LECTURE NOTES

Stochastic Analysis

Paul Thévenin

Fall 2023

University of Vienna

Contents

1 Conditional expectation and Gaussian random variables **4**
1.1 Conditional expectation 4
1.2 Multivariate characteristic functions 6
1.3 Gaussian random variables 8
1.4 Gaussian random vectors 9
1.5 Gaussian processes 10

2 Brownian motion: definitions and properties **13**
2.1 Definition 13
2.2 Regularity of the paths 14
2.3 Hölder exponent 17
2.4 Wiener measure 18
2.5 First properties of Brownian motion 19
2.6 Strong Markov property 21

3 Martingales and finite variation processes **24**
3.1 Martingales in continuous time 24
3.2 Uniform integrability 25
3.3 Convergence theorems for martingales 26
3.4 Optional stopping theorem 27
3.5 Finite variation integral 28
3.6 Local martingales 31

4 Stochastic integral **33**
4.1 Simple processes 33
4.2 Quadratic variation 34
4.3 Itô integrals 35
4.4 Covariation 40
4.5 Itô's formula 41

5 Applications of Itô's formula **45**
5.1 Lévy's characterization of Brownian motion 45
5.2 Dubins-Schwarz theorem 46
5.3 Transience and recurrence of Brownian motion 48
5.4 Exponential martingale, Girsanov's theorem and Cameron-Martin formula 49

6 Appendix: Brownian motion and the Dirichlet problem **54**

Lecture 1: 02/10/2023

Introduction

Classical analysis deals with integration and differentiation of regular deterministic functions. Under regularity assumptions on a function f, one can compute integrals of the form $\int_a^b f(s)ds$. The aim of stochastic analysis is to extend this framework to (random) functions that can be quite irregular. We want to define the integral of a random process with respect to another random process: we will see what assumptions on the random functions $(H_s, s \geq 0)$ and $(X_s, s \geq 0)$ are needed to give meaning to the expression

$$\int_a^b H_s dX_s,$$

and how to manipulate such quantities. This extension of classical analysis to random processes started with Kiyoshi Ito in the 1940's, and relies on the celebrated Ito formula. It appears nowadays a lot in financial mathematics and economics.

We will start by formally introducing the central notions of *conditional expectation* and *Gaussian variables*. This will allow us to formally define the *Brownian motion*, which is maybe the most famous random process. A third part will be devoted to the study of *martingales*, which are roughly speaking random processes with constant expectation. With all these tools in hand, we will be able to construct in a rigorous way the theory of stochastic integration, and introduce the Ito formula. In a last part, we will use it to study in more detail the Brownian motion and prove some of its very interesting properties.

These notes are largely inspired from Nathanaël Berestycki's notes, which are available on his webpage.

Notations In what follows, $(\Omega, \mathscr{F}, \mathbb{P})$ will always denote a probability space. "a.s." means almost surely, and we will use $\stackrel{(d)}{\to}$ and $\stackrel{(\mathbb{P})}{\to}$ to denote respectively convergence in distribution and convergence in probability. We write "wlog" for "without loss of generality".

1 Conditional expectation and Gaussian random variables

1.1 Conditional expectation

When one considers random experiments, usually several layers of randomness are involved. It therefore makes sense to condition on one of these layers.

Example. *In a math article, there are a random number X of typos. A reviewer spots each of them independently with probability p. Let N be the number of typos that she spots.*

(i) $\mathbb{E}[N|X]$*: given X, how many typos on average will she spot?*

(ii) $\mathbb{E}[X|N]$*: given that she spotted N, how many can we expect in total?*

Note that the answer will in general be random, and depend on the value of X. More precisely, for any r such that $\mathbb{P}(X=r)>0$, any event A, we can define

$$\mathbb{P}(A|X=r) = \frac{\mathbb{P}(A\cap\{X=r\})}{\mathbb{P}(X=r)},$$

and in particular

$$\mathbb{E}[N|X=r] = \frac{\mathbb{E}[N\mathbb{1}_{X=r}]}{\mathbb{P}(X=r)}.$$

In general, $\mathbb{E}[N|X]$ is a random variable such that on the event $\{X=r\}$ its value will be $\mathbb{E}[N|X=r]$. An important property is the following: let $Y=\mathbb{E}[N|X]$; we have for all r

$$\mathbb{E}[N\mathbb{1}_{X=r}] = \mathbb{E}[Y\mathbb{1}_{X=r}].$$

Indeed,

$$\begin{aligned}\mathbb{E}[Y\mathbb{1}_{X=r}] &= \mathbb{E}[\mathbb{E}[N|X=r]\mathbb{1}_{X=r}]\\ &= \mathbb{E}\left[\frac{\mathbb{E}[N\mathbb{1}_{X=r}]}{\mathbb{P}(X=r)}\mathbb{1}_{X=r}\right]\\ &= \mathbb{E}[N\mathbb{1}_{X=r}].\end{aligned}$$

This motivates the following (more general) definition:

Definition 1. *Let $(\Omega,\mathscr{F},\mathbb{P})$ be a probability space, and $X:\Omega\to\mathbb{R}$ a random variable such that X is integrable or $X\geq 0$ a.s.. Let $\mathscr{G}\subset\mathscr{F}$ be a σ-algebra. We say that Y is a conditional expectation of X given $\mathscr{G}$ if the following holds:*

- *Y is $\mathscr{G}$-measurable ;*
- *for any $B\in\mathscr{G}$, $\mathbb{E}[X\mathbb{1}_B]=\mathbb{E}[Y\mathbb{1}_B]$.*

Existence and uniqueness of conditional expectations are not clear a priori. This is the content of the next Theorem, which relies on the Radon-Nikodym theorem.

Theorem 1 (Radon-Nikodym theorem). *Let μ, ν be two σ-finite measures on a measurable space $(\Omega, \mathscr{A})$. Then, the following are equivalent:*

(i) ν is absolutely continuous with respect to μ ($\nu \ll \mu$), that is, for all $A \in \mathscr{A}$, if $\mu(A) = 0$ then $\nu(A) = 0$.

(ii) There exists $f : \Omega \to \mathbb{R}_+$, $\mathscr{A}$-measurable, such that $\nu(A) = \int_\Omega \mathbb{1}_A f d\mu$ for all $A \in \mathscr{A}$.

If this holds, we call $f = \frac{d\nu}{d\mu}$ the Radon-Nikodym derivative of ν with respect to μ. f is defined up to a set of μ-measure 0.

Theorem 2 (Advanced Probability). *Let X and $\mathscr{G}$ be as above. Then,*

(i) there exists a conditional expectation of X given $\mathscr{G}$.

(ii) Furthermore, if Z and Z' are two such conditional expectations, then $Z = Z'$ a.s.

(iii) Moreover, if $X \geq 0$ then $\mathbb{E}[X|\mathscr{G}] \geq 0$ a.s. and, if X is integrable, so is $\mathbb{E}[X|\mathscr{G}]$.

Proof. (i) Existence. Suppose first that $X \geq 0$. Consider $\nu : A \in \mathscr{G} \mapsto \mathbb{E}[X\mathbb{1}_A]$, and $\mu = \mathbb{P}_{|\mathscr{G}}$ (that is, see $\mathbb{P}$ as a measure on $(\Omega, \mathscr{G})$). Then, clearly $\nu \ll \mu$ and μ, ν are σ-finite. By the Radon-Nikodym theorem, there exists $Y : \Omega \to \mathbb{R}_+$ that is $\mathscr{G}$-measurable and such that $\nu(A) = \int_\Omega \mathbb{1}_A Y d\mu$ for all $A \in \mathscr{G}$. This can be rewritten

$$\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A].$$

In the general case, write $X = X^+ + X^-$ where $X^+ = X \wedge 0$ and $X^- = X - X^+$. It is easy to check that $Y^+ - Y^-$ has the right properties.

(ii) Uniqueness. Assume first that $E[|X|] < \infty$. Take two such conditional expectations Y_1, Y_2. Both are $\mathscr{G}$-measurable, so $A := \{Y_1 - Y_2 > 0\} \in \mathscr{G}$. By definition, $\mathbb{E}[Y_1\mathbb{1}_A] = \mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[Y_2\mathbb{1}_A]$, so that $\mathbb{E}[(Y_1 - Y_2)\mathbb{1}_A] = 0$. Hence, $\mathbb{P}(A) = 0$ and $Y_1 \leq Y_2$ a.s. By symmetry, $Y_2 \leq Y_1$ a.s. If $X \geq 0$, consider instead the events $A_n := \{Y_1 - Y_2 > 0, Y_1 \leq n, Y_2 \leq n\}$.

(iii) Take Y such a variable. If $X \geq 0$ a.s., take $A := \{Y < 0\}$. Since Y is $\mathscr{G}$-measurable, $A \in \mathscr{G}$. Hence, $\mathbb{E}[Y\mathbb{1}_A] = \mathbb{E}[X1_A] \geq 0$, so that $Y \geq 0$ a.s. We prove the same way that Y is integrable if X is integrable.

□

The conditional expectation satisfies the following useful properties.

Proposition 1 (Properties of conditional expectation).

(i) If X is $\mathscr{G}$-measurable, then $E[X|\mathscr{G}] = X$ a.s.

(ii) If $\mathscr{G} = \{\varnothing, \Omega\}$, then $\mathbb{E}[X|\mathscr{G}] = \mathbb{E}[X]$ a.s.

(iii) If $\sigma(X)$ and $\mathscr{G}$ are independent, then $\mathbb{E}[X|\mathscr{G}] = \mathbb{E}[X]$ a.s.

(iv) If $\mathscr{G}' \subseteq \mathscr{G}$, then $\mathbb{E}[X|\mathscr{G}'] = \mathbb{E}[\mathbb{E}[X|\mathscr{G}]|\mathscr{G}']$ (tower property)

(v) If X, Y are integrable, then $\mathbb{E}[aX + bY|\mathscr{G}] = aE[X|\mathscr{G}] + b\mathbb{E}[Y|\mathscr{G}]$ (linearity)

(vi) If U is $\mathscr{G}$-measurable and either $(U \geq 0, X \geq 0)$ or (X and UX are integrable), then $\mathbb{E}[UX|\mathscr{G}] = U\mathbb{E}[X|\mathscr{G}]$ a.s.

(vii) If $X \leq Y$ a.s. then $\mathbb{E}[X|\mathscr{G}] \leq \mathbb{E}[Y|\mathscr{G}]$ a.s. (monotonicity)

(viii) We have $|\mathbb{E}[X|\mathscr{G}]| \leq \mathbb{E}[|X||\mathscr{G}]$ a.s.

(ix) If $\sigma(X,\mathscr{G})$ and $\mathscr{H} \subseteq \mathscr{F}$ are independent, then $\mathbb{E}[X|\sigma(\mathscr{G},\mathscr{H})] = \mathbb{E}[X|\mathscr{G}]$ a.s. (adding independent information does not change anything).

(x) Let $\phi : \mathbb{R} \to \mathbb{R}$ convex. Then $\phi(\mathbb{E}[X|\mathscr{G}]) \leq \mathbb{E}[\phi(X)|\mathscr{G}]$. (Jensen's conditional inequality)

In particular, if $X \in L^p$ for $p \geq 1$, then $\mathbb{E}[X|\mathscr{G}] \in L^p$ and $|\mathbb{E}[X|\mathscr{G}]|^p \leq |X|^p$.

Proof of some items. (i) X is $\mathscr{G}$-measurable and satisfies $\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[X\mathbb{1}_A]$ for all $A \in \mathscr{G}$.

(ii) It is a consequence of (iii).

(iii) First, $\mathbb{E}[X]$ is clearly $\mathscr{G}$-measurable. Now, for $A \in \mathscr{G}$, we have

$$\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[X]\mathbb{P}(A) = \mathbb{E}[\mathbb{E}[X]\mathbb{1}_A].$$

□

Some properties of the conditional expectation also pass to the limit.

Proposition 2 (Limit theorems for conditional expectation)**.** *(i) If $0 \leq X_n \nearrow X$, then $\mathbb{E}[X_n|\mathscr{G}] \nearrow \mathbb{E}[X|\mathscr{G}]$. (Monotone convergence)*

(ii) If $X_n \to X$ a.s. and there exists Y integrable such that $|X_n| \leq Y$ for all n, then $\mathbb{E}[X_n|\mathscr{G}] \to \mathbb{E}[X|\mathscr{G}]$ a.s. (Dominated convergence)

(iii) If (X_n) are nonnegative, then $\mathbb{E}[\liminf X_n|\mathscr{G}] \leq \liminf \mathbb{E}[X_n]$.

Lecture 2: 04/10/2023

1.2 Multivariate characteristic functions

If X is a real random variable, then its characteristic function is by definition

$$\phi_X(t) = \mathbb{E}\left[e^{itx}\right] = \int_{\mathbb{R}} e^{itx}\mu(dx),$$

where μ is the law of X. It is always defined, continuous and characterizes the law of X. Indeed, if e.g. X has a continuous density f_X, then ϕ_X is the (conjugate) Fourier transform of f_X, and one can recover f_X from ϕ_X by inverse Fourier transform.

We now extend this notion to variables taking their values in $\mathbb{R}^d$, for $d \geq 1$. Denote by $\langle \cdot, \cdot \rangle$ the usual scalar product on $\mathbb{R}^d$.

Definition 2 (Multivariate characteristic function)**.** *Let $X := (X_1, \ldots, X_d)$ be a random variable in $\mathbb{R}^d$.We define its characteristic function as*

$$\begin{aligned}\phi_X : &\mathbb{R}^d \to \mathbb{R}\\ &t := (t_1, \ldots, t_d \mapsto \mathbb{E}\left[e^{i(t_1X_1+\ldots+t_dX_d)}\right] = \mathbb{E}\left[e^{i\langle t,X\rangle}\right]\\ &\qquad = \int_{\mathbb{R}^d} e^{i(t_1x_1+\ldots+t_dx_d)}\mu(dx),\end{aligned}$$

where μ is the law of X.

It enjoys the same properties as in one dimension. For example, if X and Y are independent, then

$$\phi_{X+Y}(t) = \mathbb{E}\left[e^{i\langle t,X+Y\rangle}\right] = \mathbb{E}\left[e^{i\langle t,X\rangle}e^{i\langle t,Y\rangle}\right] = \phi_X(t)\phi_Y(t).$$

An important feature is that the characteristic function characterizes the law μ.

Theorem 3. *Let μ, ν be probability distributions on $\mathbb{R}^d$. If their characteristic functions are equal, then $\mu = \nu$.*

Proof. The proof could be similar to the one-dimensional case. But here is a different argument based on the Stone-Weierstrass lemma.

We will show that, if f is continuous with compact support in $\mathbb{R}^d$, then $\int_{\mathbb{R}^d} f d\mu = \int_{\mathbb{R}^d} f d\nu$. From this, it is easy to obtain $\mu = \nu$ (monotone class theorem).

Lemma 1. *(Stone-Weierstrass) Every continuous function $f : \mathbb{R}^d \to \mathbb{R}$ with period 2π in each coordinate (that is, for all $1 \le i \le d$, all $x \in \mathbb{R}^d$, $f(x+2\pi e_i) = f(x)$) admits a uniform approximation by a linear combination of $x \mapsto \cos(\langle k,x\rangle), x \mapsto \sin(\langle k,x\rangle), k \in \mathbb{Z}_+^d$, that is, for all $\varepsilon > 0$ there exists a finite linear combination g such that $||g-f||_\infty < \varepsilon$.*

Fix f continuous with compact support. There exists M such that $||f||_\infty \le M$. For any $R > 0$, let $\tilde{f}$ be a continuous function with period $2\pi R$ in each coordinate, such that $f = \tilde{f}$ on $B(0,R)$. Fix $\varepsilon > 0$. By Stone-Weierstrass, we have $g : x \mapsto \sum_{k=1}^n a_k \cos(\frac{\langle k,x\rangle}{R}) + b_k \sin(\frac{\langle k,x\rangle}{R})$ such that $||g-\tilde{f}||_\infty < \varepsilon$.

Note that $\int_{\mathbb{R}^d} g d\mu = \int_{\mathbb{R}^d} g d\nu$ since

$$\begin{aligned}\int_{\mathbb{R}^d} g d\mu &= \sum_{k=1}^n a_k Re\left(\int_{\mathbb{R}^d} e^{i\frac{\langle k,x\rangle}{R}}\mu(dx)\right) + \sum_{k=1}^n b_k Im\left(\int_{\mathbb{R}^d} e^{i\frac{\langle k,x\rangle}{R}}\mu(dx)\right)\\ &= \sum_{k=1}^n \left[a_k Re(\phi(k/R)) + b_k Im(\phi(k/R))\right].\end{aligned}$$

On the other hand,

$$\begin{aligned}\left|\int_{\mathbb{R}^d} f d\mu - \int_{\mathbb{R}^d} g d\mu\right| &\le \int_{\mathbb{R}^d} |f-g| d\mu\\ &\le \int_{\mathbb{R}^d} |f-\tilde{f}| d\mu + \int_{\mathbb{R}^d} |\tilde{f}-g| d\mu\\ &\le M\mu(B(0,R)^c) + \varepsilon\end{aligned}$$

Hence, $|\int_{\mathbb{R}^d} f d\mu - \int_{\mathbb{R}^d} f d\nu| \le 2M\mu(B(0,R)^c) + 2\varepsilon$.

Let $R \to \infty$ and $\varepsilon \to 0$. The result follows. □

As a corollary, we get the following:

Corollary 1. *The following are equivalent:*

(i) $(X_1, \ldots, X_d)$ *are independent*

(ii) There exist functions $f_1, \ldots, f_d : \mathbb{R} \to \mathbb{R}$ such that, for all $(t_1, \ldots, t_d) \in \mathbb{R}^d$,

$$\phi_X(t_1, \ldots, t_d) = f_1(t_1) \ldots f_d(t_d).$$

Proof. See Exercice session ☐

One useful property of characteristic functions in dimension $d = 1$ is that it characterises the convergence in distribution (Lévy's continuity theorem). Actually, the same holds in all dimension.

Definition 3. *Let $(X_n)_{n\geq 1}, X$ be random variables. We say that $X_n \to X$ in distribution if, for any $f : \mathbb{R}^d \to \mathbb{R}$ continuous and bounded, we have*

$$\mathbb{E}[f(X_n)] \underset{n\to\infty}{\to} \mathbb{E}[f(X)].$$

The following then holds:

Theorem 4. *We have $X_n \to X$ in distribution if and only if*

$$\forall t \in \mathbb{R}^d, \phi_{X_n}(t) \to \phi_X(t).$$

Proof. The proof is similar to the case $d = 1$. We first prove the *tightness* of $(X_n)_{n\geq 1}$, that is, for all $\varepsilon > 0$, there exists a compact K such that $\mathbb{P}(X_n \in K) \geq 1 - \varepsilon$ for all $n \geq 1$. This provides the existence of subsequential weak limits, which have to be all equal since the characteristic function characterizes the distribution. ☐

1.3 Gaussian random variables

Definition 4. *A random variable X is said to be Gaussian with mean $m \in \mathbb{R}$ and variance $\sigma^2 > 0$ if it has density*

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{\sigma^2}}.$$

We write $X \sim \mathcal{N}(m, \sigma^2)$ ($\mathcal{N}$ stands for normal). When $m = 0$ and $\sigma^2 = 1$, we call X standard Gaussian.

Its characteristic function is explicit and nice.

Lemma 2 (Characteristic function)**.** *If $X \sim \mathcal{N}(m, \sigma^2)$, then, for $t \in \mathbb{R}$, we have $\phi_X(t) = e^{imt - \frac{\sigma^2}{2}t}$.*

Proof. Assume wlog that $m = 0$ and $\sigma^2 = 1$. Consider the function $z \in \mathbb{C} \mapsto \mathbb{E}[e^{zX}]$. First, for $z \in \mathbb{R}$ we have

$$\mathbb{E}\left[e^{zX}\right] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{zx} e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} e^{z^2/2} \int_{\mathbb{R}} e^{-(x-z)^2/2} = e^{z^2/2}.$$

This ensures that $z \mapsto \mathbb{E}[e^{zX}]$ is well-defined and differentiable (hence holomorphic) on $\mathbb{C}$. By analytic continuation, it holds for all $z \in \mathbb{C}$ that $\mathbb{E}[e^{zX}] = e^{z^2/2}$. Taking $z = it$, this ends the proof. ☐

Gaussian variables enjoy some nice properties.

Lemma 3. *Let $X \sim \mathcal{N}(m, \sigma^2)$ and $X' \sim \mathcal{N}(m', (\sigma')^2)$ be independent. Then, $X + X' \sim \mathcal{N}(m + m', \sigma^2 + (\sigma')^2)$.*

Proof. This is a consequence of the previous lemma. Indeed, by independence,

$$\mathbb{E}[e^{it(X+X')}] = \mathbb{E}[e^{itX}]\mathbb{E}[e^{itX'}] = e^{itm - \frac{\sigma^2}{2}t^2} e^{itm' - \frac{(\sigma')^2}{2}t^2} = e^{it(m+m') - \frac{\sigma^2 + (\sigma')^2}{2}}.$$

☐

1.4 Gaussian random vectors

Definition 5. *A random variable X with values in $\mathbb{R}^d$ is called a Gaussian random vector if every linear combination of its coordinates is a* $1d$ *Gaussian variable, that is:*

$$\forall u \in \mathbb{R}^d, \langle u, X\rangle \text{ is Gaussian.}$$

Example. *If $X_1, \dots, X_d$ are independent Gaussian variables, then $X := (X_1, \dots, X_d)$ is a Gaussian vector by Lemma 3.*

Example. *If $X := (X_1, \dots, X_d) \in \mathbb{R}^d$ is as above (the X_i's are independent Gaussians) and $A \in \mathbb{R}^{m\times d}$, then $Y = AX$ is a Gaussian random vector (in dimension m). Indeed, a linear combination of coordinates of Y is a linear combination of coordinates of X, so it is Gaussian.*

We define the mean of X:

$$m = \mathbb{E}[X] := \begin{pmatrix} \mathbb{E}[X_1] \\ \vdots \\ \mathbb{E}[X_d] \end{pmatrix}$$

We also introduce the bilinear form

$$q(u,v) = \sum_{i,j=1}^{d} u_i v_j Cov(X_i, X_j) = Cov(\langle u, X\rangle, \langle v, X\rangle).$$

Lecture 3: 09/10/2023

Note that $q(u,u) = Var(\langle u, X\rangle) \geq 0$, so q is positive semidefinite. The bilinear form q is encoded by the so-called *covariance matrix*

$$\Sigma_X = (Cov(X_i, X_j))_{1\leq i,j\leq d}\,.$$

Furthermore, $\langle u, X\rangle$ is Gaussian with mean $\langle u, m\rangle$ and variance $q(u,u)$.

Proposition 3. *Let $X := (X_1, \dots, X_d)$ be a Gaussian vector of mean m and covariance matrix Σ. Then, the following are equivalent:*

(i) The X_j's are independent ;

(ii) Σ is diagonal, that is, $Cov(X_i, X_j) = 0$ for all $i \neq j$.

Proof. *(i) $\Rightarrow$ (ii)*
It is true for any pair of variables (X,Y) that $Cov(X,Y) = 0$ if X and Y are independent.

(ii) $\Rightarrow$ (i)
This is where Gaussianity is used. By uniqueness of characteristic functions, it suffices to factorize the characteristic function. We have for any $t \in \mathbb{R}^d$:

$$\begin{aligned}
\mathbb{E}\left[e^{i\langle t,X\rangle}\right] &= \mathbb{E}\left[e^{i\mathcal{N}(\langle t,m\rangle, q(t,t))}\right] \\
&= e^{i\langle t,m\rangle - \frac{1}{2}q(t,t)} \quad \text{by Lemma 2} \\
&= e^{i\sum_{j=1}^{d} t_j m_j - \frac{1}{2}\sum_{j,k=1}^{d} t_j t_k Cov(X_j, X_k)} \\
&= \prod_{j=1}^{d} e^{i t_j m_j - \frac{t_j^2}{2} Var(X_j)} \quad \text{since } \Sigma \text{ is diagonal.}
\end{aligned}$$

Hence the X_j's are independent. □

Remark. *It is not enough that all entries of X are Gaussian for X to be a Gaussian vector itself. For example, let* $Y \sim \mathcal{N}(0,1)$ *and* ε *such that* $\mathbb{P}(\varepsilon = 1) = \mathbb{P}(\varepsilon = -1) = 1/2$*, with Y and* ε *independent. Then one can check that* $\varepsilon \cdot Y \sim \mathcal{N}(0,1)$*. In addition,* $Cov(Y, \varepsilon \cdot Y) = \mathbb{E}[\varepsilon \cdot Y^2] = \mathbb{E}[\varepsilon]\,\mathbb{E}[Y^2] = 0$*, but Y and* $\varepsilon \cdot Y$ *are not independent since* $|Y| = |\varepsilon \cdot Y|$ *a.s. Hence,* $X := (Y, \varepsilon \cdot Y)$ *is not Gaussian.*

Let $X \sim \mathcal{N}(m, \Sigma)$ be a Gaussian vector. Let $\phi : \mathbb{R}^d \to \mathbb{R}^d$ be the map whose matrix is Σ (in the standard basis $(e_1, \ldots, e_d)$). Since Σ is symmetric, there exists an orthonormal basis $(\varepsilon_1, \ldots, \varepsilon_d)$ in which ϕ is diagonal. Moreover, since Σ is positive, all its eigenvalues are nonnegative. Hence, we can write the eigenvalues as $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > 0 = \lambda_{r+1} = \ldots = \lambda_d$. The value $r \leq d$ is the rank of Σ and we can assume that the eigenbasis $(\varepsilon_1, \ldots, \varepsilon_d)$ has this order: $\phi(\varepsilon_i) = \lambda_i \varepsilon_i$. Suppose wlog that $m = 0$.

Theorem 5. *In the basis* $(\varepsilon_1, \ldots, \varepsilon_d)$*, we can write*

$$X = \sum_{j=1}^{d} Y_j \varepsilon_j,$$

where the Y_j*'s are independent Gaussian variables with mean* 0 *and variance* $Var(Y_j) = \lambda_j \geq 0$ *(if* $\lambda_j = 0$*, we just have* $Y_j = 0$ *a.s.).*

Proof. As $(\varepsilon_1, \ldots, \varepsilon_d)$ is a basis, there is a unique such expression. Note first that $(Y_j)_{1 \leq j \leq d}$ is a Gaussian vector. Indeed, defining P the transition matrix from the base $(\varepsilon_1, \ldots, \varepsilon_d)$ to the base $(e_1, \ldots, e_d)$, we have $Y = PX$, so Y is Gaussian. We claim that the covariance matrix C of Y is the matrix of ϕ in $(\varepsilon_1, \ldots, \varepsilon_d)$, which will be enough to conclude. We have for all u, v that $q(u,v) = Cov(\langle u, X\rangle, \langle v, X\rangle)$. In particular for $u = \varepsilon_j$ and $v = \varepsilon_k$, $q(u,v) = Cov(Y_j, Y_k)$. But, in the basis $(e_1, \ldots, e_d)$,

$$\begin{aligned} q(\varepsilon_j, \varepsilon_k) &= \sum_{j',k'} (\varepsilon_j)_{j'} (\varepsilon_k)_{k'} \Sigma_{j'k'} \\ &= \sum_{j'} (\varepsilon_j)_{j'} \left(\sum_{k'} (\varepsilon_k)_{k'} \Sigma_{j',k'} \right) \\ &= \sum_{j'} (\varepsilon_j)_{j'} (\phi(\varepsilon_k))_{j'} = \langle \varepsilon_j, \phi(\varepsilon_k) \rangle = \langle \varepsilon_j, \lambda_k \varepsilon_k \rangle. \end{aligned}$$

This is nonzero only when $j = k$ and we get $Var(Y_j) = \lambda_j$ as desired. □

Corollary 2. *Let* Σ *be a symmetric positive semidefinite matrix. Then there exists a Gaussian vector X such that* $\Sigma_X = \Sigma$*.*

Proof. Take $\lambda_1, \ldots, \lambda_d$ and P orthonormal such that ${}^tP diag(\lambda_i) P = \Sigma$. For $1 \leq i \leq d$, let $Y_i \sim \mathcal{N}(0, \lambda_i)$ so that the Y_i's are independent, and define $X = P^{-1}Y$. Then, $\Sigma_X = \Sigma$ by the computations above. □

1.5 Gaussian processes

Definition 6. *Let* $(E, \mathscr{E})$ *be a measurable space and T a set. A stochastic process in E indexed by T is a collection of random variables* $(X_t)_{t \in T}$ *defined on* $(E, \mathscr{E})$*.*

Remark. *Often, we take* $E=\mathbb{R}$ *or* $\mathbb{R}^d$, $T=\mathbb{N}$ *(discrete time) or* $T=[0,\infty)$ *(continuous time). In the latter case, we write* $(X_t)_{t\geq 0}$.

Definition 7. *A Gaussian process is a stochastic process* (X_t) *in* $\mathbb{R}$ *such that any finite linear combination of the variables* $(X_t)_{t\geq 0}$ *is Gaussian:* $\forall n\geq 1$, $\forall u_1,\dots,u_n\in\mathbb{R}$, $\forall t_1,\dots,t_n\in T$, $\sum_{i=1}^n u_i X_{t_i}$ *is Gaussian.*

As in the case of a Gaussian vector, we can define the *covariance function* of a stochastic process.

Definition 8. *Given a stochastic process* $(X_t)_{t\in T}$, *its covariance function* $C_X : T\times T\to\mathbb{R}$ *is defined as*

$$\forall s,t\in T, C_X(s,t)=Cov(X_s,X_t).$$

Definition 9. *We say that a function* $C:T\times T\to\mathbb{R}$ *is of positive type if* $\forall n\geq 1, \forall t_1,\dots,t_n\in T$, $(C(t_i,t_j))_{1\leq i,j\leq n}$ *is a positive semidefinite matrix, that is,*

$$\forall\lambda_1,\dots,\lambda_n\in\mathbb{R}, \sum_{i,j=1}^n \lambda_i\lambda_j C(t_i,t_j)\geq 0.$$

Note that a covariance function is always of positive type. Indeed,

$$\forall\lambda_1,\dots,\lambda_n\in\mathbb{R}, \sum_{i,j=1}^n \lambda_i\lambda_j Cov(X_{t_i},X_{t_j}) = Var(\sum_{i=1}^n \lambda_i X_{t_i})\geq 0.$$

Conversely, if one is given a function $C:T\times T\to\mathbb{R}$ of positive type, does there exist a stochastic process $(X_t)_{t\in T}$ on some measurable space such that $C_X=C$?

Theorem 6. *Let* $C:T\times T\to\mathbb{R}$ *(here,* T *is arbitrary) be a symmetric (that is,* $C(s,t)=C(t,s)$*) function of positive type. Then there exists a probability space and a Gaussian process indexed by* T *whose covariance function is* C.

Lecture 4: 16/10/2023

This is the consequence of a much more general theorem called Kolmogorov's extension theorem, which allows to construct random processes with prescribed finite-dimensional distributions. I will state it in the case $T=\mathbb{R}_+$ and processes with values in a Polish space E (that is, separable and complete metric space). Let $\Omega=E^{\mathbb{R}_+}$ the set of functions $\omega:\mathbb{R}_+\to E$ from $T=\mathbb{R}_+$ to E. We endow Ω with a σ-algebra $\mathscr{F}$ generated by the coordinate functions $\omega\mapsto\omega(t)$, $t\in\mathbb{R}_+$. This is the smallest σ-algebra in which, for all t, $\omega\mapsto\omega(t)$ is a random variable. Let $F(\mathbb{R}_+)$ denote the finite sets of $\mathbb{R}_+$ and, for $U\in\mathbb{R}_+$, let $\pi_U:\Omega\to E^U$ be the function which, to ω, associates $\omega_{|U}$. Furthermore, if $U\subseteq V\in F(\mathbb{R}_+)$, we write similarly $\pi_U^V : E^V\to E^U$ for the obvious restriction map.

Theorem 7 (Kolmogorov's extension theorem)**.** *Suppose that* E *is Polish, with its Borel* σ*-algebra* $\mathscr{E}$. *For every* $U\in F(\mathbb{R}_+)$, *let* μ_U *be a probability measure on* E^U. *Assume that* $\{\mu_U\}_{U\in F(\mathbb{R}_+)}$ *is consistent, that is, for any* $U\subseteq V\in F(\mathbb{R}_+)$, μ_U *is the image of* μ_V *under* π_U^V. *Then, there exists a unique probability measure on* $(\Omega,\mathscr{F})$ *such that, for all* $U\in F(\mathbb{R}_+)$, μ_U *is the image of* μ *under* π_U.

Proof of Theorem 6. In our case, observe that for all $U \in F(T)$ there exists a Gaussian vector, say X^U, with the correct covariance matrix $C_{|U}$. Furthermore, the distributions form a consistent family: projecting from V to $U \subseteq V$ is equivalent to restricting the matrix $C_{|V}$ to $C_{|U}$. By Kolmogorov's extension theorem, there exists a process indexed by T with the correct finite-dimensional distributions (Gaussian), and hence covariance function C. □

2 Brownian motion: definitions and properties

We define here the archetypical stochastic process, called the *Brownian motion*.

- named after Brown (1827), who looked at the motion of small particles of pollen in water.
- 1901, Bachelier, first mathematical model and applications to finance ;
- 1905, physical model (quantitative) by Einstein, molecules pushing randomly a macroscopic particle.
- 1923, first rigorous mathematical definition (as a random process) by Wiener.

2.1 Definition

We start by defining the so-called pre-Brownian motion, which is a stochastic process with the right characteristics. A Brownian motion will be a continuous version of a pre-Brownian motion.

Definition 10. *Let $X = (X_t)_{t\geq 0}$ be a stochastic process. We say that X is a pre-Brownian motion if the following holds:*

(i) $X_0 = 0$ *a.s.*

(ii) $\forall n \geq 1, \forall 0 \leq t_1 \leq \ldots \leq t_n$, *the variables* $X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ *are independent and Gaussian ;*

(iii) $\forall t \geq 0$, $X_t \sim \mathcal{N}(0,t)$.

Let us start by proving the existence of a pre-Brownian motion.

Proposition 4. *Let X be a stochastic process. Then, X is a pre-Brownian motion if and only if X is a centered Gaussian process with $Cov(X_s, X_t) = s \wedge t$.*

Proof. ($\Rightarrow$) Let $0 \leq t_1 \leq \ldots \leq t_n$. Then, $(X_{t_1}, \ldots, X_{t_n})$ is a linear transform of $(X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}})$, which consists in independent Gaussian variables. Hence, it is a Gaussian vector. Since $X_t \sim \mathcal{N}(0,t)$, X is also centered. Finally, for any $s \leq t$, we have

$$\begin{aligned} Cov(X_s, X_t) &= \mathbb{E}[X_s X_t] = \mathbb{E}[X_s(X_t - X_s)] - \mathbb{E}[X_s^2] \\ &= \mathbb{E}[X_s]\mathbb{E}[X_t - X_s] + \mathbb{E}[X_s^2], \end{aligned}$$

by independence of X_s and $X_t - X_s$. The first term is 0, and the second is $s := s \wedge t$.

($\Leftarrow$) Since X is Gaussian and centered, we have for all $t \geq 0$ that $X_t \sim \mathcal{N}(0, Var(X_t))$. But $Var(X_t) = \mathbb{E}[X_t^2] = t \wedge t = t$. This proves (iii). Since $X_0 \sim \mathcal{N}(0,0), X_0 = 0$ a.s. which proves (i). In order to prove (ii), fix $t_1 < \ldots < t_n$ and observe that $(X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}})$ is a Gaussian vector. Moreover, for all $i < j$, we have (with the convention $t_0 = 0$)

$$\begin{aligned} Cov(X_{t_i} - X_{t_{i-1}}, X_{t_j} - X_{t_{j-1}}) &= t_i \wedge t_j - t_{i-1} \wedge t_j - t_{j-1} \wedge t_i + t_{i-1} \wedge t_{j-1} \\ &= t_i - t_{i-1} - t_i + t_{i-1} = 0. \end{aligned}$$

Hence the variables are indeed independent. □

Corollary 3 (Wiener, 1923)**.** *A (pre)-Brownian motion exists.*

Proof. By Kolmogorov's extension theorem, it suffices to show that $C(s,t) := s \wedge t$ is a symmetric function of positive type. The symmetry is clear. To prove that it is of positive type (we have actually already done it), observe that $(C(t_i,t_j))_{1\leq i,j\leq n}$ coincides with the covariance matrix of $(Y_1,\dots,Y_n)$ where the $Y_i - Y_{i-1}$ are independent $\mathcal{N}(0,t_i - t_{i-1})$. Hence the function C is of positive type. ☐

Remark. *Properties (i),(ii),(iii) should be regarded as the definition of a Brownian motion. We will see later that the Brownian motion is the "universal continuous random walk", in the following sense. Take $(A_i)_{i\geq 1}$ i.i.d. random variables with $\mathbb{P}(A_i = 1) = P(A_i = -1) = 1/2$. Then, define the random walk $(Y_t, t \geq 0)$ as follows. $Y_0 = 0$, $Y_k = \sum_{i=1}^{k} A_i$ for any integer $k \geq 1$, and we take the linear interpolation between two consecutive integers. Then, the following holds:*

$$\left(\frac{Y_{nt}}{\sqrt{n}}\right)_{t\geq 0} \xrightarrow[n\to\infty]{(d)} (B_t)_{t\geq 0},$$

where $(B_t)_{t\geq 0}$ is a (pre)-Brownian motion. Therefore, it is natural to ask (at the limit) that the increments are independent. In addition, the central limit theorem gives $Y_{nt}/\sqrt{n} \xrightarrow[n\to\infty]{(d)} \mathcal{N}(0,t)$.

Property (ii) says that a (pre)-Brownian motion has independent increments: if $[s,t] \cap [u,v] = \emptyset$, then $X_v - X_u$ and $X_t - X_s$ are independent. It also follows from the definition that its increments are stationary, that is, the law of $X_t - X_s$ only depends on $t - s$. Indeed, one has $X_t - X_s \sim \mathcal{N}(0,t-s)$. To see this, observe that $Var(X_t - X_s) = Var(X_t) + Var(X_s) - 2Cov(X_s,X_t) = t + s - 2s \wedge t = t - s$ if $s \leq t$.

Hence, X has stationary and independent increments.

2.2 Regularity of the paths

Lecture 5: 18/10/2023

Definition 11. *Let $(X_t)_{t\in T}$ be a stochastic process. We say that X is continuous if, for all ω in Ω, $t \mapsto X_t(\omega)$ is continuous. In other words, all realizations of X are continuous.*

An important notion is the notion of *modification* of a stochastic process.

Definition 12. *Let $(X_t)_{t\in T}, (\tilde{X}_t)_{t\in T}$ be two stochastic processes. We say that $\tilde{X}$ is a modification of X if*

$$\forall t \in T, \mathbb{P}(X_t = \tilde{X}_t) = 1.$$

We say that X and $\tilde{X}$ are indistinguishable if

$$\mathbb{P}(\forall t \in T, \tilde{X}_t = X_t) = 1.$$

If the set T is uncountable, this is a stronger notion than the one of modification.

Note that both are equivalence relations. Usually, we consider stochastic processes up to indistinguishability, and consider that two indistinguishable processes are the same. Note also that, if $\tilde{X}$ is a modification of X, then $\tilde{X}$ and X have the same finite-dimensional marginal distributions.

Finally, observe that a stochastic process has at most one continuous modification. Indeed, if $\tilde{X}$ is a modification of X and both are continuous, then they are indistinguishable. If $T = \mathbb{R}_+$ for instance, $\mathbb{P}(\forall q \in \mathbb{Q} \cap \mathbb{R}_+, \tilde{X}_q = X_q) = 1$, and by continuity we conclude.

Definition 13. *The Brownian motion is a continuous modification of a pre-Brownian motion.*

We now need to show that the Brownian motion exists.

Theorem 8. *The pre-Brownian motion has a continuous modification.*

Proof: Lévy's construction of the Brownian motion, 1948. We will define such a process $(X_t)_{t\geq 0}$ by induction as a piecewise linear function on $[0,1]$. Let $(\xi_{k,n}, n \geq 0, 0 \leq k \leq 2^n)$ be a family of i.i.d. $\mathcal{N}(0,1)$ random variables.

- Define $X_0(0) = 0$, $X_0(1) = \xi_{0,0}$, and X_0 is linear on $[0,1]$.
- Let $X_1(0) = 0$, $X_1(1) = X_0(1)$, $X_1(1/2) = X_0(1/2) + \frac{\xi_{1,1}}{2}$ and X_1 is piecewise linear.
- More generally, if X_{n-1} is defined on $[0,1]$, on each dyadic interval $[\frac{k}{2^n}, \frac{k+1}{2^n}](0 \leq k \leq 2^n - 1)$, we define for k even $X_n(\frac{k}{2^n}) = X_{n-1}(\frac{k}{2^n})$, and for $k = 2j+1$ we define $X_n(\frac{k}{2^n}) = X_{n-1}(\frac{k}{2^n}) + \frac{\xi_{k,n}}{2^{(n+1)/2}}$, and X_n is piecewise linear inbetween.

We now need to prove that we get indeed, at the limit, a process with the right properties.

Step 1 :

$(X_n(\frac{k}{2^n}), 0 \leq k \leq 2^n)$ is a Gaussian vector, centered, and $\mathbb{E}[X_n(s)X_n(t) = s \wedge t]$ for $s,t \in D_n := \{\frac{k}{2^n}, 0 \leq k \leq 2^n\}$.

Observe first that, for all k, $X_n(\frac{k}{2^n})$ is a linear combination of the variables $(\xi_{j,m}, m \leq n, 0 \leq j \leq 2^m)$, which are independent centered Gaussian variables. Hence, $(X_n(\frac{k}{2^n}), 0 \leq k \leq 2^n)$ is a centered Gaussian vector.

We will only prove that they have the correct variance (that is, for $s = t$, $Var(X_n(t) = t)$. Exercise: prove that they have the correct covariance.

We prove it by induction.

For $n = 0$, $X_0(1) = \xi(0,0)$ so it is true.
For $n \geq 1$, let $t = \frac{k}{2^n}$. If k is even it is true by induction. If k is odd, let $t^{\pm} = \frac{k\pm 1}{2^n}$. observe that

$$X_n(t) = \frac{1}{2}\left(X_{n-1}(t^-) + X_{n-1}(t^+)\right) + \frac{\xi_{k,n}}{2^{(n+1)/2}}.$$

Thus, we have

$$\begin{aligned}
Var(X_n(t)) &= \frac{Var(\xi_{k,n})}{2^{n+1}} + \frac{1}{4}\left(t^- + t^+ + 2Cov(X_{n-1}(t^-), X_{n-1}(t^+))\right) \\
&= \frac{1}{2^{n+1}} + \frac{1}{4}\left(t - \frac{1}{2^n} + t + \frac{1}{2^n} + 2\left(t - \frac{1}{2^n}\right)\right) \\
&= \frac{1}{2^{n+1}} + \frac{1}{4}\left(4t - \frac{2}{2^n}\right) \\
&= t.
\end{aligned}$$

Step 2: Convergence of $(X_n, n \geq 0)$.

For all $n \geq 1$, define the event $A_n = \left\{\sup_{t\in[0,1]} |X_n(t) - X_{n-1}(t)| \geq 2^{-n/4}\right\}$. Then

$$\begin{aligned}
\mathbb{P}(A_n) &= \mathbb{P}\left(\bigcup_{j=0}^{2^{n-1}-1} \sup_{[\frac{2j}{2^n}, \frac{2j+2}{2^n}]} |X_n(t) - X_{n-1}(t)| \geq 2^{-n/4}\right) \\
&\leq \sum_{j=0}^{2^{n-1}-1} \mathbb{P}\left(\frac{|\xi_{2j+1,n}|}{2^{(n+1)/2}} \geq 2^{-n/4}\right) \\
&= 2^{n-1}\mathbb{P}\left(|\mathcal{N}(0,1)| \geq 2^{(n+2)/4}\right) \\
&= 2^n \mathbb{P}\left(\mathcal{N}(0,1) \geq 2^{(n+2)/4}\right).
\end{aligned}$$

Now, we use the bound

$$\begin{aligned}
\mathbb{P}(\mathcal{N}(0,1) \geq x) &\leq \frac{1}{x} e^{-x^2/2} \\
&\leq e^{-x^2/2}
\end{aligned}$$

for $x \geq 1$. Hence,

$$\mathbb{P}(A_n) \leq 2^n e^{-\frac{1}{2}2^{(n+2)/2}},$$

which decays doubly exponentially. Thus, by Borel-Cantelli, A_n occurs only for finitely many n. It follows that $\sum_{n\geq 0}(X_n(t) - X_{n-1}(t))$ is a convergent series, and so X_n converges uniformly a.s. (for convenience, set $X_{-1}(t) = 0$ for $t \in [0,1]$.) The limit $(X(t), t \in [0,1])$ is continuous by construction. □

One can check that $(X(t), t \in [0,1])$ is indeed a pre-Brownian motion. The idea is to approximate $t_1, \ldots, t_k \in [0,1]$ by dyadic numbers, and use the convergence of characteristic functions (due to continuity of the process $(X(t), t \in [0,1])$.

From now on, a (continuous) Brownian motion will be denoted by $(B_t)_{t\geq 0}$.

2.3 Hölder exponent

Lecture 6: 23/10/2023

In fact, the paths of the Brownian motion are even more regular. We will show that, on $[0,1]$, $t \mapsto X_t$ is a.s. $(1/2-\varepsilon)$-Hölder, for any $\varepsilon > 0$. This is the consequence of a fondamental lemma in probability, also due to Kolmogorov.

(Recall that $f : I \to \mathbb{R}$ is α-Hölder if there exists $C > 0$ such that, for all $s,t \in I$, $|f(t)-f(s)| \leq C|t-s|^{\alpha}$.)

Theorem 9 (Kolmogorov's continuity criterion)**.** *Let $X = (X_t)_{t\in[0,1]}$ be a stochastic process taking values in a complete metric space (E,d). Suppose that there exist $p > 0$, $\varepsilon > 0$, $C > 0$ such that, for all $s,t \in \mathbb{R}_+$:*

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

Then, there exists a modification $\tilde{X}$ of X such that, for any $0 < \alpha < \frac{\varepsilon}{p}$, $\tilde{X}$ is a.s. α-Hölder: a.s. there exists $K = K(\omega)$ such that

$$\forall t \in [0,1], d(X_s,X_t) \leq K|t-s|^{\alpha}.$$

Proof. For convenience, let $(E,d) = (\mathbb{R}_+, |\cdot|)$. Recall the notation $D_n = \{\frac{k}{2^n}, 0 \leq k \leq 2^n\}$. Observe that, for $t = \frac{k}{2^n} \in D_n$, by Markov's inequality:

$$\begin{aligned}
\mathbb{P}\left(\left|X_{(k+1)2^{-n}} - X_{k2^{-n}}\right| > 2^{-n\alpha}\right) &\leq \frac{\mathbb{E}\left[|X_{(k+1)2^{-n}} - X_{k2^{-n}}|^p\right]}{2^{-n\alpha p}} \\
&\leq C\frac{2^{-n(1+\varepsilon)}}{2^{-n\alpha p}} = C2^{-n}2^{-n(\varepsilon-\alpha p)}.
\end{aligned}$$

By summing over all $t \in D_n$, we get

$$p_n := \mathbb{P}\left(\sup_{0\leq k\leq 2^n-1} \left|X_{(k+1)2^{-n}} - X_{k2^{-n}}\right| > 2^{-n\alpha}\right) \leq C2^{-n(\varepsilon-\alpha p)}.$$

In particular, for $\alpha < \frac{\varepsilon}{p}$, the sum $\sum_{n\geq 0} p_n$ is finite. So, by Borel-Cantelli, a.s. there exists $n_0(\omega)$ such that, for $n \geq n_0(\omega)$, $\sup_{0\leq k\leq 2^n-1} \left|X_{(k+1)2^{-n}} - X_{k2^{-n}}\right| \leq 2^{-n\alpha}$. Hence, there exists $M(\omega) < \infty$ such that

$$\sup_{n\geq 1}\sup_{0\leq k\leq 2^n-1} \frac{\left|X_{(k+1)2^{-n}} - X_{k2^{-n}}\right|}{2^{n\alpha}} \leq M(\omega).$$

We claim that this implies that for $s,t \in D = \bigcup_{n\geq 0} D_n$, $|X_s - X_t| \leq K(\omega)|t-s|^{\alpha}$.

We do it by chaining. Choose $s,t \in D$, $s < t$.

Let r be the smallest integer such that $t-s > 2^{-r-1}$. In particular, $t-s \leq 2^{-r}$. Then, there exists $0 \leq k \leq 2^{r+1}$ and integers $\ell, m > 0$ such that, for some $(\varepsilon_i)_{1\leq i\leq \ell}, (\varepsilon'_j)_{1\leq j\leq m} \in \{0,1\}$, we have

$$\begin{aligned}
s &= k2^{-r-1} - \varepsilon_1 2^{-r-1} - \ldots - \varepsilon_\ell 2^{-r-\ell} \\
t &= k2^{-r-1} + \varepsilon'_1 2^{-r-1} + \ldots + \varepsilon'_m 2^{-r-m}.
\end{aligned}$$

Let us write, for $0 \le i \le \ell$, $s_i = k2^{-r-1} - \varepsilon_1 2^{-r-1} - \ldots - \varepsilon_i 2^{-r-i}$, and for $0 \le i \le m$, $t_i = k2^{-r-1} + \varepsilon'_1 2^{-r-1} + \ldots + \varepsilon'_i 2^{-r-i}$. By triangle inequality, we have

$$\begin{aligned}
|X_t - X_s| = |X_{t_m} - X_{s_\ell}| &\le |X_{t_0} - X_{s_0}| + \sum_{i=1}^{\ell} |X_{s_i} - X_{s_{i-1}}| + \sum_{i=1}^{m} |X_{t_i} - X_{t_{i-1}}| \\
&\le M(\omega) 2^{-(r+1)\alpha} + M(\omega) \sum_{i=1}^{\ell} 2^{-(r+i)\alpha} + M(\omega) \sum_{i=1}^{\ell} 2^{-(r+i)\alpha} \\
&\le 2^{-(r+1)\alpha} M(\omega) \left(1 + \frac{2}{1 - 2^{-\alpha}}\right) = 2^{-(r+1)\alpha} K(\omega) \\
&\le K(\omega) |t - s|^{\alpha}.
\end{aligned}$$

Since $X_{|D}$ is α-Hölder and thus uniformly continuous on a dense set, $X_{|D}$ has a unique extension $\tilde{X}$ on $[0,1]$ which is also α-Hölder. Indeed, set

$$\tilde{X}_t = \lim_{\substack{s \to t \\ s \in D}} X_s.$$

On the exceptional set where $X_{|D}$ is not α-Hölder, set $\tilde{X} \equiv 0$. Let us show that $\tilde{X}$ is a modification of X. Fix $t \in [0,1]$. Let $t_n \in D$ such that $t_n \to t$. By definition, $\tilde{X}_t = \lim_{n \to \infty} X_{t_n}$. Thus, by Fatou's lemma,

$$\begin{aligned}
\mathbb{E}\left[|X_t - \tilde{X}_t|^p\right] &\le \liminf_{n \to \infty} \mathbb{E}\left[|X_t - X_{t_n}|^p\right] \\
&\le C |t - t_n|^{1+\varepsilon} \to 0,
\end{aligned}$$

so $X_t = \tilde{X}_t$ a.s. □

Corollary 4. *If B is a Brownian motion, then B is a.s. locally α-Hölder for any $\alpha < 1/2$.*

Proof. Indeed, for any $p > 0$,

$$\mathbb{E}\left[|B_t - B_s|^p\right] \le C_p |t - s|^{p/2},$$

so B is locally α-Hölder for $\alpha < \frac{\varepsilon}{p}$ with $\varepsilon = \frac{p}{2} - 1$. To see this, just observe that

$$\mathbb{E}\left[|B_t - B_s|^p\right] = \mathbb{E}\left[|\mathcal{N}(0, t-s)|^p\right] = (t-s)^{p/2} \mathbb{E}\left[\mathcal{N}(0,1)\right].$$

Hence $\frac{\varepsilon}{p} = \frac{p/2-1}{p} = \frac{1}{2} - \frac{1}{p}$. Since p is arbitrary, any $\alpha < 1/2$ works. □

Remark. *It can be shown that B is a.s. nowhere α-Hölder for $\alpha \ge 1/2$. In particular, the Brownian motion is nowhere differentiable (Paley-Wiener-Zygmund theorem).*

2.4 Wiener measure

The fact that we have a continuous modification of (pre)-Brownian motion makes possible to define the entire path of a Brownian motion as a single random variable, taking its values in the space of continuous functions. This can be a useful point of view.

Let $\Omega^* := C(\mathbb{R}_+, \mathbb{R})$ the set of continuous functions from $\mathbb{R}_+$ to $\mathbb{R}$. We equip it with the topology of uniform convergence on all compact sets, and its Borel σ-field $\mathscr{F}^*$.

$$d(\omega,\tilde{\omega}) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{\delta_n(\omega,\tilde{\omega})}{1+\delta_n(\omega,\tilde{\omega})},$$

where

$$\delta_n(\omega,\tilde{\omega}) = \sup_{t\in[0,n]} |\omega(t) - \tilde{\omega}(t)|.$$

Let $(X_t)_{t\geq 0}$ be the canonical process on $(\Omega^*, \mathscr{F}^*)$, that is:

$$X_t : \Omega^* \to \mathbb{R}, \omega \mapsto \omega(t).$$

Remark. *We have $\mathscr{F}^* = \sigma(X_t, t \geq 0)$, the σ-algebra considered in Kolmogorov's extension theorem.*

Let B be a (continuous) Brownian motion, defined on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$. We can define a measurable map $\phi : \Omega \to \Omega^*$ by setting $\phi(\omega) = (t \mapsto B_t(\omega))$. The image $\mathbb{W}$ of $\mathbb{P}$ under ϕ defines a measure on $(\Omega^*, \mathscr{F}^*)$, which is the law of any continuous Brownian motion on $(\Omega^*, \mathscr{F}^*)$.

Definition 14. *This measure $\mathbb{W}$ is called the Wiener measure. On $(\Omega^*, \mathscr{F}^*, \mathbb{W})$, the coordinate process is by definition a Brownian motion. The measure $\mathbb{W}$ is essentially the "Lebesgue measure" on $(\Omega^*, \mathscr{F}^*)$.*

2.5 First properties of Brownian motion

Lecture 7: 28/10/2023

We can now investigate some properties of the Brownian motion. We start with the so-called *Markov property*.

Theorem 10. *Let B be a Brownian motion and $s \geq 0$. Then, the process*

$$\tilde{B} = (B_{t+s} - B_s, t \geq 0)$$

is a Brownian motion, independent of $\mathscr{F}_s := \sigma(B_u, u \leq s)$.

Proof. First, $\tilde{B}_0 = 0$. Also, $\tilde{B}$ is continuous, $\tilde{B}_t \sim \mathscr{N}(0,t)$ for all $t \geq 0$ and the increments of $\tilde{B}$ are independent Gaussian, so $\tilde{B}$ is a Brownian motion. We now need to check the independence property. Fix $m, n \geq 1$, $s_1 \leq \ldots \leq s_m \leq s$, $t_1 \leq \ldots \leq t_n$. It suffices to check that $(\tilde{B}_{t_1}, \ldots, \tilde{B}_{t_n})$ is independent of $(B_{s_1}, \ldots, B_{s_m})$. But $(B_{s_1}, \ldots, B_{s_m}, \tilde{B}_{t_1}, \ldots, \tilde{B}_{t_n})$ is a Gaussian vector and, for all $1 \leq i \leq m$, all $1 \leq j \leq n$, we have

$$Cov(B_{s_i}, \tilde{B}_{t_j}) = Cov(B_{s_i}, B_{s+t_j}) - Cov(B_{s_i}, B_s) = s_i - s_i = 0.$$

□

Corollary 5. *Conditionally given $\mathscr{F}_s$, $B' = (B_{t+s}, t \geq 0)$ is a Brownian motion started from $x = B_s$.*

Indeed, $\forall n \geq 1, \forall t_1, \ldots, t_n$,

$$\begin{aligned}\mathbb{E}[F(B'_{t_1}, \ldots, B'_{t_n}) | \mathscr{F}_s] &= \mathbb{E}[F(B_s + \tilde{B}_{t_1}, \ldots, B_s + \tilde{B}_{t_n}) | \mathscr{F}_s] \\ &= \mathbb{E}^x[F(X_{t_1}, \ldots, X_{t_n})],\end{aligned}$$

for X a Brownian motion.

Theorem 11 (Blumenthal's $0-1$ law)**.** *For any $t \geq 0$, let $\mathscr{F}_t = \sigma(B_s, s \leq t)$ and $\mathscr{F}_{t+} := \bigcap_{u>t} \mathscr{F}_u$. Then, $\mathscr{F}_{0+}$ is trivial, that is, $\forall A \in \mathscr{F}_{0+}$, $\mathbb{P}(A) \in \{0,1\}$.*

Proof. Let $A \in \mathscr{F}_{0+}$, $n \geq 1$, $t_1 \leq \ldots \leq t_n$. For any $\varepsilon > 0$, any $F : \mathbb{R}^n \to \mathbb{R}$ bounded and continuous, we have

$$\mathbb{E}[\mathbb{1}_A F(B_{t_1+\varepsilon} - B_\varepsilon, \ldots, B_{t_n+\varepsilon} - B_\varepsilon)] = \mathbb{E}[\mathbb{1}_A]\mathbb{E}[F(\tilde{B}_{t_1}, \ldots, \tilde{B}_{t_n})].$$

Let $\varepsilon \to 0$. By dominated convergence, we get

$$\mathbb{E}[\mathbb{1}_A F(B_{t_1}, \ldots, B_{t_n})] = \mathbb{P}(A)\mathbb{E}[F(B_{t_1}, \ldots, B_{t_n})]$$

Since A and $t_1, \ldots, t_n$ are arbitrary, we deduce that $\mathscr{F}_{0+}$ is independent of $(B_t)_{t\geq 0}$. But if $A \in \mathscr{F}_{0+}$, A is also measurable with respect to $(B_t)_{t\geq 0}$. Hence

$$\mathbb{P}(A) = \mathbb{P}(A \cap A) = \mathbb{P}(A)^2,$$

and $\mathbb{P}(A) \in \{0,1\}$. □

Example. *Let $\tau := \inf\{t \geq 0, B_t > 0\}$. Then $\tau = 0$ a.s. Indeed,*

$$\mathbb{P}(\tau = 0) = \lim_{t\to 0+} \mathbb{P}(\tau \leq t) \geq \limsup_{t\to 0} \mathbb{P}(B_t > 0) = 1/2,$$

so $\mathbb{P}(\tau = 0) \geq 1/2$. By Blumenthal's $0-1$ law, we have $\mathbb{P}(\tau = 0) \in \{0,1\}$. The result follows. By symmetry, $\inf\{t \geq 0, B_t < 0\} = 0$ a.s. Consequently, B has infinitely many zeros around 0.

We now turn to invariance properties of the Brownian motion.

Proposition 5 (Invariance properties)**.** *Let $B = (B_t)_{t\geq 0}$ be a Brownian motion. Then*

(i) $(-B_t)_{t\geq 0}$ is a Brownian motion.

(ii) for any $\lambda > 0$, $(\lambda^{-1/2} B_{\lambda t})_{t\geq 0}$ is a Brownian motion (scaling property, self-similarity).

(iii) The process defined by $X_0 = 0$, $X_t = tB_{1/t}$ for $t > 0$, is a Brownian motion (time inversion).

Proof. All three processes are centered Gaussian processes. We consider their covariance function. It is clear that the first one has the right covariance. For the second one, for $s,t \in \mathbb{R}_+$, we have

$$\mathbb{E}[\lambda^{-1/2} B_{\lambda s} \lambda^{-1/2} B_{\lambda t}] = \lambda^{-1}(\lambda s \wedge \lambda t) = s \wedge t.$$

Since these two processes are continuous, they are Brownian motions. The third one also clearly has the right covariance, and is continuous on $(0, +\infty)$. It remains to prove that it is continuous at the point 0 (Exercise...). □

As a consequence of (iii) and the example above, a.s. B hits 0 infinitely often in a neighbourhood of $+\infty$.

Example. *We have* $\sup B_s = +\infty$ *a.s.*
Indeed, for $t \geq 0$, $x \geq 0$, *we have*

$$\mathbb{P}(\sup_{s\in[0,t]} B_s \geq x) = \mathbb{P}(\sqrt{t}\sup_{u\in[0,1]} \tilde{B}_u \geq x),\ \text{where } \tilde{B}_u = \frac{1}{\sqrt{t}} B_{tu}$$
$$= \mathbb{P}\left(\sup_{u\in[0,1]} B_u \geq \frac{x}{\sqrt{t}}\right),$$

where the second line follows from self-similarity. Let $t \to +\infty$. *Then the left-hand side goes to* $\mathbb{P}(\sup_{s\geq 0} B_s \geq x)$ *by monotone convergence, and the right-hand side goes to* $\mathbb{P}\left(\sup_{u\in[0,1]} B_u \geq 0\right)$. *By the result above, we have that* $\mathbb{P}(\sup_{s\geq 0} B_s \geq x) = 1$ *for all* $x \geq 0$. *Hence* $\sup_{s\geq 0} B_s = +\infty$. *Therefore,* $\limsup_{s\geq 0} B_s = +\infty$, *and by symmetry* $\liminf_{s\geq 0} B_s = -\infty$.

In particular, for $a \in \mathbb{R}$, *define* $T_a = \inf\{t \geq 0, B_t = a\}$. *Then,* $T_a < \infty$ *a.s.*

2.6 Strong Markov property

We now consider a stronger version of the Markov property, using stopping times.

Definition 15. *Let* $(B_t)_{t\geq 0}$ *be a Brownian motion,* $\mathscr{F}_t := \sigma(B_u, u \leq t)$, $\mathscr{F}_\infty := \sigma(B_u, u \geq 0)$. *Let* $T : \Omega \to \mathbb{R}_+ \cup \{\infty\}$. *We say that* T *is a stopping time adapted to the filtration* $(\mathscr{F}_t, t \geq 0)$ *if, for any* $t \geq 0$, $\{T \leq t\} \in \mathscr{F}_t$.

Example. *For* $a \geq 0$, $T_a := \inf\{t \geq 0, B_t = a\}$ *is a stopping time. Indeed, for all* $t \geq 0$,

$$\{T_a \leq t\} = \{\sup_{s\in[0,t]} B_s \geq a\} = \{\sup_{s\in[0,t]\cap\mathbb{Q}} B_s \geq a\} \in \mathscr{F}_t.$$

We now define the σ-algebra associated to a stopping time.

Definition 16. *If* T *is a stopping time, we let*

$$\mathscr{F}_T := \{A \in \mathscr{F}_\infty : \forall t \in \mathbb{R}_+, A \cap \{T \leq t\} \in \mathscr{F}_t\}.$$

Example. *In particular,* T *is* $\mathscr{F}_T$*-measurable, by definition. Indeed, for all* $s > 0$, *all* $t > 0$, *we have* $\{T \leq s\} \cap \{T \leq t\} = \{T \leq s \wedge t\} \in \mathscr{F}_t$, *so* $\{T \leq s\} \in \mathscr{F}_T$. *Also,* $B_T \mathbb{1}_{T<\infty}$ *is* $\mathscr{F}_T$*-measurable. Indeed,*

$$B_T \mathbb{1}_{T<\infty} = \lim_{n\to+\infty} \sum_{i=0}^{\infty} B_{i/2^n} \mathbb{1}_{i/2^n \leq T \leq (i+1)/2^n}.$$

Lecture 8: 06/11/2023

Theorem 12 (Strong Markov property)**.** *Let* T *be a stopping time. Then, conditionally given* $T < +\infty$, $\tilde{B} := (B_{T+t} - B_T, t \geq 0)$ *is a Brownian motion, independent of* $\mathscr{F}_T$.

Proof. Suppose in a first time that $T < +\infty$ a.s., and take $F : \mathbb{R}_+^n \to \mathbb{R}$ a continuous and bounded function. Let $t_1 \leq \ldots \leq t_n$. We want to prove that

$$\mathbb{E}\left[\mathbb{1}_A F(\tilde{B}_{t_1}, \tilde{B}_{t_2}, \ldots, \tilde{B}_{t_n})\right] = \mathbb{P}(A)\mathbb{E}\left[F(B_{t_1}, \ldots, B_{t_n})\right].$$

Note that, as $m \to \infty$, almost surely:

$$\sum_{k=1}^{\infty} \mathbb{1}\left\{\frac{k-1}{2^m} < T \le \frac{k}{2^m}\right\} F\left(B_{\frac{k}{2^m}+t_1} - B_{\frac{k}{2^m}}, \dots, B_{\frac{k}{2^m}+t_n} - B_{\frac{k}{2^m}}\right) \underset{m\to\infty}{\to} F\left(\tilde{B}_{t_1}, \dots, \tilde{B}_{t_n}\right).$$

Hence,

$$\mathbb{E}\left[\mathbb{1}_A F\left(\tilde{B}_{t_1}, \dots, \tilde{B}_{t_n}\right)\right] = \lim_{m\to\infty} \sum_{k\ge 1} \mathbb{E}\left[\mathbb{1}\left\{A \cap \left\{\frac{k-1}{2^m} < T \le \frac{k}{2^m}\right\}\right\} F\left(B_{\frac{k}{2^m}+t_1} - B_{\frac{k}{2^m}}, \dots, B_{\frac{k}{2^m}+t_n} - B_{\frac{k}{2^m}}\right)\right]$$

By the simple Markov property, this is equal to

$$\begin{aligned}\lim_{m\to\infty} \sum_{k\ge 1} \mathbb{P}\left(A \cap \left\{\frac{k-1}{2^m} < T \le \frac{k}{2^m}\right\}\right) \mathbb{E}\left[F\left(B_{t_1}, \dots, B_{t_n}\right)\right] \\ = \mathbb{P}(A)\mathbb{E}\left[F\left(B_{t_1}, \dots, B_{t_n}\right)\right]\end{aligned}$$

If $T = +\infty$ with positive probability, consider the same argument with A replaced by $A \cap \{T < \infty\}$. □

Example. *Let $T_a = \inf\{t \ge 0, B_t = a\}$, for $a \in \mathbb{R}$. Let $a, b \ge 0$. Then, $T_{a+b} - T_a$ has the same law as T_b and is independent of $\{T_x, x \le a\}$. Indeed, observe that if $x \le a$ then T_x is $\mathscr{F}_{T_a}$-measurable. Applying strong Markov property to T_a provides the result.*

Hence, $(T_a, a \ge 0)$ has independent and stationary increments too. Furthermore, it has inverse Brownian scaling:

$$\left(\frac{1}{c^2} T_{ca}, a \ge 0\right) \overset{(d)}{=} (T_a, a \ge 0).$$

In particular, $T_a \overset{(d)}{=} a^2 T_1$.

Example. *Let $L = \sup\{t \in [0,1], B_t = 0\}$. Then $L < 1$ a.s. and L is not a stopping time.*

Indeed, if it was a stopping time, it would contradict the strong Markov property and the fact that there are infinitely many zeros in the neighbourhood of 0 for a standard Brownian motion.

Theorem 13 (Reflection principle)**.** *For $t \ge 0$, let $S_t := \sup_{0\le s\le t} B_s$. Then, for all $a > 0$, for all $b \in (-\infty, a]$,*

$$\mathbb{P}(S_t \ge a, B_t \le b) = \mathbb{P}(B_t \ge 2a - b).$$

Proof. We have

$$\begin{aligned}\mathbb{P}(S_t \ge a, B_t \le b) &= \mathbb{P}(T_a \le t, B_t \le b) \\ &= \mathbb{P}\left(T_a \le t, \tilde{B}_{t-T_a} \le b - a\right),\end{aligned}$$

where $\tilde{B} = (B_{T_a+s} - B_{T_a}, s \ge 0) = (B_{T_a+s} - a, s \ge 0)$. Hence, it is equal to

$$\begin{aligned}\mathbb{E}\left[\mathbb{1}_{T_a\le t}\mathbb{1}_{\tilde{B}_{t-T_a}\le b-a}\right] &= \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{T_a\le t}\mathbb{1}_{\tilde{B}_{t-T_a}\le b-a} | \mathscr{F}_{T_a}\right]\right] \\ &= \mathbb{E}\left[\mathbb{1}_{T_a\le t}\mathbb{E}\left[\mathbb{1}_{\tilde{B}_{t-T_a}\le b-a} | \mathscr{F}_{T_a}\right]\right] \\ &= \mathbb{E}\left[\mathbb{1}_{T_a\le t} h(T_a)\right],\end{aligned}$$

by the strong Markov property, where $h(s) := \mathbb{P}(B_{t-s} \le b-a)$. But by symmetry, $h(s) = \mathbb{P}(B_{t-s} \ge a-b)$. Thus, going to the other direction, we get

$$\begin{aligned}\mathbb{P}(S_t \ge a, B_t \le b) &= \mathbb{P}\left(T_a \le t, \tilde{B}_{t-T_a} \ge a-b\right) \\ &= \mathbb{P}(T_a \le t, B_t - a \ge a-b) \\ &= \mathbb{P}(T_a \le t, B_t \ge 2a-b) \\ &= \mathbb{P}(B_t \ge 2a-b) \text{ since } a \ge b.\end{aligned}$$

□

Corollary 6. *Fix $t \ge 0$. Then, $S_t \overset{(d)}{=} |B_t|$, that is, for all $a \ge 0$, $\mathbb{P}(S_t \ge a) = 2\mathbb{P}(\mathcal{N}(0,t) \ge a)$.*

Proof. The reflection principle for $b = a$ provides

$$\mathbb{P}(B_t \ge a) = \mathbb{P}(S_t \ge a, B_t \le a) = \mathbb{P}(S_t \ge a) - \mathbb{P}(S_t \ge a, B_t \ge a).$$

Since $\{B_t \ge a\} \subseteq \{S_t \ge a\}$, we obtain

$$\mathbb{P}(S_t \ge a) = 2\mathbb{P}(B_t \ge a).$$

□

Example. *Consider the density of the law of T_a.*

$$\mathbb{P}(T_a \le t) = \mathbb{P}(S_t \ge a) = \mathbb{P}(\sqrt{t}S_1 \ge a) = \mathbb{P}\left(\frac{a^2}{S_1^2} \le t\right).$$

Hence, $T_a \overset{(d)}{=} \frac{a^2}{S_1^2} \overset{(d)}{=} \frac{a^2}{|B_1|^2} \overset{(d)}{=} \frac{a^2}{|\mathcal{N}(0,1)|^2}$. The probability distribution function is then easily computed by change of variable (Exercise), and we obtain

$$f(t) = \frac{|a|}{\sqrt{2\pi t^3}} \exp\left(-\frac{a^2}{2t}\right) \mathbb{1}_{t>0}.$$

3 Martingales and finite variation processes

3.1 Martingales in continuous time

We start by defining the notion of filtration, which will be important in the whole section.

Definition 17. *We say that a family $(\mathscr{F}_t, t \geq 0)$ of σ-algebras on a space $(\Omega, \mathscr{F}, \mathbb{P})$ is a filtration if $\mathscr{F}_s \subseteq \mathscr{F}_t$ for all $s \leq t$. We define, for all $t \geq 0$, $\mathscr{F}_{t+} := \cap_{u>t} \mathscr{F}_u$, and we say that the filtration $(\mathscr{F}_t, t \geq 0)$ is right-continuous if, for all $t \geq 0$, $\mathscr{F}_t = \mathscr{F}_{t+}$. Finally, we say that it is complete if, for all $t \geq 0$, $\mathscr{F}_t$ contains also all subsets of sets of measure 0.*

In all this section, we will always consider right-continuous complete filtrations.

Lecture 9: 13/11/2023

Definition 18. *We say that a stochastic process $(X_t)_{t\geq 0}$ is a martingale (resp. submartingale, supermartingale) if:*

(i) it is adapted to the filtration, that is, for all $t \geq 0$, X_t is $\mathscr{F}_t$-measurable ;

(ii) $\mathbb{E}[|X_t|] < \infty$ for all $t \geq 0$;

(iii) for all $0 \leq s \leq t$, $\mathbb{E}[X_t|\mathscr{F}_s] = X_s$ (resp. $\geq$, $\leq$).

Remark. *Take any random variable Y such that $\mathbb{E}[|Y|] < \infty$. Define the process $X := (\mathbb{E}[Y|\mathscr{F}_s], s \geq 0)$. Then, X is a martingale by properties of the conditional expectation.*

Example.

- *B is a martingale.*

  Indeed, for $s \leq t$, $\mathbb{E}[B_t|\mathscr{F}_s] = \mathbb{E}[B_s + \tilde{B}_{t-s}|\mathscr{F}_s] = B_s + 0$.

- *$t \mapsto B_t^2 - t$ is a martingale.*

  Indeed, for $s \leq t$, we have

$$\begin{aligned}\mathbb{E}[B_t^2 - t|\mathscr{F}_s] &= \mathbb{E}[B_s^2 + (B_t - B_s)^2 + 2B_s(B_t - B_s)|\mathscr{F}_s] - t \\ &= B_s^2 + (t-s) + 0 - t = B_s^2 - s.\end{aligned}$$

- *For all $\theta \in \mathbb{R}$, $\exp\left(\theta B_t - \frac{\theta^2}{2} t\right)$ is a martingale.*

Remark. *By properties of conditional expectation, if X is a martingale and f is convex such that $\mathbb{E}[|f(X_s)|] < +\infty$ for all $s \geq 0$, then $s \to f(X_s)$ is a submartingale. In particular, B_t^2 is a submartingale. If X is a submartingale and f is convex and increasing, then $f(X)$ is a submartingale. The proof is just an application of Jensen's inequality.*

The theory of martingales goes back to Doob, who understood their almost magical properties. As a first example, we consider Doob's maximum inequality, which connects the maximum of a martingale to its final value.

Theorem 14 (Doob's maximum inequality (continuous-time))**.** *Let $(X_t)_{t\geq 0}$ be a right-continuous submartingale. Then $\forall \lambda > 0, \forall t \geq 0$:*

$$\mathbb{P}\left(\sup_{0\leq s\leq t} |X_s| \geq \lambda\right) \leq \frac{1}{\lambda} \mathbb{E}\left[|X_t|\right].$$

Proof. Fix $0 = t_0 \leq \ldots \leq t_k \leq t$. Then, $Y_n := |X_{t_{n \wedge k}}|$ $(n \geq 0)$ is a discrete time submartingale. Hence, by Doob's discrete time inequality, we have

$$\lambda \mathbb{P}\left(\max_{0 \leq n \leq k} |Y_n| \geq \lambda\right) \leq \mathbb{E}[|X_t|].$$

Now, let $D_i := \left\{\frac{k}{2^i} t, 0 \leq k \leq 2^i\right\}$. Then, $S^{(i)} := \max\{|X_s|, s \in D_i\} \nearrow \sup_{0 \leq s \leq t} |X_s|$ as $i \to \infty$, since X is right-continuous. The result follows. $\square$

Let us sketch the proof in the discrete setting.

Proof in the discrete setting. Let $(X_1, \ldots, X_n)$ be a discrete-time submartingale. Define, for all $i \geq 1$, the event $E_i : |X_i| \geq \lambda$ and $|X_j| < \lambda$ for all $j < i$. Then we have

$$\{sup_{1 \leq i \leq n} |X_i| \geq \lambda\} = \bigcup_{i=1}^n E_i,$$

the union being disjoint. Then, we get, for all i:

$$\lambda \mathbb{P}(E_i) = \int_{E_i} \lambda d\mathbb{P} \leq \int_{E_i} |X_i| d\mathbb{P} \leq \int_{E_i} \mathbb{E}[|X_n||\mathscr{F}_i] d\mathbb{P} = \int_{E_i} |X_n| d\mathbb{P}.$$

The result follows by summing over all $1 \leq i \leq n$. $\square$

In the same vein, we can obtain Doob's L^p inequality:

Theorem 15 (Doob's L^p inequality)**.** *Let X be a right-continuous submartingale, let $p > 1$ and suppose that $\mathbb{E}[|X_t|^p] < \infty$ for all $t \geq 0$. Then, we get*

$$\mathbb{E}\left[\sup_{0 \leq s \leq t} |X_s|^p\right] < \infty.$$

More precisely,

$$\left(\mathbb{E}\left[\sup_{0 \leq s \leq t} |X_t|^p\right]\right)^{1/p} \leq \frac{p}{p-1} \left(\mathbb{E}\left[|X_t|^p\right]\right)^{1/p}.$$

3.2 Uniform integrability

We now turn to some reminders on what we call uniform integrability, which is a very important property for martingales.

Definition 19. *A collection of random variables $X := (X_t)_{t \in T}$ is called uniformly integrable if they are bounded in L^1 and $I(\delta) \to 0$ as $\delta \searrow 0$, where*

$$I(\delta) = \sup\left\{\mathbb{E}\left[|X_t| \mathbb{1}_A\right] : t \in T, A \in \mathscr{F}, \mathbb{P}(A) < \delta\right\}.$$

That is, no event of small probability contributes to a significant amount of the expectation of X_t, uniformly in t.

In particular, observe that if X is bounded in L^p for some $p > 1$, then X is uniformly integrable. Indeed, $\forall t \in T, \forall A \in \mathscr{F}$ with $\mathbb{P}(A) \leq \delta$, Hölder's inequality provides, setting q such that $1/p + 1/q = 1$:

$$\begin{aligned} \mathbb{E}\left[|X_t| \mathbb{1}_A\right] &\leq ||X_t||_p \mathbb{P}(A)^{1/q} \\ &\leq M \delta^{1/q} \underset{\delta \searrow 0}{\to} 0, \end{aligned}$$

as desired.

Lemma 4. *A family $(X_t)_{t\in T}$ is uniformly integrable if and only if*

$$\sup\left\{\mathbb{E}\left[|X_t|\mathbb{1}_{|X_t|\geq K}\right], t\in T\right\} \underset{K\to\infty}{\to} 0.$$

The proof is left as an exercise. The reason why uniform integrability is so interesting is because of the following result:

Theorem 16. *Let $(X_n, n\geq 1)$, X be random variables. Then the following are equivalent:*

(i) $X_n \to X$ in probability and (X_n) are uniformly integrable;

(ii) $X_n \to X$ in L^1.

In particular, if $X_n \to X$ a.s. and (X_n) are uniformly integrable, then $\mathbb{E}[X_n]\to\mathbb{E}[X]$.

3.3 Convergence theorems for martingales

We review here some convergence results concerning martingales, submartingales and supermartingales.

Theorem 17. *Let $(X_t, t\geq 0)$ be a right-continuous submartingale such that*

$$\sup_{t\geq 0}\mathbb{E}[X_t^+] < +\infty$$

. Then, $X_\infty := \lim_{t\to\infty} X_t$ exists almost surely and is in L^1.

Here, $X^+ := X\vee 0$ is the positive part of X.

Proof. Let us prove it in the discrete time setting. We recall Doob's upcrossing inequality: let $(X_n)_{n\geq 0}$ be a discrete-time martingale. For any $a<b$, the number of upcrossings of $[a,b]$ up to time n, denoted by $U_n[a,b]$, is the supremum k (possibly infinite) such that there exist times $s_1 < t_1 < \ldots < s_k < t_k$ for which, for all $1\leq i\leq k$, $X_{s_i}\leq a < b\leq X_{t_i}$. Doob's upcrossing inequality tells that

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

Now, if $\mathbb{E}[X_n^+] = M < \infty$ for all n, we get that $\mathbb{E}[(X_n-a)^+]\leq M+|a|$. Hence, $U_\infty[a,b] := \lim_{n\to\infty} U_n[a,b]$ exists and has finite expectation by Fatou's lemma. In particular, for all $a<b$, $\mathbb{P}(U_\infty<\infty)=1$. Therefore, it follows that $X_\infty := \lim_{n\to\infty} X_n$ exists almost surely. The proof of the L^1 convergence is left as an exercise. □

Corollary 7. *Let $(X_t, t\geq 0)$ be a nonnegative right-continuous supermartingale. Then, $X_\infty := \lim_{t\to\infty} X_t$ exists almost surely and is in L^1. Moreover, $\mathbb{E}[X_\infty]\leq\mathbb{E}[X_0]$ by Fatou's lemma.*

Remark. *Being bounded in L^1 is not enough for convergence in L^1, and one needs uniform integrability. Indeed, let B be a Brownian motion starting at 1, and $T=\inf\{t\geq 0, B_t=0\}$. Then, $M_t := B_{T\wedge t}$ is a nonnegative martingale and $\mathbb{E}[M_t]=\mathbb{E}[M_0]=1$ so it is bounded in L^1 (Exercise). But $M_t\to 0$ a.s. as $t\to+\infty$, since $T<\infty$ a.s.*

Theorem 18. *Let X be a right-continuous submartingale, uniformly integrable. Then, there exists a random variable X_∞ in L^1 such that $X_t\to X_\infty$ a.s. and in L^1 (important!). Moreover, for all $t\geq 0$, $X_t\leq\mathbb{E}[X_\infty|\mathscr{F}_t]$. If X is a martingale, $X_t=\mathbb{E}[X_\infty|\mathscr{F}_t]$.*

Lecture 10: 15/11/2023

Remark. *Conversely, for any random variable* Y $\mathscr{F}_\infty$*-measurable, with* $\mathbb{E}[|Y|] < +\infty$*, we have that* $X_t := \mathbb{E}[Y|\mathscr{F}_t]$ *defines a martingale. Then,* $X_t \to Y$ *in* L^1*, so* (X_t) *is also uniformly integrable.*

Definition 20. *We say that a martingale* $(X_t, t \geq 0)$ *is closed if there exists* $Y \in L^1(\mathbb{P})$ *such that* $X_t = \mathbb{E}[Y|\mathscr{F}_t]$ *for all* $t \geq 0$*.*

We have actually proved that a right-continuous martingale is closed if and only if it is uniformly integrable.

Theorem 19 (L^p martingale convergence theorem)**.** *Let* $X := (X_t, t \geq 0)$ *be a right-continuous martingale, and* $p > 1$*. Then the following are equivalent:*

(i) *X is bounded in* L^p *(that is,* $\sup_{t\geq 0} ||X_t||_p < +\infty$*);*

(ii) X_t *converges a.s. and in* L^p *to a random variable* X_∞*;*

(iii) *There exists* $Z \in L^p$ *such that* $X_t = \mathbb{E}[Z|\mathscr{F}_t]$ *a.s.*

This does not hold when $p = 1$.

3.4 Optional stopping theorem

We now present one of the most important results concerning martingales, the optional stopping theorem.

Theorem 20 (Optional stopping theorem)**.** *Let* X *be a right-continuous closed martingale, and denote by* X_∞ *its limit. Then, if* T *is a stopping time,* $\mathbb{E}[X_\infty|\mathscr{F}_T] = X_T$ *a.s. (note that* T *can be* $+\infty$*).*

Proof. Suppose first that T can only take a countable number of values, and denote them by $\{t_k, k \geq 1\}$. Note that, for any integrable random variable Y, we have

$$\mathbb{E}[Y|\mathscr{F}_T] = \sum_{k\geq 1} \mathbb{E}[Y|\mathscr{F}_{t_k}]\mathbb{1}_{T=t_k}.$$

Indeed, if $A \in \mathscr{F}_T$, we have

$$\mathbb{E}[Y\mathbb{1}_A] = \sum_{k\geq 1} \mathbb{E}[Y\mathbb{1}_{A\cap\{T=t_k\}}],$$

and $A \cap \{T = t_k\} \in \mathscr{F}_{t_k}$ by definition. Thus,

$$\mathbb{E}[Y\mathbb{1}_A] = \sum_{k\geq 1} \mathbb{E}\left[\mathbb{E}[Y|\mathscr{F}_{t_k}]\mathbb{1}_{A\cap\{T=t_k\}}\right] = \mathbb{E}[Z\mathbb{1}_A],$$

where $Z := \sum_{k\geq 1} \mathbb{E}[Y|\mathscr{F}_{t_k}]\mathbb{1}_{T=t_k}$ is clearly $\mathscr{F}_T$-measurable. Applying it to $Y = X_\infty$, we get

$$\begin{aligned}\mathbb{E}[X_\infty|\mathscr{F}_T] &= \sum_{k\geq 1} \mathbb{E}[X_\infty|\mathscr{F}_{t_k}]\mathbb{1}_{T=t_k} \\ &= \sum_{k\geq 1} \mathbb{1}_{T=t_k} X_{t_k} = X_T, \text{ a.s.}\end{aligned}$$

In the general case, set for all n:

$$T_n := \sum_{k=1}^{\infty} \frac{k}{2^n} \mathbb{1}_{\frac{k-1}{2^n} < T \leq \frac{k}{2^n}}.$$

Then $T_n \downarrow T$ a.s. Note also that $X_{T_n} \to X_T$ a.s. and in L^1. Indeed, we have $X_{T_n} = \mathbb{E}[X_\infty | \mathscr{F}_{T_n}]$ so X_{T_n} is uniformly integrable. Moreover, if $A \in \mathscr{F}_T \subseteq \mathscr{F}_{T_n}$, we get from the discrete case that $\mathbb{E}[X_\infty \mathbb{1}_A] = \mathbb{E}[X_{T_n} \mathbb{1}_A] \to \mathbb{E}[X_T \mathbb{1}_A]$ by the L^1 convergence. Thus, $\mathbb{E}[X_\infty \mathbb{1}_A] = \mathbb{E}[X_T \mathbb{1}_A]$ as desired. □

Corollary 8. *(i) If X is a uniformly integrable right-continuous martingale and if S, T are two stopping times with $S \leq T$ a.s., then a.s.*

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

(ii) If $X$is any right-continuous martingale and if $t \geq 0$ and S, T are two bounded stopping times such that $S \leq T \leq t$ a.s., then

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

Proof. To prove (i), note that $t \mapsto X_{t \wedge T}$ is a martingale, and apply the optional stopping theorem to it at time S. For (ii), note that $\{X_s, 0 \leq s \leq t\}$ is a closed martingale since $X_s = \mathbb{E}[X_t | \mathscr{F}_s]$ for all $s \leq t$. Then, apply (i) to $t \wedge T$ and $t \wedge S$. □

Example. *Consider a standard Brownian motion B and $a, b > 0$. We want to compute the probability p that B reaches a before reaching $-b$. Let $T = T_a \wedge T_{-b}$, and observe that $(B_{t \wedge T})_{t \geq 0}$ is a bounded martingale, hence closed. Furthermore, $T < \infty$ a.s. By optional stopping theorem, we have*

$$\begin{aligned} 0 = \mathbb{E}[B_0] = \mathbb{E}[B_T] &= a\mathbb{P}(T_a < T_{-b}) + (-b)\mathbb{P}(T_{-b} < T_a) \\ &= ap - b(1-p). \end{aligned}$$

Hence, $p = \frac{b}{a+b}$.

Now, we want to compute $\mathbb{E}[T]$. We know that $(B_{t \wedge T}^2 - t \wedge T, t \geq 0)$ is a martingale. Hence, by the optional stopping theorem (or, more precisely, by Corollary 8 (i) with $S = 0$), $0 = \mathbb{E}[B_0^2 - 0] = \mathbb{E}[B_T^2 - T]$. Thus,

$$\begin{aligned} \mathbb{E}[T] = \mathbb{E}[B_T^2] &= pa^2 + (1-p)(-b)^2 \\ &= \frac{a^2 b}{a+b} + \frac{b^2 a}{a+b} = ab. \end{aligned}$$

Idea: a lot of problems are about finding the right martingale!

3.5 Finite variation integral

We start now our approach of stochastic integration, that is, integration with respect to a random process. We first define integration with respect to so-called *finite variation processes*.

Definition 21. *A function $f : D \to \mathbb{R}$ (for some $D \subseteq \mathbb{R}$) is called càdlàg (continu à droite, limite à gauche) if it is right-continuous with left limits, that is, $\forall t \in D$:*

- $f(t-) := \lim_{\substack{s \nearrow t \\ s \neq t}} f(s)$ *exists ;*
- $f(t+) := \lim_{\substack{s \searrow t \\ s \neq t}} f(s)$ *exists and is equal to* $f(t)$*.*

A stochastic process is called càdlàg if its paths are a.s. càdlàg.

This turns out to be the right framework for sotchastic integration. In particular, all continuous functions are càdlàg, but we also allow jumps.

Let $a : [0, +\infty) \to \mathbb{R}$ be càdlàg and non-decreasing. Then, we can define a unique Borel measure da by

$$da([s,t]) = a(t) - a(s), s \leq t,$$

which is the Lebesgue-Stieltjes measure associated with a. a is then the distribution function of da. For any function $h : [0, +\infty) \to \mathbb{R}$, we define the Lebesgue-Stieltjes integral

$$(h \cdot a)(t) = \int_{[0,t]} h(s) da(s).$$

This can also be extended to $a = a_1 - a_2$, where a_1, a_2 are two càdlàg non-decreasing functions, defining $h \cdot a = h \cdot a_1 - h \cdot a_2$ (provided the terms on the right-hand side are finite).

We can characterize such functions in a nice and useful way.

Lemma 5. *Let* $a : [0, \infty) \to \mathbb{R}$ *be càdlàg. Define, for all* $n \geq 0, t \geq 0$*:*

$$V_n(t) := \sum_{k=0}^{\lceil 2^n t \rceil - 1} \left| a\left(\frac{k+1}{2^n}\right) - a\left(\frac{k}{2^n}\right) \right|.$$

Then, for all $t \geq 0$, $V_n(t)$ *has a limit* $V(t)$ *(possibly infinite) as* $n \to \infty$*.*

Moreover, a *can be expressed as* $a = a_1 - a_2$ *for* a_1, a_2 *càdlàg and nondecreasing if and only if* $V(t) < \infty$ *for all* $t \geq 0$*. In this case,* V *is itself càdlàg.*

V is called the *total variation* of a. If $V(t) < \infty$, we say that a has finite variation on $[0,t]$.

Proof. If $a = a_1 - a_2$ with a_1, a_2 càdlàg and nondecreasing then clearly $V(t) < \infty$ for all t by triangular inequality. Indeed,

$$\begin{aligned}
V_n(t) &= \sum_{k=0}^{\lceil 2^n t \rceil - 1} \left| a\left(\frac{k+1}{2^n}\right) - a\left(\frac{k}{2^n}\right) \right| \\
&\leq \sum_{k=0}^{\lceil 2^n t \rceil - 1} \left| a_1\left(\frac{k+1}{2^n}\right) - a_1\left(\frac{k}{2^n}\right) \right| + \sum_{k=0}^{\lceil 2^n t \rceil - 1} \left| a_2\left(\frac{k}{2^n}\right) - a_2\left(\frac{k+1}{2^n}\right) \right| \\
&= a_1\left(\frac{\lceil 2^n t \rceil}{2^n}\right) - a_1(0) + a_2\left(\frac{\lceil 2^n t \rceil}{2^n}\right) - a_2(0) \\
&\to a_1(t) - a_1(0) + a_2(t) - a_2(0) < \infty,
\end{aligned}$$

where the last step comes from the rgiht-continuity.

For the converse, set $a_1 = \frac{1}{2}(V + a)$ and $a_2 = \frac{1}{2}(V - a)$, and check that they are càdlàg and nondecreasing. □

Lecture 11: 20/11/2023

Remark. *If a is C^1, then a has finite variation (bound on each interval given by the sup of a').*

Definition 22. *Suppose that we have a probability space with a filtration, and a càdlàg, adapted process A. We can define pathwise its total variation: $V(\omega,\cdot)$ is the total variation of $A(\omega,\cdot)$. Then V is itself càdlàg and adapted.*

Our goal now is to define the integration of random processes with respect to the (random) measure dA. To this end, we will need the notion of *previsibility*.

Definition 23. *The previsible σ-algebra on $\Omega\times[0,+\infty)$ is the σ-algebra $\mathscr{P}$ generated by $E\times(s,t]$, where $t>s$ and $E\in\mathscr{F}_s$. A previsible process H is a $\mathscr{P}$-measurable random variable $(\omega,t)\mapsto H(\omega,t)$ on $\Omega\times[0,+\infty)$.*

Lemma 6. *Suppose that X is càdlàg and adapted. Define $H: t\mapsto X_{t-}$. Then H is previsible.*

In particular, if X is continuous then it is previsible.

Proof. Note that H is adapted and left-continuous. Set $H_t^n := \sum_{k=0}^{\infty} X_{k2^{-n}} \mathbb{1}_{k2^{-n}<t\le(k+1)2^{-n}}$. Then, $H_t^n \to X_{t-} = H_t$ for all t,ω. Moreover, for all n, H^n is previsible since H_t^n is $\mathscr{F}_{k2^{-n}}$-measurable for $k2^{-n}<t\le(k+1)2^{-n}$. The result follows. ◻

Definition 24. *Let A be a càdlàg, adapted process with finite variation process V. Let H be previsible such that, for all $t\ge 0$, all ω, $\int_0^t |H_s(\omega)|dV_s(\omega)<\infty$. Then, we define the (pathwise) integral*

$$(H\cdot A)_t = \int_0^t H_s dA_s.$$

Proposition 6. *$(H\cdot A)$ is càdlàg, adapted and has finite variation.*

Proof. Details of this proof are optional.

To prove that it is càdlàg, note that $\mathbb{1}_{(0,s]} \to \mathbb{1}_{(0,t]}$ as $s\searrow t$, and $\mathbb{1}_{(0,s]}\to\mathbb{1}_{(0,t)}$ as $s\nearrow t$. Recall that $(H\cdot A)_t = \int H_s \mathbb{1}_{s\in(0,t]} dA_s$. By dominated convergence (for each $\omega\in\Omega$), we have

$$(H\cdot A)_t = \int H_s \lim_{r\searrow t} \mathbb{1}_{[0,r]} dA_s = \lim_{r\searrow t}\int H_s \mathbb{1}_{[0,r]} dA_s = \lim_{r\searrow t}(H\cdot A)_r.$$

To prove that it is adapted, suppose that $H=\mathbb{1}_{B\times(s,u]}$ for $B\in\mathscr{F}_s$ and $s<u$. Then:

$$(H\cdot A)_t = \mathbb{1}_B\left(A_{t\wedge u}-A_{t\wedge s}\right),$$

which is $\mathscr{F}_t$-measurable. Then, for $H=\mathbb{1}_C$ for $C\in\mathscr{P}$, $H\cdot A$ is adapted by a π-λ argument. Finally, approach any H by sums of such indicator functions.

To prove that it has finite variation, let $H^+ = H\vee 0, H^- = (-H)\vee 0, A^+ = \frac{1}{2}(V+A), A^- = \frac{1}{2}(V-A)$. Then,

$$H\cdot A = (H^+-H^-)\cdot(A^+-A^-) = (H^+\cdot A^+ + H^-\cdot A^-) - (H^+\cdot A^- + H^-\cdot A^+),$$

and both are nondecreasing. The result follows. ◻

3.6 Local martingales

We weaken here a bit the notion of martingale, for reasons that we will see. We introduce the notion of *local martingale*.

Definition 25. *A local martingale X is a càdlàg process such that there exists a sequence of stopping times $(T_n, n \geq 1)$ such that $T_n \uparrow +\infty$ a.s. and $X^{T_n} := (X_{t\wedge T_n}, t \geq 0)$ is a martingale for each n.*

We say that the sequence (T_n) reduces X.

We denote by $\mathscr{M}_{loc}$ the set of local martingales.

Remark. *Any martingale is a local martingale by the optional stopping theorem (taking $T_n = n$). But some local martingales are not martingales.*

Proposition 7. *Let X be adapted, right-continuous and integrable. Then the following are equivalent:*

(i) X is a martingale,

(ii) $\mathbb{E}[X_T] = \mathbb{E}[X_0]$ for any bounded stopping time T ;

(iii) X is a local martingale and the set $\{X_T, T$ stopping time, $T \leq t_0\}$ is uniformly integrable for all $t_0 \geq 0$.

Proof. (i) $\Rightarrow$ (ii) comes from Corollary 8 (second part).

(ii) $\Rightarrow$ (i): let $s < t < u$. Let $A \in \mathscr{F}_s$ and define two stopping times T, S as follows. $T = t$ if A occurs, $T = u$ otherwise ; $S = s$ if A occurs, $S = u$ otherwise. Then

$$\begin{aligned}&\mathbb{E}[X_S] = \mathbb{E}[X_0] = \mathbb{E}[X_T]\\ &\mathbb{E}[X_s \mathbb{1}_A] + \mathbb{E}[X_u \mathbb{1}_{A^c}] = \mathbb{E}[X_t \mathbb{1}_A] + \mathbb{E}[X_u \mathbb{1}_{A^c}],\end{aligned}$$

so $\mathbb{E}[X_t|\mathscr{F}_s] = X_s$ as desired.

(i) $\Rightarrow$ (iii): X is a martingale so it is a local martingale. Moreover, by optional stopping theorem, if T is bounded by t_0, we have $X_T = \mathbb{E}[X_{t_0}|\mathscr{F}_T]$, so $\{X_T : T \leq t_0$ stopping time $\}$ is uniformly integrable as desired.

(iii) $\Rightarrow$ (ii): Let T_n be a sequence that reduces X. Let T be a bounded stopping time, say $T \leq t_0$. Then, by optional stopping theorem for X^{T_n}, we have

$$\mathbb{E}[X_0] = \mathbb{E}[X_T^{T_n}] = \mathbb{E}[X_{T\wedge T_n}],$$

and $X_{T\wedge T_n}$ is uniformly integrable and converges to X_T. Hence, $\mathbb{E}[X_0] = \mathbb{E}[X_T]$. □

Corollary 9. *If M is a local martingale with $|M_t| \leq Z$ for all $t \geq 0$, where $Z \in L^1$, then M is a true martingale.*

Indeed, in this case, (iii) is easily satisfied.

Proposition 8. *Let $M \in \mathscr{M}_{loc}$ such that $M_t \geq 0$ for all t. Then, M is a (true) supermartingale.*

Proof. Take (T_n) reducing M. Then, by conditional Fatou's lemma, we have for $s < t$:

$$\mathbb{E}[M_t|\mathscr{F}_s] = \mathbb{E}[\lim_{n\to\infty} M_{t\wedge T_n}|\mathscr{F}_s] \leq \liminf_{n\to\infty} \mathbb{E}[M_{t\wedge T_n}|\mathscr{F}_s] = \liminf_{n\to\infty} M_{s\wedge T_n} = M_s.$$

□

As a consequence, if $M \in \mathscr{M}_{loc}$ is bounded, then it is a martingale.

Theorem 21. *Let M be a continuous local martingale of finite variation, such that $M_0 = 0$ a.s. Then, $M \equiv 0$ (M is indistinguishable from 0).*

Proof. Let V be the total variation of M, and $S_n := \inf\{t : V_t \geq n\}$, for $n \geq 1$. Then, S_n is a stopping time for all n and $|M_t^{S_n}| \leq |V_{t \wedge S_n}| \leq n$ for all t. Hence, $M_n^{S_n}$ is a bounded local martingale, hence it is a true martingale. Since $S_n \to \infty$ a.s., it is enough to prove that $M_n^{S_n} \equiv 0$ a.s. for all n. For this, we use the following lemma:

Lemma 7. *Let M be a martingale with $M_t \in L^2$ for all t. Then, for all $s \leq t$:*

$$\mathbb{E}[(M_t - M_s)^2 | \mathscr{F}_s] = \mathbb{E}[(M_t^2 - M_s^2) | \mathscr{F}_s].$$

Proof of Lemma 7. The left-hand side is equal to $\mathbb{E}[M_t^2 | \mathscr{F}_s] - 2M_s \mathbb{E}[M_t | \mathscr{F}_s] + M_s^2$. □

For $t \geq 0$, $1 \leq k \leq N$, set $t_n = \frac{kt}{N}$. Lemma 7 implies that, if M is of finite variation, we have for all t and all N (by conditioning on $\mathscr{F}_{t_k}$):

$$\begin{aligned}
\mathbb{E}[M_t^2] &= \mathbb{E}[\sum_{k=0}^{N-1} (M_{t_{k+1}}^2 - M_{t_k}^2)] \\
&= \mathbb{E}[\sum_{k=0}^{N-1} (M_{t_{k+1}} - M_{t_k})^2] \\
&\leq \mathbb{E}\left[\sup_k |M_{t_{k+1}} - M_{t_k}| \sum_{k=0}^{N-1} \left|M_{t_{k+1}} - M_{t_k}\right|\right].
\end{aligned}$$

The sum is bounded by $V_t \leq n$, while the supremum converges to 0 a.s. by continuity. Since this supremum is also bounded by $V_t \leq n$, we conclude by dominated convergence that $M_t = 0$ a.s.. Since M is continuous, $M \equiv 0$. □

We now introduce the class of processes that "works well with stochastic integration": *semimartingales*.

Definition 26. *We call a process X a semimartingale if*

$$X = X_0 + M + A,$$

where M is a continuous local martingale, A is a previsible process of finite variation and $M_0 = A_0 = 0$. By the previous theorem, this decomposition is unique and is called the Doob-Meyer decomposition.

4 Stochastic integral

Lecture 12: 27/11/2023

We now want to define the stochastic integrals $\int_0^t H_s dM_s$, where H is càdlàg and previsible, and M is, say, a continuous semimartingale. We know how to handle the finite variation part of a semimartingale, thus we only need to care about the (local) martingale part. To this end, a first step is to imagine that H is "simple", and to cook up a definition. Then, if we find a space on which we can approximate all càdlàg previsible processes by "simple" processes, so that the sequence of integrals is a Cauchy sequence, then we would be able to define the integral of our limiting process as the limit of the integrals of "simple" processes. This is the main idea of Itô integration.

4.1 Simple processes

We will start, for convenience, with martingales of bounded square. Later, we will see that we can get rid of this condition!

We denote by $\mathscr{M}^2$ the set of continuous martingales bounded in L^2 (that is, there exists $C > 0$ such that $\mathbb{E}[M_t^2] < C$ for all t). Recall that, if $M \in \mathscr{M}^2$, then there exists M_∞ such that $M_t \to M_\infty$ a.s. and in L^2, and that $M_t = \mathbb{E}[M_\infty|\mathscr{F}_t]$ for all t. Furthermore,

$$\mathbb{E}[\sup M_t^2] \leq 4 \sup \mathbb{E}[M_t^2] < \infty,$$

by Doob's L^2 inequality. Hence,

$$M \mapsto ||M||_2 := \sqrt{\mathbb{E}[M_\infty^2]}$$

defines a norm (by Jensen's inequality), and $\mathscr{M}^2$ endowed with this norm is a Hilbert space (that is, a space with an inner product making it complete). Then, $M \mapsto M_\infty$ is an isometry of Hilbert spaces from $\mathscr{M}^2$ to L^2.

We will first define stochastic integrals of simple processes against martingales bounded in L^2, before extending it. The idea is really to use the fact that the set of "simple processes" that we define is dense in L^2, and extend the stochastic integral by continuity.

Definition 27. *A simple process is a map $H : \Omega \times \mathbb{R}_+ \to \mathbb{R}$ of the form*

$$H_t(\omega) = \sum_{k=0}^{n-1} Z_k(\omega) \mathbb{1}_{(t_k, t_{k+1}]}(t),$$

where $n \geq 1$, $0 = t_0 < t_1 \ldots < t_n < \infty$ and Z_k is a <u>bounded</u> $\mathscr{F}_{t_k}$-measurable random variable. We denote by $\mathscr{S}$ the set of simple processes.

For $M \in \mathscr{M}_{loc}$ and $H \in \mathscr{S}$, we define the stochastic integral of H against M as

$$(H \cdot M)_t = \sum_{k=0}^{n-1} Z_k \left(M_{t_{k+1} \wedge t} - M_{t_k \wedge t} \right).$$

Proposition 9. *Let $H \in \mathscr{S}$ and $M \in \mathscr{M}^2$, and let T be a stopping time. Then:*

(i) $(H \cdot M)^T = H \cdot M^T$ *;*

(ii) $H \cdot M$ *is a martingale ;*

(iii) $H \cdot M \in \mathcal{M}^2$ *and, for all* $t \in \mathbb{R}_+ \cup \{\infty\}$*, we have the explicit bound*

$$\mathbb{E}\left[(H\cdot M)_t^2\right] = \sum_{k=1}^{n-1} \mathbb{E}\left[Z_k^2\left(M_{t_{k+1}\wedge t} - M_{t_k\wedge t}\right)^2\right]$$
$$\leq ||H||_\infty^2 \mathbb{E}\left[(M_\infty - M_0)^2\right] < \infty.$$

Proof. (i) is obvious by definition. For (ii), note that for $t_k \leq s \leq t < t_{k+1}$, we have $(H\cdot M)_t - (H\cdot M)_s = Z_k(M_t - M_s)$, so that

$$\mathbb{E}\left[(H\cdot M)_t - (H\cdot M)_s|\mathscr{F}_s\right] = \mathbb{E}[Z_k(M_t - M_s)|\mathscr{F}_s] = Z_k\mathbb{E}[M_t - M_s|\mathscr{F}_s] = 0.$$

This is also true (check) for general $s \leq t$, so $H\cdot M$ is a martingale. Moreover, we have the following orthogonality result: if $j < k$, then

$$\mathbb{E}\left[Z_j\left(M_{t_{j+1}\wedge t} - M_{t_j\wedge t}\right)Z_k\left(M_{t_{k+1}\wedge t} - M_{t_k\wedge t}\right)\right] = \mathbb{E}\left[Z_j\left(M_{t_{j+1}\wedge t} - M_{t_j\wedge t}\right)Z_k\left(M_{t_{k+1}\wedge t} - M_{t_k\wedge t}\right)|\mathscr{F}_{t_k}\right] = 0.$$

Hence,

$$\begin{aligned}
\mathbb{E}[(H\cdot M)_t^2] &= \mathbb{E}\left[\left(\sum_{k=0}^{n-1} Z_k\left(M_{t_{k+1}\wedge t} - M_{t_k\wedge t}\right)\right)^2\right] \\
&= \sum_{k=0}^{n-1}\mathbb{E}\left[Z_k^2\left(M_{t_{k+1}\wedge t} - M_{t_k\wedge t}\right)^2\right] \text{ by orthogonality} \\
&\leq ||H||_\infty^2\mathbb{E}\left[\sum_{k=0}^{n-1}\left(M_{t_{k+1}\wedge t} - M_{t_k\wedge t}\right)^2\right] \\
&= ||H||_\infty^2\mathbb{E}\left[\sum_{k=0}^{n-1}\left(M_{t_{k+1}\wedge t}^2 - M_{t_k\wedge t}^2\right)\right] \text{ by Lemma 7} \\
&= ||H||_\infty^2\mathbb{E}\left[\left(M_{t_n\wedge t}^2 - M_0^2\right)\right] \\
&= ||H||_\infty^2\mathbb{E}\left[\left(M_{t_n\wedge t} - M_0\right)^2\right] \text{ by Lemma 7} \\
&= ||H||_\infty^2\mathbb{E}\left[\mathbb{E}\left[M_\infty - M_0|\mathscr{F}_{t_n\wedge t}\right]^2\right] \text{ by the } L^2 \text{ convergence} \\
&\leq ||H||_\infty^2\mathbb{E}\left[(M_\infty - M_0)^2\right] \text{ by Jensen.}
\end{aligned}$$

The result follows. This also holds for $t = \infty$. □

4.2 Quadratic variation

Theorem 22. *For each* $M \in \mathscr{M}_{loc}$ *, there exists a unique (up to indistinguishability) nondecreasing adapted process, denoted by* $[M]$*, such that* $M^2 - [M] \in \mathscr{M}_{loc}$ *and* $[M]_0 = 0$*.* $[M]$ *is called the quadratic variation of* M*.*

Furthermore, if we define

$$[M]_t^n := \sum_{k=0}^{\lceil 2^n t\rceil - 1}\left(M_{\frac{k+1}{2^n}} - M_{\frac{k}{2^n}}\right)^2,$$

then $[M]^n \to [M]$ *uniformly on compact subsets in probability:*

$$\forall\varepsilon > 0, \forall t \geq 0, \mathbb{P}\left(\sup_{0\leq s\leq t}|[M]_s^n - [M]_s| > \varepsilon\right) \underset{n\to\infty}{\to} 0.$$

The existence of this quadratic variation will be admitted. It is a very deep and strong result.

Remark. *The uniqueness follows from the fact that $[M]$ is of finite variation, and the Doob-Meyer decomposition.*

Example. *For B a Brownian motion, $B_t^2 - t$ is a martingale, so $[B]_t = t$ for all t.*

Theorem 23. *Let $M \in \mathscr{M}^2_{loc}$ with $M_0 = 0$ a.s. Then,*

$$\mathbb{E}[M_\infty^2] = \mathbb{E}[[M]_\infty],$$

and $(M_t^2 - [M]_t, t \geq 0)$ is uniformly integrable.

Proof. Consider, for all n, the stopping time $S_n := \inf\{t, [M]_t \geq n\}$. We have that

$$\left|M_{t\wedge S_n}^2 - [M]_{t\wedge S_n}\right| \leq n + \sup_t M_t^2,$$

so $\left((M_t^2 - [M]_t)^{S_n}\right)_{t\geq 0}$ is a uniformly integrable martingale. In particular, for all t:

$$\mathbb{E}[M_{t\wedge S_n}^2] = \mathbb{E}\left[[M]_{t\wedge S_n}\right].$$

Letting $n, t \to \infty$, we get (by dominated convergence and monotone convergence)

$$\mathbb{E}[[M]_\infty] = \mathbb{E}[M_\infty^2] < \infty.$$

Since this limit is finite, $[M]_\infty \in L^1$ and $[M]_t$ is bounded by an integrable variable. Hence, $M^2 - [M]$ is uniformly integrable. ☐

4.3 Itô integrals

Proposition 10. *Let μ be a finite measure on the previsible σ-algebra $\mathscr{P}$. Then $\mathscr{S}$ is dense in $L^2(\mathscr{P}, \mu)$.*

Proof. If $H \in \mathscr{S}$ then H is bounded so $H \in L^2(\mathscr{P}, \mu)$. So $\mathscr{S} \subseteq L^2(\mathscr{P}, \mu)$. For the density, we use a monotone class argument. It suffices to show that $\mathbb{1}_A \in \overline{\mathscr{S}}$ for any $A \in \mathscr{P}$. Let $\mathscr{A} := \{A \in \mathscr{P}, \mathbb{1}_A \in \overline{\mathscr{S}}\}$. $\mathscr{A}$ is a λ-system. Indeed, if $C \subseteq D \in \mathscr{A}$ then $D \backslash C \in \mathscr{A}$, and if $C_n \in \mathscr{A}$ with (C_n) increasing, then $\bigcup_n C_n \in \mathscr{A}$. Moreover, $\mathscr{A}$ contains a generating π-system by definition. ☐

Now, fix $M \in \mathscr{M}^2$ and define a measure μ on $\mathscr{P}$ by setting, for $s < t, A \in \mathscr{F}_s$:

$$\mu\left(A \times (s,t]\right) = \mathbb{E}\left[\mathbb{1}_A\left([M]_t - [M]_s\right)\right].$$

By Caratheodory's extension theorem, this specifies a unique measure on $\mathscr{P}$. Alternatively, $\mu(d\omega \otimes dt) = d[M](\omega, dt)\mathbb{P}(d\omega)$. Thus, if H is previsible and nonnegative, then

$$\int_{\Omega\times(0,\infty)} H d\mu = \mathbb{E}\left[\int_0^\infty H_s d[M]_s\right].$$

Lecture 13: 29/11/2023

Observe that, under the expectation, we have used the usual Lebesgue-Stieltjes integral.

Definition 28. *Denote by $L^2(M)$ the space $L^2(\mathscr{P},\mu)$ (recall that μ depends on M). Then, since $M \in \mathscr{M}^2$, μ has finite mass (by Theorem 23) and $\mathscr{S}$ is dense in $L^2(M)$. Write $||H||_M$ for $||H||_{L^2(M)}$, that is,*

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

Moreover, recall that $\mathscr{M}^2$ is a Hilbert space with norm $||M||^2 = \mathbb{E}\left(M_\infty^2\right)$.

Theorem 24 (Itô isometry)**.** *For every $M \in \mathscr{M}^2$, there exists a unique isometry*

$$I : \left(L^2(M), ||\cdot||_M\right) \to \left(\mathscr{M}^2, ||\cdot||\right)$$

such that $I(H) = H \cdot M$ for all $H \in \mathscr{S}$.

Proof. Let $H := \sum_{k=0}^{n-1} Z_k \mathbb{1}_{(t_k,t_{k+1}]} \in \mathscr{S}$. By Proposition 9, we have that $H \cdot M \in \mathscr{M}^2$, with

$$\begin{aligned}
||H\cdot M||^2_{\mathscr{M}^2} = \mathbb{E}\left[(H\cdot M)^2_\infty\right] &= \mathbb{E}\left[\sum_{k=0}^{n-1} Z_k^2\left(M_{t_{k+1}} - M_{t_k}\right)^2\right] \\
&= \sum_{k=0}^{n-1}\mathbb{E}\left[Z_k^2 \mathbb{E}\left[\left(M_{t_{k+1}} - M_{t_k}\right)^2 | \mathscr{F}_{t_k}\right]\right] \\
&= \sum_{k=0}^{n-1}\mathbb{E}\left[Z_k^2 \mathbb{E}\left[M^2_{t_{k+1}} - M^2_{t_k} | \mathscr{F}_{t_k}\right]\right] \\
&= \sum_{k=0}^{n-1}\mathbb{E}\left[Z_k^2 \mathbb{E}\left[[M]_{t_{k+1}} - [M]_{t_k} | \mathscr{F}_{t_k}\right]\right] \\
&= \mathbb{E}\left[\int_0^\infty H_s^2 d[M]_s\right].
\end{aligned}$$

Thus, defining $I(H) = H\cdot M$ for $H \in \mathscr{S}$, I is an isometry from $\left(L^2(M)\cap\mathscr{S}, ||\cdot||_M\right)$ to $\left(\mathscr{M}^2, ||\cdot||\right)$.

Since $\mathscr{S}$ is dense in $L^2(M)$, we claim that there is a unique extension of I to $L^2(M)$. Let us prove it. Let $H \in L^2(M)$ and take any sequence $H_n \in \mathscr{S}$ such that $H_n \to H$ in $L^2(M)$. Then, $(I(H_n))_{n\geq 1}$ is a Cauchy sequence in $\mathscr{M}^2$. Indeed, we have

$$||I(H_n) - I(H_m)||_{\mathscr{M}^2} = ||H_n - H_m||_{L^2(M)} \underset{n,m\to\infty}{\to} 0.$$

Since $\mathscr{M}^2$ is a Hilbert space and hence complete, we have that $I(H_n)$ converges to some limit, which we call $I(H)$, independent of the approximating sequence (H_n). □

We write

$$I(H) = H\cdot M = \left(\int_0^t H_s dM_s, t \geq 0\right)$$

.

This holds as long as H is in $L^2(M)$ and $M \in \mathscr{M}^2$. The next step is to manage to extend this definition to (semi)martingales M that are not in $\mathscr{M}^2$. We do this through stopping times: indeed, it turns out that the stochastic integral commutes with stopping, as we will prove now.

Proposition 11. *Let $M \in \mathscr{M}_c^2$ and $H \in L^2(M)$. Let T be a stopping time. Then,*

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

Proof. It is easy to see that $H\mathbb{1}_{[0,T]} \in L^2(M)$. Also, by definition of $[M]$, we have that $[M^T] = [M]^T$ Therefore,

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

So, $H \in L^2(M^T)$. Let us now show our result.

Case 1 If $H \in \mathscr{S}$ and $M \in \mathscr{M}_c^2$, and T takes only finitely many values. Then, one can check that $H\mathbb{1}_{[0,T]} \in \mathscr{S}$ and $(H \cdot M)^T = (H\mathbb{1}_{[0,T]}) \cdot M$.

Case 2 If $H \in \mathscr{S}$ and $M \in \mathscr{M}_c^2$, and T is general. Then for $m,n \ge 0$, let $T_{n,m} = 2^{-n}\lceil 2^n T\rceil \wedge m$. Then, $T_{n,m}$ takes finitely many values, and $T_{n,m} \searrow T \wedge m$ as $n \to \infty$. Therefore,

$$||H\mathbb{1}_{(0,T_{n,m}]} - H\mathbb{1}_{(0,T\wedge m]}||_M^2 = \mathbb{E}\left[\int_0^\infty H_t^2 \mathbb{1}_{(T\wedge m, T_{n,m}]} d[M]_t\right] \to 0,$$

by dominated convergence. Therefore, by Itô isometry, $(H\mathbb{1}_{(0,T_{n,m}]}) \cdot M \to (H\mathbb{1}_{(0,T\wedge m]}) \cdot M$ in $\mathscr{M}_c^2$. But, also,

$$\begin{aligned}\Big((H\mathbb{1}_{(0,T_{n,m}]}) \cdot M\Big)_t &= (H\mathbb{1}_{(0,T_{n,m}]}) \cdot M \text{ by Case 1}\\ &\to (H \cdot M)_t^{T\wedge m},\end{aligned}$$

almost surely by continuity of $H \cdot M$. Therefore, $(H \cdot M)^{T\wedge m} = (H\mathbb{1}_{(0,T\wedge m]}) \cdot M$. Sending m to $+\infty$ and applying similar arguments, we conclude that $(H \cdot M)^T = (H\mathbb{1}_{(0,T]}) \cdot M$.

Case 3 If $H \in L^2(M)$, approximate by $H^n \in \mathscr{S}$.

□

Remark. *In particular we have the following consistency property: for S, T stopping times with $S \le T$, we have*

$$\left(H \cdot (M^T)\right)^S = H \cdot \left(M^{T\wedge S}\right) = H \cdot \left(M^S\right).$$

Using all this, we can now extend the definition of the stochastic integral to more general H and M. The property that we need is called *local boundedness*.

Definition 29. *We say that a previsible process H is locally bounded if there exists a sequence of stopping times $(S_n)_{n\ge 1}$ satisfying $S_{n+1} \ge S_n$ a.s., $S_n \to \infty$ a.s. and $|H^{S_n}|$ is bounded for all n, that is, there exists a deterministic constant $C_n < \infty$ such that*

$$\sup_{t\ge 0} |H_t| \mathbb{1}_{t\in[0,S_n]} \le C_n.$$

Remark. *Any adapted continuous process X is previsible and locally bounded, taking $S_n = \inf\{t, |X_t| = n\}$.*

Definition 30. *Let H be previsible and locally bounded, and let $M \in \mathscr{M}_{c,loc}$. For all $n \geq 1$, define $S_n' = \inf\{t \geq 0, |M_t| = n\}$ and $T_n = S_n \wedge S_n'$, where S_n is as above (in local boundedness). Then, we define the stochastic integral of H against M as*

$$(H \cdot M)_t := \left(H\mathbb{1}_{[0,T_n]} \cdot M^{(T_n)}\right)_t,$$

for all $t \leq T_n$.

This is called *localisation*. Roughly speaking, the stochastic integral is defined locally (in a consistent way).

Remark. *There are several points that we can highlight.*

(i) First, the right-hand side is well-defined since $M^{(T_n)} \in \mathscr{M}_c^2$ and $H\mathbb{1}_{[0,T_n]} \in L^2(M)$;

(ii) The left-hand side does not depend on n by the consistency property;

(iii) The left-hand side does not depend on $(T_n)_{n\geq 1}$, as long as it localizes H and M, that is, $H\mathbb{1}_{[0,T_n]} \in L^2(M)$ and $M^{(T_n)} \in \mathscr{M}_c^2$.

(iv) $(H \cdot M) \in \mathscr{M}_{c,loc}$, and we can use T_n as a reducing sequence of stopping times.

Lecture 14: 04/12/2023

Theorem 25. *Let H be previsible and locally bounded, and $M \in \mathscr{M}_{loc}$. Then, $(H \cdot M)$ is a local martingale and its quadratic variation is given by*

$$[H \cdot M]_t = \int_0^t H_s^2 d[M]_s.$$

In other terms, $[H \cdot M] = H^2 \cdot [M]$.

Proof. By localisation, assume that $M \in \mathscr{M}^2$ and H is bounded. Then $H \cdot M \in \mathscr{M}^2$ by the Itô isometry. For any stopping time T, we have:

$$\begin{aligned}
\mathbb{E}\left[(H \cdot M)_T^2\right] &= \mathbb{E}\left[\left((H\mathbb{1}_{[0,T]}) \cdot M\right)_\infty^2\right] \\
&= ||H\mathbb{1}_{[0,T]} \cdot M||_{\mathscr{M}^2}^2 \\
&= ||H\mathbb{1}_{[0,T]}||_{L^2(M)}^2 \text{ by Itô isometry} \\
&= \mathbb{E}\left[\int_0^T H_s^2 d[M]_s\right].
\end{aligned}$$

Now recall (Proposition 7) that a (adapted, right-continuous, integrable) process X is a martingale if and only if $\mathbb{E}[X_T] = \mathbb{E}[X_0]$ for any bounded stopping time T. Hence,

$$t \mapsto (H \cdot M)_t^2 - \int_0^t H_s^2 d[M]_s$$

is a martingale. By uniqueness of quadratic variation, and since

$$t \mapsto \int_0^t H_s^2 d[M]_s$$

is increasing, we get the result. □

Now, we go back to semimartingales. For $X = X_0 + M + A$ a continuous semimartingale. We define the quadratic variation of X as $[X] = [M]$. This is consistent with the fact that $[X]^n \to [M]$ uniformly on all compact, where

$$[X]^n_t := \sum_{k=0}^{\lceil 2^n t\rceil -1} \left(X_{(k+1)2^{-n}} - X_{k2^{-n}}\right)^2.$$

In other words, $[A] = 0$ when A is a finite variation process (check), and all the "roughness" of a semimartingale comes from the local martingale part.

Definition 31. *Let X be as above, and H be previsible and locally bounded. We define the stochastic integral of H against X as*

$$(H \cdot X)_t = \int_0^t H_s dX_s := (H \cdot M)_t + (H \cdot A)_t.$$

The first one is the Itô stochastic integral, the second one is the usual Lebesgue-Stieltjes integral. Since $H \cdot M \in \mathscr{M}_{c,loc}$ and $H \cdot A$ is of finite variation, we have that $H \cdot X$ is a continuous semimartingale.

Useful notation: we write $dZ_t = H_t dX_t$ to mean that $Z_t - Z_0 = \int_0^t H_s dX_s$.

Proposition 12 (Chain rule)**.** *Let H, K be previsible and locally bounded, and $M \in \mathscr{M}_{loc}$. Then*

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

that is,

$$H_s d\left(\int_0^s K_u dM_u\right) = H_s K_s dM_s.$$

Proof. Check it for simple processes and then take limits. □

Proposition 13. *Let X be a semimartingale and H left-continuous, adapted, locally bounded. Then:*

$$\sum_{k=0}^{\lceil 2^n t\rceil -1} H_{k2^{-n}} \left(X_{(k+1)2^{-n}} - X_{k2^{-n}}\right) \to \int_0^t H_s dX_s,$$

in probability, uniformly on all compacts.

Proof. We show it when X is a local martingale, say M. By localization, we can reduce ourselves to the case when $M \in \mathscr{M}_c^2$ and H is uniformly bounded. Define $H^n_t := H_{2^{-n}\lfloor 2^n t\rfloor}$ (constant by parts). We have:

$$(H^n \cdot M)_t = \sum_{k=0}^{\lceil 2^n t\rceil -1} H_{k2^{-n}} \left(M_{(k+1)2^{-n}} - M_{k2^{-n}}\right) + H_{2^{-n}(\lceil 2^n t\rceil -1)} \left(M_t - M_{2^{-n}\lceil 2^n t\rceil}\right).$$

Since H is bounded and H continuous, the last term goes to 0 in probability, uniformly on all compacts. Moreover, since $H^n \to H$ μ-almost everywhere (by left-continuity of H), we have

$$\begin{aligned} \|H^n - H\|_M^2 &= \mathbb{E}\left[\int_0^\infty (H^n - H)^2 d[M]_s\right] \text{ by Itô isometry} \\ &\underset{n\to\infty}{\to} 0, \end{aligned}$$

by dominated convergence. Hence, $(H^n \cdot M) \to (H \cdot M)$ in $\mathscr{M}_c^2$, which by Doob's maximal inequality implies convergence in probability, uniformly on all compacts. □

4.4 Covariation

We now define the covariation of two local martingales, which is important to extend integration by parts and change of variables to stochastic integrals. The quadratic variation $[M]$ of a local martingale M was such that $M^2 - [M]$ was a local martingale. The covariation $[M,N]$ of two local martingales M,N is such that $MN - [M,N]$ is a local martingale.

The main idea is to use the so-called *polarization identity*: for all x,y:

$$xy = \frac{(x+y)^2 - (x-y)^2}{4}.$$

Definition 32. *Let $M,N \in \mathscr{M}_{c,loc}$. We set*

$$[M,N] = \frac{[M+N] - [M-N]}{4}.$$

Proposition 14. *(i) $[M,N]$ is the unique adapted continuous finite variation process such that $[M,N]_0 = 0$ and*

$$MN - [M,N] \in \mathscr{M}_{c,loc}.$$

(ii) For $n \geq 1$ and $t \geq 0$, let

$$[M,N]_t^n := \sum_{k=0}^{\lceil 2^n t \rceil - 1} \left(M_{(k+1)2^{-n}} - M_{k2^{-n}}\right)\left(N_{(k+1)2^{-n}} - N_{k2^{-n}}\right).$$

Then, $[M,N]^n \to [M,N]$ uniformly in probability on all compacts, as $n \to \infty$.

(iii) If $M,N \in \mathscr{M}_c^2$, then $MN - [M,N]$ is a uniformly integrable martingale.

(iv) $(M,N) \mapsto [M,N]$ is a symmetric bilinear form.

Proof. (i) to (iii) are clear, follow from similar statements for the quadratic variation, using the fact that $MN = \frac{1}{4}((M+N)^2 - (M-N)^2)$ and $[M,N]^n = \frac{1}{4}([M+N]^n - [M-N]^n)$. (iv) follows directly from the uniqueness in (i). □

Remark. *For any $M \in \mathscr{M}_{c,loc}$, we have $[M,M] = [M]$.*

Theorem 26 (Kunita-Watanabe identity)**.** *Let $M,N \in \mathscr{M}_{c,loc}$ and H locally bounded and previsible. Then,*

$$[H \cdot M, N] = H \cdot [M,N] = [M, H \cdot N].$$

We will not prove it.

Remark. *As a consequence, we get $[H \cdot M] = [H \cdot M, H \cdot M] = H^2 \cdot [M]$, as we already know.*

Definition 33. *Let $X = X_0 + M + A$, $Y = Y_0 + N + B$ be two continuous semimartingales in their Doob-Meyer decompositon. We define their covariation as $[X,Y] = [M,N]$.*

The Kunita-Watanabe identity is also valid for semimartingales: $[H \cdot X, Y] = H \cdot [X,Y]$.

Proposition 15. *Let X,Y be independent continuous semimartingales. Then*

$$[X,Y]=0.$$

Remark. *This is the analogue of covariance for random variables. Also, the opposite is not true.*

Example. *Take B,B' two independent standard Brownian motions. Then, $[B,B']=0$. Indeed,*

$$[B,B']=\frac{1}{4}\left([B+B']-[B-B']\right).$$

But $B+B',B-B'$ are both distributed as $\sqrt{2}B$, so they have the same (deterministic) quadratic variation $t\mapsto 2t$ (check). Hence, $[B,B']=0$.

4.5 Itô's formula

Itô's formula is the main tool in stochastic integration. Indeed, even if it is usually impossible to compute stochastic integrals explicitly, we can still do calculus with it. We start with the equivalent of integration by parts, in the framework of stochastic processes.

Proposition 16 (Integration by parts)**.** *Let X,Y be continuous semimartingales. Then, we have*

$$X_tY_t-X_0Y_0=\int_0^t X_s dY_s+\int_0^t Y_s dX_s+[X,Y]_t,$$

that is,

$$d(X_tY_t)=X_t dY_t+Y_t dX_t+d[X,Y]_t.$$

Proof. Both sides are continuous, so it is enough to prove it for t of the form $m2^{-n}$ with m,n integers. First, observe that, for $s<t$, we have

$$X_tY_t-X_sY_s=X_s\left(Y_t-Y_s\right)+Y_s(X_t-X_s)+(X_t-X_s)(Y_t-Y_s).$$

Hence, writing $t_k:=k2^{-n}$, we have

$$X_tY_t-X_0Y_0=\sum_{k=0}^{m-1}\left[X_{t_k}\left(Y_{t_{k+1}}-Y_{t_k}\right)+Y_{t_k}\left(X_{t_{k+1}}-X_{t_k}\right)+\left(X_{t_{k+1}}-X_{t_k}\right)\left(Y_{t_{k+1}}-Y_{t_k}\right)\right].$$

We have seen that this, as $n\to\infty$, converges uniformly in probability on all compact to $(X\cdot Y)_t+(Y\cdot X)_t+[X,Y]_t$. □

We can now state Itô's formula, which is the basis of all stochastic calculus.

Theorem 27 (Itô's formula)**.** *Let $X^1,\dots,X^d$ be continuous semimartingales, and let $X:=(X^1,\dots,X^d)$. Let $f:\mathbb{R}^d\to\mathbb{R}$ be a C^2 map. Then we have*

$$f(X_t)=f(X_0)+\sum_{i=1}^d\int_0^t\frac{\partial f}{\partial x_i}(X_s)dX_s^i+\frac{1}{2}\sum_{i,j=1}^d\int_0^t\frac{\partial^2 f}{\partial x_i\partial x_j}(X_s)d[X^i,X^j]_s.$$

In particular, $f(X)$ is a continuous semimartingale.

Lecture 15: 11/12/2023

Remark. *As particular cases, for X a continuous semimartingale, we have*

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

and

$$f(t,X_t) = f(0,X_0) + \int_0^t \frac{\partial f}{\partial t}(s,X_s)ds + \int_0^t \frac{\partial f}{\partial x}(s,X_s)dX_s + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s,X_s)d[X]_s.$$

Remark. *Here is an intuitive proof. Let us perform a Taylor expansion of f. Let $t_k := k2^{-n}$.*

$$\begin{aligned} f(X_t) &= f(X_0) + \sum_{k=0}^{\lfloor 2^n t\rfloor -1} f(X_{t_{k+1}}) - f(X_{t_k}) + f(X_t) - f(X_{\lfloor 2^n t\rfloor}) \\ &\approx f(X_0) + \sum_{k=0}^{\lfloor 2^n t\rfloor -1} f'(X_{t_{k+1}})\left(X_{t_{k+1}} - X_{t_k}\right) + \frac{1}{2}f''(X_{t_{k+1}})\left(X_{t_{k+1}} - X_{t_k}\right)^2 \\ &\to f(X_0) + \int_0^t f'(X_s)ds + \frac{1}{2}\int_0^t f''(X_s)d[X]_s. \end{aligned}$$

The second term is not negligible, as it is in the usual Lebesgue integral.

Proof of Theorem 27. We prove it for $d = 1$ (that is, (1)). Let $X = X_0 + M + A$ a continuous semimartingale in its Doob-Meyer decomposition, where A has total variation process V. For $r \geq 0$, define the stopping time $T_r = \inf\{t \geq 0, |X_t| + [M]_t + V_t \geq r\}$. Then, $T_r \nearrow \infty$ as $r \to \infty$. It is therefore enough to prove the identity on $[0,T_r]$. Let $\mathscr{A} \subseteq C^2([-r,r],\mathbb{R})$ be the set of functions for which Itô's formula holds. Then

(i) $\mathscr{A}$ contains $x \mapsto 1$ and $x \mapsto x$.

(ii) $\mathscr{A}$ is a vector space.

(iii) $\mathscr{A}$ is an algebra, that is, $f,g \in \mathscr{A} \Rightarrow fg \in \mathscr{A}$.

(iv) If $f_n \to f$ in $C^2([-r,r],\mathbb{R})$ and $f_n \in \mathscr{A}$ for all n, then $f \in \mathscr{A}$.

In (iv), convergence means

$$\Delta_{n,r} := \max\left\{ \sup_{x\in[-r,r]} |f_n(x) - f(x)|, \sup_{x\in[-r,r]} |f_n'(x) - f'(x)|, \sup_{x\in[-r,r]} |f_n''(x) - f''(x)| \right\} \underset{n\to\infty}{\to} 0.$$

Roughly speaking, (i) - (iii) imply that all polynomials are in $\mathscr{A}$. The Weierstrass approximation theorem states that polynomials are dense in $C^2([-r,r],\mathbb{R})$, so that, by (iv), $\mathscr{A} = C^2([-r,r],\mathbb{R})$. This implies our result. Observe first that (i) and (ii) are clear. Let us now show (iii). Let $f,g \in \mathscr{A}$ and set $F_t = f(X_t), G_t = g(X_t)$. Since the claim holds for f and g, we know that F and G are continuous semimartingales. Hence, by integration by parts, we have

$$F_tG_t - F_0G_0 = (F\cdot G)_t + (G\cdot F)_t + [F,G]_t. \tag{2}$$

By the Kunita-Watanabe identity and Itô's formula for G, we get

$$
\begin{aligned}
(F\cdot G)_t &= (F\cdot(1\cdot G))_t = (F\cdot(g(X_s)-g(X_0))_t \\
&= \left(F\cdot\left(g'(X)\cdot X+\frac{1}{2}g''(X)\cdot[X]\right)\right)_t = \big((f(X)g'(X))\cdot X\big)_t + \left(\frac{1}{2}\big(f(X)g''(X)\big)\cdot[X]\right)_t \\
&= \int_0^t f(X_s)g'(X_s)dX_s + \frac{1}{2}\int_0^t f(X_s)g''(X_s)d[X]_s.
\end{aligned}
$$

The same holds for $(G\cdot F)_t$. Using the same way the Kunita-Watanabe identity and Itô's formula for f and g, we obtain

$$
\begin{aligned}
[F,G]_t &= \left[f'(X)\cdot X+\frac{1}{2}f''(X)\cdot[X], g'(X)\cdot X+\frac{1}{2}g''(X)\cdot[X]\right] \\
&= \left[f'(X)\cdot X, g'(X)\cdot X\right] = f'(X)\cdot[X, g'(X)\cdot X] \\
&= f'(X)\cdot(g'(X)\cdot[X]) = f'(X)g'(X)\cdot[X].
\end{aligned}
$$

Plugging these three equalities into (2) proves (iii). Indeed, we get

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

We now need to prove (iv). Let $(f_n)\in\mathscr{A}$ such that $f_n\to f$ in $C^2([-r,r],\mathbb{R})$. We consider first the finite variation part:

$$
\begin{aligned}
\int_0^{t\wedge T_r} |f_n'(X_s)-f'(X_s)|dA_s + \frac{1}{2}\int_0^{t\wedge T_r} |f_n''(X_s)-f''(X_s)|d[M]_s &\leq \Delta_{n,r}V_{t\wedge T_r} + \frac{1}{2}\Delta_{n,r}[M]_{t\wedge T_r} \\
&\leq 2r\Delta_{n,r} \underset{n\to\infty}{\to} 0.
\end{aligned}
$$

Hence, we have

$$
\int_0^{t\wedge T_r} f_n'(X_s)dA_s + \frac{1}{2}\int_0^{t\wedge T_r} f_n''(X_s)d[M]_s \underset{n\to\infty}{\to} \int_0^{t\wedge T_r} f'(X_s)dA_s + \frac{1}{2}\int_0^{t\wedge T_r} f''(X_s)d[M]_s.
$$

Let us now look at the Itô part. We have that $M^{T_r}\in\mathscr{M}_c^2$, so we have by Itô isometry

$$
\begin{aligned}
||(f_n'(X)\cdot M)^{T_r} - (f'(X)\cdot M)^{T_r}||^2_{\mathscr{M}_c^2} &= \mathbb{E}\left[\int_0^{T_r} (f_n'(X_s)-f'(X_s))^2 d[M]_s\right] \\
&\leq \Delta_{n,r}^2\mathbb{E}[[M]_{T_r}] \leq r\Delta_{n,r}^2 \underset{n\to\infty}{\to} 0.
\end{aligned}
$$

Hence, $(f_n'(X)\cdot M)^{T_r}\to(f'(X)\cdot M)^{T_r}$ in $\mathscr{M}_c^2$ as $n\to\infty$. Therefore, we can pass to the limit in Itô's formula for f_n, to get

$$
f(X_{t\wedge T_r}) = f(X_0) + \int_0^{t\wedge T_r} f'(X_s)dX_s + \frac{1}{2}\int_0^{t\wedge T_r} f''(X_s)d[X]_s.
$$

Finally, we let $r\to\infty$.

□

Example. *Let $X=B$ be a standard Brownian motion and $f: x\mapsto x^2$. Then, we obtain*

$$
B_t^2 = 2\int_0^t B_s dB_s + t.
$$

Notation and computational rules

- $$Z_t - Z_0 = \int_0^t H_s dX_s \Leftrightarrow dZ_t = H_t dX_t.$$

- $$Z_t = [X,Y]_t \Leftrightarrow dZ_t = dX_t dY_t.$$

- $$H_t(K_t dX_t) = (H_t K_t) dX_t$$

- $$H_t dX_t dY_t = (H_t dX_t) dY_t \text{ (Kunita-Watanabe)}$$

- $$d(X_t Y_t) = X_t dY_t + Y_t dX_t + dX_t dY_t \text{ (integration by parts)}$$

- $$df(X_t) = \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(X_t) dX_t^i + \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_t) dX_t^i dX_t^j \text{ (Itô's formula)}$$

5 Applications of Itô's formula

We now turn to some nice applications of Itô's formula. Most of the time, it consists in finding the right function to apply the formula to.

5.1 Lévy's characterization of Brownian motion

We can derive many interesting results from Itô's formula. Let us first look at a charaterization of Brownian motion, in all dimension.

Definition 34. *Let $d \geq 1$. A d-dimensional Brownian motion $B := (B^1, \dots, B^d)$ is a random process with values in $\mathbb{R}^d$ such that $B^1, \dots, B^d$ are i.i.d. 1-dimensional standard Brownian motions.*

Lemma 8. *Let $f(t,x) : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ be $C^{1,2}$ (continuously differentiable, and twice continuously differentiable in x). Let $B := (B^1, \dots, B^d)$ be a d-dimensional Brownian motion. Then, the process $(M_t, t \geq 0)$ defined as*

$$M_t = f(t, B_t) - f(0, B_0) - \int_0^t \left(\frac{\partial}{\partial t} + \frac{1}{2}\Delta \right) f(s, B_s)$$

is a continuous local martingale, where $\Delta := \sum_{i=1}^d \frac{\partial^2 f}{\partial x_i^2}$ is the Laplacian.

Proof. Assume first that $f \in C^{2,2}$. Then, by Itô's formula, we have (easily checked)

$$M_t = \int_0^t \sum_{i=1}^d \frac{\partial f}{\partial x_i}(s, X_s) dB_s^i \in \mathscr{M}_{c,loc}.$$

To finish the proof, approximate $f \in C^{1,2}$ by $C^{2,2}$ functions. □

Now, let us consider a filtered probability space $(\Omega, (\mathscr{F}_t)_{t \geq 0}, \mathbb{P})$ (again here, $(\mathscr{F}_t, t \geq 0)$ is assumed to be left-continuous and complete). We say that B is an $\mathscr{F}$-Brownian motion if B is a Brownian motion, adapted to $\mathscr{F}$, and for all $s \leq t$, $B_t - B_s$ is independent of $\mathscr{F}_s$. In particular, B is always a Brownian motion with respect to its own filtration.

Theorem 28 (Lévy's characterization of Brownian motion)**.** *Let $X := (X^1, \dots, X^d)$ be continuous and adapted to some filtration $\mathscr{F}$. Then, the following are equivalent:*

- *X is a d-dimensional Brownian motion ;*
- *X^i are local martingales and for all $t \geq 0, 1 \leq i, j \leq d$: $[X^i, X^j]_t = \delta_{i,j} t$.*

Proof. (i) ⇒ (ii) is already known. Let us prove (ii) ⇒ (i). To this end, it is enough to show that, for all $0 \leq s \leq t$:

$$X_t - X_s \sim \mathscr{N}(0, (t-s)Id)$$

and $X_t - X_s$ is independent of $\mathscr{F}_s$. This is equivalent to showing that for all $u \in \mathbb{R}^d$,

$$\mathbb{E}\left[e^{i\langle u, X_t - X_s \rangle} | \mathscr{F}_s \right] = e^{-\frac{1}{2}||u||^2 (t-s)}. \tag{3}$$

Fix $u \in \mathbb{R}^d$, and define $Y_t := \langle u, X_t \rangle = \sum_{i=1}^d u_i X_t^i$. Then, $Y \in \mathscr{M}_{c,loc}$ since $\mathscr{M}_{c,loc}$ is a vector space. By assumption, we have

$$\begin{aligned} [Y]_t = [Y,Y]_t &= \left[\sum_{i=1}^d u_i X_t^i, \sum_{i=1}^d u_i X_t^i \right] \\ &= \sum_{i,j=1}^d u_i u_j [X_t^i, X_t^j] = ||u||^2 t. \end{aligned}$$

Consider now

$$Z_t := e^{iY_t + \frac{1}{2}[Y]_t} = e^{i\langle u, X_t \rangle + \frac{1}{2}||u||^2 t}.$$

By Itô's formula for $U := iY + \frac{1}{2}[Y]$ and $f : x \mapsto e^x$, we get

$$\begin{aligned} dZ_t &= Z_t dU_t - \frac{1}{2} Z_t d[Y]_t \\ &= iZ_t dY_t + \frac{1}{2} Z_t d[Y]_t - \frac{1}{2} Z_t d[Y]_t \\ &= iZ_t dY_t. \end{aligned}$$

Therefore, $Z \in \mathscr{M}_{c,loc}$. Moreover, Z is bounded on any interval $[0, t_0]$ where t_0 is fixed, so it is uniformly integrable on $[0, t_0]$. Hence it is a true martingale. In particular, for all $s \leq t$, we have $\mathbb{E}[Z_t | \mathscr{F}_s] = Z_s$. This is exactly (3). □

Lecture 16: 13/12/2023

Remark. *This means that, if a continuous local martingale M satisfies that $t \mapsto M_t^2 - t$ is a continuous local martingale, then M is a Brownian motion.*

This is a particular case of a martingale problem: some stochastic processes X are characterized by the fact that some given functionals of X are (local) martingales. In the case of Brownian motion, these functionals are $F : (X_t) \mapsto (X_t)$ and $G : (X_t) \mapsto (X_t^2 - t)$.

5.2 Dubins-Schwarz theorem

The Dubins-Schwarz theorem can be summarized as follows: every (continuous) local martingale with infinite total quadratic variation is a Brownian motion, up to a change of time.

Theorem 29 (Dubins-Schwarz). *Let $M \in \mathscr{M}_{c,loc}$, adapted to some filtration $(\mathscr{F}_t, t \geq 0)$, with $M_0 = 0$ and $[M]_\infty = \infty$ a.s.. For all $s \geq 0$, set $\tau_s = \inf\{t > 0, [M]_t > s\}$, and set $B_s := M_{\tau_s}$. Then, τ_s is an $(\mathscr{F}_t, t \geq 0)$ stopping time, and for $\mathscr{G}_s = \mathscr{F}_{\tau_s}$, we have that $(B_t, t \geq 0)$ is a $(\mathscr{G}_t, t \geq 0)$ Brownian motion.*

Observe that, if M is a Brownian motion, then $\tau_s = s$ for all s, ω and the result is clear.

Proof. In the whole proof, assume by localization that $M \in \mathscr{M}_c^2$. Since $[M]_\infty = \infty$ a.s. and M is adapted and continuous, we have that τ_s is a stopping time and $\tau_s < \infty$ a.s. for all s. We also know that $s \mapsto \tau_s$ is càdlàg (check!). Since M is continuous, we get that B is càdlàg.

Lemma 9. *$s \mapsto B_s$ is continuous.*

Proof. We need to prove that, for all $s > 0$, $B_{s-} = B_s$, or equivalently that $M_{\tau_{s-}} = M_{\tau_s}$, where

$$\tau_{s-} = \lim_{\substack{u \uparrow s \\ u \neq s}} \tau_u = \inf\{t \geq 0, [M]_t \geq s\}.$$

Note that τ_{s-} is also a stopping time, with $\tau_{s-} \leq \tau_s$ for all s. We therefore only need to show that M is constant on $[\tau_{s-}, \tau_s]$, that is, intervals on which $[M]$ is constant. Note that $[M^{\tau_s}]_\infty = s$, so that $(M^2 - [M])^{\tau_s}$ is a uniformly integrable martingale. We apply the optional stopping theorem to τ_{s-}, to get

$$\mathbb{E}\left[(M^2 - [M])_{\tau_s} | \mathscr{F}_{\tau_{s-}}\right] = M^2_{\tau_{s-}} - [M]_{\tau_{s-}}.$$

Using the fact that $[M]_{\tau_{s-}} = [M]_{\tau_s} = s$, we obtain that $\mathbb{E}[M^2_{\tau_s} - M^2_{\tau_{s-}} | \mathscr{F}_{\tau_{s-}}] = 0$, so

$$\mathbb{E}\left[(M_{\tau_s} - M_{\tau_{s-}})^2\right] = 0 \text{ a.s.}$$

Hence, M is constant a.s. on $[\tau_{s-}, \tau_s]$ and $B_{s-} = M_{\tau_{s-}} = M_{\tau_s} = B_s$. Hence, B is continuous at s, almost surely. We now need to prove that a.s. B is continuous on $\mathbb{R}_+$. To this end, for all $r \in \mathbb{R}_+$, let $S_r := \inf\{t \geq r, M_t \neq M_r\}$ and $T_r := \inf\{t \geq r, [M]_t \neq [M]_r\}$. Then, for each r, $T_r = S_r$ a.s., and both are càdlàg (because M and $[M]$ are continuous). So, T and S are indistinguishable, and B is continuous. □

It is also clear that B is adapted to $\mathscr{G}$. We now show that B is a Brownian motion. Again, $M^2 - [M]$ is a uniformly integrable martingale. By the optional stopping theorem, we obtain that, for $0 \leq r < s < \infty$:

$$\mathbb{E}\left[B_s | \mathscr{G}_r\right] = \mathbb{E}[M_{\tau_s} | \mathscr{F}_{\tau_r}] = M_{\tau_r} = B_r,$$

and

$$\mathbb{E}\left[B_s^2 - s | \mathscr{G}_r\right] = \mathbb{E}\left[M^2_{\tau_s} - [M]_{\tau_s} | \mathscr{F}_{\tau_r}\right] = M^2_{\tau_r} - [M]_{\tau_r} = B_r - r.$$

Hence, $[B]_t = t$ for all t. We conclude the proof using Lévy's characterization of the Brownian motion. □

A big interest of the Dubins-Schwarz theorem is the following.

Theorem 30. *Let M be a continuous local martingale. Then, the following hold:*

(i) $\mathbb{P}(\lim_{t\to\infty} |M_t| = \infty) = 0$.

(ii) $\{\omega : \lim_{t\to\infty} M_t(\omega)$ *exists and is finite* $\} = \{\omega : [M]_\infty(\omega) < \infty\}$*, up to null events.*

(iii) $\{[M]_\infty = \infty\} = \{\limsup M_t = +\infty$ *and* $\liminf M_t = -\infty\}$*, up to null events.*

In other words, either M converges, or it oscillates.

Proof. We know by Dubins-Schwarz theorem that $M_t = B_{[M]_t}$, where B is distributed as a standard Brownian motion. Since $[M]$ is continuous, the three items follow from the same events for the Brownian motion. □

5.3 Transience and recurrence of Brownian motion

We now consider Brownian motion in dimension possibly larger than 1. One of the most important properties to check on d-dimensional Brownian motion is what is called *transience* or *recurrence*. Namely, what is the probability that the process ever returns to the origin 0?

Theorem 31. *Let B be a d-dimensional Brownian motion. Then:*

(i) if $d=1$, then B is point-recurrent, that is, for any $x \in \mathbb{R}$, $\{t, B_t = x\}$ is unbounded a.s.

(ii) if $d \geq 3$, then B is transient, that is, $\lim_{t\to\infty} ||B_t|| = +\infty$ a.s.

(iii) if $d=2$, then B is neighbourhood-recurrent, that is, for every nonempty open set $\mathscr{O} \subseteq \mathbb{R}^2$, $\{t, B_t \in \mathscr{O}\}$ is unbounded a.s. However, it is not point-recurrent, since, for all $x \in \mathbb{R}^2$, $\mathbb{P}_0\left(B_t = x \text{ for some } t > 0\right) = 0$.

Proof. Observe that (i) is a consequence of $\limsup B_t = +\infty$ and $\liminf B_t = -\infty$. We now turn to (ii). Since $||B_t||^2 \geq ||\left(B_t^1\right)^2 + \left(B_t^2\right)^2 + \left(B_t^3\right)^2||$, it suffices to prove it for $d=3$.

We claim that, in dimension 3, $f : x \mapsto \frac{1}{||x||}$ is harmonic at any $x \neq 0$, that is, $\Delta f = 0$ (check). Therefore, by Itô's formula, letting $\tau := \inf\{t, t_t = 0\}$, we get that

$$M_t := f\left(B_{t\wedge\tau}\right)$$

is a local martingale. Furthermore, $M \geq 0$ so by Proposition 8 M is a nonnegative supermartingale. Hence, it converges a.s. by Theorem 17, to some M. We only need to show that $M = 0$ a.s. To this end, observe that, on the event $\{M > 0\}$, $\{\frac{1}{||B_t||}, t \geq 0\}$ is bounded. But B^1 is unbounded a.s. so this holds with probability 0. Finally, $M = 0$ a.s. and thus $||B|| \to \infty$ a.s.

We now need to consider the case $d=2$. Without loss of generality, we will assume that $B_0 = 1$ and show that B never hits 0, but is close to it infinitely often. For $k \in \mathbb{Z}$, let $r_k := e^k$ and $J_k = \inf\{t \geq 0, ||B_t|| = r_k\}$. Set also $J := J_{-\infty} = \inf\{t \geq 0, ||B_t|| = 0\}$.

We first show that $J = +\infty$ a.s. To this end, define $T_0 = 0$ and, letting $Z_n := ||B_{T_n}||$, set $T_{n+1} := \inf\{t \geq T_n, ||B_t|| \in \{e^{-1}Z_n, eZ_n\}\}$. Then, for $k, m \geq 1$, we have that $J_{-k} < J_m$ if and only if the random walk $(\log Z_n)_{n\geq 1}$ visits $-k$ before visiting m. On the other hand, by Itô's formula, we have that $M_t := \log ||B_{t\wedge J}||$ is a local martingale, since $(x,y) \mapsto \log(x^2+y^2)$ is harmonic on $\mathbb{R}^2 \setminus \{0,0\}$. If we stop it at time T_n, it is bounded so $(M_{t\wedge T_n}, t \geq 0)$ is a UI martingale. By the optional stopping theorem applied to T_{n+1}, we have

$$\mathbb{E}[\log Z_{n+1} | \log Z_n = k] = \log Z_n = k.$$

Thus,

$$\mathbb{P}(\log Z_{n+1} = k \pm 1 | \log Z_n = k) = 1/2,$$

and $(\log Z_n)_{n\geq 0}$ is a simple symmetric random walk on $\mathbb{Z}$. Hence, again by optional stopping theorem, we have that

$$\mathbb{P}(J_{-k} < J_m) = \frac{m}{m+k} \underset{k\to\infty}{\to} 0.$$

Therefore, $\mathbb{P}(J < J_m) = 0$, for all $m \geq 0$. Since $J_m \to \infty$ a.s. as $m \to \infty$, we conclude that $J = +\infty$ a.s. Moreover, fix $\varepsilon > 0$ and choose $k \in \mathbb{Z}_+$ such that $e^{-k} < \varepsilon$. Then, by recurrence of the simple random walk on $\mathbb{Z}$, there are infinitely many n such that $\log Z_n = 0$ and, for any such n, there exists $m \geq n$ such that $\log Z_m = -k$, that is, $||B_{T_m}|| \leq \varepsilon$. The result follows. □

5.4 Exponential martingale, Girsanov's theorem and Cameron-Martin formula

Lecture 17: 15/01/2024

Girsanov's theorem is connected to the fact that, if our set $(\Omega, \mathscr{F})$ and a random variable X on Ω is given, then changing the probability measure that we put on $\mathscr{F}$ changes the distribution of X.

Example (ÉA simple example of change of measure)**.** *Let $\Omega = \{\omega_1, \omega_2\}$ with the σ-algebra $\mathscr{F} := \{\emptyset, \{\omega_1\}, \{\omega_2\}, \Omega\}$, and consider the variable X defined as $X(\omega_1) = 0, X(\omega_2) = 1$. Then, consider the probability measures $\mathbb{P}$ defined as $\mathbb{P}(\omega_1) = \mathbb{P}(\omega_2) = 1/2$, and $\mathbb{Q}$ defined as $\mathbb{Q}(\omega_1) = .4, \mathbb{Q}(\omega_2) = .6$. Then, under $\mathbb{P}$, X is a Bernoulli variable of parameter $1/2$, while under $\mathbb{Q}$, it is a Bernoulli variable of parameter $.6$.*

Hence, changing the underlying probability measure changes the laws of the random variables. Girsanov's theorem exploits this idea, and shows how semimartingales changes under a change of measure.

If I am given the entire trajectory of a Brownian motion with drift, I can figure out whether it came from a Brownian motion with drift or without drift. Indeed, if it has drift μ, then $\frac{X_t}{t} \to \mu$ a.s. as $t \to \infty$. However, it is less clear if I am only given part of this trajectory (say, up to time t). But this can be estimated. In particular, changing the drift makes the laws of these processes absolutely continuous with respect to each other. This is the content of Girsanov's (or Cameron-Martin's) theorem, which can be used to understand processes with drift.

In particular, we will see as an example (Example 5.4) that we can compute the distribution of $\inf\{t \geq 0, B_t + ct = a\}$ for $a, c \in \mathbb{R}$.

Exponential martingale

Let us first define a specific local martingale of importance, called the *exponential (local) martingale* associated to a martingale.

Definition 35 (Exponential martingale)**.** *Let $M \in \mathscr{M}_{c,loc}$ with $M_0 = 0$ a.s. The process $(Z_t, t \geq 0)$ defined as*

$$Z_t = \mathscr{E}(M)_t := \exp\left(M_t - \frac{1}{2}[M]_t\right)$$

is called the exponential (local) martingale.

To show that it is a local martingale, let $f(x,y) := \exp\left(x - \frac{1}{2}y\right)$. Then, Itô's formula provides:

$$\begin{aligned} Z_t = f(M_t, [M]_t) &= f(0,0) + \int_0^t Z_s dM_s - \int_0^t \frac{1}{2} Z_s d[M]_s + \frac{1}{2}\int_0^t Z_s d[M]_s \\ &= \int_0^t Z_s dM_s, \end{aligned}$$

or equivalently $dZ_t = Z_t dM_t$, so that $Z \in \mathscr{M}_{c,loc}$.

This exponential martingale has nice properties, which allows us to derive two important results: Girsanov's theorem, and the Cameron-Martin formula.

Let us start with a lemma.

Lemma 10 (Exponential martingale inequality)**.** *Let $M \in \mathscr{M}_{c,loc}$ with $M_0 = 0$. Then, for all $x > 0, u > 0$, we have*

$$\mathbb{P}\left(\sup_{t\geq 0} M_t \geq x, [M]_\infty \leq u\right) \leq \exp\left(-\frac{x^2}{2u}\right).$$

Proof. Fix $x \geq 0$ and set $T := \inf\{t \geq 0, M_t \geq x\}$. Fix now $\theta > 0$, and let

$$Z := \mathscr{E}(\theta M^T) = \exp(\theta M^T - \frac{1}{2}\theta^2 [M]^T).$$

Then, $Z \in \mathscr{M}_{c,loc}$ and $|Z| \leq e^{\theta x}$. Hence, $Z \in \mathscr{M}_c^2$ and converges to some Z_∞. By optional stopping theorem, we have $1 = \mathbb{E}[Z_0] = \mathbb{E}[Z_\infty]$. Finally, for any $u > 0$, we get

$$\begin{aligned}\mathbb{P}\left(\sup_{t\geq 0} M_t \geq x, [M]_\infty \leq u\right) &\leq \mathbb{P}\left(\sup_{t\geq 0} M_t^T \geq x, [M^T]_\infty \leq u\right)\\ &\leq \mathbb{P}\left(Z_\infty \geq e^{\theta x - \frac{1}{2}\theta^2 u}\right)\\ &\leq e^{-\theta x + \frac{1}{2}\theta^2 u} \text{ by Markov's inequality.}\end{aligned}$$

Finally, we observe that $\theta \mapsto \frac{1}{2}\theta^2 u - \theta x$ is minimum when $\theta u = x$. The result follows. □

The next proposition is a criterion for $\mathscr{E}(M)$ to be a uniformly integrable martingale, which will be our assumption in the statement of Girsanov's theorem.

Proposition 17. *Let $M \in \mathscr{M}_{c,loc}$ with $M_0 = 0$ be such that $[M]_\infty$ is uniformly bounded. Then, $\mathscr{E}(M)$ is a UI martingale.*

Proof. Let $C \in (0, \infty)$ such that $[M]_\infty \leq C$. By the exponential martingale inequality, we get for all $x > 0$:

$$\mathbb{P}\left(\sup_{t\geq 0} M_t \geq x\right) \leq \exp\left(-\frac{x^2}{2C}\right).$$

Now, observe that $\sup_{t\geq 0} \mathscr{E}(M)_t \leq e^{\sup_{t\geq 0} M_t}$, and

$$\begin{aligned}\mathbb{E}\left[e^{\sup_{t\geq 0} M_t}\right] &= \int_0^\infty \mathbb{P}\left(e^{\sup_{t\geq 0} M_t} \geq \lambda\right) d\lambda\\ &= \int_0^\infty \mathbb{P}\left(\sup_{t\geq 0} M_t \geq \log\lambda\right) d\lambda\\ &\leq 1 + \int_1^\infty e^{-\frac{(\log\lambda)^2}{2C}} d\lambda < \infty.\end{aligned}$$

Hence, $\sup_{t\geq 0} \mathscr{E}(M)_t$ is integrable and therefore $\mathscr{E}(M)$ is a UI martingale . □

Remark. *Let $M \in \mathscr{M}_{c,loc}$ be such that $\mathbb{E}[\sup_{t\geq 0} M_t] < \infty$. Then M is a true martingale. Indeed, fix $t \geq s \geq 0$ and $A \in \mathscr{F}_s$. Letting $T_n)$ be a sequence of stopping times reducing M, we have $0 = \mathbb{E}[\mathbb{1}_A(M_t^{T_n} - M_s^{T_n})]$. By dominated convergence, this converges to $\mathbb{E}[\mathbb{1}_A(M_t - M_s)]$, which is thus equal to 0.*

Before stating Girsanov's theorem, we briefly recall notions from measure theory (see also Theorem 1).

Definition 36. *Let $\mathbb{P}, \tilde{\mathbb{P}}$ be two probability measures on a measured space $(\Omega, \mathscr{F})$. We say that $\tilde{\mathbb{P}}$ is absolutely continuous with respect to $\mathbb{P}$ if, for all $A \in \mathscr{F}$,*

$$\mathbb{P}(A) = 0 \Rightarrow \tilde{\mathbb{P}}(A) = 0.$$

In this case, we write $\tilde{\mathbb{P}} \ll \mathbb{P}$ and there exists $f : \Omega \to \mathbb{R}_+$ (defined up to a $\mathbb{P}$-null set) that is $\mathscr{F}$-measurable, such that, for all $A \in \mathscr{F}$:

$$\tilde{\mathbb{P}}(A) = \int_{\Omega} f(\omega) \mathbb{1}_{\omega \in A} d\mathbb{P}(\omega).$$

f is called the Radon-Nikodym derivative of $\tilde{\mathbb{P}}$ with respect to $\mathbb{P}$. We write

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}_{|\mathscr{F}}} = f.$$

Theorem 32 (Girsanov's theorem)**.** *Let $M \in \mathscr{M}_{c,loc}$ with $M_0 = 0$ and such that $Z := \mathscr{E}(M)$ is a UI martingale. We can define a new probability measure $\tilde{\mathbb{P}} \ll \mathbb{P}$ on $(\Omega, \mathscr{F})$ by setting, for all $A \in \mathscr{F}$:*

$$\tilde{\mathbb{P}}(A) = \mathbb{E}_{\mathbb{P}}\left[Z_\infty \mathbb{1}_A\right].$$

Then, for all $X \in \mathscr{M}_{c,loc}(\mathbb{P})$, $X - [X,M] \in \mathscr{M}_{c,loc}(\tilde{\mathbb{P}})$. Moreover, the quadratic variations of X under $\mathbb{P}$ and $\tilde{\mathbb{P}}$ agree a.s.

Remark that we have

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = Z_\infty.$$

Proof of Girsanov's theorem. Observe first that $\tilde{\mathbb{P}}$ is well-defined: since Z is UI, Z_∞ exists a.s., $Z_\infty \geq 0$ and $\mathbb{E}[Z_\infty] = \mathbb{E}[Z_0] = 1$ by optional stopping. Thus, $\tilde{\mathbb{P}}(\Omega) = 1$. Countable additivity is obtained by monotone convergence. Furthermore, clearly, if $\mathbb{P}(A) = 0$ then $\tilde{\mathbb{P}}(A) = 0$ and we have $\tilde{\mathbb{P}} \ll \mathbb{P}$. Now, let $X \in \mathscr{M}_{c,loc}(\mathbb{P})$ and define, for all n:

$$T_n := \inf\{t \geq 0, |X_t - [X,M]_t| \geq n\}.$$

Then, $\mathbb{P}(T_n \to \infty) = 1$, so $\tilde{\mathbb{P}}(T_n \to \infty) = 1$. It suffices to show that, for all n, $Y^{T_n} := X^{T_n} - [X,M]^{T_n} \in \mathscr{M}_{c,loc}(\tilde{\mathbb{P}})$. Then without loss of generality we can assume that Y is uniformly bounded. Then, for $\mathbb{P}$, by integration by parts:

$$\begin{aligned} d(YZ)_t &= Y_t dZ_t + Z_t dY_t + d[Y,Z]_t \\ &= Y_t Z_t dM_t + Z_t(dX_t - d[X,M]_t) + d[Y,Z]_t \\ &= Y_t Z_t dM_t + Z_t(dX_t - d[X,M]_t) + Z_t d[X,M]_t \\ &= Y_t Z_t dM_t + Z_t dX_t, \end{aligned}$$

where we have used that $Z = (Z \cdot M)$, and Kunita-Watanabe. Hence, $YZ \in \mathscr{M}_{c,loc}(\mathbb{P})$.

Furthermore, Z is a UI martingale and Y is uniformly bounded so $\{Y_T Z_T, T$ stopping time$, T \leq t_0\}$ is UI, and thus YZ is a true martingale.

Hence, for all $t \geq s \geq 0$ and $A \in \mathscr{F}_s$, we have

$$\mathbb{E}_{\tilde{\mathbb{P}}}\left[\mathbb{1}_A(Y_t - Y_s)\right] = \mathbb{E}_{\mathbb{P}}\left[\mathbb{1}_A Z_\infty (Y_t - Y_s)\right] = \mathbb{E}_{\mathbb{P}}\left[\mathbb{1}_A (Z_t Y_t - Z_s Y_s)\right] = 0.$$

Thus, $Y \in \mathscr{M}_{c,loc}(\tilde{P})$.

Finally, to see why the quadratic variations are the same, recall that

$$\sum_{k=0}^{\lceil 2^n \rceil t - 1} (X_{(k+1)2^{-n}} - X_{k2^{-n}})^2$$

converges ucp (uniformly in probability on all compacts) for $\mathbb{P}$, and thus also for $\tilde{\mathbb{P}}$. □

We can get from there a lot of results. For instance, let X be a continuous $\mathbb{P}$-martingale. Then, X is a continuous $\tilde{P}$-semimartingale, whose Doob-Meyer decomposition is: $X = (X - [X,M]) + [X,M]$. Hence, the class of continuous $\mathbb{P}$-semimartingales is included in the set of continuous $\tilde{P}$-semimartingales (with equality if $Z_\infty > 0$ $\mathbb{P}$-a.s.).

Corollary 10. *Let B be a Brownian motion under $\mathbb{P}$, and $M \in \mathscr{M}_{c,loc}(\mathbb{P})$. Let $\tilde{\mathbb{P}}$ be as in Girsanov's theorem. Then, $\tilde{B} = B - [B,M]$ is a Brownian motion under $\tilde{\mathbb{P}}$.*

Proof. By Girsanov's theorem, $\tilde{B} \in \mathscr{M}_{c,loc}(\tilde{\mathbb{P}})$, and $[\tilde{B}]_t^{\tilde{P}} = [B]_t^{\mathbb{P}} = t$. The result follows by Lévy's characterization. □

We will now state a useful consequence of this result, which is the Cameron-Martin formula.

Definition 37. *Recall the definition of the Wiener space $(W, \mathscr{W}, \mu)$. Let $W := C(\mathbb{R}_+, \mathbb{R})$ the set of continuous functions from $\mathbb{R}_+$ to $\mathbb{R}$. Let $\mathscr{W} := \sigma(X_t, t \geq 0)$, where $(X_t)_{t \geq 0}$ is the canonical process on $(W, \mathscr{W})$, that is:*

$$X_t : W \to \mathbb{R}, \omega \mapsto \omega(t).$$

The Wiener measure μ is the unique measure on Ω^* such that X is a Brownian motion.

Definition 38. *Define the Cameron-Martin space*

$$\mathscr{H} := \left\{ h \in W, h(t) = \int_0^t \phi(s) ds \text{ for some } \phi \in L^2(\mathbb{R}_+) \right\}.$$

For $h \in \mathscr{H}$, write $\dot{h}$ for ϕ.

Theorem 33 (Cameron-Martin formula)**.** *Let $h \in \mathscr{H}$ and, for all $A \in \mathscr{W}$, set*

$$\mu^h(A) := \mu(\omega \in W, \omega + h \in A)$$

(in words, μ^h is the law of $X_t + h(t)$, where we recall that X is a Brownian motion on W). Then, $\mu^h \ll \mu$, and the Radon-Nikodym derivative is given by

$$\frac{d\mu^h}{d\mu} = \exp\left(\int_0^\infty \dot{h}(s) dX_s - \frac{1}{2} \int_0^\infty \dot{h}(s)^2 ds \right).$$

Lecture 18: 17/01/2024

Proof. Set, for all $t \geq 0$, $M_t = \int_0^t \dot{h}(s) dX_s = \int_0^t \phi(s) dX_s$. Then, M is a local martingale and $[M]_t = \int_0^t \phi(s)^2 ds$. Since $\phi \in L^2$, we have

$$[M]_\infty < \infty,$$

so by Proposition 17, $\mathscr{E}(M)$ is a UI martingale. By Girsanov's theorem, $\mathscr{E}(M)$ defines a new probability measure $\mathbb{P}$, with

$$\begin{aligned}\frac{d\mathbb{P}}{d\mu} &= \exp\left(M_\infty - \frac{1}{2}[M]_\infty\right) \\ &= \exp\left(\int_0^\infty \dot{h}(s) dX_s - \frac{1}{2}\int_0^\infty \dot{h}(s)^2 ds\right)\end{aligned}$$

and, under $\mathbb{P}$, $\tilde{X} := X - [X, M]$ is a Brownian motion. But, for all t:

$$[X, M]_t = [X, \phi \cdot X]_t = (\phi \cdot [X, X])_t = \int_0^t \phi(s) ds = h(t).$$

Hence, $X_t = \tilde{X}_t + h(t)$; under $\mathbb{P}$, X has the law of $\tilde{X}_t + h(t)$ and $\mathbb{P} = \mu^h$ as wanted. □

We finish with an application of the Cameron-Martin formula.

Example. *Let B be a standard Brownian motion and, for $a > 0$, set $T_a := \inf\{t \geq 0, B_t = a\}$. Let $c \in \mathbb{R}$. We are interested in the distribution of the stopping time*

$$S_a := \inf\{t \geq 0, B_t + ct = a\}.$$

Clearly, if $c = 0$, we have that $S_a = T_a$ and the result is given in Example 2.6. Thanks to the Cameron-Martin formula, we can handle the case $c \neq 0$. Fix $t \geq 0$, and apply the Cameron-Martin formula with $\phi : s \mapsto c\mathbb{1}_{0 \leq s \leq t}$, so that $h : s \mapsto c\,(s \wedge t)$. Define the function on Ω^*:

$$F : w \mapsto \mathbb{1}_{\max_{[0,t]} w(s) \geq a}.$$

Applying the Cameron-Martin formula provides

$$\begin{aligned}\mathbb{P}(S_a \leq t) &= \mathbb{E}[F(B + h)] \\ &= \mathbb{E}\left[F(B) \exp\left(\int_0^\infty \phi(s) dB_s - \frac{1}{2}\int_0^\infty \phi(s)^2 ds\right)\right] \\ &= \mathbb{E}\left[\mathbb{1}_{T_a \leq t} \exp\left(cB_t - \frac{c^2}{2} t\right)\right].\end{aligned}$$

We know that $t \mapsto \exp\left(cB_t - \frac{c^2}{2} t\right)$ is a martingale, so that by optional stopping theorem:

$$\mathbb{E}\left[\exp\left(cB_t - \frac{c^2}{2} t\right) | \mathscr{F}_{t \wedge T_a}\right] = \exp\left(cB_{t \wedge T_a} - \frac{c^2}{2}(t \wedge T_a)\right)$$

We therefore get by tower property

$$\begin{aligned}\mathbb{P}(S_a \le t) &= \mathbb{E}\left[\mathbb{1}_{T_a \le t}\exp\left(cB_{t\wedge T_a} - \frac{c^2}{2}(t \wedge T_a)\right)\right]\\ &= \mathbb{E}\left[\mathbb{1}_{T_a \le t}\exp\left(ca - \frac{c^2}{2}T_a\right)\right]\\ &= \int_0^t ds \frac{a}{\sqrt{2\pi s^3}} e^{-\frac{a^2}{2s}} e^{ca - \frac{c^2}{2}s}\\ &= \int_0^t ds \frac{a}{\sqrt{2\pi s^3}} e^{-\frac{1}{2s}(a-cs)^2}.\end{aligned}$$

So, the density of S_a is

$$s \mapsto \frac{a}{\sqrt{2\pi s^3}} e^{-\frac{1}{2s}(a-cs)^2}.$$

In particular, for $t \to \infty$, we get

$$\mathbb{P}(S_a < \infty) = \begin{cases} 1 & \text{if } c \ge 0 \\ e^{2ca} & \text{if } c \le 0. \end{cases}$$

This is consistent with the fact that a Brownian motion with negative drift converges a.s. to $-\infty$.

6 Appendix: Brownian motion and the Dirichlet problem

We end this lecture by going back to Brownian motion, and showing one of its many interesting features.

The Dirichlet problem is a famous problem in analysis, which can be formulated as follows. Fix a domain of $\mathbb{R}^d$, that is, an open connected subset of $\mathbb{R}^d$, and a function ϕ defined on its boundary $\partial D := \overline{D}\backslash\mathring{D}$. Is it possible to find a *harmonic* function f defined on D (that is, such that $\Delta f := \sum_{i=1}^d \frac{\partial^2 f}{\partial x_i^2} = 0$ on $\mathring{D}$) such that $f_{|\partial D} = \phi$?

Definition 39. *We say that a domain D satisfies the Poincaré cone condition if, for all $x \in \partial D$, there exists a nonempty open cone C with origin at x (that is, a set of the form $\{x + tu, t > 0, u \in \mathscr{U}\}$ where $\mathscr{U}$ is bounded and open) and such that, for some $r > 0$, we have*

$$C \cap B(x,r) \subset D^c.$$

Theorem 34 (Dirichlet problem)**.** *Let D be a bounded domain satisfying the Poincaré cone condition. Suppose that $\phi : \partial D \to \mathbb{R}$ is continuous. Let B be a d-dimensional Brownian motion and $J := \inf\{t \ge 0, B_t \in \partial D\}$. Then, J is an a.s. finite stopping time if $B_0 = x \in D$, and the function*

$$u : D \to \mathbb{R}, x \mapsto \mathbb{E}_x[\phi(B_J)]$$

is the unique continuous function on $\overline{D}$ satisfying

$$\begin{aligned}&\Delta u(x) = 0 \text{ for all } x \in D\\ &u(x) = \phi(x) \text{ for all } x \in \partial D.\end{aligned}$$

This theorem provides existence and uniqueness, under the only condition that ϕ has to be continuous. Furthermore, the solution is easily expressed in terms of Brownian motion.

In order to prove it, we need the following results:

Theorem 35 (Spherical averages)**.** *Let D be a domain in $\mathbb{R}^d$, and let $u : D \to \mathbb{R}$ be measurable and locally bounded. The following conditions are equivalent:*

(i) $u \in C^2(D, \mathbb{R})$ *and* $\Delta u = 0$ *on D (harmonicity)*

(ii) *For any ball $B(x,r) \subseteq D$, we have*

$$u(x) = \frac{1}{\lambda(B(x,r))} \int_{B(x,r)} u(y) d\lambda(y),$$

where λ is the d-dimensional Lebesgue measure on $\mathbb{R}^d$.

(iii) *For any ball $B(x,r) \subseteq D$, we have*

$$u(x) = \frac{1}{\sigma(\partial B(x,r))} \int_{\partial B(x,r)} u(y) d\sigma(y),$$

where σ is the surface area measure on $\partial B(x,r)$.

We will not prove this.

Theorem 36 (Maximum principle)**.** *Let $u : \mathbb{R}^d \to \mathbb{R}$ be harmonic on a domain $D \subseteq \mathbb{R}^d$. Then:*

(i) *If u attains its maximum in D, then u is constant on D;*

(ii) *If u is continuous on $\overline{D}$ and D is bounded, then*

$$\max_{x \in \overline{D}} u(x) = \max_{x \in \partial D} u(x).$$

Lecture 19: 22/01/2024

Proof. To show (i), let M be the maximum of u, and $V := \{x \in D, u(x) = M\}$. Then, V is closed in D and by Theorem 35 (ii) it is open. Since D is connected, $V = D$ and u is constant equal to M. To prove (ii), since u is continuous on $\overline{D}$ and D is bounded, u attains its maximum on $\overline{D}$. By (i), either the maximum is on ∂D, or u is constant on D. But if u is constant on D, by continuity it is constant on $\overline{D}$ and in particular attains its maximum on ∂D. □

Proof of Theorem 34.
Uniqueness of the solution
Suppose that there are two solutions u, u' continuous on $\overline{D}$. Then, $u - u'$ is harmonic on D and, by the maximum principle, we have

$$\max_{x \in \overline{D}} (u(x) - u'(x)) = \max_{x \in \partial D} (u(x) - u'(x)) = 0,$$

so $u \leq u'$ on $\overline{D}$. The same way, $u \geq u'$ and the result follows.

Existence and explicit form of the solution
Define $u : x \mapsto \mathbb{E}_x[\phi(B_T)]$. Since D is bounded and ϕ is continuous on ∂D, u is bounded on $\overline{D}$. Let us first show that u is harmonic, by showing that it satisfies Theorem 35 (ii). Let $x \in D$ and $\delta > 0$ such that $B(x,\delta) \subset D$. Let $T := \inf\{t \geq 0, B_t \notin B(x,\delta)\}$. We have that

$$\begin{aligned} u(x) &= \mathbb{E}_x[\phi(B_T)] = \mathbb{E}_x\left[\mathbb{E}_x[\phi(B_T)|\mathscr{F}_T]\right] \\ &= \mathbb{E}_x\left[\mathbb{E}_{B_T}[\phi(B_J)]\right] \text{ by strong Markov property} \\ &= \frac{1}{\sigma(\partial B(x,\delta))}\int_{\partial B(x,\delta)} u(y)d\sigma(y), \end{aligned}$$

where in the last equality we used the radial symmetry of Brownian motion (which implies that B_T is uniform on $\partial B(x,\delta)$).

Let us now show that u is continuous on $\overline{D}$. It is continuous on D, clearly, since it is harmonic. We use the Poincaré condition to prove that it is continuous at any $z \in \partial D$.

Let $z \in \partial D$ be fixed. For all $x \in D$, we have that

$$\begin{aligned} |u(x) - u(z)| &= \left|\mathbb{E}_x\left[\phi(B_J) - \phi(z)\right]\right| \\ &\leq \mathbb{E}_x[|\phi(B_J) - \phi(z)|], \end{aligned}$$

so we only need to prove that $\lim_{x\to z, x\in D}\mathbb{E}_x[|\phi(B_J) - \phi(z)|] = 0$. Let $\varepsilon > 0$. We want to find $r > 0$ such that if $x \in D$ and $|x-z| < r$, then $\mathbb{E}_x[|\phi(B_J) - \phi(z)|] < \varepsilon$.

Fix $r > 0$, and let $x \in D \cap B(z,r)$. We have:

$$\mathbb{E}_x[|\phi(B_J) - \phi(z)|] = \mathbb{E}_x[|\phi(B_J) - \phi(z)|\mathbb{1}_{J_{\partial B(z,r)} \geq J_D}] + \mathbb{E}_x[|\phi(B_J) - \phi(z)|\mathbb{1}_{J_{\partial B(z,r)} < J_D}],$$

where, for a set A, J_A is the first hitting time of A.

The first term can be easily bounded:

$$\mathbb{E}_x[|\phi(B_J) - \phi(z)|\mathbb{1}_{J_{\partial B(z,r)} \geq J_D}] \leq \sup_{|y-z|\leq r, y\in\partial D} |\phi(y) - \phi(z)|,$$

which goes to 0 as $r \to 0$ by continuity of ϕ.

The second term can be bounded brutally:

$$\mathbb{E}_x[|\phi(B_J) - \phi(z)|\mathbb{1}_{J_{\partial B(z,r)} < J_D}] \leq 2||\phi||_\infty \mathbb{P}_x(J_{\partial B(z,r)} < J_D),$$

and thus we only need to prove that $\mathbb{P}_x(J_{\partial B(z,r)} < J_D) \to 0$, uniformly in $x \in D \cap B(z,r)$.

Observe that, for any $t > 0$, we have

$$\mathbb{P}_x(J_{\partial B(z,r)} < J_D) \leq \mathbb{P}_x(J_{\partial B(z,r)} < t) + \mathbb{P}_x(J_D > t).$$

We bound the first term using the distribution of B_t, and the second term taking the limit as $x \to z$.

We have clearly that $\lim_{t\to 0}(J_{\partial B(z,r)} < t) = 0$ uniformly in, say, $z \in B(z, r/2)$. Hence, we only need to prove that, at t fixed, $\lim_{x\to z}\mathbb{P}_x(J_D > t) = 0$. This is clear by the cone property.

This ends the proof, and solves the Dirichlet problem. $\square$