $T_n \wedge n$ to obtain a u.i. martingale. Likewise, if X is continuous as will nearly always be in this book, by setting $S_n = \inf\{t : |X_t| = n\}$ and replacing T_n by $T_n \wedge S_n$, we may assume the martingales in ii) to be bounded. This will be used extensively in the sequel.

We further say that the stopping time T reduces X if $X^T 1_{[T>0]}$ is a u.i. martingale. This property can be decomposed in two parts if one introduces the process $Y_t = X_t - X_0$: T reduces X if and only if

i) X_0 is integrable on $\{T > 0\}$;

ii) Y^T is a u.i. martingale.

A common situation however is that in which X_0 is constant this explains why in the sequel we will often drop the qualifying $1_{[T>0]}$. As an exercise, the reader will show the following simple properties :

i) if T reduces X and $S \le T$, then S reduces X;

ii) the sum of two local martingales is a local martingale;

iii) if Z is a $\mathscr{F}_0$-measurable r.v. and X is a local martingale then, so is ZX; in particular, the set of local martingales is a vector space;

iv) a stopped local martingale is a local martingale;

v) a positive local martingale is a supermartingale.

We can now extend the quadratic variations to local martingales.

Theorem 4.2. *If M is a continuous local martingale, there exists a unique increasing continuous process $\langle M, M\rangle$, vanishing at zero, such that $M^2 - \langle M, M\rangle$ is a continuous local martingale. Moreover, for every t and for any sequence $\{\Delta_n\}$ of subdivisions of $[0, t]$ such that $|\Delta_n| \to 0$, the r.v.'s*

$$\sup_{s \le t} \left| T_s^{\Delta_n}(M) - \langle M, M\rangle_s \right|$$

converge to zero in probability.

Proof. Let $\{T_n\}$ be a sequence of stopping times increasing to $+\infty$ and such that $X_n = M^{T_n} 1_{[T_n>0]}$ is a bounded martingale. By Theorem 4.1, there is, for each n, a continuous process A_n with finite variations vanishing at zero and such that $X_n^2 - A_n$ is a martingale. Now, $\left(X_{n+1}^2 - A_{n+1}\right)^{T_n} 1_{[T_n>0]}$ is a martingale and is equal to $X_n^2 - A_{n+1}^{T_n} 1_{[T_n>0]}$. By the uniqueness property in Theorem 4.1, we have $A_{n+1}^{T_n} = A_n$ on $[T_n > 0]$ and we may therefore define unambiguously a process $\langle M, M\rangle$ by setting it equal to A_n on $[T_n > 0]$. Obviously, $\left(M^{T_n}\right)^2 1_{[T_n>0]} - \langle M, M\rangle^{T_n}$ is a martingale and therefore $\langle M, M\rangle$ is the sought-after process. The uniqueness follows from the uniqueness on each interval $[0, T_n]$.

To prove the second statement, let $\delta, \varepsilon > 0$ and t be fixed. One can find a stopping time S such that $M^S 1_{[S>0]}$ is bounded and $P[S \leq t] \leq \delta$. Since $T^\Delta(M)$ and $\langle M, M\rangle$ coincide with $T^\Delta\left(M^S\right)$ and $\left\langle M^S, M^S\right\rangle$ on $[0, S]$, we have

$$P\left[\sup_{s\leq t}\left|T_s^\Delta(M) - \langle M, M\rangle_s\right| > \varepsilon\right] \leq \delta + P\left[\sup_{s\leq t}\left|T_s^\Delta\left(M^S\right) - \left\langle M^S, M^s\right\rangle_s\right| > \varepsilon\right]$$

and the last term goes to zero as $|\Delta|$ tends to zero. □

Theorem 4.2 may still be further extended by polarization.

Theorem 4.3. *If M and N are two continuous local martingales, there exists a unique continuous process $\langle M, N\rangle$ in t, with finite variations, vanishing at zero and such that $MN - \langle M, N\rangle$ is a local martingale. Moreover, for any t and any sequence $\{\Delta_n\}$ of subdivisions of $[0, t]$ such that $|\Delta_n| \to 0$,*

$$\mathbb{P} - \lim \sup_{s\leq t}\left|\tilde{T}_s^{\Delta_n} - \langle M, N\rangle_s\right| = 0$$

where $\tilde{T}_s^{\Delta_n} = \sum_{t_i\in\Delta_n}\left(M_{t_{i+1}}^s - M_{t_i}^s\right)\left(N_{t_{i+1}}^s - N_{t_i}^s\right)$.

Proof. The uniqueness follows again from Proposition 4.1 after suitable stoppings. Moreover the process

$$\langle M, N\rangle = \frac{1}{4}[\langle M + N, M + N\rangle - \langle M - N, M - N\rangle]$$

is easily seen to have the desired properties. □

Definition 4.2. *The process $\langle M, N\rangle$ is called the bracket of M and N, the process $\langle M, M\rangle$ the increasing process associated with M or simply the increasing process of M.*

Proposition 4.3. *If T is a stopping time,*

$$\left\langle M^T, N^T\right\rangle = \left\langle M, N^T\right\rangle = \langle M, N\rangle^T$$

Proof. This is an obvious consequence of the last part of Theorem 4.3. As an exercise, the reader may also observe that $M^T N^T - \langle M, N\rangle^T$ and $M^T\left(N - N^T\right)$ are local martingales, hence by difference, so is $M^T N - \langle M, N\rangle^T$.

□

The properties of the bracket operation are reminiscent of those of a scalar product. The map $(M, N) \to \langle M, N\rangle$ is bilinear, symmetric and $\langle M, M\rangle \geq 0$; it is also non-degenerate as is shown by the following

Proposition 4.4. $\langle M, M\rangle = 0$ *if and only if* M *is constant, that is* $M_t = M_0$ *a.s. for every* t.

Proof. By Proposition 4.3, it is enough to consider the case of a bounded M and then by Theorem 4.1, $E\left[(M_t - M_0)^2\right] = E\left[\langle M, M\rangle_t\right]$; the result follows immediately. □

This property may be extended in the following way.

Proposition 4.5. *The intervals of constancy are the same for* M *and for* $\langle M, M\rangle$*, that is to say, for almost all* ω *'s,* $M_t(\omega) = M_a(\omega)$ *for* $a \leq t \leq b$ *if and only if* $\langle M, M\rangle_b(\omega) = \langle M, M\rangle_a(\omega)$

The following inequality will be very useful in defining stochastic integrals. It shows in particular that $d\langle M, N\rangle$ is absolutely continuous with respect to $d\langle M, M\rangle$.

Definition 4.3. *A real-valued process* H *is said to be measurable if the map* $(\omega, t) \to H_t(\omega)$ *is* $\mathscr{F}_\infty \otimes \mathscr{B}(\mathbb{R}_+)$*-measurable.*

The class of measurable processes is obviously larger than the class of progressively measurable processes.

Proposition 4.6. *For any two continuous local martingales* M *and* N *and measurable processes* H *and* K*, the inequality*

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

holds a.s. for $t \leq \infty$.

Proof. By taking increasing limits, it is enough to prove the inequality for $t < \infty$ and for bounded H and K. Moreover, it is enough to prove the inequality where the left-hand side has been replaced by

$$\left| \int_0^t H_s K_s d\langle M, N\rangle_s \right|$$

indeed, if J_s is a density of $d\langle M, N\rangle_s / |d\langle M, N\rangle|_s$ with values in $\{-1, 1\}$ and we replace H by $HJ\,\mathrm{sgn}(HK)$ in this expression, we get the left-hand side of the statement.

By a density argument, it is enough to prove that for those K 's which may be written

$$K = K_0 1_{\{0\}} + K_1 1_{]0,t_1]} + \ldots + K_n 1_{\}]t_{n-1},t_n]}$$

for a finite subdivision $\{t_0 = 0 < t_1 < \ldots < t_n = t\}$ of $[0, t]$ and bounded measurable r.v. K_i 's. By another density argument, we can also take H of the same form and with the same subdivision. □

If we now define $\langle M, N\rangle_s^t = \langle M, N\rangle_t - \langle M, N\rangle_s$, we have

$$\left|\langle M, N\rangle_s^t\right| \leq \left(\langle M, M\rangle_s^t\right)^{1/2} \left(\langle N, N\rangle_s^t\right)^{1/2}$$

Indeed, almost surely, the quantity

$$\langle M, M\rangle_s^t + 2r\langle M, N\rangle_s^t + r^2\langle N, N\rangle_s^t = \langle M + rN, M + rN\rangle_s^t$$

is non-negative for every $r \in \mathbb{Q}$, hence by continuity for every $r \in \mathbb{R}$, and our claim follows from the usual quadratic form reasoning.

As a result

$$\begin{aligned} \left| \int_0^t H_s K_s d\langle M, N\rangle_s \right| &\leq \sum_i |H_i K_i| \left|\langle M, N\rangle_{t_i}^{t_{i+1}}\right| \\ &\leq \sum_i |H_i|\,|K_i| \left(\langle M, M\rangle_{t_i}^{t_{i+1}}\right)^{1/2} \left(\langle N, N\rangle_{t_i}^{t_{i+1}}\right)^{1/2} \end{aligned}$$

and using the Cauchy-Schwarz inequality for the summation over i, this is still less than

$$\left(\sum_i H_i^2 \langle M, M\rangle_{t_i}^{t_{i+1}}\right)^{1/2} \left(\sum_i K_i^2 \langle N, N\rangle_{t_i}^{t_{i+1}}\right)^{1/2}$$
$$= \left(\int_0^t H_s^2 d\langle M, M\rangle_s\right)^{1/2} \left(\int_0^t K_s^2 d\langle N, N\rangle_s\right)^{1/2}$$

which completes the proof.

Corollary 4.1 (Kunita-Watanabe inequality)**.** *For every $p \geq 1$ and $p^{-1}+q^{-1} = 1$,*

$$E\left[\int_0^\infty |H_s|\,|K_s|\,|d\langle M, N\rangle|_s\right]$$
$$\leq \left\|\left(\int_0^\infty H_s^2 d\langle M, M\rangle_s\right)^{1/2}\right\|_p \left\|\left(\int_0^\infty K_s^2 d\langle N, N\rangle_s\right)^{1/2}\right\|_q$$

Proof. Straightforward application of Hölder's inequality. □

We now introduce a fundamental class of processes of finite quadratic variation.

Definition 4.4. *A continuous $(\mathscr{F}_t, P)$-semimartingale is a continuous process X which can be written $X = M + A$ where M is a continuous $(\mathscr{F}_t, P)$-local martingale and A a continuous adapted process of finite variation.*

As usual, we will often drop $(\mathscr{F}_t, P)$ and we will use the abbreviation cont. semi. mart. The decomposition into a local martingale and a finite variation process is unique as follows readily from Proposition 4.1; however, if a process X is a continuous semimartingale in two different filtrations $(\mathscr{F}_t)$ and $(\mathscr{G}_t)$, the decompositions may be different even if $\mathscr{F}_t \subset \mathscr{G}_t$ for each t.

Proposition 4.7. *A continuous semimartingale $X = M + A$ has a finite quadratic variation and $\langle X, X\rangle = \langle M, M\rangle$.*

Proof. If Δ is a subdivision of $[0, t]$,

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

where $\operatorname{Var}_t(A)$ *is the variation of* A *on* $[0,t]$*, and this converges to zero when* $|\Delta|$ *tends to zero because of the continuity of* M*. Likewise*

$$\lim_{|\Delta|\to 0} \sum_i \left(A_{t_{i+1}} - A_{t_i}\right)^2 = 0$$

Fundamental remark. Since the process $\langle X, X\rangle$ is the limit in probability of the sums $T^{\Delta_n}(X)$, it does not change if we replace $(\mathscr{F}_t)$ by another filtration for which X is still a semimartingale and likewise if we change P for a probability measure Q such that $Q \ll P$ and X is still a Q-semimartingale.

Definition 4.5. *If* $X = M + A$ *and* $Y = N + B$ *are two continuous semimartingales, we define the bracket of* X *and* Y *by*

$$\langle X, Y\rangle = \langle M, N\rangle = \frac{1}{4}[\langle X+Y, X+Y\rangle - \langle X-Y, X-Y\rangle]$$

Obviously, $\langle X, Y\rangle_t$ is the limit in probability of $\sum_i \left(X_{t_{i+1}} - X_{t_i}\right)\left(Y_{t_{i+1}} - Y_{t_i}\right)$, and more generally, if H is left-continuous and adapted,

$$P - \lim_{|\Delta|\to 0} \sup_{s\leq t} \left|\sum_i H_{t_i}\left(X^s_{t_{i+1}} - X^s_{t_i}\right)\left(Y^s_{t_{i+1}} - Y^s_{t_i}\right) - \int_0^s H_u d\langle X, Y\rangle_u\right| = 0$$

the proof of which is left to the reader as an exercise.

Finally, between the class of local martingales and that of bounded martingales, there are several interesting classes of processes among which the following ones will be particularly important. We will indulge in the usual confusion between processes and classes of indistinguishable processes in order to get norms and not merely semi-norms in the discussion below.

Definition 4.6. *We denote by* $\mathbb{H}^2$ *the space of* L^2*-bounded martingales, i.e. the space of* $(\mathscr{F}_t, P)$*-martingales* M *such that*

$$\sup_t E\left[M_t^2\right] < +\infty$$

We denote by H^2 the subset of L^2-bounded continuous martingales, and H_0^2 the subset of elements of H^2 vanishing at zero.

An $(\mathscr{F}_t)$-Brownian motion is not in H^2, but it is when suitably stopped, for instance at a constant time. Bounded martingales are in $\mathbb{H}^2$. Moreover, by Doob's inequality, $M_\infty^* = \sup_t |M_t|$ is in L^2 if $M \in \mathbb{H}^2$; hence M is u.i. and $M_t = E[M_\infty \mid \mathscr{T}_t]$ with $M_\infty \in L^2$. This sets up a one to one correspondence between $\mathbb{H}^2$ and $L^2(\Omega, \mathscr{F}_\infty, P)$, and we have the

Proposition 4.8. *The space $\mathbb{H}^2$ is a Hilbert space for the norm*

$$\|M\|_{\mathbb{H}^2} = E\left[M_\infty^2\right]^{1/2} = \lim_{t\to\infty} E\left[M_t^2\right]^{1/2},$$

and the set H^2 is closed in $\mathbb{H}^2$.

Proof. The first statement is obvious; to prove the second, we consider a sequence $\{M^n\}$ in H^2 converging to M in $\mathbb{H}^2$. By Doob's inequality,

$$E\left[\left(\sup_t |M_t^n - M_t|\right)^2\right] \leq 4\,\|M^n - M\|_{\mathrm{H}^2}^2$$

as a result, one can extract a subsequence for which $\sup_t |M_t^{n_k} - M_t|$ converges to zero a.s. which proves that $M \in H^2$. □

The mapping $M \to \|M_\infty^*\|_2 = E\left[(\sup_t |M_t|)^2\right]^{1/2}$ is also a norm on $\mathbb{H}^2$; it is equivalent to $\|\quad\|_{H^2}$ since obviously $\|M\|_{H^2} \leq \|M_\infty^*\|_2$ and by Doob's inequality $\|M_\infty^*\|_2 \leq 2\|M\|_{H^2}$, but it is no longer a Hilbert space norm.

We now study the quadratic variation of the elements of H^2.

Proposition 4.9. *A continuous local martingale M is in H^2 if and only if the following two conditions hold i) $M_0 \in L^2$; ii) $\langle M, M\rangle$ is integrable i.e. $E[\langle M, M\rangle_\infty] < \infty$.*

In that case, $M^2 - \langle M, M\rangle$ is uniformly integrable and for any pair $S \leq T$ of stopping times

$$E\left[M_T^2 - M_S^2 \mid \mathscr{F}_S\right] = E\left[(M_T - M_S)^2 \mid \mathscr{F}_S\right] = E\left[\langle M, M\rangle_S^T \mid \mathscr{F}_S\right]$$

Proof. Let $\{T_n\}$ be a sequence of stopping times increasing to $+\infty$ and such that $M^{T_n} 1_{[T_n>0]}$ is bounded; we have

$$E\left[M^2_{T_n\wedge t} 1_{[T_n>0]}\right] - E\left[\langle M, M\rangle_{T_n\wedge t} 1_{[T_n>0]}\right] = E\left[M_0^2 1_{[T_n>0]}\right]$$

If M is in H^2 then obviously i) holds and, since $M^*_\infty \in L^2$, we may also pass to the limit in the above equality to get

$$E\left[M^2_\infty\right] - E\left[\langle M, M\rangle_\infty\right] = E\left[M_0^2\right]$$

which proves that ii) holds. If, conversely, i) and ii) hold, the same equality yields

$$E\left[M^2_{T_n\wedge 1} 1_{[T_n>0]}\right] \le E\left[\langle M, M\rangle_\infty\right] + E\left[M_0^2\right] = K < \infty$$

and by Fatou's lemma

$$E\left[M_t^2\right] \le \liminf_n E\left[M^2_{T_n\wedge t} 1_{[T_n>0]}\right] \le K$$

which proves that the family of r.v.'s M_t is bounded in L^2. Furthermore, the same inequality shows that the set of r.v.'s $M_{T_n\wedge t} 1_{[T_n>0]}$ is bounded in L^2, hence uniformly integrable, which allows to pass to the limit in the equality

$$E\left[M_{t\wedge T_n} 1_{[T_n>0]} \mid \mathscr{F}_s\right] = M_{s\wedge T_n} 1_{[T_n>0]}$$

to get $E\left[M_t \mid \mathscr{F}_s\right] = M_s$. The process M is a L^2-bounded martingale. To prove that $M^2 - \langle M, M\rangle$ is u.i., we observe that

$$\sup_t \left|M_t^2 - \langle M, M\rangle_t\right| \le (M^*_\infty)^2 + \langle M, M\rangle_\infty$$

which is an integrable r.v. The last equalities derive immediately from the stopping theorem. □

Corollary 4.2. *If $M \in H_0^2$,*

$$\|M\|_{\mathbb{H}^2} = \left\|\langle M, M\rangle_\infty^{1/2}\right\|_2 \equiv E\left[\langle M, M\rangle_\infty\right]^{1/2}$$

Proof. If $M_0 = 0$, we have $E\left[M_\infty^2\right] = E\left[\langle M, M\rangle_\infty\right]$ as is seen in the last proof. □

We could have worked in exactly the same way on $[0, t]$ instead of $[0, \infty]$ to get the

Corollary 4.3. *If M is a continuous local martingale, the following two conditions are equivalent i) $M_0 \in L^2$ and $E\left[\langle M, M\rangle_t\right] < \infty$; ii) $\{M_s, s \leq t\}$ is an L^2-bounded martingale.*

We notice that for $M \in H^2$, simultaneously $\langle M, M\rangle_\infty$ is in L^1 and $\lim_{t\to\infty} M_t$ exists a.s. This is generalized in the following

Proposition 4.10. *A continuous local martingale M converges a.s. as t goes to infinity, on the set $\{\langle M, M\rangle_\infty < \infty\}$.*

Proof. Without loss of generality, we may assume $M_0 = 0$. Then, if $T_n = \inf\{t : \langle M, M\rangle_t \geq n\}$, the local martingale M^{T_n} is bounded in L^2 as follows from Proposition (1.23). As a result, $\lim_{t\to\infty} M_t^{T_n}$ exists a.s. But on $\{\langle M, M\rangle_\infty < \infty\}$ the stopping times T_n are a.s. infinite from some n on, which completes the proof. □

Remark 4.1. *The converse statement that $\langle M, M\rangle_\infty < \infty$ on the set where M_t converges a.s. is also true.*

2 Stochastic integral

For several reasons, it is necessary to define an integral with respect to the paths of BM. The natural idea is to consider the "Riemann sums"

$$\sum_i K_{u_i}\left(B_{t_{i+1}} - B_{t_i}\right)$$

where K is the process to integrate and u_i is a point in $[t_i, t_{i+1}]$. But it is known from integration theory that these sums do not converge pathwise because the paths of B are a.s. not of bounded variation. We will prove that the convergence holds in probability, but in a first stage we use L^2-convergence and define integration with respect to the elements of H^2. The class of integrands is the object of the following

Definition 4.7. *If $M \in H^2$, we call $\mathscr{L}^2(M)$ the space of progressively measurable processes K such that*

$$\|K\|_M^2 = E\left[\int_0^\infty K_s^2 d\langle M, M\rangle_s\right] < +\infty$$

If, for any $\Gamma \in \mathscr{B}(\mathbb{R}_+) \otimes \cdot \mathscr{F}_\infty$, we set

$$P_M(\Gamma) = E\left[\int_0^\infty 1_\Gamma(s,\omega) d\langle M, M\rangle_s(\omega)\right]$$

we define a bounded measure P_M on $\mathscr{B}(\mathbb{R}_+) \otimes \mathscr{F}_\infty$ and the space $\mathscr{L}^2(M)$ is nothing else than the space of P_M-square integrable, progressively measurable, functions. As usual, $L^2(M)$ will denote the space of equivalence classes of elements of $\mathscr{L}^2(M)$; it is of course a Hilbert space for the norm $\|\cdot\|_M$.

Since those are the processes we are going to integrate, it is worth recalling that they include all the bounded and left (or right)-continuous adapted processes and, in particular, the bounded continuous adapted processes.

Theorem 4.4. *Let $M \in H^2$; for each $K \in L^2(M)$, there is a unique element of H_0^2, denoted by $K \cdot M$, such that*

$$\langle K \cdot M, N\rangle = K \cdot \langle M, N\rangle$$

for every $N \in H^2$. The map $K \to K \cdot M$ is an isometry from $L^2(M)$ into H_0^2.

Proof. a) Uniqueness.

If L and L' are two martingales of H_0^2 such that $\langle L, N\rangle = \langle L', N\rangle$ for every $N \in H^2$, then in particular $\langle L - L', L - L'\rangle = 0$ which by Proposition 4.4 implies that $L - L'$ is constant, hence $L = L'$.

b) Existence.

Suppose first that M is in H_0^2. By the Kunita-Watanabe inequality and Corollary 4.2, for every N in H_0^2 we have

$$\left|E\left[\int_0^\infty K_s d\langle M, N\rangle_s\right]\right| \leq \|N\|_{H^2}\|K\|_M$$

the map $N \to E\left[(K \cdot \langle M, N\rangle)_\infty\right]$ is thus a linear and continuous form on the Hilbert space H_0^2 and, consequently, there is an element $K \cdot M$ in H_0^2 such that

$$E\left[(K \cdot M)_\infty N_\infty\right] = E\left[(K \cdot \langle M, N\rangle)_\infty\right] \tag{4.2}$$

for every $N \in H_0^2$. Let T be a stopping time; the martingales of H^2 being u.i., we may write

$$\begin{aligned} E\left[(K \cdot M)_T N_T\right] &= E\left[E\left[(K \cdot M)_\infty \mid \mathscr{F}_T\right] N_T\right] = E\left[(K \cdot M)_\infty N_T\right] \\ &= E\left[(K \cdot M)_\infty N_\infty^T\right] = E\left[\left(K \cdot \langle M, N^T\rangle\right)_\infty\right] \\ &= E\left[\left(K \cdot \langle M, N\rangle^T\right)_\infty\right] = E\left[(K \cdot \langle M, N\rangle)_T\right] \end{aligned}$$

which proves that $(K \cdot M)N - K \cdot \langle M, N\rangle$ is a martingale. Furthermore, by (4.2),

$$\|K \cdot M\|_{\mathbb{H}^2}^2 = E\left[(K \cdot M)_\infty^2\right] = E\left[\left(K^2 \cdot \langle M, M\rangle\right)_\infty\right] = \|K\|_M^2$$

which proves that the map $K \to K \cdot M$ is an isometry. If $N \in H^2$ instead of H_0^2, then we still have $\langle K \cdot M, N\rangle = K \cdot \langle M, N\rangle$ because the bracket of a martingale with a constant martingale is zero.

Finally, if $M \in H^2$ we set $K \cdot M = K \cdot (M - M_0)$ and it is easily checked that the properties of the statement carry over to that case.

□

Definition 4.8. *The martingale $K \cdot M$ is called the stochastic integral of K with respect to M and is also denoted by*

$$\int_0 K_s dM_s$$

It is also called the Itô integral to distinguish it from other integrals. The Itô integral is the only one among them for which the resulting process is a martingale.

We stress the fact that the stochastic integral $K \cdot M$ vanishes at 0 .

The reasons for calling $K \cdot M$ a stochastic integral will become clearer in the sequel; here is one of them. We shall denote by $\mathscr{E}$ the space of elementary processes that is the processes which can be written

$$K = K_{-1} 1_{\{0\}} + \sum_i K_i 1_{]t, t_{i+1}}$$

where $0 = t_0 < t_1 < t_2 < \ldots, \lim_i t_i = +\infty$, and the r.v.'s K_i are $\mathscr{F}_{t_i}$-measurable and uniformly bounded and $K_{-1} \in \mathscr{F}_0$. The space $\mathscr{E}$ is contained in $L^2(M)$. For $K \in \mathscr{E}$, we define the so-called elementary stochastic integral $K \cdot M$ by

$$(K \cdot M)_t = \sum_{i=0}^{n-1} K_i \left(M_{t_{i+1}} - M_{t_i} \right) + K_n \left(M_t - M_{t_n} \right)$$

whenever $t_n \leq t < t_{n+1}$. It is easily seen that $K \cdot M \in H_0^2$; moreover, considering subdivisions Δ including the t_i 's, it can be proved using the definition of the brackets, that for any $N \in H^2$, we have $\langle K \cdot M, N \rangle = K \cdot \langle M, N \rangle$.

As a result, the elementary stochastic integral coincides with the stochastic integral constructed in Theorem 4.4. This will be important later to prove a property of convergence of Riemann sums which will lead to explicit computations of stochastic integrals.

We now review some properties of the stochastic integral. The first is known as the property of associativity.

Proposition 4.11. *If $K \in L^2(M)$ and $H \in L^2(K \cdot M)$ then $HK \in L^2(M)$ and*

$$(HK) \cdot M = H \cdot (K \cdot M)$$

Proof. Since $\langle K \cdot M, K \cdot M \rangle = K^2 \cdot \langle M, M \rangle$, it is clear that HK belongs to $L^2(M)$. For $N \in H^2$, we further have

$$\langle (HK) \cdot M, N \rangle = HK \cdot \langle M, N \rangle = H \cdot (K \cdot \langle M, N \rangle)$$

because of the obvious associativity of Stieltjes integrals, and this is equal to

$$H \cdot \langle K \cdot M, N\rangle = \langle H \cdot (K \cdot M), N\rangle$$

the uniqueness in Theorem 4.4 ends the proof. □

The next result shows how stochastic integration behaves with respect to stopping; this will be all important to enlarge the scope of its definition to local martingales.

Proposition 4.12. *If T is a stopping time,*

$$K \cdot M^T = K1_{[0,T]} \cdot M = (K \cdot M)^T$$

Proof. Let us first observe that $M^T = 1_{[0,T]} \cdot M$; indeed, for $N \in H^2$,

$$\left\langle M^T, N\right\rangle = \langle M, N\rangle^T = 1_{[0,T]} \cdot \langle M, N\rangle = \left\langle 1_{[0,T]} \cdot M, N\right\rangle$$

Thus, by the preceding proposition, we have on the one hand

$$K \cdot M^T = K \cdot \left(1_{[0,T]} \cdot M\right) = K1_{[0,T]} \cdot M$$

and on the other hand

$$(K \cdot M)^T = 1_{[0,T]} \cdot (K \cdot M) = 1_{[0,T]}K \cdot M$$

which completes the proof.

□

Since the Brownian motion stopped at a fixed time t is in H^2, if K is a process which satisfies

$$E\left[\int_0^t K_s^2 ds\right] < \infty, \quad \text{for all } t$$

we can define $\int_0^t K_s dB_s$ for each t hence on the whole positive half-line and the resulting process is a martingale although not an element of H^2. This idea can of course be used for all continuous local martingales.

Definition 4.9. *If M is a continuous local martingale, we call $L^2_{\text{loc}}(M)$ the space of classes of progressively measurable processes K for which there exists a sequence (T_n) of stopping times increasing to infinity and such that*

$$\mathbb{E}\left[\int_0^{T_n} K_s^2 d\langle M, M\rangle_s\right] < +\infty$$

Observe that $L^2_{\text{loc}}(M)$ consists of all the progressive processes K such that

$$\int_0^t K_s^2 d\langle M, M\rangle_s < \infty \quad \text{for every } t$$

Proposition 4.13. *For any $K \in L^2_{\text{loc}}(M)$, there exists a unique continuous local martingale vanishing at 0 denoted $K \cdot M$ such that for any continuous local martingale N*

$$\langle K \cdot M, N\rangle = K \cdot \langle M, N\rangle$$

Proof. One can choose stopping times T_n increasing to infinity and such that M^{T_n} is in H^2 and $K^{T_n} \in L^2\left(M^{T_n}\right)$. Thus, for each n, we can define the stochastic integral $X^{(n)} = K^{T_n} \cdot M^{T_n}$. But, by Proposition 4.12, $X^{(n+1)}$ coincides with $X^{(n)}$ on $[0, T_n]$; therefore, one can define unambiguously a process $K \cdot M$ by stipulating that it is equal to $X^{(n)}$ on $[0, T_n]$. This process is obviously a continuous local martingale and, by localization, it is easily seen that $\langle K \cdot M, N\rangle = K \cdot \langle M, N\rangle$ for every local martingale N. □

Remark 4.2. *To prove that a continuous local martingale L is equal to $K \cdot M$, it is enough to check the equality $\langle L, N\rangle = K \cdot \langle M, N\rangle$ for all bounded N 's.*

Again, $K \cdot M$ is called the stochastic integral of K with respect to M and is alternatively written

$$\int_0 K_s dM_s$$

Plainly, Propositions 4.11 and 4.12 carry over to the general case after the obvious changes. Also again if $K \in \mathscr{E}$ this stochastic integral will coincide with the elementary stochastic integral. Stieltjes pathwise integrals having been previously mentioned, it is now easy to extend the definition of stochastic integrals to semimartingales.

Definition 4.10. *A progressively measurable process K is locally bounded if there exists a sequence (T_n) of stopping times increasing to infinity and constants C_n such that $\left|K^{T_n}\right| \leq C_n$.*

All continuous adapted processes K are seen to be locally bounded by taking $T_n = \inf\{t : |K_t| \geq n\}$. Locally bounded processes are in $L^2_{\text{loc}}(M)$ for every continuous local martingale M.

Definition 4.11. *If K is locally bounded and $X = M + A$ is a continuous semimartingale, the stochastic integral of K with respect to X is the continuous semimartingale*

$$K \cdot X = K \cdot M + K \cdot A$$

where $K \cdot M$ is the integral of Proposition 4.13and $K \cdot A$ is the pathwise Stieltjes integral with respect to dA. The semimartingale $K \cdot X$ is also written

$$\int_0 K_s dX_s$$

Proposition 4.14. *The map $K \to K \cdot X$ enjoys the following properties:*

i) $H \cdot (K \cdot X) = (HK) \cdot X$ for any pair H, K of locally bounded processes;

ii) $(K \cdot X)^T = \left(K 1_{[0,T]}\right) \cdot X = K \cdot X^T$ for every stopping time T;

iii) if X is a local martingale or a process of finite variation, so is $K \cdot X$;

iv) if $K \in \mathscr{E}$, then if $t_n \leq t < t_{n+1}$

$$(K \cdot X)_t = \sum_{i=0}^{n} K_i \left(X_{t_{i+1}} - X_{t_i}\right) + K_n \left(X_t - X_{t_n}\right)$$

Proof. Straightforward. □

Proposition 4.15. *For almost every ω, the function $(K \cdot X).(\omega)$ is constant on any interval $[a, b]$ on which either $K_.(\omega) = 0$ or $X.(\omega) = X_a(\omega)$.*

Proof. Only the case where X is a local martingale has to be proved and it is then an immediate consequence of Proposition 4.5 since $K^2 \cdot \langle X, X\rangle$ hence $K \cdot X$ are then constant on these intervals.

As a result, for K and $\bar{K}$ locally bounded and predictable processes and X and $\bar{X}$ semimartingales we have $(K \cdot X)_t - (K \cdot X)_a = (\bar{K} \cdot \bar{X})_t - (\bar{K} \cdot \bar{X})_a$ a.s. on any interval $[a, b]$ on which $K = \bar{K}$ and $X. - X_a = \bar{X}. - \bar{X}_a$; this follows from the equality

$$K \cdot X - \bar{K} \cdot \bar{X} = K \cdot (X - \bar{X}) + (K - \bar{K}) \cdot \bar{X}$$

□

We now turn to a very important property of stochastic integrals, namely the counterpart of the Lebesgue dominated convergence theorem.

Theorem 4.5. *Let X be a continuous semimartingale. If (K^n) is a sequence of locally bounded processes converging to zero pointwise and if there exists a locally bounded process K such that $|K^n| \leq K$ for every n, then $(K^n \cdot X)$ converges to zero in probability, uniformly on every compact interval.*

Proof. The convergence property which can be stated

$$P - \lim_{n \to \infty} \sup_{s \leq t} |(K^n \cdot X)_s| = 0$$

is clear if X is a process of finite variation. If X is a local martingale and if T reduces X, then $(K^n)^T$ converges to zero in $L^2\left(X^T\right)$ and by Theorem 4.4 $(K^n \cdot X)^T$ converges to zero in H^2. The desired convergence is then easily established by the same argument as in Theoren 4.2.

□

The next result on "Riemann sums" is crucial in the following section.

Proposition 4.16. *If K is left-continuous and locally bounded, and (Δ^n) is a sequence of subdivisions of $[0, t]$ such that $|\Delta^n| \to 0$, then*

$$\int_0^t K_s dX_s = P - \lim_{n \to \infty} \sum_{t_i \in \Delta^n} K_{t_i}\left(X_{t_{i+1}} - X_{t_i}\right)$$

Proof. If K is bounded, the right-hand side sums are the stochastic integrals of the elementary processes $\sum K_{t_i} 1_{t_i, t_{i+1}]}$ which converge pointwise to K and are bounded by $\|K\|_\infty$; therefore, the result follows from the preceding theorem. The general case is obtained by the use of localization. □

3 Itô Formula

This section is fundamental. It is devoted to a "change of variables" formula for stochastic integrals which makes them easy to handle and thus leads to explicit computations.

Another way of viewing this formula is to say that we are looking for functions which operate on the class of continuous semimartingales, that is, functions F such that $F(X_t)$ is a continuous semimartingale whatever the continuous semimartingale X is. We begin with the special case $F(x) = x^2$.

Proposition 4.17 (Integration by parts formula). *If X and Y are two continuous semimartingales, then*

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle X, Y \rangle_t$$

In particular,

$$X_t^2 = X_0^2 + 2\int_0^t X_s dX_s + \langle X, X \rangle_t$$

Proof. It is enough to prove the particular case which implies the general one by polarization. If Δ is a subdivision of $[0, t]$, we have

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

letting $|\Delta|$ tend to zero and using, on one hand the definition of $\langle X, X \rangle$, on the other hand Proposition 4.16, we get the desired result. □

If X and Y are of finite variation, this formula boils down to the ordinary integration by parts formula for Stieltjes integrals. The same will be true for the following change of variables formula. Let us also observe that if M is a local martingale, we have, as a result of the above formula,

$$M_t^2 - \langle M, M \rangle_t = M_0^2 + 2\int_0^t M_s dM_s$$

we already knew that $M^2 - \langle M, M\rangle$ is a local martingale but the above formula gives us an explicit expression of this local martingale. In the case of BM, we have

$$B_t^2 - t = 2\int_0^t B_s dB_s$$

which can also be seen as giving us an explicit value for the stochastic integral in the right member. The reader will observe the difference with the ordinary integrals in the appearance of the term t. This is due to the quadratic variation.

All this is generalized in the following theorem. We first lay down the

Definition 4.12. *A d-dimensional vector local martingale (resp. vector continuous semimartingale) is a $\mathbb{R}^d$-valued process $X = \left(X^1, \ldots, X^d\right)$ such that each X^i is a local martingale (resp. cont. semimart.). A complex local martingale (resp. complex cont. semimart.) is a $\mathbb{C}$-valued process whose real and imaginary parts are local martingales (resp. cont. semimarts.).*

Theorem 4.6 (Itô's formula)**.** *Let $X = \left(X^1, \ldots, X^d\right)$ be a continuous vector semimartingale and $F \in C^2\left(\mathbb{R}^d, \mathbb{R}\right)$; then, $F(X)$ is a continuous semimartingale and*

$$F(X_t) = F(X_0) + \sum_i \int_0^t \frac{\partial F}{\partial x_i}(X_s)\, dX_s^i + \frac{1}{2}\sum_{i,j}\int_0^t \frac{\partial^2 F}{\partial x_i \partial x_j}(X_s)\, d\left\langle X^i, X^j\right\rangle_s$$

Proof. If F is a function for which the result is true, then for any i, the result is true for $G(x_1, \ldots, x_d) = x_i F(x_1, \ldots, x_d)$; this is a straightforward consequence of the integration by parts formula. The result is thus true for polynomial functions. By stopping, it is enough to prove the result when X takes its values in a compact set K of $\mathbb{R}^d$. But on K, any F in $C^2\left(\mathbb{R}^d, \mathbb{R}\right)$ is the limit in $C^2(K, \mathbb{R})$ of polynomial functions. By the ordinary and stochastic dominated convergence theorems (Theorem 4.5), the theorem is established.

□

Remark 4.3. *1. The differentiability properties of F may be somewhat relaxed. For instance, if some of the X^i 's are of finite variation, F needs only be of class C^1 in the corresponding coordinates; the proof* *goes through just the same. In particular, if X is a continuous semimartingale and A has finite variations, and if $\partial^2 F/\partial x^2$ and $\partial F/\partial y$ exist and are continuous, then*

$$\begin{aligned} F\left(X_t, A_t\right) =& F\left(X_0, A_0\right) + \int_0^t \frac{\partial F}{\partial x}\left(X_s, A_s\right) dX_s + \int_0^t \frac{\partial F}{\partial y}\left(X_s, A_s\right) dA_s \\ &+ \frac{1}{2}\int_0^t \frac{\partial^2 F}{\partial x^2}\left(X_s, A_s\right) d\langle X, X\rangle_s \end{aligned}$$

2. *One gets another obvious extension when F is defined only on an open set but X takes a.s. its values in this set. We leave the details to the reader as an exercise.*

3. *Itô's formula may be written in "differential" form*

$$dF\left(X_t\right) = \sum_i \frac{\partial F}{\partial x_i}\left(X_t\right) dX_t^i + \frac{1}{2}\sum_{i,j} \frac{\partial^2 F}{\partial x_i \partial x_j}\left(X_t\right) d\left\langle X^i, X^j\right\rangle_t$$

   More generally, if X is a vector semimartingale, $dY_t = \sum_i H_t^i dX_t^i$ will mean

$$Y_t = Y_0 + \sum_i \int_0^t H_s^i dX_s^i$$

   In this setting, Itô's formula may be read as "the chain rule for stochastic differentials".

4. *Itô's formula shows precisely that the class of semimartingales is invariant under composition with C^2-functions, which gives another reason for the introduction of semimartingales. If M is a local martingale, or even a martingale, $F(M)$ is usually not a local martingale but only a semimartingale.*

5. *Let ϕ be a C^1-function with compact support in $]0,1[$. It is of finite variation, hence may be looked upon as a semimartingale and the integration by parts formula yields*

$$X_1\phi(1) = X_0\phi(0) + \int_0^1 \phi(s) dX_s + \int_0^1 X_s \phi'(s) ds + \langle X, \phi\rangle_1$$

   which reduces to

$$\int_0^1 \phi(s)dX_s = -\int_0^1 X_s\phi'(s)ds.$$

In the following proposition, we introduce the class of exponential local martingales which turns out to be very important. For the time being, they provide us with many new examples of local martingales.

Proposition 4.18. *If f is a complex valued function, defined on $\mathbb{R}\times\mathbb{R}_+$, and such that $\frac{\partial^2 f}{\partial x^2}$ and $\frac{\partial f}{\partial y}$ exist, are continuous and satisfy $\frac{\partial f}{\partial y}+\frac{1}{2}\frac{\partial^2 f}{\partial x^2}=0$, then for any cont. local mart. M, the process $f\left(M_t, \langle M, M\rangle_t\right)$ is a local martingale. In particular for any $\lambda\in\mathbb{C}$, the process*

$$\mathscr{E}^\lambda(M)_t = \exp\left\{\lambda M_t - \frac{\lambda^2}{2}\langle M, M\rangle_t\right\}$$

is a local martingale. For $\lambda = 1$, we write simply $\mathscr{E}(M)$ and speak of the exponential of M.

Proof. This follows at once by making $A = \langle M, M\rangle$ in the first remark below Theorem 4.6. □

Another application of the Itô formula is the following

Theorem 4.7 (P. Lévy's characterization theorem)**.** *For a $(\mathscr{F}_t)$-adapted continuous d-dimensional process X vanishing at 0, the following three conditions are equivalent:*

i) X is an $\mathscr{F}_t$-Brownian motion;

ii) X is a continuous local martingale and $\left\langle X^i, X^j\right\rangle_t = \delta_{ij}t$ for every $1\le i,\, j\le d$;

iii) X is a continuous local martingale and for every d-uple $f=(f_1,\dots,f_d)$ of functions in $L^2(\mathbb{R}_+)$, the process

$$\mathscr{E}_t^{if} = \exp\left\{i\sum_k\int_0^t f_k(s)dX_s^k + \frac{1}{2}\sum_k\int_0^t f_k^2(s)ds\right\}$$

is a complex martingale.

Proof. That i) implies ii) has already been seen. Furthermore if ii) holds, Proposition 4.18 applied with $\lambda = i$ and $M_t = \sum_k \int_0^t f_k(s) dX_s^k$ implies that $\mathscr{E}^{if}$ is a local martingale; since it is bounded, it is a complex martingale.

Let us finally assume that iii) holds. Then, if $f = \xi 1_{[0,T]}$ for an arbitrary ξ in $\mathbb{R}^d$ and $T > 0$, the process

$$\mathscr{E}_t^{\mathrm{if}} = \exp\left\{ i\,(\xi, X_{t\wedge T}) + \frac{1}{2}|\xi|^2 (t \wedge T) \right\}$$

is a martingale. For $A \in \mathscr{F}_s$, $s < t < T$, we get

$$E\left[1_A \exp\left\{ i\,(\xi, X_t - X_s)\right\}\right] = P(A) \exp\left(-\frac{|\xi|^2}{2}(t-s) \right)$$

(Here, and below, we use the notation (x, y) for the Euclidean scalar product of x and y in $\mathbb{R}^d$, and $|x|^2 = (x, x)$.)

Since this is true for any $\xi \in \mathbb{R}^d$, the increment $X_t - X_s$ is independent of $\mathscr{F}_s$ and has a Gaussian distribution with variance $(t - s)$; hence i) holds. $\square$

Chapter 5

Stochastic Differential Equation

1 Introduction

Let us recall the Markov property.

Definition 5.1. *Let d be a positive integer. A d-dimensional Markov family is an adapted process $X = \{X_t, \mathscr{F}_t; t \geq 0\}$ on some $(\Omega, \mathscr{F})$, together with a family of probability measures $\{P^x\}_{x \in \mathbb{R}^d}$ on $(\Omega, \mathscr{F})$, such that*

(a) for each $F \in \mathscr{F}$, the mapping $x \mapsto P^x(F)$ is universally measurable; (See Karatzas and Shreeve p 73 for a formal definition.)

(b) $P^x[X_0 = x] = 1, \forall x \in \mathbb{R}^d$;

(c) for $x \in \mathbb{R}^d, s, t \geq 0$ and $\Gamma \in \mathscr{B}\left(\mathbb{R}^d\right)$,

$$P^x\left[X_{t+s} \in \Gamma \mid \mathscr{F}_s\right] = P^x\left[X_{t+s} \in \Gamma \mid X_s\right], \quad P^x - \text{a.s}$$

(d) for $x \in \mathbb{R}^d, s, t \geq 0$ and $\Gamma \in \mathscr{B}\left(\mathbb{R}^d\right)$,

$$P^x\left[X_{t+s} \in \Gamma \mid X_s = y\right] = P^y\left[X_t \in \Gamma\right], \quad P^x X_s^{-1}\text{-a.e. } y.$$

We remark that Brownian motions are Markovian.

Theorem 5.1. *A d-dimensional Brownian motion is a Markov process. A d-dimensional Brownian family is a Markov family.*

Some ideas for the proof of the Theorem.

Let us suppose now that we observe a Brownian motion with initial distribution μ up to time $s, 0 \leq s < t$. In particular, we see the value of B_s, which we call y. Conditioned on these observations, what is the probability that B_t is in some set $\Gamma \in \mathscr{B}\left(\mathbb{R}^d\right)$? Now $B_t = (B_t - B_s) + B_s$, and the increment $B_t - B_s$ is independent of the observations up to time s and is distributed just as B_{t-s} is under P^0. On the other hand, B_s does depend on the observations; indeed, we are conditioning on $B_s = y$. It follows that the sum $(B_t - B_s) + B_s$ is distributed as B_{t-s} is under P^y. Two points then become clear. First, knowledge of the whole past up to time s provides no more useful information about B_t than knowing the value of B_s; in other words,

$$P^{\mu}\left[B_t \in \Gamma \mid \mathscr{F}_s\right] = P^{\mu}\left[B_t \in \Gamma \mid B_s\right], \quad 0 \leq s < t, \Gamma \in \mathscr{B}\left(\mathbb{R}^d\right).$$

Second, conditioned on $B_s = y, B_t$ is distributed as B_{t-s} is under P^y; i.e.,

$$P^{\mu}\left[B_t \in \Gamma \mid B_s = y\right] = P^{y}\left[B_{t-s} \in \Gamma\right], \quad 0 \leq s < t, \Gamma \in \mathscr{B}\left(\mathbb{R}^d\right).$$

2 Diffusion

We explore in this chapter questions of existence and uniqueness for solutions to stochastic differential equations and offer a study of their properties. This endeavor is really a study of diffusion processes. Loosely speaking, the term diffusion is attributed to a Markov process which has continuous sample paths and can be characterized in terms of its infinitesimal generator.

In order to fix ideas, let us consider a d-dimensional Markov family $X = \{X_t, \mathscr{F}_t; 0 \leq t < \infty\}, (\Omega, \mathscr{F}), \{P^x\}_{x \in \mathrm{R}^d}$, and assume that X has continuous paths. We suppose, further, that the relation

$$\lim_{t \downarrow 0} \frac{1}{t}\left[E^x f\left(X_t\right) - f(x)\right] = (\mathscr{A} f)(x); \quad \forall x \in \mathbb{R}^d \tag{5.1}$$

holds for every f in a suitable subclass of the space $C^2\left(\mathbb{R}^d\right)$ of real-valued, twice continuously differentiable functions on $\mathbb{R}^d$; the operator $\mathscr{A} f$

in (5.1) is given by

$$(\mathscr{A}f)(x) \triangleq \frac{1}{2}\sum_{i=1}^{d}\sum_{k=1}^{d} a_{ik}(x)\frac{\partial^2 f(x)}{\partial x_i \partial x_k} + \sum_{i=1}^{d} b_i(x)\frac{\partial f(x)}{\partial x_i} \tag{5.2}$$

for suitable Borel-measurable functions $b_i, a_{ik} : \mathbb{R}^d \to \mathbb{R}, 1 \leq i, k \leq d$. The left hand side of (5.1) is the infinitesimal generator of the Markov family, applied to the test function f. On the other hand, the operator in (5.2) is called the second order differential operator associated with the drift vector $b = (b_1, \ldots, b_d)$ and the diffusion matrix $a = \{a_{ik}\}_{1\leq i,k\leq d}$, which is assumed to be symmetric and nonnegative-definite for every $x \in \mathbb{R}^d$.

The drift and diffusion coefficients can be interpreted heuristically in the following manner: fix $x \in \mathbb{R}^d$ and let $f_i(y) \triangleq y_i, f_{ik}(y) \triangleq (y_i - x_i)(y_k - x_k)$;

$y \in \mathbb{R}^d$. Assuming that holds (5.1) for these test functions, we obtain

$$E^x\left[X_i^{(i)} - x_i\right] = tb_i(x) + o(t) \tag{5.3}$$

$$E^x\left[\left(X_i^{(i)} - x_i\right)\left(X_i^{(k)} - x_k\right)\right] = ta_{ik}(x) + o(t) \tag{5.4}$$

as $t \downarrow 0$, for $1 \leq i, k \leq d$. In other words, the drift vector $b(x)$ measures locally the mean velocity of the random motion modeled by X, and $a(x)$ approximates the rate of change in the covariance matrix of the vector $X_t - x$, for small values of $t > 0$.

Definition 5.2. *Let* $X = \{X_t, \mathscr{F}_t; 0 \leq t < \infty\}, (\Omega, \mathscr{F}), \{P^x\}_{x\in\mathbb{R}^d}$ *be a* d *dimensional Markov family, such that*

(i) X *has continuous sample paths;*

(ii) relation(5.1) holds for every $f \in C^2\left(\mathbb{R}^d\right)$ *which is bounded and has bounded first- and second-order derivatives;*

(iii) relations (5.3), (5.4) hold for every $x \in \mathbb{R}^d$*; and*

(iv) X *is a strong Markov Family.*

Then X *is called a (Kolmogorov-Feller) diffusion process.*

The methodology of stochastic differential equations was suggested by P. Lévy as an "alternative," probabilistic approach to diffusions and was carried out in a masterly way by K. Itô. Suppose that we have a continuous, adapted d-dimensional process $X = \{X_t, \mathscr{F}_t; 0 \leq t < \infty\}$ which satisfies, for every $x \in \mathbb{R}^d$, the stochastic integral equation

$$X_i^{(i)} = x_i + \int_0^t b_i(X_s)\,ds + \sum_{j=1}^r \int_0^t \sigma_{ij}(X_s)\,dW_s^{(j)}; \quad 0 \leq t < \infty, 1 \leq i \leq d \tag{5.5}$$

on a probability space $(\Omega, \mathscr{F}, P^x)$, where $W = \{W_t, \mathscr{F}_t; 0 \leq t < \infty\}$ is a Brownian motion in $\mathbb{R}^r$ and the coefficients $b_i, \sigma_{ij} : \mathbb{R}^d \to \mathbb{R}; 1 \leq i \leq d, 1 \leq j \leq r$ are Borel-measurable. Then it is reasonable to expect that, under certain conditions, (5.1) - (5.4) will indeed be valid, with

$$a_{ik}(x) \triangleq \sum_{j=1}^r \sigma_{ij}(x)\sigma_{kj}(x)$$

We leave the verification of the following fact as an exercise for the reader.

Assume that the coefficients b_i, σ_{ij} are bounded and continuous, and the $\mathbb{R}^d$-valued process X satisfies (5.5). Show that (5.3) , (5.4) hold for every $x \in \mathbb{R}^d$, and that (5.1) holds for every $f \in C^2\left(\mathbb{R}^d\right)$ which is bounded and has bounded first- and second-order derivatives.

3 Strong solutions

In this section we introduce the concept of a stochastic differential equation with respect to Brownian motion and its solution in the so-called strong sense. We discuss the questions of existence and uniqueness of such solutions, as well as some of their elementary properties.

Let us start with Borel-measurable functions $b_i(t,x), \sigma_{ij}(t,x); 1 \leq i \leq d$, $1 \leq j \leq r$, from $[0,\infty) \times \mathbb{R}^d$ into $\mathbb{R}$, and define the $(d \times 1)$ drift vector $b(t,x) = \{b_i(t,x)\}_{1 \leq i \leq d}$ and the $(d \times r)$ dispersion matrix $\sigma(t,x) = \{\sigma_{ij}(t,x)\}_{\substack{1 \leq i \leq d \\ 1 \leq j \leq r}}$. The intent is to assign a meaning to the stochastic differential equation

$$dX_t = b(t, X_t)\,dt + \sigma(t, X_t)\,dW_t$$

written componentwise as

$$dX_t^{(i)} = b_i\,(t, X_t)\,dt + \sum_{j=1}^{r} \sigma_{ij}\,(t, X_t)\,dW_t^{(j)}; \quad 1 \le i \le d, \tag{5.6}$$

where $W = \{W_i; 0 \le t < \infty\}$ is an r-dimensional Brownian motion and $X = \{X_t; 0 \le t < \infty\}$ is a suitable stochastic process with continuous sample paths and values in $\mathbb{R}^d$, the "solution" of the equation. The drift vector $b(t, x)$ and the dispersion matrix $\sigma(t, x)$ are the coefficients of this equation; the $(d \times d)$ matrix $a(t, x) \triangleq \sigma(t, x)\sigma^T(t, x)$ with elements

$$a_{ik}(t, x) \triangleq \sum_{j=1}^{r} \sigma_{ij}(t, x)\sigma_{kj}(t, x); \quad 1 \le i, k \le d$$

will be called the diffusion matrix.

In order to develop the concept of strong solution, we choose a probability space $(\Omega, \mathscr{F}, P)$ as well as an r-dimensional Brownian motion $W = \{W_t, \mathscr{F}_t^W; 0 \le t < \infty\}$ on it. We assume also that this space is rich enough to accommodate a random vector ξ taking values in $\mathbb{R}^d$, independent of $\mathscr{F}_\infty^W$, and with given distribution

$$\mu(\Gamma) = P[\xi \in \Gamma]; \quad \Gamma \in \mathscr{B}\left(\mathbb{R}^d\right).$$

We consider the left-continuous filtration

$$\mathscr{G}_t \triangleq \sigma(\xi) \vee \mathscr{F}_t^W = \sigma\left(\xi, W_s; 0 \le s \le t\right); \quad 0 \le t < \infty,$$

as well as the collection of null sets

$$\mathscr{N} \triangleq \left\{N \subseteq \Omega; \exists G \in \mathscr{G}_\infty \text{ with } N \subseteq G \text{ and } P(G) = 0\right\},$$

and create the augmented filtration

$$\mathscr{F}_t \triangleq \sigma\left(\mathscr{G}_t \cup \mathscr{N}\right), \quad 0 \le t < \infty; \quad \mathscr{F}_\infty \triangleq \sigma\left(\bigcup_{t \ge 0} \mathscr{F}_t\right). \tag{5.7}$$

Obviously, $\{W_t, \mathscr{G}_t;\, 0 \le t < \infty\}$ is an r-dimensional Brownian motion, and then so is $\{W_t, \mathscr{F}_t;\, 0 \le t < \infty\}$. It follows also, that the filtration $\{\mathscr{F}_t\}$ satisfies the usual conditions.

Definition 5.3. *A strong solution of the stochastic differential equation (5.6), on the given probability space $(\Omega, \mathscr{F}, P)$ and with respect to the fixed Brownian motion W and initial condition ξ, is a process $X = \{X_t; 0 \le t < \infty\}$ with continuous sample paths and with the following properties:*

(i) X is adapted to the filtration $\{\mathscr{F}_t\}$ of (5.7),

(ii) $P\left[X_0 = \xi\right] = 1$,

(iii) $P\left[\int_0^t \left\{|b_i\left(s, X_s\right)| + \sigma_{ij}^2\left(s, X_s\right)\right\} ds < \infty\right] = 1$ holds for every $1 \le i \le d$, $1 \le j \le r$ and $0 \le t < \infty$, and

(iv) the integral version of (5.6)

$$X_t = X_0 + \int_0^t b\left(s, X_s\right) ds + \int_0^t \sigma\left(s, X_s\right) dW_s; \quad 0 \le t < \infty, \tag{5.8}$$

or equivalently,

$$\begin{aligned} X_i^{(i)} = X_0^{(i)} + \textstyle\int_0^t b_i\left(s, X_s\right) ds + \sum_{j=1}^r \int_0^t \sigma_{ij}\left(s, X_s\right) dW_s^{(j)} \\ 0 \le t < \infty, 1 \le i \le d \end{aligned} \tag{5.9}$$

holds almost surely.

Remark 5.1. *Above the drift and the diffusion depends on (t, x). But one can also write*

$$dX_t = F(X_t) dZ_t,$$

where $Z_t = (t, W_t)$, and consider Z as a semi-martingale. It is the point of view in the book by Protter.

The processes defined in Definition 5.3 are called Itô processes.

Definition 5.4. *Let the drift vector $b(t, x)$ and dispersion matrix $\sigma(t, x)$ be given. Suppose that, whenever W is an r-dimensional Brownian motion on some $(\Omega, \mathscr{F}, P), \xi$ is an independent, d-dimensional random vector, $\{\mathscr{F}_t\}$ is given by (5.7), and $X, \tilde{X}$ are two strong solutions of (5.6) relative to W with initial condition ξ, then $P\left[X_t = \tilde{X}_t; 0 \le t < \infty\right] = 1$. Under these conditions, we say that strong uniqueness holds for the pair (b, σ).*

We sometimes abuse the terminology by saying that strong uniqueness holds for equation (5.6), even though the condition of strong uniqueness requires us to consider every r-dimensional Brownian motion, not just a particular one.

Theorem 5.2. *Suppose that the coefficients $b(t,x), \sigma(t,x)$ are locally Lipschitz continuous in the space variable; i.e., for every integer $n \geq 1$ there exists a constant $K_n > 0$ such that for every $t \geq 0, \|x\| \leq n$ and $\|y\| \leq n$:*

$$\|b(t,x) - b(t,y)\| + \|\sigma(t,x) - \sigma(t,y)\| \leq K_n \|x - y\|. \tag{5.10}$$

Then strong uniqueness holds for equation (5.6).

Remark 5.2. *We use the following notation. For every $(d \times r)$ matrix σ, we write*

$$\|\sigma\|^2 \triangleq \sum_{i=1}^{d} \sum_{j=1}^{r} \sigma_{ij}^2.$$

Proof. Let us suppose that X and $\tilde{X}$ are both strong solutions, defined for all $t \geq 0$, of (5.6) relative to the same Brownian motion W and the same initial condition ξ, on some $(\Omega, \mathscr{F}, P)$. We define the stopping times $\tau_n = \inf\{t \geq 0; \|X_t\| \geq n\}$ for $n \geq 1$, as well as their tilded counterparts, and we set $s_n \triangleq \tau_n \wedge \tilde{\tau}_n$. Clearly $\lim_{n\to\infty} s_n = \infty$, a.s. P, and

$$\begin{aligned} X_{t\wedge s_n} - \tilde{X}_{t\wedge s_n} = &\int_0^{t\wedge s_n} \left\{ b\left(u, X_u\right) - b\left(u, \tilde{X}_u\right) \right\} du \\ &+ \int_0^{t\wedge s_n} \left\{ \sigma\left(u, X_u\right) - \sigma\left(u, \tilde{X}_u\right) \right\} dW_u \end{aligned}$$

Using the vector inequality $\|v_1 + \cdots + v_k\|^2 \leq k^2 \left(\|v_1\|^2 + \cdots + \|v_k\|^2 \right)$, the Hölder inequality for Lebesgue integrals and (5.6), we may write for $0 \leq t \leq T$:

$$
\begin{aligned}
E\left\|X_{t\wedge s_n}-\tilde{X}_{t\wedge s_n}\right\|^2 \leq & 4E\left[\int_0^{t\wedge s_n}\left\|b\left(u,X_u\right)-b\left(u,\tilde{X}_u\right)\right\|du\right]^2 \\
& +4E\sum_{i=1}^{d}\left[\sum_{j=1}^{r}\int_0^{t\wedge s_n}\left(\sigma_{ij}\left(u,X_u\right)-\sigma_{ij}\left(u,\tilde{X}_u\right)\right)dW_u^{(j)}\right]^2 \\
\leq & 4tE\int_0^{t\wedge s_n}\left\|b\left(u,X_u\right)-b\left(u,\tilde{X}_u\right)\right\|^2du \\
& +4E\int_0^{t\wedge s_n}\left\|\sigma\left(u,X_u\right)-\sigma\left(u,\tilde{X}_u\right)\right\|^2du \\
\leq & 4(T+1)K_n^2\int_0^t E\left\|X_{u\wedge s_n}-\tilde{X}_{u\wedge s_n}\right\|^2du
\end{aligned}
$$

We now apply Gronwall inequality with $g(t) \triangleq E\left\|X_{t\wedge s_n}-\tilde{X}_{t\wedge s_n}\right\|^2$ to conclude that $\{X_{t\wedge s_n}; 0\le t<\infty\}$ and $\left\{\tilde{X}_{t\wedge s_n}; 0\le t<\infty\right\}$ are modifications of one another, and thus are indistinguishable. Letting $n\to\infty$, we see that the same is true for $\{X_t; 0\le t<\infty\}$ and $\left\{\tilde{X}_t; 0\le t<\infty\right\}$. □

Theorem 5.3. *Suppose that the coefficients $b(t,x),\sigma(t,x)$ satisfy the global Lipschitz and linear growth conditions*

$$\|b(t,x)-b(t,y)\|+\|\sigma(t,x)-\sigma(t,y)\|\leq K\|x-y\| \tag{5.11}$$

$$\|b(t,x)\|^2+\|\sigma(t,x)\|^2\leq K^2\left(1+\|x\|^2\right) \tag{5.12}$$

for every $0\le t<\infty, x\in\mathbb{R}^d, y\in\mathbb{R}^d$, where K is a positive constant. On some probability space $(\Omega,\mathscr{F},P)$, let ξ be an $\mathbb{R}^d$-valued random vector, independent of the r-dimensional Brownian motion $W=\left\{W_t,\mathscr{F}_t^W; 0\le t<\infty\right\}$, and with finite second moment:

$$E\|\xi\|^2<\infty.$$

Let $\{\mathscr{F}_t\}$ be as in (5.7). Then there exists a continuous, adapted process $X=\{X_t,\mathscr{F}_t; 0\le t<\infty\}$ which is a strong solution of equation (5.6) relative to W, with initial condition ξ. Moreover, this process is square-integrable:

for every $T > 0$, there exists a constant C, depending only on K and T, such that

$$E\left\|X_t\right\|^2 \leq C\left(1+E\|\xi\|^2\right)e^{Ct}; \quad 0 \leq t \leq T. \tag{5.13}$$

The idea of the proof is to mimic the deterministic situation and to construct recursively, by a Picard iteration, a sequence of successive approximations by setting $X_t^{(0)} \equiv \xi$ and

$$X_t^{(k+1)} \triangleq \xi + \int_0^t b\left(s, X_s^{(k)}\right) ds + \int_0^t \sigma\left(s, X_s^{(k)}\right) dW_s; \quad 0 \leq t < \infty \tag{5.14}$$

for $k \geq 0$. These processes are obviously continuous and adapted to the filtration $\{\mathscr{F}_t\}$. The hope is that the sequence $\left\{X^{(k)}\right\}_{k=1}^{\infty}$ will converge to a solution of equation (5.6).

Let us start with the observation which will ultimately lead to (5.13).

For every $T > 0$, there exists a positive constant C depending only on K and T, such that for the iterations in (5.14) we have

$$E\left\|X_t^{(k)}\right\|^2 \leq C\left(1+E\|\xi\|^2\right)e^{Ct}; \quad 0 \leq t \leq T, k \geq 0 \tag{5.15}$$

Proof. We have $X_t^{(k+1)} - X_t^{(k)} = B_t + M_t$ from (5.14) where

$$B_t \triangleq \int_0^t \left\{b\left(s, X_s^{(k)}\right) - b\left(s, X_s^{(k-1)}\right)\right\} ds,$$

$$M_t \triangleq \int_0^t \left\{\sigma\left(s, X_s^{(k)}\right) - \sigma\left(s, X_s^{(k-1)}\right)\right\} dW_s.$$

Thanks to the inequalities (5.12) and (5.15), the process $\left\{M_t = \left(M_t^{(1)}, \ldots, M_t^{(d)}\right),\right.$

$\left.\mathscr{F}_t; 0 \leq t < \infty\right\}$ is seen to be a vector of square-integrable martingales such that $\exists 0 < \Lambda_1 < \infty$

$$E\left[\max_{0\le s\le t}\|M_s\|^2\right] \le \Lambda_1 E\int_0^t \left\|\sigma\left(s, X_s^{(k)}\right) - \sigma\left(s, X_s^{(k-1)}\right)\right\|^2 ds \tag{5.16}$$

$$\le \Lambda_1 K^2 E\int_0^t \left\|X_s^{(k)} - X_s^{(k-1)}\right\|^2 ds \tag{5.17}$$

where we have used that $\langle M\rangle_t = \int_0^t \left\|\sigma\left(s, X_s^{(k)}\right) - \sigma\left(s, X_s^{(k-1)}\right)\right\|^2 ds$ and Doob inequality.

On the other hand, we have $E\|B_t\|^2 \le K^2 t\int_0^t E\left\|X_s^{(k)} - X_s^{(k-1)}\right\|^2 ds$, and therefore, with $L = 4K^2(\Lambda_1 + T)$,

$$E\left[\max_{0\le s\le t}\left\|X_s^{(k+1)} - X_s^{(k)}\right\|^2\right] \le L\int_0^t E\left\|X_s^{(k)} - X_s^{(k-1)}\right\|^2 ds; \quad 0\le t\le T \tag{5.18}$$

Inequality (5.18) can be iterated to yield the successive upper bounds

$$E\left[\max_{0\le s\le t}\left\|X_s^{(k+1)} - X_s^{(k)}\right\|^2\right] \le C^*\frac{(Lt)^k}{k!}; \quad 0\le t\le T \tag{5.19}$$

where $C^* = \max_{0\le t\le T} E\left\|X_t^{(1)} - \xi\right\|^2$, a finite quantity because of (5.15). Relation (5.19) and the Čebyšev inequality now give

$$P\left[\max_{0\le t\le T}\left\|X_t^{(k+1)} - X_i^{(k)}\right\| > \frac{1}{2^{k+1}}\right] \le 4C^*\frac{(4LT)^k}{k!}; \quad k = 1, 2, \dots$$

and this upper bound is the general term in a convergent series. From the Borel-Cantelli lemma, we conclude that there exists an event $\Omega^* \in \mathscr{F}$ with $P(\Omega^*) = 1$ and an integer-valued random variable $N(\omega)$ such that for every $\omega \in \Omega^*$: $\max_{0\le t\le T}\left\|X_t^{(k+1)}(\omega) - X_t^{(k)}(\omega)\right\| \le 2^{-(k+1)}, \forall k \ge N(\omega)$. Consequently,

$$\max_{0\le t\le T}\left\|X_t^{(k+m)}(\omega) - X_t^{(k)}(\omega)\right\| \le 2^{-k}, \quad \forall m\ge 1, k\ge N(\omega)$$

We see then that the sequence of sample paths $\left\{X_t^{(k)}(\omega); 0 \le t \le T\right\}_{k=1}^{\infty}$ is convergent in the supremum norm on continuous functions, from which follows the existence of a continuous limit $\{X_t(\omega); 0 \le t \le T\}$ for all $\omega \in \Omega^*$. Since T is arbitrary, we have the existence of a continuous process $X = \{X_t;\ 0 \le t < \infty\}$ with the property that for P-a.e. ω, the sample paths $\left\{X_{\cdot}^{(k)}(\omega)\right\}_{k=1}^{\infty}$ converge to $X_{\cdot}(\omega)$, uniformly on compact subsets of $[0, \infty)$. Inequality (5.13) is a consequence of (5.15) and Fatou's lemma. From (5.13) and (5.12) we have condition (iii) of Definition 5.3. Conditions (i) and (ii) are also clearly satisfied by X. We can check that the $X_t \stackrel{def}{=} \lim_{k\to\infty} X_t^{(k)}$ satisfies the requirement (iv) of Definition 5.3.

$\square$