Advanced Probability

Perla Sousi*

January 23, 2023

Contents

*University of Cambridge, Cambridge, UK; p.sousi@statslab.cam.ac.uk

1 Conditional expectation

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, i.e. Ω is a set, $\mathcal{F}$ is a σ-algebra on Ω and $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$.

Definition 1.1. $\mathcal{F}$ is a σ-algebra on Ω if it satisfies:

1. $\Omega \in \mathcal{F}$
2. If $A \in \mathcal{F}$, then also the complement is in $\mathcal{F}$, i.e.,$A^c \in \mathcal{F}$.
3. If $(A_n)_{n\geq 1}$ is a collection of sets in $\mathcal{F}$, then $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Definition 1.2. $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$ if it satisfies:

1. $\mathbb{P} : \mathcal{F} \to [0, 1]$, i.e. it is a set function
2. $\mathbb{P}(\Omega) = 1$ and $\mathbb{P}(\varnothing) = 0$
3. If $(A_n)_{n\geq 1}$ is a collection of pairwise disjoint sets in $\mathcal{F}$, then $\mathbb{P}(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$.

Let $A, B \in \mathcal{F}$ be two events with $\mathbb{P}(B) > 0$. Then the conditional probability of A given the event B is defined by

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Definition 1.3. The Borel σ-algebra, $\mathcal{B}(\mathbb{R})$, is the σ-algebra generated by the open sets in $\mathbb{R}$, i.e., it is the intersection of all σ-algebras containing the open sets of $\mathbb{R}$. More formally, let $\mathcal{O}$ be the open sets of $\mathbb{R}$, then

$$\mathcal{B}(\mathbb{R}) = \cap\{\mathcal{E} : \mathcal{E} \text{ is a } \sigma\text{-algebra containing } \mathcal{O}\}.$$

Informally speaking, consider the open sets of $\mathbb{R}$, do all possible operations, i.e., unions, intersections, complements, and take the smallest σ-algebra that you get.

Definition 1.4. X is a random variable, i.e., a measurable function with respect to $\mathcal{F}$, if $X : \Omega \to \mathbb{R}$ is a function with the property that for all open sets V the inverse image $X^{-1}(V) \in \mathcal{F}$.

Remark 1.5. If X is a random variable, then the collection of sets

$$\{B \subseteq \mathbb{R} : X^{-1}(B) \in \mathcal{F}\}$$

is a σ-algebra (check!) and hence it must contain $\mathcal{B}(\mathbb{R})$.

Definition 1.6. For a collection $\mathcal{A}$ of subsets of Ω we write $\sigma(\mathcal{A})$ for the smallest σ-algebra that contains $\mathcal{A}$, i.e.

$$\sigma(\mathcal{A}) = \cap\{\mathcal{E} : \mathcal{E} \quad \text{is a } \sigma\text{-algebra containing } \mathcal{A}\}.$$

Let $(X_i)_{i\in I}$ be a collection of random variables. Then we define

$$\sigma(X_i : i \in I) = \sigma\left(\{\omega \in \Omega : X_i(\omega) \in B\} : i \in I, B \in \mathcal{B}\right),$$

i.e. this is the smallest σ-algebra that makes $(X_i)_{i\in I}$ measurable.

Let $A \in \mathcal{F}$. The indicator function $\mathbf{1}(A)$ is defined via

$$\mathbf{1}(A)(x) = \mathbf{1}(x \in A) = \begin{cases} 1, & \text{if } x \in A; \\ 0, & \text{otherwise.} \end{cases}$$

Recall the definition of expectation. First for positive simple random variables, i.e., linear combinations of indicator random variables, we define

$$\mathbb{E}\left[\sum_{i=1}^{n} c_i \mathbf{1}(A_i)\right] := \sum_{i=1}^{n} c_i \mathbb{P}(A_i),$$

where c_i are positive constants and A_i are measurable events. Next, let X be a non-negative random variable. Then X is the increasing limit of positive simple variables. For example

$$X_n(\omega) = 2^{-n}\lfloor 2^n X(\omega) \rfloor \wedge n \uparrow X(\omega) \text{ as } n \to \infty.$$

So we define

$$\mathbb{E}[X] :=\uparrow \lim \mathbb{E}[X_n].$$

Finally, for a general random variable X, we can write $X = X^+ - X^-$, where $X^+ = \max(X, 0)$ and $X^- = \max(-X, 0)$ and we define

$$\mathbb{E}[X] := \mathbb{E}[X^+] - \mathbb{E}[X^-],$$

if at least one of $\mathbb{E}[X^+], \mathbb{E}[X^-]$ is finite. We call the random variable X *integrable*, if it satisfies $\mathbb{E}[|X|] < \infty$.

Let X be a random variable with $\mathbb{E}[|X|] < \infty$. Let A be an event in $\mathcal{F}$ with $\mathbb{P}(A) > 0$. Then the conditional expectation of X given A is defined by

$$\mathbb{E}[X|A] = \frac{\mathbb{E}[X\mathbf{1}(A)]}{\mathbb{P}(A)},$$

Our goal is to extend the definition of conditional expectation to σ-algebras. So far we only have defined it for events and it was a number. Now, the conditional expectation is going to be a random variable, measurable with respect to the σ-algebra with respect to which we are conditioning.

1.1 Discrete case

Let X be integrable, i.e., $\mathbb{E}[|X|] < \infty$. Let's start with a σ-algebra which is generated by a countable family of disjoint events $(B_i)_{i \in I}$ with $\cup_i B_i = \Omega$, i.e., $\mathcal{G} = \sigma(B_i, i \in I)$. It is easy to check that $\mathcal{G} = \{\cup_{i \in J} B_i : J \subseteq I\}$.

The natural thing to do is to define a new random variable $X' = \mathbb{E}[X|\mathcal{G}]$ as follows

$$X' = \sum_{i \in I} \mathbb{E}[X|B_i]\mathbf{1}(B_i).$$

What does this mean? Let $\omega \in \Omega$. Then $X'(\omega) = \sum_{i\in I} \mathbb{E}[X|B_i]\mathbf{1}(\omega \in B_i)$. Note that we use the convention that $\mathbb{E}[X|B_i] = 0$, if $\mathbb{P}(B_i) = 0$

It is very easy to check that

$$X' \text{ is } \mathcal{G} - \text{measurable} \tag{1.1}$$

and integrable, since

$$\mathbb{E}[|X'|] \le \sum_{i\in I} E|X\mathbf{1}(B_i)| = \mathbb{E}[|X|] < \infty.$$

Let $G \in \mathcal{G}$. Then it is straightforward to check that

$$\mathbb{E}[X\mathbf{1}(G)] = \mathbb{E}[X'\mathbf{1}(G)]. \tag{1.2}$$

1.2 Existence and uniqueness

Before stating the existence and uniqueness theorem on conditional expectation, let us quickly recall the notion of an event happening almost surely (a.s.), the Monotone convergence theorem and $\mathcal{L}^p$ spaces.

Let $A \in \mathcal{F}$. We will say that A happens a.s., if $\mathbb{P}(A) = 1$.

Theorem 1.7. [**Monotone convergence theorem**] *Let $(X_n)_n$ be random variables such that $X_n \ge 0$ for all n and $X_n \uparrow X$ as $n \to \infty$ a.s. Then*

$$\mathbb{E}[X_n] \uparrow \mathbb{E}[X] \text{ as } n \to \infty.$$

Theorem 1.8. [**Dominated convergence theorem**] *If $X_n \to X$ and $|X_n| \le Y$ for all n a.s., for some integrable random variable Y, then*

$$\mathbb{E}[X_n] \to \mathbb{E}[X].$$

Let $p \in [1, \infty)$ and f a measurable function in $(\Omega, \mathcal{F}, \mathbb{P})$. We define the norm

$$\|f\|_p = (\mathbb{E}[|f|^p])^{1/p}$$

and we denote by $\mathcal{L}^p = \mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P})$ the set of measurable functions f with $\|f\|_p < \infty$. For $p = \infty$, we let

$$\|f\|_\infty = \inf\{\lambda : |f| \le \lambda \text{ a.e.}\}$$

and $\mathcal{L}^\infty$ the set of measurable functions with $\|f\|_\infty < \infty$.

Formally, $\mathcal{L}^p$ is the collection of equivalence classes, where two functions are equivalent if they are equal almost everywhere (a.e.). In practice, we will represent an element of $\mathcal{L}^p$ by a function, but remember that equality in $\mathcal{L}^p$ means equality a.e..

Theorem 1.9. *The space $(\mathcal{L}^2, \|\cdot\|_2)$ is a Hilbert space with $\langle f, g\rangle = \mathbb{E}[fg]$. If $\mathcal{H}$ is a closed subspace, then for all $f \in \mathcal{L}^2$, there exists a unique (in the sense of a.e.) $g \in \mathcal{H}$ such that $\|f - g\|_2 = \inf_{h\in\mathcal{H}} \|f - h\|_2$ and $\langle f - g, h\rangle = 0$ for all $h \in \mathcal{H}$.*

Remark 1.10. We call g the orthogonal projection of f on $\mathcal{H}$.

Theorem 1.11. *Let X be an integrable random variable and let $\mathcal{G} \subseteq \mathcal{F}$ be a σ-algebra. Then there exists a random variable Y such that:*

(a) *Y is $\mathcal{G}$-measurable;*

(b) *Y is integrable and $\mathbb{E}[X\mathbf{1}(A)] = \mathbb{E}[Y\mathbf{1}(A)]$ for all $A \in \mathcal{G}$.*

Moreover, if Y' also satisfies (a) *and* (b)*, then $Y = Y'$ a.s..*

We call Y (*a version of*) *the conditional expectation of X given $\mathcal{G}$* and write $Y = \mathbb{E}[X|\mathcal{G}]$ a.s.. In the case $\mathcal{G} = \sigma(G)$ for some random variable G, we also write $Y = \mathbb{E}[X|G]$ a.s..

Remark 1.12. We could replace (b) in the statement of the theorem by requiring that for all bounded $\mathcal{G}$-measurable random variables Z we have

$$\mathbb{E}[XZ] = \mathbb{E}[YZ].$$

Remark 1.13. In Section 1.4 we will show how to construct explicit versions of the conditional expectation in certain simple cases. In general, we have to live with the indirect approach provided by the theorem.

Proof of Theorem 1.11. (*Uniqueness.*) Suppose that both Y and Y' satisfy (a) and (b). Then, clearly the event $A = \{Y > Y'\} \in \mathcal{G}$ and by (b) we have

$$\mathbb{E}[(Y - Y')\mathbf{1}(A)] = \mathbb{E}[X\mathbf{1}(A)] - \mathbb{E}[X\mathbf{1}(A)] = 0,$$

hence we get that $Y \leq Y'$ a.s. Similarly we can get $Y \geq Y'$ a.s.

(*Existence.*) We will prove existence in three steps.

1st step: Suppose that $X \in \mathcal{L}^2$. The space $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ with inner product defined by $\langle U, V\rangle = \mathbb{E}[UV]$ is a Hilbert space by Theorem 1.9 and $\mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ is a closed subspace. (*Remember that $\mathcal{L}^2$ convergence implies convergence in probability and convergence in probability implies convergence a.s. along a subsequence (see for instance [2, A13.2])*).

Thus $\mathcal{L}^2(\mathcal{F}) = \mathcal{L}^2(\mathcal{G}) + \mathcal{L}^2(\mathcal{G})^{\perp}$, and hence, we can write X as $X = Y + Z$, where $Y \in \mathcal{L}^2(\mathcal{G})$ and $Z \in \mathcal{L}^2(\mathcal{G})^{\perp}$. If we now set $Y = \mathbb{E}[X|\mathcal{G}]$, then (a) is clearly satisfied. Let $A \in \mathcal{G}$. Then

$$\mathbb{E}[X\mathbf{1}(A)] = \mathbb{E}[Y\mathbf{1}(A)] + \mathbb{E}[Z\mathbf{1}(A)] = \mathbb{E}[Y\mathbf{1}(A)],$$

since $\mathbb{E}[Z\mathbf{1}(A)] = 0$.

Note that from the above definition of conditional expectation for random variables in $\mathcal{L}^2$, we get that

$$\text{if } X \geq 0, \text{ then } Y = \mathbb{E}[X|\mathcal{G}] \geq 0 \text{ a.s.}, \tag{1.3}$$

since note that $\{Y < 0\} \in \mathcal{G}$ and

$$\mathbb{E}[X\mathbf{1}(Y < 0)] = \mathbb{E}[Y\mathbf{1}(Y < 0)].$$

Notice that the left hand side is nonnegative, while the right hand side is non-positive, implying that $\mathbb{P}(Y < 0) = 0$.

2nd step: Suppose that $X \geq 0$. For each n we define the random variables $X_n = X \wedge n \leq n$, and hence $X_n \in \mathcal{L}^2$. Thus from the first part of the existence proof we have that for each n there exists a $\mathcal{G}$-measurable random variable Y_n satisfying for all $A \in \mathcal{G}$

$$\mathbb{E}[Y_n \mathbf{1}(A)] = \mathbb{E}[(X \wedge n)\mathbf{1}(A)]. \tag{1.4}$$

Since the sequence $(X_n)_n$ is increasing, from (1.3) we get that almost surely $(Y_n)_n$ is increasing. If we now set $Y = \limsup_{n\to\infty} Y_n$, then clearly Y is $\mathcal{G}$-measurable and almost surely $Y = \uparrow \lim_{n\to\infty} Y_n$. By the monotone convergence theorem in (1.4) we get for all $A \in \mathcal{G}$

$$\mathbb{E}[Y\mathbf{1}(A)] = \mathbb{E}[X\mathbf{1}(A)], \tag{1.5}$$

since $X_n \uparrow X$, as $n \to \infty$.

In particular, if $\mathbb{E}[X]$ is finite, then $\mathbb{E}[Y]$ is also finite.

3rd step: Finally, for a general random variable $X \in \mathcal{L}^1$ (not necessarily positive) we can apply the above construction to $X^+ = \max(X, 0)$ and $X^- = \max(-X, 0)$ and then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X^+|\mathcal{G}] - \mathbb{E}[X^-|\mathcal{G}]$ satisfies (a) and (b). □

Remark 1.14. Note that the 2nd step of the above proof gives that if $X \geq 0$, then there exists a $\mathcal{G}$-measurable random variable Y such that

$$\text{for all } A \in \mathcal{G},\ \mathbb{E}[X\mathbf{1}(A)] = \mathbb{E}[Y\mathbf{1}(A)],$$

i.e., all the conditions of Theorem 1.11 are satisfied except for the integrability one.

Definition 1.15. Sub-σ-algebras $\mathcal{G}_1, \mathcal{G}_2, \ldots$ of $\mathcal{F}$ are called independent, if whenever $G_i \in \mathcal{G}_i$ ($i \in \mathbb{N}$) and $i_1, \ldots, i_n$ are distinct, then

$$\mathbb{P}(G_{i_1} \cap \ldots \cap G_{i_n}) = \prod_{k=1}^{n} \mathbb{P}(G_{i_k}).$$

When we say that a random variable X is independent of a σ-algebra $\mathcal{G}$, it means that $\sigma(X)$ is independent of $\mathcal{G}$.

The following properties are immediate consequences of Theorem 1.11 and its proof.

Proposition 1.16. *Let $X, Y \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{G} \subset \mathcal{F}$ be a σ-algebra. Then*

1. $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$
2. *If X is $\mathcal{G}$-measurable, then $\mathbb{E}[X|\mathcal{G}] = X$ a.s..*
3. *If X is independent of $\mathcal{G}$, then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$ a.s..*
4. *If $X \geq 0$ a.s., then $\mathbb{E}[X|\mathcal{G}] \geq 0$ a.s..*

5. *For any* $\alpha, \beta \in \mathbb{R}$ *we have* $\mathbb{E}[\alpha X + \beta Y|\mathcal{G}] = \alpha\mathbb{E}[X|\mathcal{G}] + \beta\mathbb{E}[Y|\mathcal{G}]$ *a.s..*

6. $|\mathbb{E}[X|\mathcal{G}]| \leq \mathbb{E}[|X||\mathcal{G}]$ *a.s..*

The basic convergence theorems for expectation have counterparts for conditional expectation. We first recall the theorems for expectation.

Theorem 1.17. [**Fatou's lemma**] *If* $X_n \geq 0$ *for all* n*, then*

$$\mathbb{E}[\liminf_n X_n] \leq \liminf_n \mathbb{E}[X_n].$$

Theorem 1.18. [**Jensen's inequality**] *Let* X *be an integrable random variable and let* $\varphi : \mathbb{R} \to \mathbb{R}$ *be a convex function. Then*

$$\mathbb{E}[\varphi(X)] \geq \varphi(\mathbb{E}[X]).$$

Proposition 1.19. *Let* $\mathcal{G} \subset \mathcal{F}$ *be a* σ*-algebra.*

1. ***Conditional monotone convergence theorem:*** *If* $(X_n)_{n\geq 0}$ *is an increasing sequence of non-negative random variables with a.s. limit* X*, then*

$$\mathbb{E}[X_n|\mathcal{G}] \nearrow \mathbb{E}[X|\mathcal{G}] \text{ as } n \to \infty, \text{ a.s..}$$

2. ***Conditional Fatou's lemma:*** *If* $X_n \geq 0$ *for all* n*, then*

$$\mathbb{E}\left[\liminf_{n\to\infty} X_n|\mathcal{G}\right] \leq \liminf_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}] \text{ a.s..}$$

3. ***Conditional dominated convergence theorem:*** *If* $X_n \to X$ *and* $|X_n| \leq Y$ *for all* n *a.s., for some integrable random variable* Y*, then*

$$\lim_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}] \text{ a.s..}$$

4. ***Conditional Jensen's inequality:*** *If* X *is an integrable random variable and* $\varphi : \mathbb{R} \to (-\infty, \infty]$ *is a convex function such that either* $\varphi(X)$ *is integrable or* φ *is non-negative, then*

$$\mathbb{E}[\varphi(X)|\mathcal{G}] \geq \varphi(\mathbb{E}[X|\mathcal{G}]) \text{ a.s..}$$

In particular, for all $1 \leq p < \infty$

$$\|\mathbb{E}[X|\mathcal{G}]\|_p \leq \|X\|_p.$$

Proof. 1. Let Y_n be a version of $\mathbb{E}[X_n|\mathcal{G}]$. Since $0 \leq X_n \nearrow X$ as $n \to \infty$, we have that almost surely Y_n is an increasing sequence and $Y_n \geq 0$. Let $Y = \limsup_{n\to\infty} Y_n$. We want to show that $Y = \mathbb{E}[X|\mathcal{G}]$ a.s.. Clearly Y is $\mathcal{G}$-measurable, as the limsup of $\mathcal{G}$-measurable random variables. Also, by the monotone convergence theorem we have for all $A \in \mathcal{G}$

$$\mathbb{E}[X\mathbf{1}(A)] = \lim_{n\to\infty} \mathbb{E}[X_n\mathbf{1}(A)] = \lim_{n\to\infty} \mathbb{E}[Y_n\mathbf{1}(A)] = \mathbb{E}[Y\mathbf{1}(A)].$$

2. The sequence $\inf_{k\geq n} X_k$ is increasing in n and $\lim_{n\to\infty} \inf_{k\geq n} X_k = \liminf_{n\to\infty} X_n$. Thus, by the conditional monotone convergence theorem we get

$$\lim_{n\to\infty} \mathbb{E}[\inf_{k\geq n} X_k|\mathcal{G}] = \mathbb{E}[\liminf_{n\to\infty} X_n|\mathcal{G}].$$

Clearly, $\mathbb{E}[\inf_{k\geq n} X_k|\mathcal{G}] \leq \inf_{k\geq n} \mathbb{E}[X_k|\mathcal{G}]$. Passing to the limit gives the desired inequality.

3. Since $X_n + Y$ and $Y - X_n$ are positive random variables for all n, applying conditional Fatou's lemma we get

$$\mathbb{E}[X+Y|\mathcal{G}] = \mathbb{E}[\liminf(X_n+Y)|\mathcal{G}] \leq \liminf_{n\to\infty} \mathbb{E}[X_n+Y|\mathcal{G}] \text{ and}$$
$$\mathbb{E}[Y-X|\mathcal{G}] = \mathbb{E}[\liminf(Y-X_n)|\mathcal{G}] \leq \liminf_{n\to\infty} \mathbb{E}[Y-X_n|\mathcal{G}].$$

Hence, we obtain that

$$\liminf_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}] \geq \mathbb{E}[X|\mathcal{G}] \text{ and } \limsup_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}] \leq \mathbb{E}[X|\mathcal{G}].$$

4. A convex function is the supremum of countably many affine functions: (see for instance [2, §6.6])

$$\varphi(x) = \sup_i (a_i x + b_i), x \in \mathbb{R}.$$

So for all i we have $\mathbb{E}[\varphi(X)|\mathcal{G}] \geq a_i \mathbb{E}[X|\mathcal{G}] + b_i$ a.s. Now using the fact that the supremum is over a countable set we get that

$$\mathbb{E}[\varphi(X)|\mathcal{G}] \geq \sup_i (a_i \mathbb{E}[X|\mathcal{G}] + b_i) = \varphi(\mathbb{E}[X|\mathcal{G}]) \text{ a.s.}$$

In particular, for $1 \leq p < \infty$,

$$\|\mathbb{E}[X|\mathcal{G}]\|_p^p = \mathbb{E}[|\mathbb{E}[X|\mathcal{G}]|^p] \leq \mathbb{E}[\mathbb{E}[|X|^p|\mathcal{G}]] = \mathbb{E}[|X|^p] = \|X\|_p^p.$$

□

Conditional expectation has the *tower property*:

Proposition 1.20. *Let $\mathcal{H} \subset \mathcal{G}$ be σ-algebras and $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$. Then*

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}] \text{ a.s..}$$

Proof. Clearly, $\mathbb{E}[X|\mathcal{H}]$ is $\mathcal{H}$-measurable and for all $A \in \mathcal{H}$ we have

$$\mathbb{E}[\mathbb{E}[X|\mathcal{H}]\mathbf{1}(A)] = \mathbb{E}[X\mathbf{1}(A)] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbf{1}(A)],$$

since A is also $\mathcal{G}$-measurable. □

We can always *take out what is known*:

Proposition 1.21. *Let $X \in \mathcal{L}^1$ and $\mathcal{G}$ a σ-algebra. If Y is bounded and $\mathcal{G}$-measurable, then*

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

Proof. Since Y is $\mathcal{G}$ measurable, clearly $Y\mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}$ measurable. Let $A \in \mathcal{G}$. Then by Remark 1.12 we have

$$\mathbb{E}[Y\mathbb{E}[X|\mathcal{G}]\mathbf{1}(A)] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}](\mathbf{1}(A)Y)] = \mathbb{E}[YX\mathbf{1}(A)],$$

which implies that $\mathbb{E}[YX|\mathcal{G}] = Y\mathbb{E}[X|\mathcal{G}]$ a.s.. □

Before stating the next proposition, we quickly recall the definition of a π-system and the uniqueness of extension theorem for probability measures agreeing on a π-system generating a σ-algebra.

Definition 1.22. Let $\mathcal{A}$ be a set of subsets of Ω. We call $\mathcal{A}$ a π-system if for all $A, B \in \mathcal{A}$, the intersection $A \cap B \in \mathcal{A}$.

Theorem 1.23. [**Uniqueness of extension**] *Let μ_1, μ_2 be two measures on $(E, \mathcal{E})$, where $\mathcal{E}$ is a σ-algebra on E . Suppose that $\mu_1 = \mu_2$ on a π-system $\mathcal{A}$ generating $\mathcal{E}$ and that $\mu_1(E) = \mu_2(E) < \infty$. Then $\mu_1 = \mu_2$ on $\mathcal{E}$.*

Proposition 1.24. *Let X be integrable and $\mathcal{G}, \mathcal{H} \subset \mathcal{F}$ be σ-algebras. If $\sigma(X, \mathcal{G})$ is independent of $\mathcal{H}$, then*

$$\mathbb{E}[X|\sigma(\mathcal{G}, \mathcal{H})] = \mathbb{E}[X|\mathcal{G}] \text{ a.s..}$$

Proof. We can assume that $X \geq 0$. The general case will follow by writing $X = X^+ - X^-$. Let $A \in \mathcal{G}$ and $B \in \mathcal{H}$. Then

$$\begin{aligned}\mathbb{E}[\mathbf{1}(A \cap B)\mathbb{E}[X|\sigma(\mathcal{H}, \mathcal{G})]] &= \mathbb{E}[\mathbf{1}(A \cap B)X] = \mathbb{E}[X\mathbf{1}(A)]\mathbb{P}(B) \\ &= \mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbf{1}(A)]\mathbb{P}(B) = \mathbb{E}[\mathbf{1}(A \cap B)\mathbb{E}[X|\mathcal{G}]],\end{aligned}$$

where we used the independence assumption in the second and last equality. Let $Y = \mathbb{E}[X|\sigma(\mathcal{H}, \mathcal{G})]$ a.s., then $Y \geq 0$ a.s.. We can now define the measures

$$\mu(F) = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbf{1}(F)] \text{ and } \nu(F) = \mathbb{E}[Y\mathbf{1}(F)], \text{ for all } F \in \mathcal{F}.$$

Then we have that μ and ν agree on the π-system $\{A \cap B : A \in \mathcal{G}, B \in \mathcal{H}\}$ which generates $\sigma(\mathcal{G}, \mathcal{H})$. Also, by the integrability assumption, $\mu(\Omega) = \nu(\Omega) < \infty$. Hence, they agree everywhere on $\sigma(\mathcal{G}, \mathcal{H})$ and this finishes the proof. □

Warning! If in the above proposition the independence assumption is weakened and we just assume that $\sigma(X)$ is independent of $\mathcal{H}$ and $\mathcal{G}$ is independent of $\mathcal{H}$, then the conclusion does not follow. See example sheet!

1.3 Product measure and Fubini's theorem

A measure space $(E, \mathcal{E}, \mu)$ is called *σ-finite*, if there exists a collection of sets $(S_n)_{n\geq 0}$ in $\mathcal{E}$ such that $\cup_n S_n = E$ and $\mu(S_n) < \infty$ for all n.

Let $(E_1, \mathcal{E}_1, \mu_1)$ and $(E_2, \mathcal{E}_2, \mu_2)$ be two σ-finite measure spaces. The set

$$\mathcal{A} = \{A_1 \times A_2 : A_1 \in \mathcal{E}_1, A_2 \in \mathcal{E}_2\}$$

is a π-system of subsets of $E = E_1 \times E_2$. Define the product σ-algebra

$$\mathcal{E}_1 \otimes \mathcal{E}_2 = \sigma(\mathcal{A}).$$

Set $\mathcal{E} = \mathcal{E}_1 \otimes \mathcal{E}_2$.

Theorem 1.25. [**Product measure**] *Let $(E_1, \mathcal{E}_1, \mu_1)$ and $(E_2, \mathcal{E}_2, \mu_2)$ be two σ-finite measure spaces. There exists a unique measure $\mu = \mu_1 \otimes \mu_2$ on $\mathcal{E}$ such that*

$$\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$$

for all $A_1 \in \mathcal{E}_1$ and $A_2 \in \mathcal{E}_2$.

Theorem 1.26. [**Fubini's theorem**] *Let $(E_1, \mathcal{E}_1, \mu_1)$ and $(E_2, \mathcal{E}_2, \mu_2)$ be two σ-finite measure spaces.*

Let f be $\mathcal{E}$-measurable and non-negative. Then

$$\mu(f) = \int_{E_1} \left(\int_{E_2} f(x_1, x_2)\, \mu_2(dx_2) \right) \mu_1(dx_1). \tag{1.6}$$

If f is integrable, then

1. *$x_2 \to f(x_1, x_2)$ is μ_2-integrable for μ_1-almost all x_1,*
2. *$x_1 \to \int_{E_2} f(x_1, x_2)\, \mu_2(dx_2)$ is μ_1-integrable and formula (1.6) for $\mu(f)$ holds.*

1.4 Examples of conditional expectation

Definition 1.27. A random vector $(X_1, \ldots, X_n) \in \mathbb{R}^n$ is called a *Gaussian random vector* iff for all $a_1, \ldots, a_n \in \mathbb{R}$ the random variable $\sum_{i=1}^n a_i X_i$ has a Gaussian distribution.

A real-valued process $(X_t, t \geq 0)$ is called a *Gaussian process* iff for every $t_1 < t_2 < \ldots < t_n$ the random vector $(X_{t_1}, \ldots, X_{t_n})$ is a Gaussian random vector.

1.4.1 Gaussian case

Let (X, Y) be a Gaussian random vector in $\mathbb{R}^2$. Set $\mathcal{G} = \sigma(Y)$. In this example, we are going to compute $X' = \mathbb{E}[X|\mathcal{G}]$.

Since X' must be $\mathcal{G}$-measurable and $\mathcal{G} = \sigma(Y)$, by [2, A3.2.] we have that $X' = f(Y)$, for some Borel function f. Let us try X' of the form $X' = aY + b$, for $a, b \in \mathbb{R}$ that we will determine.

Since $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$, we must have that

$$a\mathbb{E}[Y] + b = \mathbb{E}[X]. \tag{1.7}$$

Also, we must have that

$$\mathbb{E}[(X - X')Y] = 0 \Rightarrow \operatorname{Cov}(X - X', Y) = 0 \Rightarrow \operatorname{Cov}(X, Y) = a \operatorname{var}(Y). \tag{1.8}$$

So, if a satisfies (1.8), then

$$\operatorname{Cov}(X - X', Y) = 0$$

and since $(X - X', Y)$ is Gaussian, we get that $X - X'$ and Y are independent. Hence, if Z is $\sigma(Y)$-measurable, then using also (1.7) we get that

$$\mathbb{E}[(X - X')Z] = 0.$$

Therefore we proved that $\mathbb{E}[X|\mathcal{G}] = aY + b$, for a, b satisfying (1.7) and (1.8).

1.4.2 Conditional density functions

Suppose that X and Y are random variables having a joint density function $f_{X,Y}(x, y)$ in $\mathbb{R}^2$. Let $h : \mathbb{R} \to \mathbb{R}$ be a Borel function such that $h(X)$ is integrable.

In this example we want to compute $\mathbb{E}[h(X)|Y] = \mathbb{E}[h(X)|\sigma(Y)]$.

The random variable Y has a density function f_Y, given by

$$f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x, y)\, dx.$$

Let g be bounded and measurable. Then we have that

$$\mathbb{E}[h(X)g(Y)] = \int_{\mathbb{R}} \int_{\mathbb{R}} h(x)g(y)f_{X,Y}(x, y)\, dx\, dy = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} h(x) \frac{f_{X,Y}(x, y)}{f_Y(y)}\, dx \right) g(y)f_Y(y)\, dy,$$

where we agree say that $0/0 = 0$. If we now set

$$\varphi(y) = \int_{\mathbb{R}} h(x) \frac{f_{X,Y}(x, y)}{f_Y(y)}\, dx, \text{ if } f_Y(y) > 0$$

and 0 otherwise, then we get that

$$\mathbb{E}[h(X)|Y] = \varphi(Y) \text{ a.s.}$$

We interpret this result by saying that

$$\mathbb{E}[h(X)|Y] = \int_{\mathbb{R}} h(x)\nu(Y, dx),$$

where $\nu(y, dx) = f_Y(y)^{-1} f_{X,Y}(x, y)\mathbf{1}(f_Y(y) > 0)dx = f_{X|Y}(x|y)dx$. The measure $\nu(y, dx)$ is called the *conditional distribution* of X given $Y = y$, and $f_{X|Y}(x|y)$ is the *conditional density function* of X given $Y = y$. Notice this function of x, y is defined only up to a zero-measure set.

2 Discrete-time martingales

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(E, \mathcal{E})$ be a measurable space. (We will mostly consider $E = \mathbb{R}, \mathbb{R}^d, \mathbb{C}$. Unless otherwise indicated, it is to be understood from now on that $E = \mathbb{R}$.)

Let $X = (X_n)_{n \geq 0}$ be a sequence of random variables taking values in E. We call X a *stochastic process* in E.

A *filtration* $(\mathcal{F}_n)_n$ is an increasing family of sub-σ-algebras of $\mathcal{F}$, i.e., $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$, for all n. We can think of $\mathcal{F}_n$ as the information available to us at time n. Every process has a natural filtration $(\mathcal{F}_n^X)_n$, given by

$$\mathcal{F}_n^X = \sigma(X_k, k \leq n).$$

The process X is called *adapted* to the filtration $(\mathcal{F}_n)_n$, if X_n is $\mathcal{F}_n$-measurable for all n. Of course, every process is adapted to its natural filtration. We say that X is *integrable* if X_n is integrable for all n.

Definition 2.1. Let $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \geq 0}, \mathbb{P})$ be a filtered probability space. Let $X = (X_n)_{n \geq 0}$ be an adapted integrable process taking values in $\mathbb{R}$.

- X is a **martingale** if $\mathbb{E}[X_n | \mathcal{F}_m] = X_m$ a.s., for all $n \geq m$.
- X is a **supermartingale** if $\mathbb{E}[X_n | \mathcal{F}_m] \leq X_m$ a.s., for all $n \geq m$.
- X is a **submartingale** if $\mathbb{E}[X_n | \mathcal{F}_m] \geq X_m$ a.s., for all $n \geq m$.

Note that every process which is a martingale (resp. super, sub) with respect to the given filtration is also a martingale (resp. super, sub) with respect to its natural filtration by the tower property of conditional expectation.

Example 2.2. Let $(\xi_i)_{i \geq 1}$ be a sequence of i.i.d. random variables with $\mathbb{E}[\xi_1] = 0$. Then it is easy to check that $X_n = \sum_{i=1}^n \xi_i$ is a martingale.

Example 2.3. Let $(\xi_i)_{i \geq 1}$ be a sequence of i.i.d. random variables with $\mathbb{E}[\xi_1] = 1$. Then the product $X_n = \prod_{i=1}^n \xi_i$ is a martingale.

2.1 Stopping times

Definition 2.4. Let $(\Omega, \mathcal{F}, (\mathcal{F}_n)_n, \mathbb{P})$ be a filtered probability space. A **stopping time** T is a random variable $T : \Omega \to \mathbb{Z}_+ \cup \{\infty\}$ such that $\{T \leq n\} \in \mathcal{F}_n$, for all n.

Equivalently, T is a stopping time if $\{T = n\} \in \mathcal{F}_n$, for all n. Indeed,

$$\{T = n\} = \{T \leq n\} \setminus \{T \leq n-1\} \in \mathcal{F}_n.$$

Conversely,

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

Example 2.5.

- Constant times are trivial stopping times.
- Let $(X_n)_{n\geq 0}$ be an adapted process taking values in $\mathbb{R}$. Let $A \in \mathcal{B}(\mathbb{R})$. The first entrance time to A is
$$T_A = \inf\{n \geq 0 : X_n \in A\}$$
with the convention that $\inf(\varnothing) = \infty$, so that $T_A = \infty$, if X never enters A. This is a stopping time, since
$$\{T_A \leq n\} = \cup_{k\leq n}\{X_k \in A\} \in \mathcal{F}_n.$$
- The last exit time though, $T_A = \sup\{n \leq 10 : X_n \in A\}$, is **not** always a stopping time.

As an immediate consequence of the definition, one gets:

Proposition 2.6. *Let $S, T, (T_n)_n$ be stopping times on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n), \mathbb{P})$. Then also $S \wedge T, S \vee T, \inf_n T_n, \sup_n T_n, \liminf_n T_n, \limsup_n T_n$ are stopping times.*

Proof. Note that in discrete time everything follows straight from the definitions. But when one considers continuous time processes, then right continuity of the filtration is needed to ensure that the limits are indeed stopping times. □

Definition 2.7. Let T be a stopping time on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n), \mathbb{P})$. Define the σ-algebra $\mathcal{F}_T$ via

$$\mathcal{F}_T = \{A \in \mathcal{F} : A \cap \{T \leq t\} \in \mathcal{F}_t, \text{ for all } t\}.$$

Intuitively $\mathcal{F}_T$ is the information available at time T.

It is easy to check that if $T = t$, then T is a stopping time and $\mathcal{F}_T = \mathcal{F}_t$.

For a process X, we set $X_T(\omega) = X_{T(\omega)}(\omega)$, whenever $T(\omega) < \infty$. We also define the *stopped process* X^T by $X^T_t = X_{T\wedge t}$.

Proposition 2.8. *Let S and T be stopping times and let $X = (X_n)_{n\geq 0}$ be an adapted process. Then*

1. *if $S \leq T$, then $\mathcal{F}_S \subseteq \mathcal{F}_T$,*
2. *$X_T \mathbf{1}(T < \infty)$ is an $\mathcal{F}_T$-measurable random variable,*
3. *X^T is adapted,*
4. *if X is integrable, then X^T is integrable.*

Proof.

1. Straightforward from the definition.
2. Let $A \in \mathcal{E}$. Then
$$\{X_T \mathbf{1}(T < \infty) \in A\} \cap \{T \leq t\} = \bigcup_{s=1}^{t} \{X_s \in A\} \cap \{T = s\} \in \mathcal{F}_t,$$
since X is adapted and $\{T = s\} = \{T \leq s\} \setminus \{T \leq s-1\} \in \mathcal{F}_s$.

3. For every t we have that $X_{T\wedge t}$ is $\mathcal{F}_{T\wedge t}$-measurable, hence by (1) $\mathcal{F}_t$-measurable since $T \wedge t \leq t$.

4. We have

$$\mathbb{E}[|X_{T\wedge t}|] = \mathbb{E}\left[\sum_{s=0}^{t-1} |X_s| \mathbf{1}(T=s)\right] + \mathbb{E}\left[\sum_{s=t}^{\infty} |X_t| \mathbf{1}(T=s)\right] \leq \sum_{s=0}^{t} \mathbb{E}[|X_t|] < \infty.$$

$\square$

2.2 Optional stopping

Theorem 2.9. [**Optional stopping**] *Let $X = (X_n)_{n\geq 0}$ be a martingale.*

1. *If T is a stopping time, then X^T is also a martingale, so in particular $\mathbb{E}[X_{T\wedge t}] = \mathbb{E}[X_0]$, for all t.*

2. *If $S \leq T$ are bounded stopping times, then $\mathbb{E}[X_T|\mathcal{F}_S] = X_S$ a.s..*

3. *If $S \leq T$ are bounded stopping times, then $\mathbb{E}[X_T] = \mathbb{E}[X_S]$.*

4. *If there exists an integrable random variable Y such that $|X_n| \leq Y$ for all n, and T is a stopping time which is finite a.s., then*

$$\mathbb{E}[X_T] = \mathbb{E}[X_0].$$

5. *If X has bounded increments, i.e., $\exists M > 0 : \forall n \geq 0, |X_{n+1} - X_n| \leq M$ a.s., and T is a stopping time with $\mathbb{E}[T] < \infty$, then*

$$\mathbb{E}[X_T] = \mathbb{E}[X_0].$$

Proof. 1. Notice that by the tower property of conditional expectation, it suffices to check that $\mathbb{E}[X_{T\wedge t}|\mathcal{F}_{t-1}] = X_{T\wedge(t-1)}$ a.s.. We can write

$$\begin{aligned}\mathbb{E}[X_{T\wedge t}|\mathcal{F}_{t-1}] &= \mathbb{E}\left[\sum_{s=0}^{t-1} X_s \mathbf{1}(T=s)|\mathcal{F}_{t-1}\right] + \mathbb{E}\left[X_t \mathbf{1}(T > t-1)|\mathcal{F}_{t-1}\right] \\ &= X_T \mathbf{1}(T \leq t-1) + \mathbf{1}(T > t-1) X_{t-1},\end{aligned}$$

since $\{T > t-1\} \in \mathcal{F}_{t-1}$ and $\mathbb{E}[X_t|\mathcal{F}_{t-1}] = X_{t-1}$ a.s. by the martingale property.

2. Suppose that $T \leq n$ a.s.. Since $S \leq T$, we can write

$$\begin{aligned}X_T &= (X_T - X_{T-1}) + (X_{T-1} - X_{T-2}) + \ldots + (X_{S+1} - X_S) + X_S \\ &= X_S + \sum_{k=0}^{n} (X_{k+1} - X_k) \mathbf{1}(S \leq k < T).\end{aligned}$$

Let $A \in \mathcal{F}_S$. Then

$$\mathbb{E}[X_T \mathbf{1}(A)] = \mathbb{E}[X_S \mathbf{1}(A)] + \sum_{k=0}^{n} \mathbb{E}[(X_{k+1} - X_k)\mathbf{1}(S \le k < T)\mathbf{1}(A)] = \mathbb{E}[X_S \mathbf{1}(A)],$$

since $\{S \le k < T\} \cap A \in \mathcal{F}_k$, for all k and X is a martingale.

3. Taking expectations in 2 gives the equality in expectation.

4. See example sheet.

5. See example sheet.

□

Remark 2.10. Note that Theorem 2.9 is true if X is a super-martingale or sub-martingale with the respective inequalities in the statements.

Remark 2.11. Let $(\xi_k)_k$ be i.i.d. random variables taking values ± 1 with probability $1/2$. Then $X_n = \sum_{k=0}^n \xi_k$ is a martingale. Let $T = \inf\{n \ge 0 : X_n = 1\}$. Then T is a stopping time and $\mathbb{P}(T < \infty) = 1$. However, although from Theorem 2.9 we have that $\mathbb{E}[X_{T\wedge t}] = \mathbb{E}[X_0]$, for all t, it holds that $1 = \mathbb{E}[X_T] \neq \mathbb{E}[X_0] = 0$.

For non-negative supermartingales, Fatou's lemma gives:

Proposition 2.12. *Suppose that X is a non-negative supermartingale. Then for any stopping time T which is finite a.s. we have*

$$\mathbb{E}[X_T] \le \mathbb{E}[X_0].$$

2.3 Gambler's ruin

Let $(\xi_i)_{i\ge 1}$ be an i.i.d. sequence of random variables taking values ± 1 with probabilities $\mathbb{P}(\xi_1 = +1) = \mathbb{P}(\xi_1 = -1) = 1/2$. Define $X_n = \sum_{i=1}^n \xi_i$, for $n \ge 1$, and $X_0 = 0$. This is called the simple symmetric random walk in $\mathbb{Z}$. For $c \in \mathbb{Z}$ we write

$$T_c = \inf\{n \ge 0 : X_n = c\},$$

i.e. T_c is the first hitting time of the state c, and hence is a stopping time. Let $a, b > 0$. We will calculate the probability that the random walk hits $-a$ before b, i.e. $\mathbb{P}(T_{-a} < T_b)$.

As mentioned earlier in this section X is a martingale. Also, $|X_{n+1} - X_n| \le 1$ for all n. We now write $T = T_{-a} \wedge T_b$. We will first show that $\mathbb{E}[T] < \infty$.

It is easy to see that T is bounded from above by the first time that there are $a+b$ consecutive $+1$'s. The probability that the first $\xi_1, \ldots, \xi_{a+b}$ are all equal to $+1$ is $2^{-(a+b)}$. If the first block of $a+b$ variables ξ fail to be all $+1$'s, then we look at the next block of $a+b$, i.e. $\xi_{a+b+1}, \ldots, \xi_{2(a+b)}$. The probability that this block consists only of $+1$'s is again $2^{-(a+b)}$ and this event is independent of the previous one. Hence T can be bounded from above by a geometric random variable of success probability $2^{-(a+b)}$ times $a+b$. Therefore we get

$$\mathbb{E}[T] \le (a+b)2^{a+b}.$$

We thus have a martingale with bounded increments and a stopping time with finite expectation. Hence, from the optional stopping theorem (5), we deduce that

$$\mathbb{E}[X_T] = \mathbb{E}[X_0] = 0.$$

We also have

$$\mathbb{E}[X_T] = -a\mathbb{P}(T_{-a} < T_b) + b\mathbb{P}(T_b < T_{-a}) \quad \text{and} \quad \mathbb{P}(T_{-a} < T_b) + \mathbb{P}(T_b < T_{-a}) = 1,$$

and hence we deduce that

$$\mathbb{P}(T_{-a} < T_b) = \frac{b}{a+b}.$$

2.4 Martingale convergence theorem

Theorem 2.13. [A.s. martingale convergence theorem] *Let $X = (X_n)_n$ be a supermartingale which is bounded in L^1, i.e., $\sup_n \mathbb{E}[|X_n|] < \infty$. Then $X_n \to X_\infty$ a.s. as $n \to \infty$, for some $X_\infty \in L^1(\mathcal{F}_\infty)$, where $\mathcal{F}_\infty = \sigma(\mathcal{F}_n, n \ge 0)$.*

Usually when we want to prove convergence of a sequence, we have an idea of what the limit should be. In the case of the martingale convergence theorem though, we do not know the limit. And, indeed in most cases, we just know the existence of the limit. In order to show the convergence in the theorem, we will employ a beautiful trick due to Doob, which counts the number of upcrossings of every interval with rational endpoints.

Corollary 2.14. *Let $X = (X_n)_n$ be a non-negative supermartingale. Then X converges a.s. towards an a.s. finite limit.*

Proof. Since X is non-negative we get that

$$\mathbb{E}[|X_n|] = \mathbb{E}[X_n] \le \mathbb{E}[X_0] < \infty,$$

hence X is bounded in L^1. □

Let $x = (x_n)_n$ be a sequence of real numbers. Let $a < b$ be two real numbers. We define $T_0(x) = 0$ and inductively for $k \ge 0$

$$S_{k+1}(x) = \inf\{n \ge T_k(x) : x_n \le a\} \quad \text{and} \quad T_{k+1}(x) = \inf\{n \ge S_{k+1}(x) : x_n \ge b\} \tag{2.1}$$

with the usual convention that $\inf \varnothing = \infty$.

We also define $N_n([a,b],x) = \sup\{k \ge 0 : T_k(x) \le n\}$, i.e., the number of upcrossings of the interval $[a,b]$ by the sequence x by time n. As $n \to \infty$ we have

$$N_n([a,b],x) \uparrow N([a,b],x) = \sup\{k \ge 0 : T_k(x) < \infty\},$$

i.e., the total number of upcrossings of the interval $[a,b]$.

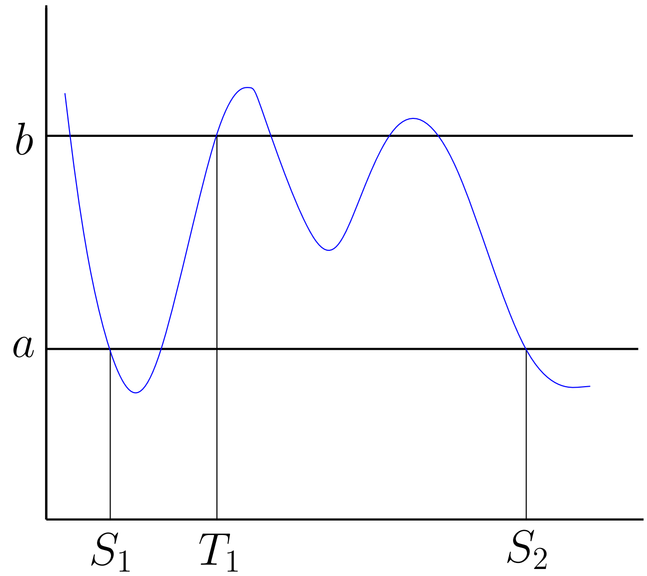


Figure 1. Upcrossings.

Before stating and proving Doob's upcrossing inequality, we give an easy lemma that will be used in the proof of Theorem 2.13.

Lemma 2.15. *A sequence of real numbers $x = (x_n)_n$ converges in $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ if and only if $N([a,b],x) < \infty$ for all rationals $a < b$.*

Proof. Suppose that x converges. Then if for some $a < b$ we had that $N([a,b],x) = \infty$, that would imply that $\liminf_n x_n \le a < b \le \limsup_n x_n$, which is a contradiction.

Next suppose that x does not converge. Then $\liminf_n x_n < \limsup_n x_n$ and so taking $a < b$ rationals between these two numbers gives that $N([a,b],x) = \infty$. □

Theorem 2.16. [Doob's upcrossing inequality] *Let X be a supermartingale and $a < b$ be two real numbers. Then for all $n \ge 0$*

$$(b-a)\mathbb{E}[N_n([a,b],X)] \le \mathbb{E}[(X_n - a)^-].$$

Proof. We will omit the dependence on X from T_k and S_k and we will write $N = N_n([a,b],X)$ to simplify notation. By the definition of the times (T_k) and (S_k), it is clear that for all k

$$X_{T_k} - X_{S_k} \ge b - a. \tag{2.2}$$

We have

$$\sum_{k=1}^{n}(X_{T_k \wedge n} - X_{S_k \wedge n}) = \sum_{k=1}^{N}(X_{T_k} - X_{S_k}) + \sum_{k=N+1}^{n}(X_n - X_{S_k \wedge n})\mathbf{1}(N < n) \tag{2.3}$$

$$= \sum_{k=1}^{N}(X_{T_k} - X_{S_k}) + (X_n - X_{S_{N+1}})\mathbf{1}(S_{N+1} \le n), \tag{2.4}$$

since the only term contributing in the second sum appearing on the right hand side of (2.3) is $N+1$, by the definition of N. Indeed, if $S_{N+2} \le n$, then that would imply that $T_{N+1} \le n$, which would contradict the definition of N.

Using induction on k, it is easy to see that $(T_k)_k$ and $(S_k)_k$ are sequences of stopping times. Hence for all n, we have that $S_k \wedge n \le T_k \wedge n$ are bounded stopping times and thus by the Optional stopping theorem, Theorem 2.9 we get that $\mathbb{E}[X_{S_k \wedge n}] \ge \mathbb{E}[X_{T_k \wedge n}]$, for all k.

Therefore, taking expectations in (2.3) and (2.4) and using (2.2) we get

$$0 \ge \mathbb{E}\left[\sum_{k=1}^{n}(X_{T_k\wedge n} - X_{S_k\wedge n})\right] \ge (b-a)\mathbb{E}[N] - \mathbb{E}[(X_n - a)^-],$$

since $(X_n - X_{S_{N+1}})\mathbf{1}(S_{N+1} \le n) \ge -(X_n - a)^-$. Rearranging gives the desired inequality. □

Proof of Theorem 2.13. Let $a < b \in \mathbb{Q}$. By Doob's upcrossing inequality, Theorem 2.16 we get that

$$\mathbb{E}[N_n([a,b],X)] \le (b-a)^{-1}\mathbb{E}[(X_n-a)^-] \le (b-a)^{-1}\mathbb{E}[|X_n| + a].$$

By monotone convergence theorem, since $N_n([a,b],X) \uparrow N([a,b],X)$ as $n \to \infty$, we get that

$$\mathbb{E}[N([a,b],X)] \le (b-a)^{-1}(\sup_n \mathbb{E}[|X_n|] + a) < \infty,$$

by the assumption on X being bounded in L^1. Therefore, we get that $N([a,b],X) < \infty$ a.s. for every $a < b \in \mathbb{Q}$. Hence,

$$\mathbb{P}\left(\bigcap_{a<b\in\mathbb{Q}}\{N([a,b[,X) < \infty\}\right) = 1.$$

Writing $\Omega_0 = \bigcap_{a<b\in\mathbb{Q}}\{N([a,b],X) < \infty\}$, we have that $\mathbb{P}(\Omega_0) = 1$ and by Lemma 2.15 on Ω_0 we have that X converges to a possibly infinite limit X_∞. So we can define

$$X_\infty = \begin{cases} \lim_{n\to\infty} X_n, & \text{on } \Omega_0, \\ 0, & \text{on } \Omega \setminus \Omega_0. \end{cases}$$

Then X_∞ is $\mathcal{F}_\infty$-measurable and by Fatou's lemma and the assumption on X being in L^1 we get

$$\mathbb{E}[|X_\infty|] = \mathbb{E}[\liminf_n |X_n|] \le \liminf_n \mathbb{E}[|X_n|] < \infty.$$

Hence $X_\infty \in L^1$ as required. □

2.5 Doob's inequalities

Theorem 2.17. [Doob's maximal inequality] *Let $X = (X_n)_n$ be a non-negative submartingale. Writing $X_n^* = \sup_{0\le k\le n} X_k$ we have*

$$\lambda\mathbb{P}(X_n^* \ge \lambda) \le \mathbb{E}[X_n\mathbf{1}(X_n^* \ge \lambda)] \le \mathbb{E}[X_n].$$

Proof. Let $T = \inf\{k \geq 0 : X_k \geq \lambda\}$. Then $T \wedge n$ is a bounded stopping time, hence by the Optional stopping theorem, Theorem 2.9, we have

$$\mathbb{E}[X_n] \geq \mathbb{E}[X_{T\wedge n}] = \mathbb{E}[X_T \mathbf{1}(T \leq n)] + \mathbb{E}[X_n \mathbf{1}(T > n)] \geq \lambda \mathbb{P}(T \leq n) + \mathbb{E}[X_n \mathbf{1}(T > n)].$$

It is clear that $\{T \leq n\} = \{X_n^* \geq \lambda\}$. Hence we get

$$\lambda \mathbb{P}(X_n^* \geq \lambda) \leq \mathbb{E}[X_n \mathbf{1}(T \leq n)] = \mathbb{E}[X_n \mathbf{1}(X_n^* \geq \lambda)] \leq \mathbb{E}[X_n].$$

□

Theorem 2.18. [Doob's L^p inequality] *Let X be a martingale or a non-negative submartingale. Then for all $p > 1$ letting $X_n^* = \sup_{k \leq n} |X_k|$ we have*

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

Proof. If X is a martingale, then by Jensen's inequality $|X|$ is a non-negative submartingale. So it suffices to consider the case where X is a non-negative submartingale.

Fix $k < \infty$. We now have

$$\begin{aligned}
\mathbb{E}[(X_n^* \wedge k)^p] &= \mathbb{E}\left[\int_0^k p x^{p-1} \mathbf{1}(X_n^* \geq x)\, dx\right] = \int_0^k p x^{p-1} \mathbb{P}(X_n^* \geq x)\, dx \\
&\leq \int_0^k p x^{p-2} \mathbb{E}[X_n \mathbf{1}(X_n^* \geq x)]\, dx = \frac{p}{p-1} \mathbb{E}\left[X_n (X_n^* \wedge k)^{p-1}\right] \\
&\leq \frac{p}{p-1} \|X_n\|_p \|X_n^* \wedge k\|_p^{p-1},
\end{aligned}$$

where in the second and third equalities we used Fubini's theorem, for the first inequality we used Theorem 2.17 and for the last inequality we used Hölder's inequality. Rearranging, we get

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

Letting $k \to \infty$ and using monotone convergence completes the proof. □

2.6 L^p convergence for $p > 1$

Theorem 2.19. *Let X be a martingale and $p > 1$, then the following statements are equivalent:*

1. *X is bounded in $\mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P}) : \sup_{n \geq 0} \|X_n\|_p < \infty$*
2. *X converges a.s. and in $\mathcal{L}^p$ to a random variable X_∞*
3. *There exists a random variable $Z \in \mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P})$ such that*

$$X_n = \mathbb{E}[Z | \mathcal{F}_n] \text{ a.s.}$$

Proof. **1 $\Longrightarrow$ 2** Suppose that X is bounded in $\mathcal{L}^p$. Then by Jensen's inequality, X is also bounded in $\mathcal{L}^1$. Hence by Theorem 2.13 we have that X converges to a finite limit X_∞ a.s. By Fatou's lemma we have

$$\mathbb{E}[|X_\infty|^p] = \mathbb{E}[\liminf_n |X_n|^p] \leq \liminf_n \mathbb{E}[|X_n|^p] \leq \sup_{n\geq 0} \|X_n\|_p^p < \infty.$$

By Doob's $\mathcal{L}^p$ inequality, Theorem 2.18 we have that

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

where recall that $X_n^* = \sup_{k\leq n} |X_k|$. If we now let $n \to \infty$, then by monotone convergence we get that

$$\|X_\infty^*\|_p \leq \frac{p}{p-1} \sup_{n\geq 0} \|X_n\|_p.$$

Therefore

$$|X_n - X_\infty| \leq 2X_\infty^* \in \mathcal{L}^p$$

and dominated convergence theorem gives that X_n converges to X_∞ in $\mathcal{L}^p$.

2 $\Longrightarrow$ 3 We set $Z = X_\infty$. Clearly $Z \in \mathcal{L}^p$. We will now show that $X_n = \mathbb{E}[Z|\mathcal{F}_n]$ a.s. If $m \geq n$, then by the martingale property we can write

$$\|X_n - \mathbb{E}[X_\infty|\mathcal{F}_n]\|_p = \|\mathbb{E}[X_m - X_\infty|\mathcal{F}_n]\|_p \leq \|X_m - X_\infty\|_p \to 0 \text{ as } m \to \infty. \tag{2.5}$$

Hence $X_n = \mathbb{E}[X_\infty|\mathcal{F}_n]$ a.s.

3 $\Longrightarrow$ 1 This is immediate by the conditional Jensen's inequality. $\square$

Remark 2.20. A martingale of the form $\mathbb{E}[Z|\mathcal{F}_n]$ (it is a martingale by the tower property) with $Z \in \mathcal{L}^p$ is called a martingale closed in $\mathcal{L}^p$.

Corollary 2.21. *Let $Z \in \mathcal{L}^p$ and $X_n = \mathbb{E}[Z|\mathcal{F}_n]$ a martingale closed in $\mathcal{L}^p$. If $\mathcal{F}_\infty = \sigma(\mathcal{F}_n, n \geq 0)$, then we have*

$$X_n \to X_\infty = \mathbb{E}[Z|\mathcal{F}_\infty] \text{ as } n \to \infty \text{ a.s. and in } \mathcal{L}^p.$$

Proof. By the above theorem we have that $X_n \to X_\infty$ as $n \to \infty$ a.s. and in $\mathcal{L}^p$. It only remains to show that $X_\infty = \mathbb{E}[Z|\mathcal{F}_\infty]$ a.s. Clearly X_∞ is $\mathcal{F}_\infty$-measurable. Let $A \in \cup_{n\geq 0}\mathcal{F}_n$. Then $A \in \mathcal{F}_N$ for some N and

$$\mathbb{E}[Z\mathbf{1}(A)] = \mathbb{E}[\mathbb{E}[Z|\mathcal{F}_\infty]\mathbf{1}(A)] = \mathbb{E}[X_N\mathbf{1}(A)] \to \mathbb{E}[X_\infty\mathbf{1}(A)] \text{ as } N \to \infty.$$

So this shows that for all $A \in \cup_{n\geq 0}\mathcal{F}_n$ we have

$$\mathbb{E}[X_\infty\mathbf{1}(A)] = \mathbb{E}[\mathbb{E}[Z|\mathcal{F}_\infty]\mathbf{1}(A)].$$

But $\cup_{n\geq 0}\mathcal{F}_n$ is a π-system generating $\mathcal{F}_\infty$, and hence we get the equality for all $A \in \mathcal{F}_\infty$. $\square$

2.7 Uniformly integrable martingales

Definition 2.22. A collection $(X_i, i \in I)$ of random variables is called *uniformly integrable* (UI) if

$$\sup_{i\in I} \mathbb{E}[|X_i|\mathbf{1}(|X_i| > \alpha)] \to 0 \text{ as } \alpha \to \infty.$$

Equivalently, (X_i) is UI, if (X_i) is bounded in $\mathcal{L}^1$ and

$$\forall \varepsilon > 0,\ \exists \delta > 0:\ \forall A \in \mathcal{F},\ \mathbb{P}(A) < \delta \Rightarrow \sup_{i\in I} \mathbb{E}[|X_i|\mathbf{1}(A)] < \varepsilon.$$

Remember that a UI family is bounded in $\mathcal{L}^1$. The converse is not true.

If a family is bounded in $\mathcal{L}^p$, for some $p > 1$, then it is UI.

Theorem 2.23. *Let $X \in \mathcal{L}^1$. Then the class*

$$\{\mathbb{E}[X|\mathcal{G}] : \mathcal{G} \text{ a sub-}\sigma\text{-algebra of } \mathcal{F}\}$$

is uniformly integrable.

Proof. Since $X \in \mathcal{L}^1$, we have that for every $\varepsilon > 0$ there exists a $\delta > 0$ such that whenever $\mathbb{P}(A) \le \delta$, then

$$\mathbb{E}[|X|\mathbf{1}(A)] \le \varepsilon. \tag{2.6}$$

We now choose $\lambda < \infty$ so that $\mathbb{E}[|X|] \le \lambda\delta$. For any sub-$\sigma$-algebra $\mathcal{G}$ we have

$$\mathbb{E}[|\mathbb{E}[X|\mathcal{G}]|] \le \mathbb{E}[|X|].$$

Writing $Y = \mathbb{E}[X|\mathcal{G}]$ we have by Markov's inequality $\mathbb{P}(|Y| \ge \lambda) \le \mathbb{E}[|X|]/\lambda \le \delta$. Finally from (2.6) and the fact that $\{|Y| \ge \lambda\} \in \mathcal{G}$ we have

$$\mathbb{E}[|Y|\mathbf{1}(|Y| \ge \lambda)] \le \mathbb{E}[|X|\mathbf{1}(|Y| \ge \lambda)] \le \varepsilon.$$

□

Lemma 2.24. *Let $(X_n)_n, X \in \mathcal{L}^1$ and $X_n \to X$ as $n \to \infty$ a.s. Then*

$$X_n \xrightarrow{\mathcal{L}^1} X \text{ as } n \to \infty \text{ iff } (X_n)_{n\ge 0} \text{ is UI.}$$

Proof. See [2, Theorem 13.7]. □

Definition 2.25. A martingale $(X_n)_{n\ge 0}$ is called a UI martingale if it is a martingale and the collection of random variables $(X_n)_{n\ge 0}$ is a UI family.

Theorem 2.26. *Let X be a martingale. The following statements are equivalent.*

1. *X is a uniformly integrable martingale*
2. *X_n converges a.s. and in $\mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ to a limit X_∞*

3. There exists $Z \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ so that $X_n = \mathbb{E}[Z|\mathcal{F}_n]$ a.s. for all $n \geq 0$.

Proof. **1 $\Longrightarrow$ 2** Since X is UI, it follows that it is bounded in $\mathcal{L}^1$, and hence from Theorem 2.13 we get that X_n converges a.s. towards a finite limit X_∞ as $n \to \infty$. Since X is UI, [2, Theorem 13.7] gives the $\mathcal{L}^1$ convergence.

2 $\Longrightarrow$ 3 We set $Z = X_\infty$. Clearly $Z \in \mathcal{L}^1$. We will now show that $X_n = \mathbb{E}[Z|\mathcal{F}_n]$ a.s. For all $m \geq n$ by the martingale property we have

$$\|X_n - \mathbb{E}[X_\infty|\mathcal{F}_n]\|_1 = \|\mathbb{E}[X_m - X_\infty|\mathcal{F}_n]\|_1 \leq \|X_m - X_\infty\|_1 \to 0 \text{ as } m \to \infty.$$

3 $\Longrightarrow$ 1 Notice that by the tower property of conditional expectation, $\mathbb{E}[Z|\mathcal{F}_n]$ is a martingale. The uniform integrability follows from Theorem 2.23. □

Remark 2.27. As in Corollary 2.21, if X is a UI martingale, then $\mathbb{E}[Z|\mathcal{F}_\infty] = X_\infty$, where $\mathcal{F}_\infty = \sigma(\mathcal{F}_n, n \geq 0)$.

Remark 2.28. If X is a UI supermartingale (resp. submartingale), then X_n converges a.s. and in $\mathcal{L}^1$ to a limit X_∞, so that $\mathbb{E}[X_\infty|\mathcal{F}_n] \leq X_n$ (resp. $\geq$) for every n.

Example 2.29. Let $(X_i)_i$ be i.i.d. random variables with $\mathbb{P}(X_1 = 0) = \mathbb{P}(X_1 = 2) = 1/2$. Then $Y_n = X_1 \cdots X_n$ is a martingale bounded in $\mathcal{L}^1$ and it converges to 0 as $n \to \infty$ a.s. But $\mathbb{E}[Y_n] = 1$ for all n, and hence it does not converge in $\mathcal{L}^1$.

If X is a UI martingale and T is a stopping time, which could also take the value ∞, then we can unambiguously define

$$X_T = \sum_{n=0}^{\infty} X_n \mathbf{1}(T = n) + X_\infty \mathbf{1}(T = \infty).$$

Theorem 2.30. [Optional stopping for UI martingales] Let X be a UI martingale and let S and T be stopping times with $S \leq T$. Then

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

Proof. We will first show that $\mathbb{E}[X_\infty|\mathcal{F}_T] = X_T$ a.s. for any stopping time T. We will now check that $X_T \in \mathcal{L}^1$. Since $|X_n| \leq \mathbb{E}[|X_\infty||\mathcal{F}_n]$, we have

$$\mathbb{E}[|X_T|] = \sum_{n=0}^{\infty} \mathbb{E}[|X_n|\mathbf{1}(T = n)] + \mathbb{E}[|X_\infty|\mathbf{1}(T = \infty)] \leq \sum_{n \in \mathbb{Z}_+ \cup \{\infty\}} \mathbb{E}[|X_\infty|\mathbf{1}(T = n)] = \mathbb{E}[|X_\infty|].$$

Let $B \in \mathcal{F}_T$. Then

$$\mathbb{E}[\mathbf{1}(B)X_T] = \sum_{n \in \mathbb{Z}_+ \cup \{\infty\}} \mathbb{E}[\mathbf{1}(B)\mathbf{1}(T = n)X_n] = \sum_{n \in \mathbb{Z}_+ \cup \{\infty\}} \mathbb{E}[\mathbf{1}(B)\mathbf{1}(T = n)X_\infty] = \mathbb{E}[\mathbf{1}(B)X_\infty],$$

where for the second equality we used that $\mathbb{E}[X_\infty|\mathcal{F}_n] = X_n$ a.s. Also, clearly X_T is $\mathcal{F}_T$-measurable, and hence

$$\mathbb{E}[X_\infty|\mathcal{F}_T] = X_T \text{ a.s.}$$

Now using the tower property of conditional expectation, we get for stopping times $S \leq T$, since $\mathcal{F}_S \subseteq \mathcal{F}_T$

$$\mathbb{E}[X_T|\mathcal{F}_S] = \mathbb{E}[\mathbb{E}[X_\infty|\mathcal{F}_T]|\mathcal{F}_S] = \mathbb{E}[X_\infty|\mathcal{F}_S] = X_S \text{ a.s.}$$

□

2.8 Backwards martingales

Let $\ldots \subseteq \mathcal{G}_{-2} \subseteq \mathcal{G}_{-1} \subseteq \mathcal{G}_0$ be a sequence of sub-σ-algebras indexed by $\mathbb{Z}_-$. Given such a filtration, a process $(X_n, n \leq 0)$ is called a *backwards martingale*, if it is adapted to the filtration, $X_0 \in \mathcal{L}^1$ and for all $n \leq -1$ we have

$$\mathbb{E}[X_{n+1}|\mathcal{G}_n] = X_n \text{ a.s.}$$

By the tower property of conditional expectation we get that for all $n \leq 0$

$$\mathbb{E}[X_0|\mathcal{G}_n] = X_n \text{ a.s.} \tag{2.7}$$

Since $X_0 \in \mathcal{L}^1$, from (2.7) and Theorem 2.23 we get that X is uniformly integrable. This is a nice property that backwards martingales have: they are automatically UI.

Theorem 2.31. *Let X be a backwards martingale, with $X_0 \in \mathcal{L}^p$ for some $p \in [1, \infty)$. Then X_n converges a.s. and in $\mathcal{L}^p$ as $n \to -\infty$ to the random variable $X_{-\infty} = \mathbb{E}[X_0|\mathcal{G}_{-\infty}]$, where $\mathcal{G}_{-\infty} = \cap_{n \leq 0} \mathcal{G}_n$.*

Proof. We will first adapt Doob's up-crossing inequality, Theorem 2.16, in this setting. Let $a < b$ be real numbers and $N_{-n}([a, b], X)$ be the number of up-crossings of the interval $[a, b]$ by X between times $-n$ and 0 as defined at the beginning of Section 2.4.

If we write $\mathcal{F}_k = \mathcal{G}_{-n+k}$, for $0 \leq k \leq n$, then $\mathcal{F}_k$ is an increasing filtration and the process $(X_{-n+k}, 0 \leq k \leq n)$ is an $\mathcal{F}$-martingale. Then $N_{-n}([a, b], X)$ is the number of up-crossings of the interval $[a, b]$ by X_{-n+k} between times 0 and n. Thus applying Doob's up-crossing inequality to X_{-n+k} we get that

$$(b - a)\mathbb{E}[N_{-n}([a, b], X)] \leq \mathbb{E}[(X_0 - a)^-].$$

Letting $n \to \infty$ we have that $N_{-n}([a, b], X)$ increases to the total number of up-crossings of X from a to b and thus we deduce that

$$X_m \to X_{-\infty} \text{ as } m \to -\infty \text{ a.s.},$$

for some random variable $X_{-\infty}$, which is $\mathcal{G}_{-\infty}$-measurable, since the σ-algebras $\mathcal{G}_n$ are decreasing.

Since $X_0 \in \mathcal{L}^p$, it follows that $X_n \in \mathcal{L}^p$, for all $n \leq 0$. Also, by Fatou's lemma, we get that $X_{-\infty} \in \mathcal{L}^p$. Now by conditional Jensen's inequality we obtain

$$|X_n - X_{-\infty}|^p = |\mathbb{E}[X_0 - X_{-\infty}|\mathcal{G}_n]|^p \leq \mathbb{E}[|X_0 - X_{-\infty}|^p|\mathcal{G}_n].$$

But the latter family of random variables, $(\mathbb{E}[|X_0 - X_{-\infty}|^p|\mathcal{G}_n])_n$ is UI, by Theorem 2.23 again. Hence also $(|X_n - X_{-\infty}|^p)_n$ is UI, and thus by [2, Theorem 13.7], we conclude that $X_n \to X_{-\infty}$ as $n \to -\infty$ in $\mathcal{L}^p$.

In order to show that $X_{-\infty} = \mathbb{E}[X_0|\mathcal{G}_{-\infty}]$ a.s., it only remains to show that if $A \in \mathcal{G}_{-\infty}$, then

$$\mathbb{E}[X_0 \mathbf{1}(A)] = \mathbb{E}[X_{-\infty} \mathbf{1}(A)].$$

Since $A \in \mathcal{G}_n$, for all $n \leq 0$, we have by the martingale property that

$$\mathbb{E}[X_0 \mathbf{1}(A)] = \mathbb{E}[X_n \mathbf{1}(A)].$$

Letting $n \to -\infty$ in the above equality and using the $\mathcal{L}^1$ convergence of X_n to $X_{-\infty}$ finishes the proof. $\square$

2.9 Applications of martingales

Theorem 2.32. **[Kolmogorov's 0-1 law]** *Let $(X_i)_{i\geq 1}$ be a sequence of i.i.d. random variables. Let $\mathcal{F}_n = \sigma(X_k, k \geq n)$ and $\mathcal{F}_\infty = \cap_{n\geq 0}\mathcal{F}_n$. Then $\mathcal{F}_\infty$ is trivial, i.e. every $A \in \mathcal{F}_\infty$ has probability $\mathbb{P}(A) \in \{0, 1\}$.*

Proof. Let $\mathcal{G}_n = \sigma(X_k, k \leq n)$ and $A \in \mathcal{F}_\infty$. Since $\mathcal{G}_n$ is independent of $\mathcal{F}_{n+1}$, we have that

$$\mathbb{E}[\mathbf{1}(A)|\mathcal{G}_n] = \mathbb{P}(A) \text{ a.s.}$$

Theorem 2.26 gives that $\mathbb{E}[\mathbf{1}(A)|\mathcal{G}_n]$ converges to $\mathbb{E}[\mathbf{1}(A)|\mathcal{G}_\infty]$ a.s. as $n \to \infty$, where $\mathcal{G}_\infty = \sigma(\mathcal{G}_n, n \geq 0)$. Hence we deduce that

$$\mathbb{E}[\mathbf{1}(A)|\mathcal{G}_\infty] = \mathbf{1}(A) = \mathbb{P}(A) \text{ a.s.},$$

since $\mathcal{F}_\infty \subseteq \mathcal{G}_\infty$. Therefore

$$\mathbb{P}(A) \in \{0, 1\}.$$

□

Theorem 2.33. **[Strong law of large numbers]** *Let $(X_i)_{i\geq 1}$ be a sequence of i.i.d. random variables in $\mathcal{L}^1$ with $\mu = \mathbb{E}[X_1]$. Let $S_n = X_1 + \ldots + X_n$, for $n \geq 1$ and $S_0 = 0$. Then $S_n/n \to \mu$ as $n \to \infty$ a.s. and in $\mathcal{L}^1$.*

Proof. Let $\mathcal{G}_n = \sigma(S_n, S_{n+1}, \ldots) = \sigma(S_n, X_{n+1}, \ldots)$. We will now show that $(M_n)_{n\leq -1} = (S_{-n}/(-n))_{n\leq -1}$ is a $(\mathcal{F}_n)_{n\leq -1} = (\mathcal{G}_{-n})_{n\leq -1}$ backwards martingale. We have for $m \leq -1$

$$\mathbb{E}\left[M_{m+1}\middle|\mathcal{F}_m\right] = \mathbb{E}\left[\frac{S_{-m-1}}{-m-1}\middle|\mathcal{G}_{-m}\right]. \tag{2.8}$$

Setting $n = -m$, since X_n is independent of $X_{n+1}, X_{n+2}, \ldots$, we obtain

$$\mathbb{E}\left[\frac{S_{n-1}}{n-1}\middle|\mathcal{G}_n\right] = \mathbb{E}\left[\frac{S_n - X_n}{n-1}\middle|\mathcal{G}_n\right] = \frac{S_n}{n-1} - \mathbb{E}\left[\frac{X_n}{n-1}\middle|S_n\right]. \tag{2.9}$$

By symmetry, notice that $\mathbb{E}[X_k|S_n] = \mathbb{E}[X_1|S_n]$ for all k. Indeed, for any $A \in \mathcal{B}(\mathbb{R})$ we have that $\mathbb{E}[X_k\mathbf{1}(S_n \in A)]$ does not depend on k. Clearly

$$\mathbb{E}[X_1|S_n] + \ldots + \mathbb{E}[X_n|S_n] = \mathbb{E}[S_n|S_n] = S_n,$$

and hence $\mathbb{E}[X_n|S_n] = S_n/n$ a.s. Finally putting everything together we get

$$\mathbb{E}\left[\frac{S_{n-1}}{n-1}\middle|\mathcal{G}_n\right] = \frac{S_n}{n-1} - \frac{S_n}{n(n-1)} = \frac{S_n}{n} \text{ a.s.}$$

Thus, by the backwards martingale convergence theorem, we deduce that $\frac{S_n}{n}$ converges as $n \to \infty$ a.s. and in $\mathcal{L}^1$ to a random variable, say $Y = \lim S_n/n$. Obviously for all k

$$Y = \lim \frac{X_{k+1} + \ldots + X_{k+n}}{n},$$

and hence Y is $\mathcal{T}_k = \sigma(X_{k+1}, \ldots)$-measurable, for all k, hence it is $\cap_k \mathcal{T}_k$-measurable. By Kolmogorov's 0-1 law, Theorem 2.32, we conclude that there exists a constant $c \in \mathbb{R}$ such that $\mathbb{P}(Y = c) = 1$. But

$$c = \mathbb{E}[Y] = \lim \mathbb{E}[S_n/n] = \mu.$$

□

Theorem 2.34. [Kakutani's product martingale theorem] *Let $(X_n)_{n\geq 0}$ be a sequence of independent non-negative random variables of mean 1. We set*

$$M_0 = 1 \text{ and } M_n = X_1 X_2 \ldots X_n, \; n \in \mathbb{N}.$$

Then $(M_n)_{n\geq 0}$ is a non-negative martingale and $M_n \to M_\infty$ a.s. as $n \to \infty$ for some random variable M_∞. We set $a_n = \mathbb{E}[\sqrt{X_n}]$, then $a_n \in (0, 1]$. Moreover,

1. *if $\prod_n a_n > 0$, then $M_n \to M_\infty$ in $\mathcal{L}^1$ and $\mathbb{E}[M_\infty] = 1$,*
2. *if $\prod_n a_n = 0$, then $M_\infty = 0$ a.s.*

Proof. Clearly $(M_n)_n$ is a positive martingale and $\mathbb{E}[M_n] = 1$, for all n, since the random variables (X_i) are independent and of mean 1. Hence, by the a.s. martingale convergence theorem, we get that M_n converges a.s. as $n \to \infty$ to a finite random variable M_∞. By Cauchy-Schwarz $a_n \leq 1$ for all n.

We now define

$$N_n = \frac{\sqrt{X_1 \ldots X_n}}{a_1 \ldots a_n}, \text{ for } n \geq 1.$$

Then N_n is a non-negative martingale that is bounded in $\mathcal{L}^1$, and hence converges a.s. towards a finite limit N_∞ as $n \to \infty$.

1. We have

$$\sup_{n\geq 0} \mathbb{E}[N_n^2] = \sup_{n\geq 0} \frac{1}{\left(\prod_{i=1}^n a_i\right)^2} = \frac{1}{(\prod_n a_n)^2} < \infty, \tag{2.10}$$

under the assumption that $\prod_n a_n > 0$. Since $M_n = N_n^2 (\prod_{i=1}^n a_i)^2 \leq N_n^2$ for all n, we get

$$\mathbb{E}[\sup_{k\leq n} M_k] \leq \mathbb{E}[\sup_{k\leq n} N_k^2] \leq 4\mathbb{E}[N_n^2],$$

where the last inequality follows by Doob's $\mathcal{L}^2$-inequality, Theorem 2.18. Hence by Monotone convergence and (2.10) we deduce

$$\mathbb{E}[\sup_n M_n] < \infty,$$

and since $M_n \leq \sup_n M_n$ we conclude that M_n is UI, and hence it also converges in $\mathcal{L}^1$ towards M_∞. Finally since $\mathbb{E}[M_n] = 1$ for all n, it follows that $\mathbb{E}[M_\infty] = 1$.

2. We have $M_n = N_n^2 (\prod_{i=1}^n a_i)^2 \to 0$, as $n \to \infty$, since $\prod_n a_n = 0$ and N_∞ exists and is finite a.s. by the a.s. martingale convergence theorem. Hence $M_\infty = 0$ a.s. □

2.9.1 Martingale proof of the Radon-Nikodym theorem

Theorem 2.35. **[Radon-Nikodym theorem]** *Let $\mathbb{P}$ and Q be two probability measures on the measurable space $(\Omega, \mathcal{F})$. Assume that $\mathcal{F}$ is countably generated, i.e. there exists a collection of sets $(F_n : n \in \mathbb{N})$ such that*

$$\mathcal{F} = \sigma(F_n : n \in \mathbb{N}).$$

Then the following statements are equivalent:

(a) $\mathbb{P}(A) = 0$ implies that $Q(A) = 0$ for all $A \in \mathcal{F}$ (and in this case we say that Q is absolutely continuous with respect to $\mathbb{P}$ and write $Q \ll \mathbb{P}$).

(b) $\forall \varepsilon > 0, \exists \delta > 0, \forall A \in \mathcal{F}, \mathbb{P}(A) \leq \delta$ implies that $Q(A) \leq \varepsilon$.

(c) There exists a non-negative random variable X such that

$$Q(A) = \mathbb{E}[X\mathbf{1}(A)], \;\; A \in \mathcal{F}.$$

Remark 2.36. The random variable X which is unique $\mathbb{P}$-a.s., is called (a version of) the *Radon-Nikodym derivative* of Q with respect to $\mathbb{P}$. We write $X = dQ/d\mathbb{P}$ a.s. The theorem extends immediately to finite measures by scaling, then to σ-finite measures by breaking the space Ω into pieces where the measures are finite. Also we can lift the assumption that the σ-algebra $\mathcal{F}$ is countably generated and the details for that can be found in [2, Chapter 14].

***Proof*.** We will first show that **(a) implies (b)**. If (b) does not hold, then we can find $\varepsilon > 0$ such that for all $n \geq 1$ there exists a set A_n with $\mathbb{P}(A_n) \leq 1/n^2$ and $Q(A_n) \geq \varepsilon$. Then by the Borel-Cantelli lemma we get that

$$\mathbb{P}(A_n \text{ i.o.}) = 0$$

Therefore from (a) we will get that $Q(A_n \text{ i.o.}) = 0$. But

$$Q(A_n \text{ i.o.}) = Q(\cap_n \cup_{k \geq n} A_k) = \lim_{n \to \infty} Q(\cup_{k \geq n} A_k) \geq \varepsilon,$$

which is a contradiction, so (a) implies (b).

Next we will show that **(b) implies (c)**. We consider the following filtration:

$$\mathcal{F}_n = \sigma(F_k, k \leq n).$$

If we write $\mathcal{A}_n = \{H_1 \cap \ldots \cap H_n : H_i = F_i \text{ or } F_i^{\mathrm{c}}\}$, then it is easy to see that

$$\mathcal{F}_n = \sigma(\mathcal{A}_n).$$

Note that the sets in $\mathcal{A}_n$ are disjoint. We now let $X_n : \Omega \to [0, \infty)$ be the random variable defined as follows

$$X_n(\omega) = \sum_{A \in \mathcal{A}_n} \frac{Q(A)}{\mathbb{P}(A)} \mathbf{1}(\omega \in A).$$

Since the sets in $\mathcal{A}_n$ are disjoint, we get that

$$Q(A) = \mathbb{E}[X_n \mathbf{1}(A)], \text{ for all } A \in \mathcal{F}_n.$$

We will use the notation

$$X_n = \frac{dQ}{d\mathbb{P}} \text{ on } \mathcal{F}_n.$$

It is easy to check that $(X_n)_n$ is a non-negative martingale with respect to the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n), \mathbb{P})$. Indeed, if $A \in \mathcal{F}_n$, then

$$\mathbb{E}[X_{n+1}\mathbf{1}(A)] = Q(A) = \mathbb{E}[X_n \mathbf{1}(A)], \text{ for all } A \in \mathcal{F}_n.$$

Also (X_n) is bounded in $\mathcal{L}^1$, since $\mathbb{E}[X_n] = Q(\Omega) = 1$. Hence by the a.s. martingale convergence theorem, it converges a.s. towards a random variable X_∞ as $n \to \infty$.

We will now show that (X_n) is a uniformly integrable martingale. Set $\lambda = 1/\delta$. Then by Markov's inequality

$$\mathbb{P}(X_n \geq \lambda) \leq \frac{\mathbb{E}[X_n]}{\lambda} = \frac{1}{\lambda} = \delta.$$

Therefore by (b)

$$\mathbb{E}[X_n \mathbf{1}(X_n \geq \lambda)] = Q(\{X_n \geq \lambda\}) \leq \varepsilon,$$

which proves the uniform integrability. Thus by the convergence theorem for UI martingales, Theorem 2.26, we get that X_n converges to X_∞ as $n \to \infty$ in $\mathcal{L}^1$ and $\mathbb{E}[X_\infty] = 1$. So for all $A \in \mathcal{F}_n$ we have

$$\mathbb{E}[X_n \mathbf{1}(A)] = \mathbb{E}[X_\infty \mathbf{1}(A)].$$

Hence if we now define a new probability measure $\widetilde{Q}(A) = \mathbb{E}[X_\infty \mathbf{1}(A)]$, then $Q(A) = \widetilde{Q}(A)$ for all $A \in \cup_n \mathcal{F}_n$. But $\cup_n \mathcal{F}_n$ is a π-system that generates the σ- algebra $\mathcal{F}$, and hence

$$Q = \widetilde{Q} \text{ on } \mathcal{F},$$

which implies (c).

The implication **(c)** $\implies$ **(a)** is straightforward. $\square$

3 Continuous-time random processes

3.1 Definitions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. So far we have considered stochastic processes in discrete time only. In this section the time index set is going to be the whole positive real line, $\mathbb{R}_+$. As in Section 2, we define a filtration $(\mathcal{F}_t)_t$ to be an increasing collection of sub σ-algebras of $\mathcal{F}$, i.e. $\mathcal{F}_t \subseteq \mathcal{F}_{t'}$, if $t \leq t'$. A collection of random variables $(X_t : t \in \mathbb{R}_+)$ is called a stochastic process. Usually as in Section 2, X will take values in $\mathbb{R}$ or $\mathbb{R}^d$. X is called *adapted* to the filtration $(\mathcal{F}_t)$, if X_t is $\mathcal{F}_t$-measurable for all t. A stopping time T is a random variable taking values in $[0, \infty]$ such that $\{T \leq t\} \in \mathcal{F}_t$, for all t.

When we consider processes in discrete time, if we equip $\mathbb{N}$ with the σ-algebra $\mathcal{P}(\mathbb{N})$ that contains all the subsets of $\mathbb{N}$, then the process

$$(\omega, n) \mapsto X_n(\omega)$$

is clearly measurable with respect to the product σ-algebra $\mathcal{F} \otimes \mathcal{P}(\mathbb{N})$.

Back to continuous time, if we fix $t \in \mathbb{R}_+$, then $\omega \mapsto X_t(\omega)$ is a random variable. But, the mapping $(\omega, t) \mapsto X_t(\omega)$ has no reason to be measurable with respect to $\mathcal{F} \otimes \mathcal{B}(\mathbb{R})$ ($\mathcal{B}(\mathbb{R})$ is the Borel σ-algebra) unless some regularity conditions are imposed on X. Also, if $A \subseteq \mathbb{R}$, then the first hitting time of A,

$$T_A = \inf\{t : X_t \in A\}$$

is not in general a stopping time as the set

$$\{T_A \leq t\} = \cup_{0 \leq s \leq t}\{T = s\} \notin \mathcal{F}_t \text{ in general,}$$

since this is an uncountable union.

A quite natural requirement is that for a fixed ω the mapping $t \mapsto X_t(\omega)$ is continuous in t. Then, indeed the mapping $(\omega, t) \mapsto X_t(\omega)$ is measurable. More generally we will consider processes that are right-continuous and admit left limits everywhere a.s. and we will call such processes *càdlàg* from the french continu à droite limité à gauche. Continuous and càdlàg processes are determined by their values in a countable dense subset of $\mathbb{R}_+$, for instance $\mathbb{Q}_+$.

Note that if a process $X = (X_t)_{t \in (0,1]}$ is continuous, then the mapping

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

is measurable with respect to $\mathcal{F} \otimes \mathcal{B}((0, 1])$. To see this, note that by the continuity of X in t we can write

$$X_t(\omega) = \lim_{n \to \infty} \sum_{k=0}^{2^n - 1} \mathbf{1}(t \in (k2^{-n}, (k+1)2^{-n}])X_{k2^{-n}}(\omega).$$

For each n it is easy to see that

$$(\omega, t) \mapsto \sum_{k=0}^{2^n - 1} \mathbf{1}(t \in (k2^{-n}, (k+1)2^{-n}])X_{k2^{-n}}(\omega)$$

is $\mathcal{F} \otimes \mathcal{B}((0, 1])$-measurable. Hence $X_t(\omega)$ is $\mathcal{F} \otimes \mathcal{B}((0, 1])$-measurable, as a limit of measurable functions.

We let $C(\mathbb{R}_+, E)$ ($D(\mathbb{R}_+, E)$) be the space of continuous (cadlag) functions $x : \mathbb{R}_+ \to E$ endowed with the product σ-algebra that makes the projections $\pi_t : X \mapsto X_t$ measurable for every t. Note that $E = \mathbb{R}$ or $\mathbb{R}^d$ in this course.

For a stopping time T we define as before

$$\mathcal{F}_T = \{A \in \mathcal{F} : A \cap \{T \leq t\} \in \mathcal{F}_t \text{ for all } t\}.$$

For a cadlag process X we set $X_T(\omega) = X_{T(\omega)}(\omega)$, whenever $T(\omega) < \infty$ and again as before we define the *stopped* process X^T by $X^T_t = X_{T \wedge t}$.

Proposition 3.1. *Let S and T be stopping times and X a cadlag adapted process. Then*

1. *$S \wedge T$ is a stopping time,*
2. *if $S \leq T$, then $\mathcal{F}_S \subseteq \mathcal{F}_T$,*
3. *$X_T \mathbf{1}(T < \infty)$ is an $\mathcal{F}_T$-measurable random variable,*
4. *X^T is adapted.*

Proof. 1,2 follow directly from the definition like in the discrete time case. We will only show 3. Note that 4 follows from 3, since $X_{T\wedge t}$ will then be $\mathcal{F}_{T\wedge t}$-measurable, and hence $\mathcal{F}_t$-measurable, since by 2, $\mathcal{F}_{T\wedge t} \subseteq \mathcal{F}_t$.

Note that a random variable Z is $\mathcal{F}_T$ measurable if and only if $Z\mathbf{1}(T \leq t)$ is $\mathcal{F}_t$-measurable for all t. It follows directly by the definition that if Z is $\mathcal{F}_T$-measurable, then $Z\mathbf{1}(T \leq t)$ is $\mathcal{F}_t$-measurable for all t. For the other implication, note that if $Z = c\mathbf{1}(A)$, then the claim is true. This extends to all finite linear combinations of indicators, since if $Z = \sum_{i=1}^n c_i \mathbf{1}(A_i)$, where the constants c_i are positive, then we can write Z as a linear combination of indicators of disjoint sets and then the claim follows easily. Finally for any positive random variable Z we can approximate it by $Z_n = 2^{-n}\lfloor 2^n Z \rfloor \wedge n \uparrow Z$ as $n \to \infty$. Then the claim follows for each Z_n, since if $Z\mathbf{1}(T \leq t)$ is $\mathcal{F}_t$-measurable, then also $Z_n\mathbf{1}(T \leq t)$ is $\mathcal{F}_t$-measurable, for all t. Finally the limit of $\mathcal{F}_T$-measurable random variables is $\mathcal{F}_T$-measurable.

So in order to prove that $X_T\mathbf{1}(T < \infty)$ is $\mathcal{F}_T$-measurable, we will show that $X_T\mathbf{1}(T \leq t)$ is $\mathcal{F}_t$-measurable for all t. We can write

$$X_T\mathbf{1}(T \leq t) = X_T\mathbf{1}(T < t) + X_t\mathbf{1}(T = t).$$

Clearly, the random variable $X_t\mathbf{1}(T = t)$ is $\mathcal{F}_t$-measurable. It only remains to show that $X_T\mathbf{1}(T < t)$ is $\mathcal{F}_t$-measurable. If we let $T_n = 2^{-n}\lceil 2^n T \rceil$, then it is easy to see that T_n is a stopping time that takes values in the set $\mathcal{D}_n = \{k2^{-n} : k \in \mathbb{N}\}$. Indeed

$$\{T_n \leq t\} = \{\lceil 2^n T \rceil \leq 2^n t\} = \{T \leq 2^{-n}\lfloor 2^n t \rfloor\} \in \mathcal{F}_{2^{-n}\lfloor 2^n t \rfloor} \subseteq \mathcal{F}_t.$$

By the cadlag property of X and the convergence $T_n \downarrow T$ we get that

$$X_T\mathbf{1}(T < t) = \lim_{n\to\infty} X_{T_n \wedge t}\mathbf{1}(T < t).$$

Since T_n takes only countably many values, we have

$$X_{T_n\wedge t}\mathbf{1}(T < t) = \sum_{d \in \mathcal{D}_n, d \leq t} X_d \mathbf{1}(T_n = d) + X_t\mathbf{1}(T_n > t)\mathbf{1}(T < t).$$

But T_n is a stopping time wrt the filtration $(\mathcal{F}_t)$, and hence we see that $X_{T_n\wedge t}\mathbf{1}(T < t)$ is $\mathcal{F}_t$-measurable for all n and this finishes the proof. $\square$

Example 3.2. Note that when the time index set is $\mathbb{R}_+$, then hitting times are not always stopping times. Let J be a random variable that takes values $+1$ or -1 each with probability $1/2$. Consider now the following process

$$X_t = \begin{cases} t, & \text{if } t \in [0,1]; \\ 1 + J(t-1), & \text{if } t > 1. \end{cases}$$

Let $\mathcal{F}_t = \sigma(X_s, s \leq t)$ be the natural filtration of X. Then if $A = (1,2)$ and we consider $T_A = \inf\{t \geq 0 : X_t \in A\}$, then clearly

$$\{T_A \leq 1\} \notin \mathcal{F}_1.$$

If we impose some regularity conditions on the process or the filtration though, then we get stopping times like in the next two propositions.

Proposition 3.3. *Let A be a closed set and let X be a continuous adapted process. Then the first hitting time of A,*

$$T_A = \inf\{t \geq 0 : X_t \in A\},$$

is a stopping time.

***Proof*.** It suffices to show that

$$\{T_A \leq t\} = \left\{ \inf_{s \in \mathbb{Q}, s \leq t} d(X_s, A) = 0 \right\}, \tag{3.1}$$

where $d(x, A)$ stands for the distance of x from the set A. If $T_A = s \leq t$, then there exists a sequence s_n of times such that $X_{s_n} \in A$ and $s_n \downarrow s$ as $n \to \infty$. By continuity of X, we then deduce that $X_{s_n} \to X_s$ as $n \to \infty$ and since A is closed, we must have that $X_s \in A$. Thus we showed that $X_{T_A} \in A$. We can now find a sequence of rationals q_n such that $q_n \uparrow T_A$ as $n \to \infty$ and since $d(X_{T_A}, A) = 0$ we get that $d(X_{q_n}, A) \to 0$ as $n \to \infty$.

Suppose now that $\inf_{s \in \mathbb{Q}, s \leq t} d(X_s, A) = 0$. Then there exists a sequence $s_n \in \mathbb{Q}$, $s_n \leq t$, for all n such that

$$d(X_{s_n}, A) \to 0 \text{ as } n \to \infty.$$

We can extract a converging subsequence of $s_n \to s$ and by continuity of X we get that $X_{s_n} \to X_s$ as $n \to \infty$. Since $d(X_s, A) = 0$ and A is a closed set, we conclude that $X_s \in A$, and hence $T_A \leq t$. □

Definition 3.4. Let $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ be a filtration. For each t we define

$$\mathcal{F}_{t+} = \bigcap_{s > t} \mathcal{F}_s.$$

If $\mathcal{F}_{t+} = \mathcal{F}_t$ for all t, then we call the filtration $(\mathcal{F}_t)$ right-continuous.

Proposition 3.5. *Let A be an open set and X a continuous process. Then*

$$T_A = \inf\{t \geq 0 : X_t \in A\}$$

is a stopping time with respect to the filtration $(\mathcal{F}_{t+})$.

Proof. First we show that for all t, the event $\{T_A < t\} \in \mathcal{F}_t$. Indeed,by the continuity of X and the fact that A is open we get that

$$\{T_A < t\} = \bigcup_{q\in\mathbb{Q},q<t} \{X_q \in A\} \in \mathcal{F}_t,$$

since it is a countable union.

Since we can write

$$\{T_A \le t\} = \bigcap_n \{T < t + 1/n\}$$

we get that $\{T_A \le t\} \in \mathcal{F}_{t+}$. □

3.2 Martingale regularization theorem

As we discussed at the beginning of the section, we can view a stochastic process indexed by $\mathbb{R}_+$ as a random variable with values in the space of functions $\{f : \mathbb{R}_+ \to E\}$ endowed with the product σ-algebra that makes the projections $f \mapsto f(t)$ measurable. The law of the process X is the measure μ that is defined as

$$\mu(A) = \mathbb{P}(X \in A),$$

where A is in the product σ-algebra. However the measure μ is not easy to work with. Instead we consider simpler objects that we define below.

Given a probability measure μ on $D(\mathbb{R}_+, E)$ we consider the probability measure μ_J, where $J \subset \mathbb{R}_+$ is a finite set, defined as the law of $(X_t, t \in J)$. The probability measures (μ_J) are called the finite dimensional distributions of μ. By a π-system uniqueness argument, μ is uniquely determined by its finite-dimensional distributions. Indeed the set

$$\{\cap_{s\in J}\{X_s \in A_s\} : J \text{ is finite }, A_s \in \mathcal{B}(\mathbb{R})\}$$

is a π-system generating the product σ-algebra. So, when we want to specify the law of a cadlag process, it suffices to describe its finite-dimensional distributions. Of course we have no *a priori* reason to believe there exists a cadlag process whose finite-dimensional distributions coincide with a given family of measures $(\mu^J : J \subseteq \mathbb{R}_+, J \text{ finite})$.

Even if we know the law of a process, this does not give us much information about the sample path properties of the process. Namely, there could be different processes with the same finite marginal distributions. This motivates the following definition:

Definition 3.6. Let X and X' be two processes defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We say that X' is a version of X if $X_t = X'_t$ a.s. for every t.

Remark 3.7. Note that two versions of the same process have the same finite marginal distributions. But they do not share the same sample path properties.

Example 3.8. Let $X = (X_t)_{t\in[0,1]}$ be the process that is identical to 0 for all t. Then obviously the finite marginal distributions will be Dirac measures at 0. Now let U be a uniform random variable on $[0, 1]$. We define $X'_t = \mathbf{1}(U = t)$. Then clearly the finite marginal distributions of X' are Dirac measures at 0, and hence it is a version of X. However it is not continuous and furthermore

$$\mathbb{P}(X'_t = 0 \ \ \forall t \in [0,1]) = 0.$$

In this section we are going to show two theorems that guarantee the existence of a continuous or cadlag version of a process.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ be a filtered probability space. Let $\mathcal{N}$ be the collection of sets in $\mathcal{F}$ of measure 0. We define the filtration

$$\widetilde{\mathcal{F}}_t = \sigma(\mathcal{F}_{t+}, \mathcal{N}).$$

Definition 3.9. If a filtration satisfies $\widetilde{\mathcal{F}}_t = \mathcal{F}_t$ for all t, then we say that $(\mathcal{F}_t)$ satisfies the *usual conditions.*

Before stating the next theorem, note that the definitions of martingales (resp. supermartingales and submartingales) are the same in continuous time as the ones given for discrete time processes.

Theorem 3.10. **[Martingale regularization theorem]** *Let $(X_t)_{t\geq 0}$ be a martingale with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$. Then there exists a cadlag process $\widetilde{X}$ which is a martingale with respect to $(\widetilde{\mathcal{F}}_t)$ and satisfies*

$$X_t = \mathbb{E}[\widetilde{X}_t | \mathcal{F}_t] \ \ a.s.$$

for all $t \geq 0$. If the filtration $(\mathcal{F}_t)$ satisfies the usual conditions, then $\widetilde{X}$ is a cadlag version of X.

Before proving the theorem we state and prove an easy result about functions which is analogous to Lemma 2.15 which was used in the proof of the a.s. martingale convergence theorem.

Lemma 3.11. *Let $f : \mathbb{Q}_+ \to \mathbb{R}$ be a function defined on the positive rational numbers. Suppose that for all $a < b$ and $a, b \in \mathbb{Q}$ and all bounded $I \subseteq \mathbb{Q}_+$ the function f is bounded on I and the number of upcrossings of the interval $[a, b]$ during the time intervals I by f is finite, i.e. $N([a, b], I, f) < \infty$, where $N([a, b], I, f)$ is defined as*

$$\sup \{n \geq 0 : \exists \ 0 \leq s_1 < t_1 < \ldots < s_n < t_n, s_i, t_i \in I, f(s_i) < a, f(t_i) > b, 1 \leq i \leq n\}.$$

Then for every $t \in \mathbb{R}_+$ the right and left limits of f exist and are finite, i.e.

$$\lim_{s \downarrow t} f(s), \lim_{s \uparrow t} f(s) \ \text{exist and are finite.}$$

Proof. First note that if (s_n) is a sequence of rationals decreasing to t, then by Lemma 2.15 we get that the limit $\lim_n f(s_n)$ exists. Similarly if s'_n is a sequence increasing to t, then the limit $\lim_n f(s'_n)$ exists. So far we showed that for any sequence converging to t from above (or below) the limit exists. It remains to show that the limit is the same along any sequence decreasing to t. To see this, note that if s_n is a sequence decreasing to t and q_n is another sequence decreasing to t and $\lim_n f(s_n) \neq \lim_n f(q_n)$, then we can combine the two sequences and get a decreasing sequence (a_n) converging to t such that $\lim_n f(a_n)$ does not exist, which is a contradiction, since we already showed that for every decreasing sequence the limit exists. Finally the limits from above or below are finite, which follows by the assumption that f is bounded on any bounded subset of $\mathbb{Q}_+$. $\square$

Proof of Theorem 3.10. The goal is to define $\widetilde{X}$ as follows:

$$\widetilde{X}_t = \lim_{s \downarrow t, s \in \mathbb{Q}_+} X_s$$

on a set of measure 1 and 0 elsewhere.

So first we need to check that the limit above exists a.s. and is finite. In order to do so, we are going to use Lemma 3.11. Therefore we first show that X is bounded on bounded subsets I of $\mathbb{Q}_+$. Let I be such a subset. Consider $J = \{j_1, \ldots, j_n\} \subseteq I$, where $j_1 < j_2 < \ldots < j_n$. Then the process $(X_j)_{j \in J}$ is a discrete time martingale. By Doob's maximal inequality we obtain

$$\lambda \mathbb{P}(\max_{j \in J} |X_j| > \lambda) \leq \mathbb{E}[|X_{j_n}|] \leq \mathbb{E}[|X_K|],$$

where $K > \sup I$. So taking a monotone limit over J finite subsets of I with union the set I, then we get that

$$\lambda \mathbb{P}(\sup_{t \in I} |X_t| > \lambda) \leq \mathbb{E}[|X_K|].$$

Therefore by letting $\lambda \to \infty$ this shows that

$$\mathbb{P}(\sup_{t \in I} |X_t| < \infty) = 1.$$

Let $a < b$ be rational numbers. Then we have $N([a,b], I, X) = \sup_{J \subset I, \text{ finite}} N([a,b], J, X)$. Let $J = \{a_1, \ldots, a_n\}$ (in increasing order again) be a finite subset of I. Then $(X_{a_i})_{i \leq n}$ is a martingale and Doob's upcrossing lemma gives that

$$(b-a)\mathbb{E}[N([a,b], J, X)] \leq \mathbb{E}[(X_{a_n} - a)^-] \leq \mathbb{E}[(X_K - a)^-] \tag{3.2}$$

By monotone convergence again, if we let $I_M = \mathbb{Q}_+ \cap [0, M]$, we then get that for all M

$$N([a,b], I_M, X) < \infty \text{ a.s.}$$

Thus if we now let

$$\Omega_0 = \cap_{M \in \mathbb{N}} \cap_{a<b, a,b \in \mathbb{Q}} \{N([a,b], I_M, X) < \infty\} \cap \{\sup_{t \in I_M} |X_t| < \infty\},$$

then we obtain that $\mathbb{P}(\Omega_0) = 1$. For $\omega \in \Omega_0$ by Lemma 3.11 the following limits exist in $\mathbb{R}$:

$$X_{t+}(\omega) = \lim_{s \downarrow t, s \in \mathbb{Q}} X_s(\omega), \; t \geq 0$$
$$X_{t-}(\omega) = \lim_{s \uparrow t, s \in \mathbb{Q}} X_s(\omega), \; t > 0.$$

Hence we can now define for $t \geq 0$,

$$\widetilde{X}_t = \begin{cases} X_{t+}, & \text{on } \Omega_0; \\ 0, & \text{otherwise.} \end{cases}$$

Then clearly $\widetilde{X}$ is $\widetilde{\mathcal{F}}$ adapted, since $\widetilde{\mathcal{F}}$ contains also the events of 0 probability.

Let t_n be a sequence in $\mathbb{Q}$ such that $t_n \downarrow t$ as $n \to \infty$. Then

$$\widetilde{X}_t = \lim_{n\to\infty} X_{t_n}.$$

Notice that the process $(X_{t_n} : n \geq 1)$ is a backwards martingale, and hence it converges a.s. and in $\mathcal{L}^1$ as $n \to \infty$. Therefore,

$$\mathbb{E}[X_{t_i}|\mathcal{F}_t] \to \mathbb{E}[\widetilde{X}_t|\mathcal{F}_t] \text{ in } \mathcal{L}^1.$$

But $\mathbb{E}[X_{t_i}|\mathcal{F}_t] = X_t$. Therefore

$$X_t = \mathbb{E}[\widetilde{X}_t|\mathcal{F}_t] \text{ a.s..} \tag{3.3}$$

It remains to show the martingale property of $\widetilde{X}$. Let $s < t$ and s_n a sequence in $\mathbb{Q}$ such that $s_n \downarrow s$ and $s_0 < t$. Then

$$\widetilde{X}_s = \lim X_{s_n} = \lim \mathbb{E}[X_t|\mathcal{F}_{s_n}].$$

Now note that $(\mathbb{E}[X_t|\mathcal{F}_{s_n}])$ is a backwards martingale and hence it converges a.s. and in $\mathcal{L}^1$ to $\mathbb{E}[X_t|\mathcal{F}_{s+}]$. Therefore

$$\widetilde{X}_s = \mathbb{E}[X_t|\mathcal{F}_{s+}] \text{ a.s.} \tag{3.4}$$

If $s < t$, then by the tower property and (3.4) and (3.3) we get that

$$\mathbb{E}[\widetilde{X}_t|\mathcal{F}_{s+}] = \widetilde{X}_s \text{ a.s.}$$

Notice that if $\mathcal{G}$ is any σ-algebra and X is an integrable random variable, then

$$\mathbb{E}[X|\mathcal{G} \vee \mathcal{N}] = \mathbb{E}[X|\mathcal{G}] \text{ a.s.}$$

Finally we get that $\mathbb{E}[\widetilde{X}_t|\widetilde{\mathcal{F}}_s] = \widetilde{X}_s$ a.s., which shows that $\widetilde{X}$ is a martingale with respect to the filtration $\widetilde{\mathcal{F}}$.

The only thing that remains to prove is the cadlag property.

Suppose that for some $\omega \in \Omega_0$ we have that $\widetilde{X}$ is not right continuous. Then this means that there exists a sequence (s_n) such that $s_n \downarrow t$ as $n \to \infty$ and

$$|\widetilde{X}_{s_n} - \widetilde{X}_t| > \varepsilon,$$

for some $\varepsilon > 0$. By the definition of $\widetilde{X}$ for $\omega \in \Omega_0$, there exists a sequence of rational numbers (s'_n) such that $s'_n > s_n$, $s'_n \downarrow t$ as $n \to \infty$ and

$$|\widetilde{X}_{s_n} - X_{s'_n}| \leq \frac{\varepsilon}{2.}$$

Therefore, we get that

$$|X_{s'_n} - \widetilde{X}_t| > \frac{\varepsilon}{2},$$

which is a contradiction, since $X_{s'_n} \to \widetilde{X}_t$ as $n \to \infty$.

The proof that $\widetilde{X}$ has left limits is left as an exercise (*hint:* use the finite up-crossing property of X on rationals). □

Example 3.12. Let ξ, η be independent random variables taking values $+1$ or -1 with equal probability. We now define

$$X_t = \begin{cases} 0, & \text{if } t < 1; \\ \xi, & \text{if } t = 1; \\ \xi + \eta, & \text{if } t > 1. \end{cases}$$

We also define $\mathcal{F}_t$ to be the natural filtration, i.e. $\mathcal{F}_t = \sigma(X_s, s \leq t)$. Then clearly, X is a martingale relative to the filtration $(\mathcal{F}_t)$, but it is not right continuous at 1. Also, it is easy to see that $\mathcal{F}_1 = \sigma(\xi)$ but $\mathcal{F}_{1+} = \sigma(\xi, \eta)$. We now define

$$\widetilde{X}_t = \begin{cases} 0, & \text{if } t < 1; \\ \xi + \eta, & \text{if } t \geq 1. \end{cases}$$

It is easy to check that $X_t = \mathbb{E}[\widetilde{X}_t|\mathcal{F}_t]$ a.s. for all t and $\widetilde{X}$ is a martingale with respect to the filtration $(\mathcal{F}_{t+})$. It is obvious that $\widetilde{X}$ is cadlag. Note though that $\widetilde{X}$ is not a version of X, since $X_1 \neq \widetilde{X}_1$.

From now on when we work with martingales in continuous time, we will always consider their cadlag version, provided that the filtration satisfies the usual conditions.

3.3 Convergence and Doob's inequalities in continuous time

In this section we will give the continuous time analogues of Doob's inequalities and the convergence theorems for martingales.

Theorem 3.13. **[A.s. martingale convergence]** *Let $(X_t : t \geq 0)$ be a cadlag martingale which is bounded in $\mathcal{L}^1$. Then $X_t \to X_\infty$ a.s. as $t \to \infty$, for some $X_\infty \in \mathcal{L}^1(\mathcal{F}_\infty)$.*

***Proof*.** If $N([a, b], I_M, X)$ stands for the number of up-crossings of the interval $[a, b]$ as defined in Lemma 3.11, then from (3.2) in the proof of the martingale regularization theorem, we get that

$$(b - a)\mathbb{E}[N([a, b], I_M, X)] \leq a + \sup_{t\geq 0} \mathbb{E}[|X_t|] < \infty,$$

since X is bounded in $\mathcal{L}^1$. Hence, if we take the limit as $M \to \infty$ then we get that

$$N([a, b], \mathbb{Q}_+, X) < \infty \quad \text{a.s.}$$

Therefore, the set

$$\Omega_0 = \cap_{a<b, a,b\in\mathbb{Q}}\{N([a, b], \mathbb{Q}_+, X) < \infty\}$$

has probability 1. On Ω_0 it is easy to see that X_q converges as $q \to \infty$ and $q \in \mathbb{Q}_+$. Indeed, as in the proof of Lemma 2.15, if X_q did not converge, then $\limsup X_q \neq \liminf X_q$ and this would contradict the finite number of up-crossings of the interval $[a, b]$, where $\liminf < a < b < \limsup$. Thus $(X_q)_{q \in \mathbb{Q}_+}$ converges a.s. as $q \to \infty, q \in \mathbb{Q}_+$, to X_∞. We will now use the cadlag property of X to deduce that

$$X_t \to X_\infty \quad \text{as} \quad t \to \infty.$$

Since $X_q \to X_\infty$ as $q \to \infty, q \in \mathbb{Q}_+$, for each $\varepsilon > 0$, there exists q_0 such that

$$|X_q - X_\infty| < \frac{\varepsilon}{2}, \quad \text{for all } q > q_0.$$

By right continuity, we get that for $t > q_0$ there exists a rational q such that $q > t$ and

$$|X_t - X_q| < \frac{\varepsilon}{2}.$$

Hence we conclude that

$$|X_t - X_\infty| \leq \varepsilon.$$

□

Theorem 3.14. [Doob's maximal inequality] *Let $(X_t : t \geq 0)$ be a cadlag martingale and $X_t^* = \sup_{s \leq t} |X_s|$. Then, for all $\lambda \geq 0$ and $t \geq 0$*

$$\lambda \mathbb{P}(X_t^* \geq \lambda) \leq \mathbb{E}[|X_t|].$$

Proof. Notice that by the cadlag property we have

$$\sup_{s \leq t} |X_s| = \sup_{s \in \{t\} \cup ([0,t] \cap \mathbb{Q}_+)} |X_s|.$$

The rest of the proof follows in the same way as the first part of the proof of Theorem 3.10 □

Theorem 3.15. [Doob's $\mathcal{L}^p$-inequality] *Let $(X_t : t \geq 0)$ be a cadlag martingale. Setting $X_t^* = \sup_{s \leq t} |X_s|$, then for all $p > 1$ we have*

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

Theorem 3.16. [$\mathcal{L}^p$ martingale convergence theorem] *Let X be a cadlag martingale and $p > 1$, then the following statements are equivalent:*

1. *X is bounded in $\mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P}) : \sup_{t \geq 0} \|X_t\|_p < \infty$*
2. *X converges a.s. and in $\mathcal{L}^p$ to a random variable X_∞*
3. *There exists a random variable $Z \in \mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P})$ such that*

$$X_t = \mathbb{E}[Z | \mathcal{F}_t] \text{ a.s.}$$

Theorem 3.17. [UI martingale convergence theorem] *Let X be a cadlag martingale. Then X is UI if and only if X converges a.s. and in $\mathcal{L}^1$ to X_∞ and this if and only if X is closed.*

Theorem 3.18. [Optional stopping theorem] *Le X be a cadlag UI martingale. Then for every stopping times $S \le T$, we have*

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

Proof. Let $A \in \mathcal{F}_S$. We need to show tbat

$$\mathbb{E}[X_T \mathbf{1}(A)] = \mathbb{E}[X_S \mathbf{1}(A)].$$

Let $T_n = 2^{-n}\lceil 2^n T \rceil$ and $S_n = 2^{-n}\lceil 2^n S \rceil$. Then $T_n \downarrow T$ and$S_n \downarrow S$ as $n \to \infty$ and by the right continuity of X we get that

$$X_{S_n} \to X_S \quad \text{and} \quad X_{T_n} \to X_T \quad \text{as } n \to \infty.$$

Also, from the discrete time optional stopping theorem we have that $X_{T_n} = \mathbb{E}[X_\infty|\mathcal{F}_{T_n}]$ and thus we see that X_{T_n} is UI. Hence it converges to X_T as $n \to \infty$ also in $\mathcal{L}^1$. By the discrete time optional stopping theorem for UI martingales we have

$$\mathbb{E}[X_{T_n}|\mathcal{F}_{S_n}] = X_{S_n} \text{ a.s.} \tag{3.5}$$

Since $A \in \mathcal{F}_S$ the definition of S_n implies that $A \in \mathcal{F}_{S_n}$. Hence from (3.5) we obtain that

$$\mathbb{E}[X_{T_n} \mathbf{1}(A)] = \mathbb{E}[X_{S_n} \mathbf{1}(A)]$$

Letting $n \to \infty$ and using the $\mathcal{L}^1$ convergence of X_{T_n} to X_T and of X_{S_n} to X_S we have

$$\mathbb{E}[X_T \mathbf{1}(A)] = \mathbb{E}[X_S \mathbf{1}(A)].$$

□

3.4 Kolmogorov's continuity criterion

Let $\mathcal{D}_n = \{k2^{-n} : 0 \le k \le 2^n\}$ be the set of dyadic rationals of level n and $\mathcal{D} = \cup_{n \ge 0} \mathcal{D}_n$.

Theorem 3.19. [Kolmogorov's continuity criterion] *Let $(X_t)_{t \in \mathcal{D}}$ be a stochastic process with real values. Suppose there exists $p > 0, \varepsilon > 0$ so that*

$$\mathbb{E}[|X_t - X_s|^p] \le c|t-s|^{1+\varepsilon}, \text{ for all } s, t \in \mathcal{D},$$

for some constant $c < \infty$. Then for every $\alpha \in (0, \varepsilon/p)$, the process $(X_t)_{t \in \mathcal{D}}$ is α-Hölder continuous, i.e. there exists a random variable K_α such that

$$|X_t - X_s| \le K_\alpha |s-t|^\alpha, \text{ for all } s, t \in \mathcal{D}.$$

Proof. By Markov's inequality and the assumption we have

$$\mathbb{P}\left(|X_{k2^{-n}} - X_{(k+1)2^{-n}}| \geq 2^{-n\alpha}\right) \leq c2^{n\alpha p}2^{-n-n\varepsilon}.$$

By the union bound we have

$$\mathbb{P}\left(\max_{0\leq k<2^n} |X_{k2^{-n}} - X_{(k+1)2^{-n}}| \geq 2^{-n\alpha}\right) \leq c2^{-n(\varepsilon-p\alpha)}.$$

By Borel-Cantelli, since $\alpha \in (0, \varepsilon/p)$, we deduce

$$\max_{0\leq k<2^n} |X_{k2^{-n}} - X_{(k+1)2^{-n}}| \leq 2^{-n\alpha}, \text{ for all } n \text{ sufficiently large.}$$

Therefore, there exists a random variable M such that

$$\sup_{n\geq 0} \max_{0\leq k<2^n} \frac{|X_{k2^{-n}} - X_{(k+1)2^{-n}}|}{2^{-n\alpha}} \leq M < \infty. \tag{3.6}$$

We will now show that there exists a random variable $M' < \infty$ a.s. so that for every $s, t \in \mathcal{D}$ we have

$$|X_t - X_s| \leq M'|t-s|^{\alpha}.$$

Let $s, t \in \mathcal{D}$ and let r be the unique integer such that

$$2^{-(r+1)} < t - s \leq 2^{-r}.$$

Then there exists k such that $s < k2^{-(r+1)} < t$. Set $\alpha = k2^{-r+1}$, then $0 < t - \alpha < 2^{-r}$. So we have that

$$t - \alpha = \sum_{k\geq r+1} \frac{x_j}{2^j},$$

where $x_j \in \{0, 1\}$ for all j (in fact this is a finite sum because $t - \alpha$ is dyadic). Similarly we can write

$$\alpha - s = \sum_{j\geq r+1} \frac{y_j}{2^j},$$

where $y_j \in \{0, 1\}$ for all j. Thus we see that we can write the interval $[s, t)$ as a disjoint union of dyadic intervals of length 2^{-n} for $n \geq r+1$ and where at most 2 such intervals have the same length. Therefore,

$$|X_s - X_t| \leq \sum_{d,n} |X_d - X_{d+2^{-n}}|,$$

where $d, d+2^{-n}$ in the summation above are the endpoints of the intervals in the decomposition of $[s, t)$. Hence using (3.6) we obtain that for all $s, t \in \mathcal{D}$

$$|X_s - X_t| \leq 2 \sum_{n\geq r+1} M2^{-n\alpha} = 2M\frac{2^{-(r+1)\alpha}}{1-2^{-\alpha}}.$$

Thus, if we set $M' = 2M/(1-2^{-\alpha})$, then we get that for $s, t \in \mathcal{D}$

$$|X_s - X_t| \leq M'2^{-(r+1)\alpha} \leq M'|t-s|^{\alpha}.$$

Therefore we get that $(X_t)_{t\in\mathcal{D}}$ is α-Hölder continuous a.s. $\square$

4 Weak convergence

4.1 Definitions

Let (M, d) be a metric space endowed with its Borel σ-algebra. All the measures that we will consider in this section will be measures on such a measurable space.

Definition 4.1. Let $(\mu_n, n \geq 0)$ be a sequence of probability measures on a metric space (M, d). We say that μ_n converges *weakly* to μ and write $\mu_n \Rightarrow \mu$ if $\mu_n(f) \to \mu(f)$ as $n \to \infty$ for all bounded continuous functions f on M, where $\mu(f) = \int_M f \, d\mu$.

Notice that by the definition μ is also a probability measure, since $\mu(1) = 1$.

Example 4.2. Let $(x_n)_{n \geq 0}$ be a sequence in a metric space M that converges to x as $n \to \infty$. Then δ_{x_n} converges weakly to δ_x as $n \to \infty$, since if f is any continuous function, then $f(x_n) \to f(x)$ as $n \to \infty$.

Example 4.3. Let $M = [0, 1]$ with the Euclidean metric and $\mu_n = n^{-1} \sum_{0 \leq k \leq n-1} \delta_{k/n}$. Then $\mu_n(f)$ is the Riemann sum $n^{-1} \sum_{0 \leq k \leq n-1} f(k/n)$ and it converges to $\int_0^1 f(x) \, dx$ if f is continuous, which shows that μ_n converges weakly to Lebesgue measure on $[0, 1]$.

Remark 4.4. Notice that if A is a Borel set, then it is not always true that $\mu_n(A) \to \mu(A)$ as $n \to \infty$, when $\mu_n \Rightarrow \mu$. Indeed, let $x_n = 1/n$ and $\mu_n = \delta_{x_n}$. Then $\mu_n \Rightarrow \delta_0$, but $\mu_n(A) = 1$ for all n, when A is the open set $(0, 1)$, and $\delta_0(A) = 0$.

Theorem 4.5. *Let $(\mu_n)_{n \geq 0}$ be a sequence of probability measures. The following are equivalent:*

(a) $\mu_n \Rightarrow \mu$ *as* $n \to \infty$,

(b) $\liminf_n \mu_n(G) \geq \mu(G)$ *for all open sets* G,

(c) $\limsup_n \mu_n(A) \leq \mu(A)$ *for all closed sets* A,

(d) $\lim_n \mu_n(A) = \mu(A)$ *for all sets* A *with* $\mu(\partial A) = 0$.

Proof. (a) $\Longrightarrow$ (b). Let G be an open set with non-empty complement G^c. For every positive M we now define

$$f_M(x) = 1 \wedge (M d(x, G^c)).$$

Then f_M is a continuous and bounded function and for all M we have $f_M(x) \leq \mathbf{1}(x \in G)$. Also $f_M \uparrow \mathbf{1}(G)$ as $M \to \infty$, since G^c is a closed set. Since f_M is continuous and bounded we have

$$\mu_n(f_M) \to \mu(f_M) \text{ as } n \to \infty.$$

Hence

$$\liminf_n \mu_n(G) \geq \liminf_n \mu_n(f_M) = \mu(f_M).$$

Now using monotone convergence as $M \to \infty$ we get

$$\liminf_n \mu_n(G) \geq \mu(G).$$

(b) $\iff$ (c). This is obvious by taking complements.

(b),(c) $\implies$ (d). Let $\mathring{A}$ and $\bar{A}$ denote the interior and the closure of the set A respectively. Since

$$\mu(\partial A) = \mu(\bar{A} \setminus \mathring{A}) = 0,$$

we get that $\mu(\mathring{A}) = \mu(A) = \mu(\bar{A})$. Hence,

$$\limsup_n \mu_n(\bar{A}) \leq \mu(A) \leq \liminf_n \mu_n(\mathring{A})$$

and since $\mathring{A} \subseteq \bar{A}$ this gives the result.

(d) $\implies$ (a). Let $f : M \to \mathbb{R}_+$ be a continuous bounded non-negative function. Using Fubini's theorem we get

$$\int_M f(x)\mu_n(dx) = \int_M \mu_n(dx) \int_0^\infty \mathbf{1}(t \leq f(x))\, dt = \int_0^K \mu_n(\{f \geq t\})\, dt,$$

where K is an upper bound for f. We will now show that for Lebesgue almost all t we have

$$\mu(\partial\{f \geq t\}) = 0. \tag{4.1}$$

Notice that $\partial\{f \geq t\} \subseteq \{f = t\}$, since $\{f \geq t\}$ is a closed set by the continuity of f and $\{f > t\}$ is an open set contained in the interior. However, there can be at most a countable set of numbers t such that $\mu(\{f = t\}) > 0$, because

$$\{t : \mu(\{f = t\}) > 0\} = \cup_{n\geq 1}\{t : \mu(\{f = t\}) \geq n^{-1}\}$$

and the n-th set on the right has at most n elements. Hence this proves (4.1).

Therefore by (d) and dominated convergence on $\int_0^K \mu_n(\{f \geq t\})\, dt$ we get that

$$\mu_n(f) \to \mu(f) \text{ as } n \to \infty.$$

The extension to the case of a function f not necessarily positive is immediate. □

For a finite non-negative measure μ on $\mathbb{R}$ we define its *distribution function*

$$F_\mu(x) = \mu((-\infty, x]),\ x \in \mathbb{R}.$$

As a consequence of the theorem above we will now prove the following:

Proposition 4.6. *Let $(\mu_n)_n$ be a sequence of probability measures in $\mathbb{R}$. The following are equivalent:*

(a) *μ_n converges weakly to μ as $n \to \infty$,*

(b) *for every $x \in \mathbb{R}$ such that F_μ is continuous at x, the distribution functions $F_{\mu_n}(x)$ converges to $F_\mu(x)$ as $n \to \infty$.*

Proof. (a)$\Longrightarrow$(b). Let x be a continuity point of F_μ. Then

$$\mu(\partial(-\infty,x]) = \mu(\{x\}) = \mu((-\infty,x]) - \lim_{n\to\infty}\mu((-\infty,x-1/n]) = F_\mu(x) - \lim_{n\to\infty}F_\mu(x-1/n) = 0,$$

since x is a continuity point of F_μ. Thus we get that

$$F_{\mu_n}(x) = \mu_n((-\infty,x]) \to \mu((-\infty,x]),$$

by the 4-th equivalence in Theorem 4.5.

(b)$\Longrightarrow$(a). First of all note that a distribution function is increasing, and hence has only countably many points of discontinuity.

Let G be an open set in $\mathbb{R}$. Then we can write $G = \cup_k (a_k,b_k)$, where the intervals (a_k,b_k) are disjoint. We thus have that $\mu_n(G) = \sum_k \mu_n((a_k,b_k))$. For each interval (a,b) we have

$$\mu_n((a,b)) = F_{\mu_n}(b-) - F_{\mu_n}(a) \geq F_{\mu_n}(b') - F_{\mu_n}(a'),$$

where a',b' are continuity points of F_μ (remember there are only countably many points of discontinuity – set of continuity points is dense) satisfying

$$a < a' < b' < b.$$

Therefore

$$\liminf_n \mu_n((a,b)) \geq F_\mu(b') - F_\mu(a') = \mu((a',b'))$$

and hence if we let $a' \downarrow a$ and $b' \uparrow b$ along continuity points of F_μ, then

$$\liminf_n \mu_n((a,b)) \geq \mu(a,b). \tag{4.2}$$

Finally we deduce

$$\liminf_n \mu_n(G) = \liminf_n \sum_k \mu_n((a_k,b_k)) \geq \sum_k \liminf_n \mu_n((a_k,b_k)) \geq \sum_k \mu((a_k,b_k)) = \mu(G),$$

where the first inequality follows from Fatou's lemma and the second one from (4.2). □

Definition 4.7. Let $(X_n)_n$ be a sequence of random variables taking values in a metric space (M,d) but defined on possibly different probability spaces $(\Omega_n\mathcal{F}_n,\mathbb{P}_n)$. We say that X_n converges *in distribution* to a random variable X defined on the probability space $(\Omega,\mathcal{F},\mathbb{P})$ if the law of X_n converges weakly to the law of X as $n\to\infty$. Equivalently, if for all functions $f : M \to \mathbb{R}$ continuous and bounded

$$\mathbb{E}_{\mathbb{P}_n}[f(X_n)] \to \mathbb{E}_{\mathbb{P}}[f(X)] \text{ as } n\to\infty.$$

Proposition 4.8 (a)**.** *Let $(X_n)_n$ be a sequence of random variables that converges to X in probability as $n\to\infty$. Then X_n converges to X in distribution to X as $n\to\infty$.*

(b)*. Let $(X_n)_n$ be a sequence of random variables that converges to a constant c in distribution as $n\to\infty$. Then X_n converges to c in probability as $n\to\infty$.*

Proof. See example sheet. □

Example 4.9. [Central limit theorem] Let $(X_n)_n$ be a sequence of i.i.d. random variables in $\mathcal{L}^2$ with $m = \mathbb{E}[X_1]$ and $\sigma^2 = \mathrm{var}(X_1)$. We set $S_n = X_1 + \ldots + X_n$. Then the central limit theorem states that the normalized sums $(S_n - nm)/\sigma\sqrt{n}$ converge in distribution to a Gaussian $\mathcal{N}(0,1)$ random variable as $n\to\infty$.

4.2 Tightness

Definition 4.10. A sequence of probability measures $(\mu_n)_n$ on a metric space M is said to be *tight* if for every $\varepsilon > 0$, there exists a compact subset $K \subseteq M$ such that

$$\sup_n \mu_n(M \setminus K) \leq \varepsilon.$$

Remark 4.11. Note that if a metric space M is compact, then every sequence of measures is tight.

Theorem 4.12. [Prohorov's theorem] *Let $(\mu_n)_n$ be a tight sequence of probability measures on a metric space M. Then there exists a subsequence (n_k) and a probability measure μ on M such that*

$$\mu_{n_k} \Rightarrow \mu.$$

Proof. We will prove the theorem in the case when $M = \mathbb{R}$. Let $F_n = F_{\mu_n}$ be the distribution function corresponding to the measure μ_n. We will first show that there exists a subsequence n_k and a non-decreasing function F such that $F_{n_k}(x)$ converges to $F(x)$ for all $x \in \mathbb{Q}$. To prove that we will use a standard *extraction argument*.

Let $(x_1, x_2, \ldots)$ be an enumeration of $\mathbb{Q}$. Then $(F_n(x_1))_n$ is a sequence in $[0, 1]$, and hence it has a converging subsequence. Let the converging subsequence be $F_{n_k^{(1)}}(x_1)$ and the limit $F(x_1)$. Then $(F_{n_k^{(1)}}(x_2))_k$ is a sequence in $[0, 1]$ and thus also has a converging subsequence. If we continue in this way, we get for each $i \geq 1$ a sequence $n_k^{(i)}$ so that $F_{n_k^{(i)}}(x_j)$ converges to a limit $F(x_j)$ for all $j = 1, \ldots, i$. Then the diagonal sequence $m_k = n_k^{(k)}$satisfies that $F_{m_k}(x)$ converges for all $x \in \mathbb{Q}$ to $F(x)$ as $k \to \infty$. Since the distribution functions $F_n(x)$ are non-decreasing in x, then we get that $F(x)$ is also non-decreasing in x.

By the monotonicity of F we can define for all $x \in \mathbb{R}$

$$F(x) = \lim_{q \downarrow x, q \in \mathbb{Q}} F(q).$$

The definition of F gives that it is right continuous and the monotonicity property gives that left limits exist, hence F is cadlag.

We will next show that if t is a point of continuity of F, i.e. $F(t) = F(t-)$, then

$$\lim_{k \to \infty} F_{m_k}(t) = F(t).$$

Let $s_1 < t < s_2$ with $s_1, s_2 \in \mathbb{Q}$ and such that $|F(s_i) - F(t)| < \varepsilon/2$ for $i = 1, 2$. Note that such rational numbers s_1 and s_2 exist since t is a continuity point of F. Then using the monotonicity property of F_{m_k} we get that for k large enough

$$F(t) - \varepsilon < F(s_1) - \frac{\varepsilon}{2} < F_{m_k}(s_1) \leq F_{m_k}(t) \leq F_{m_k}(s_2) < F(s_2) + \frac{\varepsilon}{2} < F(t) + \varepsilon.$$

By tightness, for every $\varepsilon > 0$, there exists N such that

$$\mu_n([-N, N]^c) \leq \varepsilon \ \ \forall n.$$

Note that we can choose N so that both $N$and $-N$ are continuity points of F (F is monotone). Therefore it follows that

$$F(-N) \leq \varepsilon \text{ and } 1 - F(N) \leq \varepsilon.$$

Hence we see that

$$\lim_{x \to -\infty} F(x) = 0 \text{ and } \lim_{x \to \infty} F(x) = 1.$$

Finally we need to show that there exists a measure μ such that $F = F_\mu$. To this end, we define

$$\mu((a, b]) = F(b) - F(a).$$

Then μ can be extended to a Borel probability measure by Carathéodory's extension theorem and $F = F_\mu$. Another way to construct the measure μ is given in [2, Section 3.12].

Proposition 4.6 now finishes the proof. $\square$

4.3 Characteristic functions

Definition 4.13. Let X be a random variable taking values in $\mathbb{R}^d$ with law μ. We define the characteristic function $\varphi = \varphi_X$ by

$$\varphi(u) = \mathbb{E}[e^{i\langle u, X\rangle}] = \int_{\mathbb{R}^d} e^{i\langle u, x\rangle} \mu(dx), \ u \in \mathbb{R}^d.$$

Remark 4.14. The characteristic function of a random variable X is clearly a continuous function on $\mathbb{R}^d$ and $\varphi(0) = 1$.

The characteristic function φ_X determines the law of a random variable X, in the sense that if $\varphi_X(u) = \varphi_Y(u)$ for all u, then $\mathcal{L}(X) = \mathcal{L}(Y)$. To prove this see the Probability and Measure notes by James Norris, Theorem 7.2.2.

Theorem 4.15. [Lévy's convergence theorem] *Let $(X_n)_{n\geq 0}, X$ be random variables in $\mathbb{R}^d$. Then*

$$\mathcal{L}(X_n) \Rightarrow \mathcal{L}(X) \ \textit{ if and only if } \ \varphi_{X_n}(\xi) \to \varphi_X(\xi) \ \forall \xi \in \mathbb{R}^d.$$

We will prove the more general result:

Theorem 4.16. [Lévy] *1. If $\mathcal{L}(X_n) \Rightarrow \mathcal{L}(X)$ as $n \to \infty$, then $\varphi_{X_n}(\xi) \to \varphi_X(\xi)$ as $n \to \infty$ for all $\xi \in \mathbb{R}^d$.*
2. If $(X_n)_{n\geq 0}$ is a sequence of random variables in $\mathbb{R}^d$ such that there exists $\psi : \mathbb{R}^d \to \mathbb{C}$ continuous at 0 with $\psi(0) = 1$ and such that $\varphi_{X_n}(\xi) \to \psi(\xi)$ as $n \to \infty$ for all $\xi \in \mathbb{R}^d$, then $\psi = \varphi_X$, for some X and $\mathcal{L}(X_n) \Rightarrow \mathcal{L}(X)$ as $n \to \infty$.

Before giving the proof of Lévy's theorem we state and prove a useful lemma:

Lemma 4.17. *If X is a random variable in $\mathbb{R}^d$, then for all $K > 0$*

$$\mathbb{P}(\|X\|_\infty > K) \leq C(K/2)^d \int_{[-K^{-1}, K^{-1}]^d} (1 - \Re\varphi_X(\xi))\, d\xi,$$

where $C = (1 - \sin 1)^{-1}$.

Proof. Let μ be the distribution of X. Then by Fubini's theorem we have

$$\int_{[-\lambda,\lambda]^d} \Re\varphi_X(u)\,du = \int_{[-\lambda,\lambda]^d} \Re\left(\int e^{i\langle u,x\rangle} d\mu(x)\right)\,du = \Re\int \prod_{j=1}^{d} \int_{[-\lambda,\lambda]} e^{iu_j x_j}\,du_j\,d\mu(x)$$
$$= \Re\int \prod_{j=1}^{d}\left(\frac{1}{ix_j}\left(e^{i\lambda x_j} - e^{-i\lambda x_j}\right)\right)\,d\mu(x) = \int \prod_{j=1}^{d}\left(\frac{2\sin(\lambda x_j)}{x_j}\right)\,d\mu(x).$$

Therefore we have

$$\frac{1}{\lambda^d}\int_{[-\lambda,\lambda]^d}(1-\Re\varphi_X(u))\,du = 2^d\int_{\mathbb{R}^d} d\mu(x)\left(1-\prod_{j=1}^{d}\frac{\sin(\lambda x_j)}{\lambda x_j}\right). \tag{4.3}$$

It is easy to check that if $x \geq 1$, then

$$|\sin x| \leq x\sin 1,$$

and hence the function $f:\mathbb{R}^d \to \mathbb{R}$ given by $f(u) = \prod_{j=1}^{d}\sin u_j/u_j$ satisfies $|f(u)| \leq \sin 1$ when $\|u\|_\infty \geq 1$. Thus for $C = (1-\sin 1)^{-1}$ we have

$$\mathbf{1}(\|u\|_\infty \geq 1) \leq C(1-f(u)).$$

Hence, we have

$$\mathbb{P}(\|X\|_\infty \geq K) = \mathbb{P}\left(\left\|\frac{X}{K}\right\|_\infty \geq 1\right) \leq C\mathbb{E}\left[1-\prod_{j=1}^{d}\frac{\sin(K^{-1}X_j)}{K^{-1}X_j}\right]$$
$$= C\int_{\mathbb{R}^d} d\mu(x)\left(1-\prod_{j=1}^{d}\frac{\sin(K^{-1}x_j)}{K^{-1}x_j}\right).$$

Equation (4.3) now finishes the proof. □

Proof of Theorem 4.16. 1. If X_n converges in distribution to X as $n\to\infty$, then for all f continuous and bounded, writing $\mu_n = \mathcal{L}(X_n)$ and $\mu = \mathcal{L}(X)$, we have

$$\mu_n(f) \to \mu(f) \quad \text{as} \quad n\to\infty.$$

Take $f(x) = e^{i\langle\xi,x\rangle}$. Then f is clearly continuous and bounded, and hence

$$\varphi_{X_n}(\xi) = \mu_n(e^{i\langle\xi,\cdot\rangle}) \to \mu(e^{i\langle\xi,\cdot\rangle}) = \varphi_X(\xi).$$

2. We will first show that the sequence $(\mathcal{L}(X_n))$ is tight. From Lemma 4.17 we have that for all $K>0$

$$\mathbb{P}(\|X_n\|_\infty > K) \leq C_d K^d \int_{[-K^{-1},K^{-1}]^d}(1-\Re\varphi_{X_n}(u))\,du.$$

By the assumption and since $|1-\Re\varphi_{X_n}(u)| \leq 2$ for all n using the dominated convergence theorem we have

$$\lim_n K^d\int_{[-K^{-1},K^{-1}]^d}(1-\Re\varphi_{X_n}(u))\,du = K^d\int_{[-K^{-1},K^{-1}]^d}(1-\Re\psi(u))\,du.$$

Since ψ is continuous at 0, if we take K large enough we can make this limit $< \varepsilon/(2C_d)$ and so for all n large enough

$$\mathbb{P}(\|X_n\|_\infty > K) \leq \varepsilon.$$

If we now take K even larger, then the above inequality holds for all n showing the tightness of the family $(\mathcal{L}(X_n))$.

By Prohorov's theorem there exists a subsequence (X_{n_k}) that converges in distribution to some random variable X. So $\varphi_{X_{n_k}}$ converges pointwise to φ_X, and hence $\varphi_X = \psi$, which shows that ψ is a characteristic function.

We will finally show that X_n converges in distribution to X. If not, then there would exist a subsequence (m_k) and a continuous and bounded function f such that for some $\varepsilon > 0$ and all k

$$|\mathbb{E}[f(X_{m_k})] - \mathbb{E}[f(X)]| > \varepsilon. \tag{4.4}$$

But since the laws of (X_{m_k}) are tight, we can extract a subsequence (ℓ_k) along which (X_{ℓ_k}) converges in distribution to some Y, which would imply that $\psi = \varphi_Y$ and thus Y would have the same distribution as X, contradicting (4.4). □

5 Large deviations

5.1 Introduction

Let $\{X_i\}$ be a sequence of i.i.d. random variables with $\mathbb{E}[X_1] = \bar{x}$ and we set $S_n = \sum_{i=1}^n X_i$. By the central limit theorem (assuming $\mathrm{var}(X_i) = \sigma^2 < \infty$) we have

$$\mathbb{P}(S_n \geq n\bar{x} + a\sigma\sqrt{n}) \to \mathbb{P}(Z \geq a) \quad \text{as} \quad n \to \infty,$$

where $Z \sim \mathcal{N}(0,1)$.

Large deviations: What are the asymptotics of $\mathbb{P}(S_n \geq an)$ as $n \to \infty$, for $a > \bar{x}$?

Example 5.1. Let X_i be i.i.d. distributed as $\mathcal{N}(0,1)$. Then

$$\mathbb{P}(S_n \geq an) = \mathbb{P}(X_1 \geq a\sqrt{n}) \sim \frac{1}{a\sqrt{2\pi n}} e^{-a^2 n/2},$$

where we write $f(x) \sim g(x)$ if $f(x)/g(x) \to 1$ as $x \to \infty$. So

$$-\frac{1}{n} \log \mathbb{P}(S_n \geq an) \to I(a) = \frac{a^2}{2} \text{ as } n \to \infty.$$

In general we have

$$\mathbb{P}(S_{n+m} \geq a(n+m)) \geq \mathbb{P}(S_n \geq an)\mathbb{P}(S_m \geq am),$$

so $b_n = -\log \mathbb{P}(S_n \geq an)$ satisfies that

$$b_{n+m} \leq b_n + b_m,$$

and hence this implies the existence of the limit (exercise)

$$\lim \frac{b_n}{n} = \lim -\frac{1}{n} \log \mathbb{P}(S_n \geq an) = I(a).$$

Note that if $\mathbb{P}(X_1 \leq a_0) = 1$, then we will only consider $a \leq a_0$, since clearly $\mathbb{P}(S_n \geq na) = 0$ for $a > a_0$.

5.2 Cramer's theorem

We will now obtain a bound for $\mathbb{P}(S_n \geq na)$ using the moment generating function of X_1. For $\lambda \geq 0$ we set

$$M(\lambda) = \mathbb{E}[e^{\lambda X_1}],$$

which could also be infinite. We define

$$\Psi(\lambda) = \log M(\lambda).$$

Note that $\Psi(0) = 0$ and by Markov's inequality for $\lambda \geq 0$

$$\mathbb{P}(S_n \geq na) = \mathbb{P}(e^{\lambda S_n} \geq e^{\lambda na}) \leq e^{-\lambda na}\mathbb{E}[e^{\lambda S_n}] = \left(e^{-\lambda a}M(\lambda)\right)^n = \exp(-n(\lambda a - \Psi(\lambda)). \quad (5.1)$$

We now define the **Legendre transform** of Ψ:

$$\Psi^*(a) = \sup_{\lambda \geq 0}(\lambda a - \Psi(\lambda)) \geq -\Psi(0) = 0.$$

Then (5.1) yields

$$\mathbb{P}(S_n \geq an) \leq e^{-n\Psi^*(a)} \ \forall n,$$

whence

$$\liminf_{n\to\infty} -\frac{1}{n}\log \mathbb{P}(S_n \geq an) \geq \Psi^*(a). \tag{5.2}$$

Theorem 5.2. [**Cramer's theorem**] *Let (X_i) be i.i.d. random variables with $\mathbb{E}[X_1] = \bar{x}$ and $S_n = \sum_{i=1}^n X_i$. Then*

$$\lim_{n\to\infty} -\frac{1}{n}\log \mathbb{P}(S_n \geq na) = \Psi^*(a) \quad \text{for } a \geq \bar{x}.$$

Before proving the theorem we state and prove a preliminary lemma.

Lemma 5.3. *The functions $M(\lambda)$ and $\Psi(\lambda)$ are continuous in $D = \{\lambda : M(\lambda) < \infty\}$ and differentiable in $\mathring{D}$ with*

$$M'(\lambda) = \mathbb{E}[X_1 e^{\lambda X_1}] \text{ and } \Psi'(\lambda) = \frac{M'(\lambda)}{M(\lambda)} \text{ for } \lambda \in \mathring{D}.$$

Proof. Continuity follows immediately from the dominated convergence theorem.

Note that D is a (possibly infinite) interval i.e. if $\lambda_1 < \lambda < \lambda_2$ and $\lambda_1, \lambda_2 \in D$, then also $\lambda \in D$, since $e^{\lambda x} \leq e^{\lambda_1 x} + e^{\lambda_2 x}$ for all x.

To show that it is differentiable, note that

$$\frac{M(\lambda + h) - M(\lambda)}{h} = \mathbb{E}\left[\frac{e^{(\lambda+h)X} - e^{\lambda X}}{h}\right]$$

and for $2|h| < \min_{i=1,2}|\lambda - \lambda_i| = 2\varepsilon$ we have

$$\left|\frac{e^{(\lambda+h)x} - e^{\lambda x}}{h}\right| = |x|e^{\tilde{\lambda}} \leq (e^{\lambda_1 x} + e^{\lambda_2 x})\varepsilon^{-1},$$

where $\tilde{\lambda}$ is in $[\lambda_1, \lambda_2]$ if $|h| < \min_i |\lambda - \lambda_i| = 2\varepsilon$. $\square$

Proof of Theorem 5.2. The direction

$$\lim_{n\to\infty} -\frac{1}{n}\log \mathbb{P}(S_n \geq na) \geq \Psi^*(a)$$

follows from (5.2).

Replacing X_i by $\tilde{X}_i = X_i - a$ yields

$$\mathbb{P}(S_n \geq na) = \mathbb{P}(\tilde{S}_n \geq 0)$$

and $\widetilde{M}(\lambda) = \mathbb{E}[e^{\lambda \widetilde{X}_1}] = e^{-\lambda a} M(\lambda)$ so

$$\widetilde{\Psi}(\lambda) = \log \widetilde{M}(\lambda) = \Psi(\lambda) - \lambda a.$$

Thus we need to show that

$$-\frac{1}{n} \log \mathbb{P}(\widetilde{S}_n \geq 0) \to \Psi^*(0) = \sup_{\lambda \geq 0} [-\Psi(\lambda)].$$

In view of (5.2) what remains is (dropping tildes)

$$\liminf \frac{1}{n} \log \mathbb{P}(S_n \geq 0) \geq \inf_{\lambda \geq 0} \Psi(\lambda) \tag{5.3}$$

when $\bar{x} < 0$.

If $\mathbb{P}(X_1 \leq 0) = 1$, then

$$\inf_{\lambda \geq 0} \Psi(\lambda) \leq \lim_{\lambda \to \infty} \Psi(\lambda) = \log \mu(0),$$

where $\mu = \mathcal{L}(X_1)$, so (5.3) holds in this case. Thus we may assume that $\mathbb{P}(X_1 > 0) > 0$.

Next consider the case $M(\lambda) < \infty$ for all λ. Define a new law μ_θ where

$$\frac{d\mu_\theta}{d\mu}(x) = \frac{e^{\theta x}}{M(\theta)}, \quad \text{so} \quad \mathbb{E}_\theta[f(X_1)] = \int f(x) \frac{e^{\theta x}}{M(\theta)}\, d\mu(x).$$

More generally

$$\mathbb{E}_\theta[F(X_1, \ldots, X_n)] = \int F(x_1, \ldots, x_n) \prod_{i=1}^{n} e^{\theta x_i} \frac{d\mu(x_1) \ldots d\mu(x_n)}{M(\theta)^n}$$

holds when $F(x_1, \ldots, x_n) = \prod_{i=1}^n f_i(x_i)$, and hence for all bounded measurable F.

The dominated convergence theorem gives that $g(\theta) = \mathbb{E}_\theta[X_1]$ is continuous and $g(0) = \bar{x} < 0$, while

$$\lim_{\theta \uparrow \infty} g(\theta) = \lim_{\theta \uparrow \infty} \frac{\int x e^{\theta x} d\mu}{\int e^{\theta x} d\mu} > 0,$$

since $\mu(0, \infty) > 0$. Thus we can find $\theta > 0$ such that $\mathbb{E}_\theta[X_1] = 0$.

We now have

$$\mathbb{P}(S_n \geq 0) \geq \mathbb{P}(S_n \in [0, \varepsilon n]) \geq \mathbb{E}\left[e^{\theta(S_n - \varepsilon n)} \mathbf{1}(S_n \in [0, \varepsilon n])\right] = M(\theta)^n \mathbb{P}_\theta(S_n \in [0, \varepsilon n]) e^{-\theta \varepsilon n}.$$

By the central limit theorem we have that $\mathbb{P}_\theta(S_n \in [0, \varepsilon n]) \to 1/2$ as $n \to \infty$ so

$$\liminf_{n \to \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq 0) \geq \Psi(\theta) - \theta \varepsilon.$$

Letting $\varepsilon \downarrow 0$ proves (5.3) in the case where $M(\lambda) < \infty$ for all λ.

Now we are going to prove the theorem in the general case. Let $\mu_n = \mathcal{L}(S_n)$ and ν the law of X_1 conditioned on $\{|X_1| \leq K\}$ and ν_n the law of $S_n = \sum_{i=1}^n X_i$ conditioned on the event $\cap_{i=1}^n \{|X_i| \leq K\}$. Then we have that $\mu_n[0,\infty) \geq \nu_n[0,\infty)\mu[-K,K]^n$. We write $\Psi_K(\lambda) = \log \int_{-K}^K e^{\lambda x}\, d\mu(x)$ and observe that

$$\log \int_{-\infty}^{\infty} e^{\lambda x}\, d\nu(x) = \Psi_K(\lambda) - \log \mu[-K,K].$$

Therefore

$$\liminf \frac{1}{n} \log \mu_n[0,\infty) \geq \log \mu[-K,K] + \liminf \frac{1}{n} \log \nu_n[0,\infty) \geq \inf_{\lambda \geq 0} \Psi_K(\lambda) = J_K.$$

Note that $\Psi_K \uparrow \Psi$ as $K \to \infty$, so $J_K \uparrow J$ as $K \to \infty$, for some J, and

$$\liminf_{n\to\infty} \frac{1}{n} \log \mu_n[0,\infty) \geq J. \tag{5.4}$$

Since $J_K \leq \Psi_K(0) \leq \Psi(0) = 0$, so we have $J \leq 0$.

For large K we have that $\mu[0,K] > 0$, and hence $J_K > -\infty$ whence $J > -\infty$. By the continuity of Ψ_K (Lemma 5.3) the level sets $\{\lambda : \Psi_K(\lambda) \leq J\}$ are non-empty compact nested sets, so there exists

$$\lambda_0 \in \bigcap_K \{\lambda : \Psi_K(\lambda) \leq J\}.$$

Therefore we obtain

$$\Psi(\lambda_0) = \lim_K \Psi_K(\lambda_0) \leq J,$$

and hence by (5.4) we get

$$\liminf_{n\to\infty} \frac{1}{n} \log \mu_n[0,\infty) \geq \Psi(\lambda_0) \geq \inf_{\lambda \geq 0} \Psi(\lambda)$$

as claimed. □

5.3 Examples

Example 5.4. Let $X \sim \mathcal{N}(0,1)$, then

$$M(\lambda) = \int \frac{1}{\sqrt{2\pi}} e^{\lambda x - x^2/2}\, dx = e^{\lambda^2/2}, \text{ so } \Psi(\lambda) = \frac{\lambda^2}{2}.$$

In order to minimize $a\lambda - \Psi(\lambda)$ we need to solve $a = \Psi'(\lambda) = \lambda$, and hence

$$\Psi^*(a) = a^2 - \frac{a^2}{2} = \frac{a^2}{2}.$$

Example 5.5. Let $X \sim \text{Exp}(1)$. If $\lambda < 1$, then

$$M(\lambda) = \int e^{\lambda x - x}\, dx = \frac{1}{1-\lambda}.$$

So for $\lambda < 1$ we have $\Psi(\lambda) = -\log(1-\lambda)$ and for $\lambda \geq 1$ we have $M(\lambda) = \infty$, and thus $\Psi(\lambda) = \infty$. Solving $a = \Psi'(\lambda)$ gives that $a = \frac{1}{1-\lambda}$ or equivalently that $\lambda = 1 - \frac{1}{a}$, and hence

$$\Psi^*(a) = a - 1 - \log a.$$

Example 5.6. Let $X \sim \text{Poisson}(1)$. Then

$$M(\lambda) = \sum_{k=0}^{\infty} \frac{1}{k!} e^{\lambda k - 1} = e^{e^{\lambda} - 1},$$

so $\Psi(\lambda) = e^{\lambda} - 1$. Solving $a = \Psi'(\lambda)$ gives that $a = e^{\lambda}$, and hence

$$\Psi^*(a) = a \log a - a + 1.$$

6 Brownian motion

6.1 History and definition

Brownian motion is named after R. Brown who observed in 1827 the erratic motion of small particles in water. A physical model was developed by Einstein in 1905 and the mathematical construction is due to N. Wiener in 1923. He used a random Fourier series to construct Brownian motion. Our treatment follows later ideas of Lévy and Kolmogorov.

Definition 6.1. Let $B = (B_t)_{t \geq 0}$ be a continuous process in $\mathbb{R}^d$. We say that B is a *Brownian motion* in $\mathbb{R}^d$ started from $x \in \mathbb{R}^d$ if

(i) $B_0 = x$ a.s.,

(ii) $B_t - B_s \sim \mathcal{N}(0, (t-s)I_d)$, for all $s < t$,

(iii) B has independent increments, independent of B_0.

Remark 6.2. We say that $(B_t)_{t \geq 0}$ is a *standard Brownian motion* if $x = 0$.

Conditions (ii) and (iii) uniquely determine the law of a Brownian motion. In the next section we will show that Brownian motion exists.

Example 6.3. Suppose tht $(B_t, t \geq 0)$ is a standard Brownian motion and U is an independent random variable uniformly distributed on $[0,1]$. Then the process $(\widetilde{B}_t, t \geq 0)$ defined by

$$\widetilde{B}_t = \begin{cases} B_t, & \text{if } t \neq U; \\ 0, & \text{if } t = U \end{cases}$$

has the same finite-dimensional distributions as Brownian motion, but is discontinuous if $B(U) \neq 0$, which happens with probability one, and hence it is not a Brownian motion.

6.2 Wiener's theorem

Theorem 6.4. [Wiener's theorem] *There exists a Brownian motion on some probability space.*

Proof. We will first prove the theorem in dimension $d = 1$ and we will construct a process $(B_t, 0 \le t \le 1)$ and then extend it to the whole of $\mathbb{R}_+$ and to higher dimensions.

Let $\mathcal{D}_0 = \{0, 1\}$ and $\mathcal{D}_n = \{k2^{-n}, 0 \le k \le 2^n\}$ for $n \ge 1$, and $\mathcal{D} = \cup_{n\ge 0}\mathcal{D}_n$ be the set of dyadic rational numbers in $[0, 1]$. Let $(Z_d, d \in \mathcal{D})$ be a sequence of independent random variables distributed according to $\mathcal{N}(0, 1)$ on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We will first construct $(B_d, d \in \mathcal{D})$ inductively.

First set $B_0 = 0$ and $B_1 = Z_1$. Inductively, given that we have constructed $(B_d, d \in \mathcal{D}_{n-1})$ satisfying the conditions of the definition, we build $(B_d, d \in \mathcal{D}_n)$ as follows:
Take $d \in \mathcal{D}_n \setminus \mathcal{D}_{n-1}$ and let $d_- = d - 2^{-n}$ and $d_+ = d + 2^{-n}$, so that d_-, d_+ are consecutive dyadic numbers in $\mathcal{D}_{n-1}$. We write

$$B_d = \frac{B_{d_-} + B_{d_+}}{2} + \frac{Z_d}{2^{(n+1)/2}}.$$

Then we have

$$B_d - B_{d_-} = \frac{B_{d_+} - B_{d_-}}{2} + \frac{Z_d}{2^{(n+1)/2}} \quad \text{and} \quad B_{d_+} - B_d = \frac{B_{d_+} - B_{d_-}}{2} - \frac{Z_d}{2^{(n+1)/2}}. \tag{6.1}$$

Setting $N_d = \frac{B_{d_+} - B_{d_-}}{2}$ and $N_{d'} = \frac{Z_d}{2^{(n+1)/2}}$, we see by the induction hypothesis that N_d and $N_{d'}$ are independent centred Gaussian random variables with variance 2^{-n-1}. Therefore

$$\mathrm{Cov}(N_d + N_{d'}, N_d - N_{d'}) = \mathrm{var}(N_d) - \mathrm{var}(N_{d'}) = 0,$$

and hence the two new increments $B_d - B_{d_-}$ and $B_{d_+} - B_d$, being Gaussian, are independent.

Indeed, all increments $(B_d - B_{d-2^{-n}})$ for $d \in \mathcal{D}_n$ are independent. To see this it suffices to show that they are pairwise independent, as the vector of increments is Gaussian. Above we showed that increments over consecutive intervals are independent. If they are defined over intervals that are not consecutive, then notice that the increment is equal to half the increment of the previous scale plus an independent Gaussian random variable by (6.1), and hence this shows the claimed independence.

We have thus defined a process $(B_d, d \in \mathcal{D})$ satisfying the properties of Brownian motion. Let $s \le t \in \mathcal{D}$ and notice that for every $p > 0$, since $B_t - B_s \sim \mathcal{N}(0, t - s)$, we have

$$\mathbb{E}[|B_t - B_s|^p] = |t - s|^{p/2}\mathbb{E}[|N|^p],$$

where $N \sim \mathcal{N}(0, 1)$. Since N has moments of all orders, it follows by Kolmogorov's continuity criterion, Theorem 3.19, that $(B_d, d \in \mathcal{D})$ is α-Hölder continuous for all $\alpha < 1/2$ a.s. Hence in order to extend to the whole of $[0, 1]$ we simply let for $t \in [0, 1]$

$$B_t = \lim_{i\to\infty} B_{d_i},$$

where d_i is a sequence in $\mathcal{D}$ converging to t. It follows easily that $(B_t, t \in [0,1])$ is α-Hölder continuous for all $\alpha < 1/2$ a.s.

Finally we will check that $(B_t, t \in [0,1])$ has the properties of Brownian motion. We will first prove the independence of the increments property. Let $0 = t_0 < t_1 < \ldots < t_k$ and let $0 = t_0^n \leq t_1^n \leq \ldots \leq t_k^n$ be dyadic rational numbers such that $t_i^n \to t_i$ as $n \to \infty$ for each i.

By continuity $(B_{t_1^n}, \ldots, B_{t_k^n})$ converges a.s. to $(B_{t_1}, \ldots, B_{t_k})$ as $n \to \infty$, while on the other hand the increments $(B_{t_j^n} - B_{t_{j-1}^n}, 1 \leq j \leq k)$ are independent Gaussian random variables with variances $(t_j^n - t_{j-1}^n, 1 \leq j \leq k)$. Then as $n \to \infty$ we have

$$\mathbb{E}\left[\exp\left(i\sum_{j=1}^k u_j(B_{t_j^n} - B_{t_{j-1}^n})\right)\right] = \prod_{j=1}^k e^{-(t_j^n - t_{j-1}^n)u_j^2/2} \to \prod_{j=1}^k e^{-(t_j - t_{j-1})u_j^2/2}.$$

By Lévy's convergence theorem we now see that the increments converge in distribution to independent Gaussian random variables with respective variances $t_j - t_{j-1}$, which is thus the distribution of $(B_{t_j} - B_{t_{j-1}}, 1 \leq j \leq k)$ as desired.

To finish the proof we will construct Brownian motion indexed by $\mathbb{R}_+$. To this end, take a sequence $(B_t^i, t \in [0,1])$ for $i = 0, 1, \ldots$ of independent Brownian motions and glue them together, more precisely by

$$B_t = B_{t-\lfloor t \rfloor}^{\lfloor t \rfloor} + \sum_{i=0}^{\lfloor t \rfloor - 1} B_1^i.$$

This defines a continuous random process $B : [0, \infty) \to \mathbb{R}$ and it is easy to see from what we have already shown that B satisfies the properties of a Brownian motion.

Finally to construct Brownian motion in $\mathbb{R}^d$ we take d independent Brownian motions in 1 dimension, $B^1, \ldots, B^d$, and set $B_t = (B_t^1, \ldots, B_t^d)$. Then it is straightforward to check that B has the required properties. $\square$

Remark 6.5. The proof above gives that the Brownian paths are a.s. α-Hölder continuous for all $\alpha < 1/2$. However, a.s. there exists no interval $[a, b]$ with $a < b$ such that B is Hölder continuous with exponent $\alpha \geq 1/2$ on $[a, b]$. See example sheet for the last fact.

6.3 Invariance properties

The following invariance properties of Brownian motion will be used a lot.

Proposition 6.6. *Let B be a standard Brownian motion in $\mathbb{R}^d$.*

1. *If U is an orthogonal matrix, then $UB = (UB_t, t \geq 0)$ is again a standard Brownian motion. In particular, $-B$ is a standard Brownian motion.*

2. *If $\lambda > 0$, then $(\lambda^{-1/2} B_{\lambda t}, t \geq 0)$ is a standard Brownian motion (scaling property).*

3. *For every $t \geq 0$, the shifted process $(B_{t+s} - B_s, t \geq 0)$ is a standard Brownian motion independent of $\mathcal{F}_t^B$ (simple Markov property).*

Theorem 6.7. [**Time inversion**] *Suppose that $(B_t, t \geq 0)$ is a standard Brownian motion. Then the process $(X_t, t \geq 0)$ defined by*

$$X_t = \begin{cases} 0, & \text{if } t = 0; \\ tB_{1/t}, & \text{for } t > 0 \end{cases}$$

is also a standard Brownian motion.

Proof. The finite dimensional distributions $(B_{t_1}, \ldots, B_{t_n})$ of Brownian motion are Gaussian random vectors and are therefore characterized by their means $\mathbb{E}[B_{t_i}] = 0$ and covariances $\mathrm{Cov}(B_{t_i}, B_{t_j}) = t_i$ for $0 \leq t_i \leq t_j$.

So it suffices to show that the process X is a continuous Gaussian process with the same means and covariances as Brownian motion. Clearly the vector $(X_{t_1}, \ldots, X_{t_n})$ is a centred Gaussian vector. The covariances for $s \leq t$ are given by

$$\mathrm{Cov}(X_s, X_t) = st\,\mathrm{Cov}(B_{1/s}, B_{1/t}) = st\frac{1}{t} = s.$$

Hence X and B have the same finite marginal distributions. The paths $t \mapsto X_t$ are clearly continuous for $t > 0$, so it remains to show that they are also continuous for $t = 0$. First notice that since X and B have the same finite marginal distributions we get that $(X_t, t \geq 0, t \in \mathbb{Q})$ has the same law as a Brownian motion and hence

$$\lim_{t \downarrow 0, t \in \mathbb{Q}} X_t = 0 \text{ a.s.}$$

Since $\mathbb{Q}_+$ is dense and X is continuous for $t > 0$ we get that

$$0 = \lim_{t \downarrow 0, t \in \mathbb{Q}} X_t = \lim_{t \downarrow 0} X_t \text{ a.s.}$$

□

Corollary 6.8. [**Law of large numbers**] *Almost surely,* $\lim_{t \to \infty} \frac{B_t}{t} = 0$.

Proof. Let X_t be as defined in Theorem 6.7. Then

$$\lim_{t \to \infty} \frac{B_t}{t} = \lim_{t \to \infty} X(1/t) = X(0) = 0 \text{ a.s.}$$

□

Remark 6.9. Of course one can show the above result directly using the strong law of large numbers, i.e. $\lim_{n \to \infty} B_n/n = 0$. The one needs to show that B does not oscillate too much between n and $n+1$. See example sheet.

Definition 6.10. We define $(\mathcal{F}_t^B, t \geq 0)$ to be the natural filtration of $(B_t, t \geq 0)$ and $\mathcal{F}_s^+$ the slightly augmented σ-algebra defined by

$$\mathcal{F}_s^+ = \bigcap_{t > s} \mathcal{F}_t^B.$$

Remark 6.11. By the simple Markov property of Brownian motion $B_{t+s} - B_s$ is independent of $\mathcal{F}_s^B$. Clearly $\mathcal{F}_s^B \subset \mathcal{F}_s^+$ for all s, since in $\mathcal{F}_s^+$ we allow an additional infinitesimal glance into the future. But the next theorem shows that $B_{t+s} - B_s$ is still independent of $\mathcal{F}_s^+$.

Theorem 6.12. *For every $s \geq 0$ the process $(B_{t+s} - B_s, t \geq 0)$ is independent of $\mathcal{F}_s^+$.*

Proof. Let (s_n) be a strictly decreasing sequence converging to s as $n \to \infty$. By continuity

$$B_{t+s} - B_s = \lim_{n\to\infty} B_{s_n+t} - B_{s_n} \text{ a.s.}$$

Let $A \in \mathcal{F}_s^+$ and $t_1, \ldots, t_m \geq 0$. For any F continuous and bounded on$(\mathbb{R}^d)^m$ we have by the dominated convergence theorem

$$\mathbb{E}[F((B_{t_1+s} - B_s, \ldots, B_{t_m+s} - B_s))\mathbf{1}(A)] = \lim_{n\to\infty} \mathbb{E}[F((B_{t_1+s_n} - B_{s_n}, \ldots, B_{t_m+s_n} - B_{s_n}))\mathbf{1}(A)].$$

Since $A \in \mathcal{F}_s^+$, we have that $A \in \mathcal{F}_{s_n}^B$ for all n, and hence by the simple Markov property we obtain that for all n

$$\mathbb{E}[F((B_{t_1+s_n} - B_{s_n}, \ldots, B_{t_m+s_n} - B_{s_n}))\mathbf{1}(A)] = \mathbb{P}(A)\mathbb{E}[F((B_{t_1+s_n} - B_{s_n}, \ldots, B_{t_m+s_n} - B_{s_n}))].$$

Therefore, taking the limit again we deduce that

$$\mathbb{E}[F((B_{t_1+s} - B_s, \ldots, B_{t_m+s} - B_s))\mathbf{1}(A)] = \mathbb{E}[F((B_{t_1+s} - B_s, \ldots, B_{t_m+s} - B_s))]\mathbb{P}(A),$$

and hence proving the claimed independence. $\square$

Theorem 6.13. [Blumenthal's 0-1 law] *The σ-algebra $\mathcal{F}_0^+$ is trivial, i.e. if $A \in \mathcal{F}_0^+$, then $\mathbb{P}(A) \in \{0, 1\}$.*

Proof. Let $A \in \mathcal{F}_0^+$. Then $A \in \sigma(B_t, t \geq 0)$, and hence by Theorem 6.12 we obtain that A is independent of $\mathcal{F}_0^+$, i.e. it is independent of itself:

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

which gives that $\mathbb{P}(A) \in \{0, 1\}$. $\square$

Theorem 6.14. *Suppose that $(B_t)_{t\geq 0}$ is a standard Brownian motion in 1 dimension. Define $\tau = \inf\{t > 0 : B_t > 0\}$ and $\sigma = \inf\{t > 0 : B_t = 0\}$. Then*

$$\mathbb{P}(\tau = 0) = \mathbb{P}(\sigma = 0) = 1.$$

Proof. For all n we have

$$\{\tau = 0\} = \bigcap_{k\geq n} \{\exists\, 0 < \varepsilon < 1/k : B_\varepsilon > 0\}$$

and thus $\{\tau = 0\} \in \mathcal{F}_{1/n}^B$ for all n, and hence

$$\{\tau = 0\} \in \mathcal{F}_0^+.$$

Therefore, $\mathbb{P}(\tau = 0) \in \{0, 1\}$. It remains to show that it has positive probability. Clearly, for all $t > 0$ we have

$$\mathbb{P}(\tau \le t) \ge \mathbb{P}(B_t > 0) = \frac{1}{2}.$$

Hence by letting $t \downarrow 0$ we get that $\mathbb{P}(\tau = 0) \ge 1/2$ and this finishes the proof. In exactly the same way we get that

$$\inf\{t > 0 : B_t < 0\} = 0 \text{ a.s.}$$

Since B is a continuous function, by the intermediate value theorem, we deduce that

$$\mathbb{P}(\sigma = 0) = 1.$$

□

Proposition 6.15. *For $d = 1$ and $t \ge 0$ let $S_t = \sup_{0 \le s \le t} B_s$ and $I_t = \inf_{0 \le s \le t} B_s$.*

1. Then for every $\varepsilon > 0$ we have

$$S_\varepsilon > 0 \quad \textit{and} \quad I_\varepsilon < 0 \quad \textit{a.s.}$$

In particular, a.s. there exists a zero of B in any interval of the form $(0, \varepsilon)$, for all $\varepsilon > 0$.

2. A.s. we have

$$\sup_{t \ge 0} B_t = -\inf_{t \ge 0} B_t = +\infty.$$

Proof. 1. For all $t > 0$ we have that

$$\mathbb{P}(S_t > 0) \ge \mathbb{P}(B_t > 0) = \frac{1}{2}.$$

Thus, if t_n is a sequence of real numbers decreasing to 0 as $n \to \infty$, then by Fatou's inequality

$$\mathbb{P}(B_{t_n} > 0 \text{ i.o.}) = \mathbb{P}(\limsup\{B_{t_n} > 0\}) \ge \limsup_n \mathbb{P}(B_{t_n} > 0) = \frac{1}{2}.$$

Clearly, the event $\{B_{t_n} > 0 \text{ i.o.}\}$ is in $\mathcal{F}_0^+$ since it is $\mathcal{F}_{t_k}^B$-measurable for all k (notice that for all k it does not depend on $B_{t_1}, \ldots, B_{t_k}$). By Blumenthal's 0-1 law we get that

$$\mathbb{P}(B_{t_n} > 0 \text{ i.o.}) = 1,$$

and hence $S_\varepsilon > 0$ a.s. for all $\varepsilon > 0$.

The same is true for the infimum by considering $-B$ which is again a standard Brownian motion.

2. By scaling invariance of Brownian motion we get that

$$S_\infty = \sup_{t \ge 0} B_t = \sup_{t \ge 0} B_{\lambda t} \overset{(d)}{=} \sup_{t \ge 0} \sqrt{\lambda} B_t.$$

Hence $S_\infty \overset{(d)}{=} \alpha S_\infty$ for all $\alpha > 0$. Thus for all $x > 0$ the probability $\mathbb{P}(S_\infty \ge x)$ is a constant c, and hence

$$\mathbb{P}(S_\infty \ge 0) = c.$$

But we have already showed that $\mathbb{P}(S_\infty \ge 0) = 1$. Therefore, for all x we have

$$\mathbb{P}(S_\infty \ge x) = 1,$$

which gives that $\mathbb{P}(S_\infty = \infty) = 1$. □

Proposition 6.16. *Let C be a cone in $\mathbb{R}^d$ with non-empty interior and origin at 0, i.e. a set of the form $\{tu : t > 0, u \in A\}$, where A is a non-empty open subset of the unit sphere of $\mathbb{R}^d$. If*

$$H_C = \inf\{t > 0 : B_t \in C\}$$

is the first hitting time of C, then $H_C = 0$ a.s.

Proof. Since the cone C is invariant under multiplication by a positive scalar, by the scaling invariance property of Brownian motion we get that for all t

$$\mathbb{P}(B_t \in C) = \mathbb{P}(B_1 \in C).$$

Since C has non-empty interior, it is straightforward to check that

$$\mathbb{P}(B_1 \in C) > 0$$

and then we can finish the proof using Blumenthal's 0-1 law as in the proposition above. □

6.4 Strong Markov property

Let $(\mathcal{F}_t)_{t\geq 0}$ be a filtration. We say that a Brownian motion B is an $(\mathcal{F}_t)$-Brownian motion if B is adapted to $(\mathcal{F}_t)$ and $(B_{s+t} - B_s, t \geq 0)$ is independent of $\mathcal{F}_s$ for every $s \geq 0$.

In Proposition 3.3 we saw that the first hitting time of a closed set by a continuous process is always a stopping time. This is not true in general though for an open set. However, if we consider the right continuous filtration, i.e. $(\mathcal{F}_{t+})$, then we showed in Proposition 3.5 that the first hitting time of an open set by a continuous process is always an $(\mathcal{F}_{t+})$ stopping time. So, in what follows we will be considering the right continuous filtration. As this filtration is larger, this choice produces more stopping times.

Theorem 6.17. [**Strong Markov property**] *Let T be an a.s. finite stopping time. Then the process*

$$(B_{T+t} - B_T, t \geq 0)$$

is a standard Brownian motion independent of $\mathcal{F}_T^+$.

Proof. We will first prove the theorem for the stopping times $T_n = 2^{-n}\lceil 2^n T\rceil$ that discretely approximate T from above. We write $B_t^{(k)} = B_{t+k2^{-n}} - B_{k2^{-n}}$ which is a Brownian motion and B_* for the process defined by

$$B_*(t) = B_{t+T_n} - B_{T_n}.$$

We will first show that B_* is a Brownian motion independent of $\mathcal{F}_{T_n}^+$. Let $E \in \mathcal{F}_{T_n}^+$. For every event $\{B_* \in A\}$ we have

$$\begin{aligned}\mathbb{P}(\{B_* \in A\} \cap E) &= \sum_{k=0}^{\infty} \mathbb{P}(\{B^{(k)} \in A\} \cap E \cap \{T_n = k2^{-n}\}) \\ &= \sum_{k=0}^{\infty} \mathbb{P}(B^{(k)} \in A)\mathbb{P}(E \cap \{T_n = k2^{-n}\}),\end{aligned}$$

since by the simple Markov property $\{B^{(k)} \in A\}$ is independent of $\mathcal{F}^+_{k2^{-n}}$ and $E \cap \{T_n = k2^{-n}\} \in \mathcal{F}^+_{k2^{-n}}$. Since $B^{(k)}$ is a Brownian motion, we have $\mathbb{P}(B^{(k)} \in A) = \mathbb{P}(B \in A)$ does not depend on k, and hence

$$\mathbb{P}(\{B_* \in A\} \cap E) = \mathbb{P}(B \in A)\mathbb{P}(E).$$

Taking E to be the whole space gives that B_* is a Brownian motion, and hence

$$\mathbb{P}(\{B_* \in A\} \cap E) = \mathbb{P}(B_* \in A)\mathbb{P}(E)$$

for all A and E, thus showing the claimed independence.

By the continuity of Brownian motion we get that

$$B_{t+s+T} - B_{s+T} = \lim_{n\to\infty}(B_{s+t+T_n} - B_{s+T_n}).$$

The increments $(B_{t+s+T_n} - B_{s+T_n})$ are normally distributed with 0 mean variance equal to t. Thus for any $s \geq 0$ the increments $B_{t+s+T} - B_{s+T}$ are also normally distributed with 0 mean and variance t. As the process $(B_{t+T} - B_T, t \geq 0)$ is a.s. continuous, it is a Brownian motion. It only remains to show that it is independent of $\mathcal{F}^+_T$.

Let $A \in \mathcal{F}^+_T$ and $t_1, \ldots, t_k \geq 0$. We will show that for any function $F : (\mathbb{R}^d)^k \to \mathbb{R}$ continuous and bounded we have

$$\mathbb{E}[\mathbf{1}(A)F((B_{t_1+T} - B_T, \ldots, B_{t_k+T} - B_T))] = \mathbb{P}(A)\mathbb{E}[F((B_{t_1+T} - B_T, \ldots, B_{t_k+T} - B_T))].$$

Using the continuity again and the dominated convergence theorem, we get that

$$\mathbb{E}[\mathbf{1}(A)F((B_{t_1+T} - B_T, \ldots, B_{t_k+T} - B_T))] = \lim_{n\to\infty}\mathbb{E}[\mathbf{1}(A)F((B_{t_1+T_n} - B_{T_n}, \ldots, B_{t_k+T_n} - B_{T_n}))].$$

Since $T_n > T$, it follows that $A \in \mathcal{F}^+_{T_n}$. But we already showed that the process $(B_{t+T_n} - B_{T_n}, t \geq 0)$ is independent of $\mathcal{F}^+_{T_n}$, hence using the continuity and dominated convergence one more time gives the claimed independence. $\square$

Remark 6.18. Let $\tau = \inf\{t \geq 0 : B_t = \max_{0\leq s\leq 1} B_s\}$. It is intuitively clear that τ is not a stopping time. To prove that, first show that $\tau < 1$ a.s. The increment $B_{t+\tau} - B_\tau$ is negative in a small neighbourhood of 0, which contradicts the strong Markov property.

6.5 Reflection principle

Theorem 6.19. **[Reflection principle]** *Let T be an a.s. finite stopping time and $(B_t, t \geq 0)$ a standard Brownian motion. Then the process $(\widetilde{B}_t, t \geq 0)$ defined by*

$$\widetilde{B}_t = B_t\mathbf{1}(t \leq T) + (2B_T - B_t)\mathbf{1}(t > T)$$

is also a standard Brownian motion and we call it Brownian motion reflected at T.

Proof. By the strong Markov property, the process

$$B^{(T)} = (B_{T+t} - B_T, t \geq 0)$$

is a standard Brownian motion independent of $(B_t, 0 \leq t \leq T)$. Also the process

$$-B^{(T)} = (B_T - B_{t+T}, t \geq 0)$$

is a standard Brownian motion independent of $(B_t, 0 \leq t \leq T)$. Therefore, the pair $((B_t, 0 \leq t \leq T), B^{(T)})$ has the same law as $((B_t, 0 \leq t \leq T), -B^{(T)})$.

We now define the concatenation operation at time T between two continuous paths X and Y by

$$\Psi_T(X, Y)(t) = X_t \mathbf{1}(t \leq T) + (X_T + Y_{t-T}) \mathbf{1}(t > T).$$

Applying Ψ_T to B and $B^{(T)}$ gives us the Brownian motion B, while applying it to B and $-B^{(T)}$ gives us the process $\widetilde{B}$.

Let $\mathcal{A}$ be the product σ-algebra on the space $\mathcal{C}$ of continuous functions on $[0, \infty)$. It is easy to see that Ψ_T is a measurable mapping from $(\mathcal{C} \times \mathcal{C}, \mathcal{A} \otimes \mathcal{A})$ to $(\mathcal{C}, \mathcal{A})$ (by approximating T by discrete stopping times).

Hence, B and $\widetilde{B}$ have the same law. $\square$

Corollary 6.20. **[Reflection principle]** *Let B be a standard Brownian motion in 1 dimension and $b > 0$ and $a \leq b$ Then writing $S_t = \sup_{0 \leq s \leq t} B_s$ we have that for every $t \geq 0$*

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

Proof. For any $x > 0$ we define $T_x = \inf\{t \geq 0 : B_t = x\}$. Since $S_\infty = \infty$ ($S_\infty = \sup_{t \geq 0} B_t$) a.s. we have that $T_x < \infty$ a.s.

By the continuity of Brownian motion, we have that $B_{T_x} = x$ a.s. Clearly $\{S_t \geq b\} = \{T_b \leq t\}$. By the reflection principle applied to T_b we get

$$\mathbb{P}(S_t \geq b, B_t \leq a) = \mathbb{P}(T_b \leq t, 2b - B_t \geq 2b - a) = \mathbb{P}(T_b \leq t, \widetilde{B}_t \geq 2b - a),$$

since $\widetilde{B}_t = 2b - B_t$ when $t \geq T_b$.

Since $a \leq b$, the event $\{\widetilde{B}_t \geq 2b - a\}$ is contained in the event $\{T_b \leq t\}$. Hence we get

$$\mathbb{P}(S_t \geq b, B_t \leq a) = \mathbb{P}(\widetilde{B}_t \geq 2b - a) = \mathbb{P}(B_t \geq 2b - a),$$

where the last equality follows again by the reflection principle ($\widetilde{B}$ is a standard Brownian motion). $\square$

Corollary 6.21. *For every $t \geq 0$ the variables S_t and $|B_t|$ have the same law.*

Proof. Let $a > 0$. Then by Corollary 6.20 we get that

$$\mathbb{P}(S_t \geq a) = \mathbb{P}(S_t \geq a, B_t \leq a) + \mathbb{P}(S_t \geq a, B_t > a) = 2\mathbb{P}(B_t \geq a) = \mathbb{P}(|B_t| \geq a),$$

since the event $\{B_t > a\}$ is contained in $\{S_t \geq a\}$. $\square$

Exercise 6.22. Let $x > 0$ and $T_x = \inf\{t > 0 : B_t = x\}$. Then the random variable T_x has the same law as $(x/B_1)^2$.

6.6 Martingales for Brownian motion

Proposition 6.23. *Let $(B_t, t \geq 0)$ be a standard Brownian motion in 1 dimension. Then*

(i) *the process $(B_t, t \geq 0)$ is an $(\mathcal{F}_t^+)$-martingale,*

(ii) *the process $(B_t^2 - t, t \geq 0)$ is an $(\mathcal{F}_t^+)$-martingale.*

Proof. (i) Let $s \leq t$, then

$$\mathbb{E}[B_t - B_s | \mathcal{F}_s^+] = \mathbb{E}[B_t - B_s] = 0,$$

since the increment $B_t - B_s$ is independent of $\mathcal{F}_s^+$ by Theorem 6.12.
(ii) The process is adapted to the filtration $(\mathcal{F}_t^+)$ and if $s \leq t$, then

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

□

Using the above proposition, one can show that

Proposition 6.24. *Let B be a standard Brownian motion in 1 dimension and $x, y > 0$. Then*

$$\mathbb{P}(T_{-y} < T_x) = \frac{x}{x+y} \quad \text{and} \quad \mathbb{E}[T_x \wedge T_{-y}] = xy.$$

Proposition 6.25. *Let B be a standard Brownian motion in d dimensions. Then for each $u = (u_1, \ldots, u_d) \in \mathbb{R}^d$ the process*

$$M_t^u = \exp\left(\langle u, B_t\rangle - \frac{|u|^2 t}{2}\right), t \geq 0$$

is an $(\mathcal{F}_t^+)$ martingale.

Proof. Integrability follows since $\mathbb{E}[\exp(\langle u, B_t\rangle)] = \exp\left(t \sum_{i=1}^d u_i^2/2\right)$ for all $t \geq 0$. Let $s \leq t$, then

$$\mathbb{E}[M_t^u | \mathcal{F}_s^+] = e^{-|u|^2 t/2} \mathbb{E}\left[\exp(\langle u, B_t - B_s + B_s\rangle) | \mathcal{F}_s^+\right] = e^{-|u|^2 t/2} \exp(\langle u, B_s\rangle) \mathbb{E}\left[\exp(\langle u, B_t - B_s\rangle)\right],$$

where the last equality follows from Theorem 6.12. Since the increment $B_t - B_s$ is distributed according to $\mathcal{N}(0, (t-s)I_d)$ we get that

$$\mathbb{E}[M_t^u | \mathcal{F}_s^+] = M_s^u,$$

and hence proving the martingale property. □

We saw above that if $f(x) = x^2$, then the right term to subtract from $f(B_t)$ in order to make it a martingale is t. More generally now, we are interested in finding what we need to subtract from f in order to obtain a martingale. Before stating the theorem for Brownian motion, let's look at a discrete time analogue for a simple random walk on the integers. Let (S_n) be the random walk. Then

$$\begin{aligned}\mathbb{E}[f(S_{n+1})|S_1, \ldots, S_n] - f(S_n) &= \frac{1}{2}(f(S_n + 1) - 2f(S_n) + f(S_n - 1)) \\ &= \frac{1}{2}\widetilde{\Delta} f(S_n),\end{aligned}$$

where $\widetilde{\Delta} f(x) := f(x + 1) - 2f(x) + f(x - 1)$. Hence

$$f(S_n) - \frac{1}{2}\sum_{k=0}^{n-1} \widetilde{\Delta} f(S_k)$$

defines a discrete time martingale. In the Brownian motion case we expect a similar result with $\widetilde{\Delta}$ replaced by its continuous analogue, the Laplacian

$$\Delta f(x) = \sum_{i=1}^{d} \frac{\partial^2 f}{\partial x_i^2}.$$

Theorem 6.26. *Let B be a Brownian motion in $\mathbb{R}^d$. Let $f(t, x) : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ be continuously differentiable in the variable t and twice continuously differentiable in the variable x. Suppose in addition that f and its derivatives up to second order are bounded. Then the following process*

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

is an $(\mathcal{F}_t^+)$-martingale.

Proof. Integrability follows trivially by the assumptions on the boundedness of f and its derivatives.

We will now show the martingale property. Let $0 \leq t$. Then

$$\begin{aligned}M_{t+s} - M_s &= f(t + s, B_{t+s}) - f(s, B_s) - \int_s^{s+t} \left(\frac{\partial}{\partial r} + \frac{1}{2}\Delta\right) f(r, B_r)\, dr \\ &= f(t + s, B_{t+s}) - f(s, B_s) - \int_0^t \left(\frac{\partial}{\partial r} + \frac{1}{2}\Delta\right) f(r + s, B_{r+s})\, dr.\end{aligned}$$

Since $B_{t+s} - B_s$ is independent of $\mathcal{F}_s^+$ by Theorem 6.12 and B_s is $\mathcal{F}_s^+$-measurable, writing $p_s(z, y) = (2\pi s)^{-d/2} e^{-|z-y|^2/(2s)}$ for the transition density in time s, we have (**check!**)

$$\mathbb{E}[f(t + s, B_{t+s})|\mathcal{F}_s^+] = \mathbb{E}[f(t + s, B_{t+s} - B_s + B_s)|\mathcal{F}_s^+] = \int_{\mathbb{R}^d} f(t + s, B_s + x) p_t(0, x)\, dx.$$

Now notice that by the boundedness assumption on f and all its derivatives

$$\mathbb{E}\left[\int_0^t \left(\frac{\partial}{\partial r} + \frac{1}{2}\Delta\right) f(r+s, B_{r+s})\, dr \Big| \mathcal{F}_s^+\right] = \int_0^t \mathbb{E}\left[\left(\frac{\partial}{\partial r} + \frac{1}{2}\Delta\right) f(r+s, B_{r+s}) \Big| \mathcal{F}_s^+\right] dr.$$

(**Check!** using Fubini's theorem and the definition of conditional expectation.) Using again the fact that $B_{t+s} - B_s$ is independent of $\mathcal{F}_s^+$, we get

$$\mathbb{E}\left[\left(\frac{\partial}{\partial r} + \frac{1}{2}\Delta\right) f(r+s, B_{r+s} - B_s + B_s) \Big| \mathcal{F}_s^+\right] = \int_{\mathbb{R}^d} \left(\frac{\partial}{\partial r} + \frac{1}{2}\Delta\right) f(r+s, x+B_s) p_r(0,x)\, dx.$$

By the boundedness of f and its derivatives, using the dominated convergence theorem we deduce

$$\int_0^t \int_{\mathbb{R}^d} \left(\frac{\partial}{\partial r} + \frac{1}{2}\Delta\right) f(r+s, x+B_s) p_r(0,x)\, dx\, dr = \lim_{\varepsilon \downarrow 0} \int_\varepsilon^t \int_{\mathbb{R}^d} \left(\frac{\partial}{\partial r} + \frac{1}{2}\Delta\right) f(r+s, x+B_s) p_r(0,x) dx dr.$$

Using integration by parts twice in this last integral and Fubini's theorem we have that it is equal to

$$\int_{\mathbb{R}^d} (f(t+s, B_s+x) p_t(0,x) - f(\varepsilon+s, x+B_s) p_\varepsilon(0,x))\, dx - \int_{\mathbb{R}^d} \int_\varepsilon^t \frac{\partial}{\partial r} p_r(0,x) f(r+s, x+B_s)\, dr\, dx$$
$$+ \int_\varepsilon^t \int_{\mathbb{R}^d} \frac{1}{2} \Delta p_r(0,x) f(r+s, x+B_s)\, dx\, dr.$$

The transition density $p_r(0,x)$ satisfies the heat equation, i.e. $(\partial_r - \Delta/2)p = 0$, and hence this last expression is equal to

$$\int_{\mathbb{R}^d} (f(t+s, B_s+x) p_t(0,x) - f(\varepsilon+s, x+B_s) p_\varepsilon(0,x))\, dx.$$

Now notice that as $\varepsilon \downarrow 0$ we get

$$\lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d} f(\varepsilon+s, x+B_s) p_\varepsilon(0,x))\, dx = f(s, B_s),$$

since the limit above is equal to $\lim_{\varepsilon \to 0} \mathbb{E}[f(s+\varepsilon, B_{s+\varepsilon}) | \mathcal{F}_s^+]$ which by the continuity of the Brownian motion and of f and by the conditional dominated convergence theorem is equal to $f(s, B_s)$.

Therefore we showed that

$$\mathbb{E}[M_{t+s} - M_s | \mathcal{F}_s^+] = 0 \text{ a.s.}$$

and this finishes the proof. □

6.7 Recurrence and transience

We note that if a Brownian motion starts from $x \in \mathbb{R}^d$, i.e. $B_0 = x$, then B can be written as

$$B_t = x + \widetilde{B}_t,$$

where $\widetilde{B}$ is a standard Brownian motion.

We will write $\mathbb{P}_x$ to indicate that the Brownian motion starts from x, i.e. under $\mathbb{P}_x$ the process $(B_t - x, t \geq 0)$ is a standard Brownian motion.

Theorem 6.27. *Let B be a Brownian motion in $d \geq 1$ dimensions.*

(i) *If $d = 1$, then B is point-recurrent, in the sense that for all x a.s. the set*

$$\{t \geq 0 : B_t = x\}$$

is unbounded.

(ii) *If $d = 2$, then B is neighbourhood recurrent, in the sense that for every x, z under $\mathbb{P}_x$-a.s. the set*

$$\{t \geq 0 : |B_t - z| \leq \varepsilon\}$$

is unbounded for every $\varepsilon > 0$.

However, B does not hit points, i.e. for every $x \in \mathbb{R}^d$

$$\mathbb{P}_0(\exists t > 0 : B_t = x) = 0.$$

(iii) *If $d \geq 3$, then B is transient, in the sense that*

$$|B_t| \to \infty \quad as \quad t \to \infty \ \mathbb{P}_0\text{-}a.s.$$

Proof. (i) This is a consequence of Proposition 6.15, since

$$\limsup_{t\to\infty} B_t = \infty = -\liminf_{t\to\infty} B_t.$$

(ii) Note that it suffices to show the claim for $z = 0$.

Let $\varphi \in \mathcal{C}_b^2(\mathbb{R}^2)$ be such that

$$\varphi(y) = \log|y|, \quad \text{for } \varepsilon \leq |y| \leq R,$$

where $R > \varepsilon > 0$. Note that $\Delta\varphi(y) = 0$ for $\varepsilon \leq |y| \leq R$. Let the Brownian motion start from x, i.e. $B_0 = x$ with $\varepsilon < |x| < R$.

By Theorem 6.26 the process

$$M = \left(\varphi(B_t) - \int_0^t \frac{1}{2}\Delta\varphi(B_s)\, ds\right)_t$$

is a martingale.

We now set $S = \inf\{t \geq 0 : |B_t| = \varepsilon\}$ and $T_R = \inf\{t \geq 0 : |B_t| = R\}$. Then $H = S \wedge T_R$ is an a.s. finite stopping time and $(M_{t\wedge H})_{t\geq 0} = (\log|B_{t\wedge H}|, t \geq 0)$ is a bounded martingale. By the optional stopping theorem, since $H < \infty$ a.s., we thus obtain that

$$\mathbb{E}_x[\log|B_H|] = \log|x|$$

or equivalently,

$$\log(\varepsilon)\mathbb{P}_x(S < T_R) + \log(R)\mathbb{P}_x(T_R < S) = \log|x|,$$

which gives that

$$\mathbb{P}_x(S < T_R) = \frac{\log R - \log |x|}{\log R - \log \varepsilon}. \tag{6.2}$$

Letting $R \to \infty$ we have that $T_R \to \infty$ a.s. and hence $\mathbb{P}_x(S < \infty) = 1$, which shows that

$$\mathbb{P}_x(|B_t| \le \varepsilon, \text{ for some } t > 0) = 1.$$

Applying the Markov property at time n we get

$$\begin{aligned}\mathbb{P}_x(|B_t| \le \varepsilon, \text{ for some } t > n) &= \mathbb{P}_x(|B_{t+n} - B_n + B_n| \le \varepsilon, \text{ for some } t > 0)\\ &= \int_{\mathbb{R}^2} \mathbb{P}_0(|B_t + y| \le \varepsilon, \text{ for some } t > 0)\mathbb{P}_x(B_n \in dy)\\ &= \int_{\mathbb{R}^2} \mathbb{P}_y(|B_t| \le \varepsilon, \text{ for some } t > 0)\mathbb{P}_x(B_n \in dy).\end{aligned}$$

($\mathbb{P}_x(B_n \in dy)$ is the law of B_n under $\mathbb{P}_x$.) Since we showed above that for all z

$$\mathbb{P}_z(|B_t| \le \varepsilon, \text{ for some } t > 0) = 1,$$

we deduce that $\mathbb{P}_x(|B_t| \le \varepsilon, \text{ for some } t > n) = 1$ for all x.

Therefore the set $\{t \ge 0 : |B_t| \le \varepsilon\}$ is unbounded $\mathbb{P}_x$-a.s.

Letting $\varepsilon \to 0$ in (6.2) gives that the probability of hitting 0 before hitting the boundary of the ball around 0 of radius R is 0. Therefore, letting $R \to \infty$ gives that the probability of ever hitting 0 is 0, i.e. for all $x \neq 0$

$$\mathbb{P}_x(B_t = 0, \text{ for some } t > 0) = 0.$$

We only need to show now that

$$\mathbb{P}_0(B_t = 0, \text{ for some } t > 0) = 0.$$

Applying again the Markov property at $a > 0$ we get

$$\begin{aligned}\mathbb{P}_0(B_t = 0, \text{ for some } t \ge a) &= \int_{\mathbb{R}^2} \mathbb{P}_0(B_{t+a} - B_a + y = 0, \text{ for some } t > 0)\mathbb{P}_0(B_a \in dy)\\ &= \int_{\mathbb{R}^2} \mathbb{P}_y(B_t = 0, \text{ for some } t > 0)\frac{1}{(2\pi a)^{d/2}} e^{-|y|^2/(2a)}\, dy = 0\end{aligned}$$

since for all $y \neq 0$ we have already proved that $\mathbb{P}_y(B_t = 0, \text{ for some } t > 0) = 0$.

Thus, since $\mathbb{P}_0(B_t = 0, \text{ for some } t \ge a) = 0$ for all $a > 0$, letting $a \to 0$ we deduce that

$$\mathbb{P}_0(B_t = 0, \text{ for some } t > 0) = 0.$$

(iii) Since the first three components of a Brownian motion in $\mathbb{R}^d$ form a Brownian motion in $\mathbb{R}^3$, it suffices to treat the case $d = 3$. As we did above, let f be a function $f \in \mathcal{C}_b^2(\mathbb{R}^3)$ such that

$$f(y) = \frac{1}{|y|}, \quad \text{for } \ \varepsilon \le |y| \le R.$$

Note that $\Delta f(y) = 0$ for $\varepsilon \le |y| \le R$. Let $B_0 = x$ with $\varepsilon \le |x| \le R$. If we define again S and T_R as above the same argument shows that

$$\mathbb{P}_x(S < T_R) = \frac{|x|^{-1} - R^{-1}}{\varepsilon^{-1} - R^{-1}}.$$

As $R \to \infty$ this converges to $\varepsilon/|x|$ which is the probability of ever visiting the ball centred at 0 and of radius ε when starting from $|x| \ge \varepsilon$.

We will now show that

$$\mathbb{P}_0(|B_t| \to \infty \text{ as } t \to \infty) = 1.$$

Let $T_r = \inf\{t > 0 : |B_t| = r\}$ for $r > 0$. We define the events

$$A_n = \{|B_t| > n \text{ for all } t \ge T_{n^3}\}.$$

By the unboundedness of Brownian motion, it is clear that

$$\mathbb{P}_0(T_{n^3} < \infty) = 1.$$

Applying the strong Markov property at the time T_{n^3} we obtain

$$\begin{aligned}\mathbb{P}_0(A_n^c) &= \mathbb{P}_0\left(|B_{t+T_{n^3}} - B_{T_{n^3}} + B_{T_{n^3}}| \le n \text{ for some } t \ge 0\right) \\ &= \mathbb{E}_0[\mathbb{P}_{B_{T_{n^3}}}(T_n < \infty)] = \frac{n}{n^3} = \frac{1}{n^2}.\end{aligned}$$

Since the right hand side is summable, by the Borel-Cantelli lemma we get that only finitely many of the sets A_n^c occur, which implies that $|B_t|$ diverges to ∞ as $t \to \infty$. $\square$

6.8 Brownian motion and the Dirichlet problem

Definition 6.28. We call a connected open subset D of $\mathbb{R}^d$ a domain. We say that D satisfies the Poincaré cone condition at $x \in \partial D$ (boundary of D) if there exists a non-empty open cone C with origin at x and such that $C \cap \mathcal{B}(x, r) \subset D^c$ for some $r > 0$.

Theorem 6.29. [Dirichlet problem] *Let D be a bounded domain in $\mathbb{R}^d$ such that every boundary point satisfies the Poincaré cone condition. Suppose that φ is a continuous function on ∂D. We let $\tau(\partial D) = \inf\{t \ge 0 : B_t \in \partial D\}$, which is an almost surely finite stopping time when starting in D. Then the function $u : \bar{D} \to \mathbb{R}$ given by*

$$u(x) = \mathbb{E}_x[\varphi(B_{\tau(\partial D)})], \quad \textit{for } x \in \bar{D},$$

is the unique continuous function satisfying

$$\begin{aligned}&\Delta u = 0 \textit{ on } D \\ &u(x) = \varphi(x) \quad \textit{for } x \in \partial D.\end{aligned}$$

Before solving the Dirichlet problem we state a well-known result and for the proof we refer the reader to [1, Theorem 3.2].

Theorem 6.30. *Let D be a domain in $\mathbb{R}^d$ and $u : D \to \mathbb{R}$ measurable and locally bounded. The following conditions are equivalent:*

(i) *u is twice continuously differentiable and $\Delta u = 0$,*

(ii) *for any ball $\mathcal{B}(x, r) \subset D$ we have*

$$u(x) = \frac{1}{\mathcal{L}(\mathcal{B}(x, r))} \int_{\mathcal{B}(x,r)} u(y)\, dy,$$

(iii) *for any ball $\mathcal{B}(x, r) \subset D$ we have*

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

where $\sigma_{x,r}$ is the surface area measure on $\partial\mathcal{B}(x, r)$.

Definition 6.31. A function satisfying one of the equivalent conditions of Theorem 6.30 is called harmonic in D.

The next theorem and corollary folowing it will be used in the uniqueness part of the proof of Theorem 6.29.

Theorem 6.32. **[Maximum principle]** *Suppose that $u : \mathbb{R}^d \to \mathbb{R}$ is a harmonic function on a domain $D \subset \mathbb{R}^d$.*
(i) *If u attains its maximum in D, then u is a constant on D.*
(ii) *If u is continuous on $\bar{D}$ and D is bounded, then*

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

Proof. (i) Let M be the maximum. Then the set $V = \{x \in D : u(x) = M\}$ is relatively closed in D (if x_n is a sequence of points in V converging to $x \in D$, then $x \in V$), since u is continuous. Since D is open, for any $x \in V$ there exists $r > 0$ such that $\mathcal{B}(x, r) \subset D$. From Theorem 6.30 we have

$$M = u(x) = \frac{1}{\mathcal{L}(\mathcal{B}(x, r))} \int_{\mathcal{B}(x,r)} u(y)\, dy \leq M.$$

We thus deduce that $u(y) = M$ for almost all $y \in \mathcal{B}(x, r)$. But since u is continuous, this gives that $u(y) = M$ for all $y \in \mathcal{B}(x, r)$. Therefore, $\mathcal{B}(x, r) \subset V$. Hence V is also open and by assumption non-empty. But since D is connected, we must have that $V = D$. Hence u is constant on D.

(ii) Since u is continuous and $\bar{D}$ is closed and bounded, u attains a maximum on $\bar{D}$. By (i), the maximum has to be attained on ∂D. □

Corollary 6.33. *Suppose that $u_1, u_2 : \mathbb{R}^d \to \mathbb{R}$ are functions harmonic on a bounded domain D and continuous on $\bar{D}$. If u_1 and u_2 agree on ∂D, then they are identical.*

Proof. By Theorem 6.32 (ii) applied to $u_1 - u_2$ we obtain that

$$\max_{x\in\bar{D}}(u_1(x) - u_2(x)) = \max_{x\in\partial D}(u_1(x) - u_2(x)) = 0,$$

and hence we obtain that $u_1(x) \le u_2(x)$ for all $x \in \bar{D}$. In the same way $u_2(x) \le u_1(x)$ for all $x \in \bar{D}$. Hence $u_1 = u_2$ on $\bar{D}$. $\square$

Proof of Theorem 6.29. Since the domain D is bounded, we get that u is bounded. We will first show that $\Delta u = 0$ on D, by showing that u satisfifes condition (iii) of Theorem 6.30.

Let $x \in D$. Then there exists $\delta > 0$ such that $\bar{\mathcal{B}}(x,\delta) \subset D$. Let $\tau = \inf\{t > 0 : B_t \notin \mathcal{B}(x,\delta)\}$. Then this is an a.s. finite stopping time, and hence applying the strong Markov property at τ we get

$$\begin{aligned} u(x) &= \mathbb{E}_x[\varphi(B_{\tau_{\partial D}})] = \mathbb{E}_x[\mathbb{E}_x[\varphi(B_{\tau_{\partial D}})|\mathcal{F}_\tau]] = \mathbb{E}_x[\mathbb{E}_{B_\tau}[\varphi(B_{\tau_{\partial D}})]] \\ &= \mathbb{E}_x[u(B_\tau)] = \frac{1}{\sigma(\partial\mathcal{B}(x,r))}\int_{\partial\mathcal{B}(x,r)} u(y)\, d\sigma_{x,r}(y). \end{aligned}$$

The uniqueness now follows from Corollary 6.33.

It remains to show that u is continuous on $\bar{D}$. Clearly u is continuous on D. So we only need to show that u is continuous on ∂D. Let $z \in \partial D$. Since the domain D satisfies the Poincaré cone condition, there exists $h > 0$ and a non-empty open cone C_z with origin at z such that $C_z \cap \mathcal{B}(z,h) \subset D^c$.

Since φ is continuous on ∂D, we get that for every $\varepsilon > 0$, there exists $0 < \delta \le h$ such that if $|y - z| \le \delta$ and $y \in \partial D$, then $|\varphi(y) - \varphi(z)| < \varepsilon$.

Let x be such that $|x - z| \le 2^{-k}\delta$, for some $k > 0$. Then we have

$$\begin{aligned} |u(x) - u(z)| &= |\mathbb{E}_x[\varphi(B_{\tau_{\partial D}})] - \varphi(z)| \le \mathbb{E}_x[|\varphi(B_{\tau_{\partial D}}) - \varphi(z)|] \\ &\le \varepsilon\mathbb{P}_x(\tau_{\partial D} < \tau_{\partial\mathcal{B}(z,\delta)}) + 2\|\varphi\|_\infty \mathbb{P}_x(\tau_{\partial\mathcal{B}(z,\delta)} < \tau_{\partial D}) \\ &\le \varepsilon\mathbb{P}_x(\tau_{\partial D} < \tau_{\partial\mathcal{B}(z,\delta)}) + 2\|\varphi\|_\infty \mathbb{P}_x(\tau_{\partial\mathcal{B}(z,\delta)} < \tau_{C_z}). \end{aligned}$$

Now we note that

$$\mathbb{P}_x(\tau_{\partial\mathcal{B}(z,\delta)} < \tau_{C_z}) \le a^k,$$

for some $a < 1$. Thus by choosing k large enough, we can get this last probability as small as we like, and hence this completes the proof of continuity. $\square$

We will now give an example where the domain does not satisfy the conditions of Theorem 6.29 and the function u as defined there fails to solve the Dirichlet problem.

Example 6.34. Let v be a solution of the Dirichlet problem on $\mathcal{B}(0,1)$ with boundary condition $\varphi : \partial\mathcal{B}(0,1) \to \mathbb{R}$. We now let $D = \{x \in \mathbb{R}^2 : 0 < |x| < 1\}$ be the punctured disc. We will show that the function $u(x) = \mathbb{E}_x[\varphi(B_{\tau_{\partial D}})]$ given by Theorem 6.29 fails to solve the problem on D with boundary condition $\varphi : \partial\mathcal{B}(0,1) \cup \{0\}$ if $\varphi(0) \ne v(0)$. Indeed, since planar Brownian motion does not hit points, the first hitting time of $\partial D = \partial\mathcal{B}(0,1) \cup \{0\}$ is equal a.s. to the first hitting time of $\partial\mathcal{B}(0,1)$. Therefore,

$$u(0) = \mathbb{E}_0[\varphi(B_{\tau_{\partial D}})] = v(0) \ne \varphi(0).$$

6.9 Donsker's invariance principle

In this section we will show that Brownian motion is the scaling limit of random walks with steps of 0 mean and finite variance. This can be seen as a generalization of the central limit theorem to processes.

For a function $f \in \mathcal{C}([0,1],\mathbb{R})$ we define its uniform norm $\|f\| = \sup_t |f(t)|$. The uniform norm makes $\mathcal{C}([0,1],\mathbb{R})$ into a metric space so we can consider weak convergence of probability measures. The associated Borel σ-algebra coincides with the σ-algebra generated by the coordinate functions.

Theorem 6.35. **[Donsker's invariance principle]** *Let $(X_n, n \geq 1)$ be a sequence of $\mathbb{R}$-valued integrable independent random variables with common law μ such that*

$$\int x \, d\mu(x) = 0 \quad \text{and} \quad \int x^2 \, d\mu(x) = \sigma^2 \in (0,\infty).$$

Let $S_0 = 0$ and $S_n = X_1 + \ldots + X_n$ and define a continuous process that interpolates linearly between values of S, namely

$$S_t = (1 - \{t\})S_{[t]} + \{t\}S_{[t]+1}, \;\; t \geq 0,$$

where $[t]$ denotes the integer part of t and $\{t\} = t - [t]$. Then $S^{[N]} := ((\sigma^2 N)^{-1/2} S_{Nt}, 0 \leq t \leq 1)$ converges in distribution to a standard Brownian motion between times 0 and 1, i.e. for every bounded continuous function $F : \mathcal{C}([0,1],\mathbb{R}) \to \mathbb{R}$,

$$\mathbb{E}[F(S^{[N]})] \to \mathbb{E}[F(B)] \quad \text{as } N \to \infty.$$

Remark 6.36. Note that from Donsker's theorem we can infer that $N^{-1/2}\sup_{0\leq n \leq N} S_n$ converges to $\sup_{0\leq t\leq 1} B_t$ in distribution as $N \to \infty$, since the function $f \mapsto \sup f$ is a continuous operation on $\mathcal{C}([0,1],\mathbb{R})$.

The proof of Theorem 6.35 that we will give uses a coupling of the random walk with the Brownian motion, called the Skorokhod embedding theorem. It is however specific to dimension $d = 1$.

Theorem 6.37. **[Skorokhod embedding for random walks]** Let μ be a probability measure on $\mathbb{R}$ of mean 0 and variance $\sigma^2 < \infty$. Then there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration $(\mathcal{F}_t)_{t\geq 0}$, on which is defined a Brownian motion $(B_t)_{t\geq 0}$ and a sequence of stopping times

$$0 = T_0 \leq T_1 \leq T_2 \leq \ldots$$

such that, setting $S_n = B_{T_n}$,

(i) $(T_n)_{n\geq 0}$ is a random walk with steps of mean σ^2,

(ii) $(S_n)_{n\geq 0}$ is a random walk with step distribution μ.

Proof. Define Borel measures $\mu_\pm$ on $[0,\infty)$ by

$$\mu_\pm(A) = \mu(\pm A),\ A \in \mathcal{B}([0,\infty)).$$

There exists a probability space on which are defined a Brownian motion $(B_t)_{t\geq 0}$ and a sequence $((X_n, Y_n) : n \in \mathbb{N})$ of independent random variables in $\mathbb{R}^2$ with law ν given by

$$\nu(dx, dy) = C(x+y)\mu_-(dx)\mu_+(dy)$$

where C is a suitable normalizing constant. Set $\mathcal{F}_0 = \sigma(X_n, Y_n : n \in \mathbb{N})$ and $\mathcal{F}_t = \sigma(\mathcal{F}_0, \mathcal{F}_t^B)$. Set $T_0 = 0$ and define inductively for $n \geq 0$

$$T_{n+1} = \inf\{t \geq T_n : B_t - B_{T_n} = -X_{n+1} \text{ or } Y_{n+1}\}.$$

Then T_n is a stopping time for all n. Note that, since μ has mean 0, we must have

$$C\int_{-\infty}^{0} (-x)\mu(dx) = C\int_{0}^{\infty} y\mu(dy) = 1.$$

Write $T = T_1$,$X = X_1$ and $Y = Y_1$.

By Proposition 6.24, conditional on $X = x$ and $Y = y$, we have $T < \infty$ a.s. and

$$\mathbb{P}(B_T = Y|X, Y) = X/(X+Y) \quad \text{and} \quad \mathbb{E}[T|X, Y] = XY.$$

So, for $A \in \mathcal{B}([0,\infty))$,

$$\mathbb{P}(B_T \in A) = \int_A \int_0^\infty \frac{x}{x+y} C(x+y)\mu_-(dx)\mu_+(dy)$$

so $\mathbb{P}(B_T \in A) = \mu(A)$. A similar argument shows this identity holds also for $A \in \mathcal{B}((-\infty, 0])$. Next

$$\begin{aligned}\mathbb{E}[T] &= \int_0^\infty \int_0^\infty xyC(x+y)\mu_-(dx)\mu_+(dy)\\ &= \int_{-\infty}^0 (-x)^2\mu(dx) + \int_0^\infty y^2\mu(dy) = \sigma^2.\end{aligned}$$

Now by the strong Markov property for each $n \geq 0$ the process $(B_{T_n+t} - B_{T_n})_{t\geq 0}$ is a Brownian motion, independent of $\mathcal{F}_{T_n}^B$. So by the above argument $B_{T_{n+1}} - B_{T_n}$ has law μ, $T_{n+1} - T_n$ has mean σ^2, and both are independent of $\mathcal{F}_{T_n}^B$. The result follows. $\square$

Proof of Theorem 6.35. We assume for this proof that $\sigma = 1$. This is enough by scaling.

Let $(B_t)_{t\geq 0}$ be a Brownian motion and $(T_n)_{n\geq 1}$ be the sequence of stopping times as constructed in Theorem 6.37. Then B_{T_n} is a random walk with the same distribution as S_n. Let $(S_t)_{t\geq 0}$ be the linear interpolation between the values of (S_n).

For each $N \geq 1$ we set

$$B_t^{(N)} = \sqrt{N} B_{N^{-1}t},$$

which by the scaling invariance property of Brownian motion is again a Brownian motion. We now perform the Skorokhod embedding construction with $(B_t)_{t\geq 0}$ replaced by $(B_t^{(N)})_{t\geq 0}$, to obtain stopping times $T_n^{(N)}$. We then set $S_n^{(N)} = B^{(N)}_{T_n^{(N)}}$ and interpolate linearly to form $(S_t^{(N)})_{t\geq 0}$. Clearly, for all N we have

$$\left((T_n^{(N)})_{n\geq 0}, (S_t^{(N)})_{t\geq 0}\right) \sim ((T_n)_{n\geq 0}, (S_t)_{t\geq 0}).$$

Next we set $\widetilde{T}_n^{(N)} = N^{-1}T_n^{(N)}$ and $\widetilde{S}_t^{(N)} = N^{-1/2}S_{Nt}^{(N)}$. Then

$$(\widetilde{S}_t^{(N)})_{t\geq 0} \sim (S_t^{[N]})_{t\geq 0}$$

and $\widetilde{S}_{n/N}^{(N)} = B_{\widetilde{T}_n^{(N)}}$ for all n. We need to show that for all bounded continuous functions $F : \mathcal{C}([0,1],\mathbb{R}) \to \mathbb{R}$ that as $N \to \infty$

$$\mathbb{E}[F(S^{[N]})] \to \mathbb{E}[F(B)].$$

In fact we will show that for all $\varepsilon > 0$ we have

$$\mathbb{P}\left(\sup_{0\leq t\leq 1}\left|\widetilde{S}_t^{(N)} - B_t\right| > \varepsilon\right) \to 0.$$

Since F is continuous, this implies that $F(\widetilde{S}^{(N)}) \to F(B)$ in probability, which by bounded convergence is enough.

Since T_n is a random walk with increments of mean 1 by the strong law of large numbers we have that a.s.

$$\frac{T_n}{n} \to 1 \text{ as } n \to \infty.$$

So as $N \to \infty$ we have that a.s.

$$N^{-1}\sup_{n\leq N}|T_n - n| \to 0 \text{ as } n \to \infty.$$

Hence for all $\delta > 0$ we have that as $N \to \infty$

$$\mathbb{P}\left(\sup_{n\leq N}\left|\widetilde{T}_n^{(N)} - n/N\right| > \delta\right) \to 0.$$

Since $\widetilde{S}_{n/N}^{(N)} = B_{\widetilde{T}_n^{(N)}}$ for all n we have that for every $n/N \leq t \leq (n+1)/N$ there exists $\widetilde{T}_n^{(N)} \leq u \leq \widetilde{T}_{n+1}^{(N)}$ such that $\widetilde{S}_t^{(N)} = B_u$. This follows by the intermediate value theorem and the fact that $(\widetilde{S}_t^{(N)})$ is the linear interpolation between the values of S_n . Hence we have

$$\begin{aligned}\{|\widetilde{S}_t^{(N)} - B_t| > \varepsilon \text{ for some } t \in [0,1]\} &\subseteq \{|\widetilde{T}_n^{(N)} - n/N| > \delta \text{ for some } n \leq N\}\\ &\quad \cup \{|B_u - B_t| > \varepsilon \text{ for some } t \in [0,1] \text{ and } |u-t| \leq \delta + 1/N\}\\ &= A_1 \cup A_2.\end{aligned}$$

The paths of $(B_t)_{t\geq 0}$ are uniformly continuous on $[0,1]$. So for any $\varepsilon > 0$ we can find $\delta > 0$ so that $\mathbb{P}(A_2) \leq \varepsilon/2$ whenever $N \geq 1/\delta$. Then by choosing N even larger we can ensure that $\mathbb{P}(A_1) \leq \varepsilon/2$ also. Hence $\widetilde{S}^{(N)} \to B$ uniformly on $[0,1]$ in probability as required. □

Remark 6.38. From the proof above we see that we can construct the Brownian motion and the random walk on the same space so that as $N \to \infty$

$$\mathbb{P}\left(\sup_{0\leq t\leq 1}|S_t^{[N]} - B_t| > \varepsilon\right) \to 0.$$

6.10 Zeros of Brownian motion

Theorem 6.39. *Let $(B_t)_{t\geq 0}$ be a one dimensional Brownian motion and*

$$\mathit{Zeros} = \{t \geq 0 : B_t = 0\}$$

is the zero set. Then, almost surely, Zeros *is a closed set with no isolated points.*

Proof. Since Brownian motion is continuous almost surely, the zero set is closed a.s. To prove that no point is isolated we do the following: for each rational $q \in [0, \infty)$ we consider the first zero after q, i.e.

$$\tau_q = \inf\{t \geq q : B_t = 0\}.$$

Note that τ_q is an almost surely finite stopping time. Since Zeros is a closed set, this infimum is almost surely a minimum. By the strong Markov property, applied to τ_q, we have that for each q, almost surely τ_q is not an isolated zero from the right. But since the rational numbers is a countable set we get that almost surely for all rational q, the zero τ_q is not isolated from the right.

The next thing to prove is that the remaining points of Zeros are not isolated from the left. We claim that for any $0 < t$ in the zero set which is different from τ_q for all rational q is not an isolated point from the left. Take a sequence $q_n \uparrow t$ with $q_n \in \mathbb{Q}$. Define $t_n = \tau_{q_n}$. Clearly $q_n \leq t_n < t$ and so $t_n \uparrow t$. Thus t is not isolated from the left. □

Theorem 6.40. *Fix $t \geq 0$. Then, almost surely, Brownian motion in one dimension is not differentiable at t.*

Proof. Exercise. □

But also a much stronger statement is true, namely

Theorem 6.41. **[Paley, Wiener and Zygmund 1933]** *Almost surely, Brownian motion in one dimension is nowhere differentiable.*

7 Poisson random measures

7.1 Construction and basic properties

For $\lambda \in (0, \infty)$ we say that a random variable X in $\mathbb{Z}^+$ is *Poisson of parameter* λ and write $X \sim P(\lambda)$ if

$$\mathbb{P}(X = n) = e^{-\lambda}\lambda^n/n!$$

We also write $X \sim P(0)$ to mean $X \equiv 0$ and write $X \sim P(\infty)$ to mean $X \equiv \infty$.

Proposition 7.1. **[Addition property]** *Let $N_k, k \in \mathbb{N}$, be independent random variables, with $N_k \sim P(\lambda_k)$ for all k. Then*

$$\sum_k N_k \sim P(\sum_k \lambda_k).$$

Proposition 7.2. [**Splitting property**] *Let $N, Y_n, n \in \mathbb{N}$, be independent random variables, with $N \sim P(\lambda), \lambda < \infty$ and $\mathbb{P}(Y_n = j) = p_j$, for $j = 1, \dots, k$ and all n. Set*

$$N_j = \sum_{n=1}^{N} \mathbf{1}(Y_n = j).$$

Then $N_1, \dots, N_k$ are independent random variables with $N_j \sim P(\lambda p_j)$ for all j.

Proof. Left as an exercise. □

Let $(E, \mathcal{E}, \mu)$ be a σ-finite measure space. A *Poisson random measure with intensity* μ is a map

$$M : \Omega \times \mathcal{E} \to \mathbb{Z}_+ \cup \{\infty\}$$

satisfying, for all sequences $(A_k : k \in \mathbb{N})$ of disjoint sets in $\mathcal{E}$,

(i) $M(\cup_k A_k) = \sum_k M(A_k)$,

(ii) $M(A_k), k \in \mathbb{N}$, are independent random variables,

(iii) $M(A_k) \sim P(\mu(A_k))$ for all k.

Denote by E^* the set of $\mathbb{Z}_+ \cup \{\infty\}$-valued measures on $\mathcal{E}$ and define, for $A \in \mathcal{E}$,

$$X : E^* \times \mathcal{E} \to \mathbb{Z}_+ \cup \{\infty\}, \quad X_A : E^* \to \mathbb{Z}_+ \cup \{\infty\}$$

by

$$X(m, A) = X_A(m) = m(A).$$

Set $\mathcal{E}^* = \sigma(X_A : A \in \mathcal{E})$.

Theorem 7.3. *There exists a unique probability measure μ^* on $(E^*, \mathcal{E}^*)$ such that under μ^* X is a Poisson random measure with intensity μ.*

Proof. (*Uniqueness.*) For disjoint sets $A_1, \dots, A_k \in \mathcal{E}$ and $n_1, \dots, n_k \in \mathbb{Z}_+$, set

$$A^* = \{m \in E^* : m(A_1) = n_1, \dots, m(A_k) = n_k\}.$$

Then, for any measure μ^* making X a Poisson random measure with intensity μ,

$$\mu^*(A^*) = \prod_{j=1}^{k} e^{-\mu(A_j)} \mu(A_j)^{n_j} / n_j!$$

Since the set of such sets A^* is a π-system generating $\mathcal{E}^*$, this implies that μ^* is uniquely determined on $\mathcal{E}^*$.

(*Existence.*) Consider first the case where $\lambda = \mu(E) < \infty$. There exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which are defined independent random variables N and $Y_n, n \in \mathbb{N}$, with $N \sim P(\lambda)$ and $Y_n \sim \mu/\lambda$ for all n. Set

$$M(A) = \sum_{n=1}^{N} \mathbf{1}(Y_n \in A), \quad A \in \mathcal{E}. \tag{7.1}$$

It is easy to check, by the Poisson splitting property, that M is a Poisson random measure with intensity μ. Indeed, for disjoint $A_1, \ldots, A_k$ in $\mathcal{E}$ with finite μ measures, we let $X_n = j$ whenever $Y_n \in A_j$, so that $M(A_j), 1 \le j \le k$ are independent $P(\mu(A_j)), 1 \le j \le k$ random variables.

More generally, if $(E, \mathcal{E}, \mu)$ is σ-finite, then there exist disjoint sets $E_k \in \mathcal{E}, k \in \mathbb{N}$, such that $\cup_k E_k = E$ and $\mu(E_k) < \infty$ for all k. We can construct, on some probability space, independent Poisson random measures $M_k, k \in \mathbb{N}$, with M_k having intensity $\mu|_{E_k}$. Set

$$M(A) = \sum_{k \in \mathbb{N}} M_k(A \cap E_k), \ A \in \mathcal{E}.$$

It is easy to check, by the Poisson addition property, that M is a Poisson random measure with intensity μ. The law μ^* of M on E^* is then a measure with the required properties. $\square$

The above construction gives the following important property of Poisson random measures.

Proposition 7.4. *Let M be a Poisson random measure on E with intensity μ, and let $A \in \mathcal{E}$ be such that $\mu(A) < \infty$. Then $M(A)$ has law $P(\mu(A))$, and given $M(A) = k$, the restriction $M|_A$ has same law as $\sum_{i=1}^k \delta_{X_i}$, where $(X_1, \ldots, X_k)$ are independent with law $\mu(\cdot \cap A)/\mu(A)$. Moreover, if $A, B \in \mathcal{E}$ are disjoint, then the restrictions $M|_A, M|_B$ are independent.*

Exercise 7.5. Let $E = \mathbb{R}_+$ and $\mu = \theta \mathbf{1}(t \ge 0)\, dt$. Let M be a Poisson random measure on $\mathbb{R}_+$ with intensity measure μ and let$(T_n)_{n \ge 1}$ and $T_0 = 0$ be a sequence of random variables such that $(T_n - T_{n-1}, n \ge 1)$ are independent exponential random variables with parameter $\theta > 0$. Then

$$\left(N_t = \sum_{n \ge 1} \mathbf{1}(T_n \le t), \ t \ge 0 \right) \quad \text{and} \quad (N'_t = M([0,t]), \ t \ge 0)$$

have the same distribution.

7.2 Integrals with respect to a Poisson random measure

Theorem 7.6. *Let M be a Poisson random measure on E with intensity μ. Then for $f \in \mathcal{L}^1(\mu)$, then so is $M(f)$ defined by*

$$M(f) = \int_E f(y) M(dy)$$

and

$$\mathbb{E}[M(f)] = \int_E f(y)\mu(dy), \quad \mathrm{var}(M(f)) = \int_E f(y)^2 \mu(dy).$$

Let $f : E \to \mathbb{R}_+$ be a measurable function. Then for $u > 0$

$$\mathbb{E}\left[e^{-uM(f)}\right] = \exp\left\{ -\int_E (1 - e^{-uf(y)})\mu(dy) \right\}.$$

Let $f : E \to \mathbb{R}$ be in $\mathcal{L}^1(\mu)$. Then for any u

$$\mathbb{E}\left[e^{iuM(f)}\right] = \exp\left\{ \int_E (e^{iuf(y)} - 1)\mu(dy) \right\}.$$

Proof. First assume that $f = \mathbf{1}(A)$, for $A \in \mathcal{E}$. Then $M(A)$ is a random variable by definition of M and this extends to any finite linear combination of indicators. Since any measurable non-negative function is the increasing limit of finite linear combinations of such indicator functions, we obtain by monotone convergence that $M(f)$ is a random variable as a limit of random variables.

Let $E_n, n \geq 0$ be a measurable partition of E into sets of finite μ-measure. A similar approximation argument shows that $M(f\mathbf{1}(E_n)), n \geq 0$ are independent random variables.

Let $f \in \mathcal{L}^1(\mu)$. We will first show the formula for the expectation and the variance. If $f = \mathbf{1}(A)$, then this is clear. This extends to finite linear combinations and to any non-negative measurable functions by approximation. For a general f, we do the standard procedure, separating into $f = f^+ - f^-$ and use the fact that $M(f^+)$ and $M(f^-)$ are independent.

Since by Proposition 7.4 given $M(E_n) = k$, the restriction $M|_{E_n}$ has the same law as $\sum_{i=1}^k \delta_{X_i}$, where $(X_1, \ldots, X_k)$ are independent with law $\mu(\cdot \cap E_n)/\mu(E_n)$, we get

$$\begin{aligned}
\mathbb{E}[\exp(-uM(f\mathbf{1}(E_n)))] &= \sum_{k=0}^{\infty} \mathbb{E}[\exp(-uM(f))|M(E_n) = k]\mathbb{P}(M(E_n) = k) \\
&= \sum_{k=0}^{\infty} e^{-\mu(E_n)} \frac{\mu(E_n)^k}{k!} \left(\int_{E_n} e^{-uf(x)} \frac{\mu(dx)}{\mu(E_n)} \right)^k \\
&= e^{-\mu(E_n)} \exp\left(\int_{E_n} e^{-uf(x)} \mu(dx) \right) \\
&= \exp\left(- \int_{E_n} \mu(dx)(1 - \exp(-uf(x))) \right).
\end{aligned}$$

Since the random variables $M(f\mathbf{1}(E_n))$ are independent over $n \geq 0$, we can take products over $n \geq 0$ and by monotone convergence we obtain the wanted formula.

To establish the formula in the case where $f \in \mathcal{L}^1(\mu)$, follows by the same kind of arguments. We first establish the formula for $f\mathbf{1}(E_n)$ in place of f. Then to obtain the result, we must show that

$$\int_{A_n} \mu(dx)(e^{iuf(x)} - 1) \to \int_E \mu(dx)(e^{iuf(x)} - 1) \quad \text{as } n \to \infty,$$

where $A_n = E_0 \cup \ldots \cup E_n$. But since $|e^{ix} - 1| \leq |x|$ for all x, we have that

$$|e^{iuf(x)} - 1| \leq |uf(x)|,$$

whence the function under consideration is integrable with respect to μ, which by dominated convergence gives the result. $\square$

7.3 Poisson Brownian motions

In this section we are going to consider Poisson random measures in $\mathbb{R}^d$ for $d \geq 1$ with intensity measure given by $\mu = \lambda \, dx$, i.e. multiples of the Lebesgue measure in d dimensions.

Let Π be a Poisson random measure in $\mathbb{R}^d$ of intensity λ (this means λ times Lebesgue measure). Note that the construction of Theorem 7.3 gives that Π can be written as

$$\Pi = \sum_{i=1}^{\infty} \delta_{X_i},$$

where X_i are random variables, since the Lebesgue measure of the whole space is infinite.

We will sometimes say Poisson point process to mean a Poisson random measure in $\mathbb{R}^d$.

Proposition 7.7. **[Thinning property]** *Let $\Pi = \{X_i\}$ be a Poisson point process in $\mathbb{R}^d$ of intensity λ. For each point X_i we perform an independent experiment and we keep it with probability $p(X_i)$ and we remove it with the complementary probability, where $p : \mathbb{R}^d \to [0,1]$ is a measurable function. Thus we define a new process Φ that contains the points X_i that we kept. The process Φ is a Poisson random measure in $\mathbb{R}^d$ with intensity $\mu(A) = \lambda \int_A p(x)\,dx$.*

Proof. The independence property follows easily from the independence of Π. We will now show that for any set A with finite volume we have $\Phi(A) \sim P(\mu(A))$, where μ is the intensity measure given in the statement. By Proposition 7.4 we have

$$\begin{aligned}
\mathbb{P}(\Phi(A) = k) &= \sum_{n \geq k} \mathbb{P}(\Pi(A) = n, \Phi(A) = k) \\
&= \sum_{n \geq k} e^{-\lambda \mathrm{vol}(A)} \frac{(\lambda \mathrm{vol}(A))^n}{n!} \binom{n}{k} \left(\int_A p(x) \frac{dx}{\mathrm{vol}(A)} \right)^k \left(\int_A (1 - p(x)) \frac{dx}{\mathrm{vol}(A)} \right)^{n-k} \\
&= e^{-\lambda \mathrm{vol}(A)} \frac{\lambda^k}{k!} \left(\int_A p(x)\,dx \right)^k \sum_{n \geq k} \frac{\lambda^{n-k}}{(n-k)!} \left(\int_A (1 - p(x))\,dx \right)^{n-k} \\
&= e^{-\lambda \mathrm{vol}(A)} \frac{\lambda^k}{k!} \left(\int_A p(x)\,dx \right)^k \exp\left(\lambda \int_A (1 - p(x))\,dx \right) \\
&= \exp\left(-\lambda \int_A p(x)\,dx \right) \frac{(\lambda \int_A p(x)\,dx)^k}{k!}.
\end{aligned}$$

$\square$

Proposition 7.8. *Let $\Pi = \{X_i\}$ be a Poisson point process in $\mathbb{R}^d$ of intensity λ. Let (Y_i) be i.i.d. random variables with law ν. Define the measure $\Phi = \sum_i \delta_{X_i + Y_i}$. Then Φ is again a Poisson point process of the same intensity λ as Π.*

Proof. It suffices to check that for any $u > 0$ and $f : \mathbb{R}^d \to \mathbb{R}_+$ we have

$$\mathbb{E}\left[e^{-u\Phi(f)}\right] = \exp\left(\lambda \int_{\mathbb{R}^d} (e^{-uf(x)} - 1)\,dx \right).$$

We can write

$$\mathbb{E}\left[e^{-u\Phi(f)}\right] = \mathbb{E}\left[e^{-u \sum_i f(X_i + Y_i)}\right]$$

and conditioning on $\{X_i\}$ and using the independence of the (Y_i)'s we obtain

$$\begin{aligned}
\mathbb{E}\left[e^{-u\Phi(f)}\right] &= \mathbb{E}\left[\mathbb{E}\left[e^{-u\sum_i f(X_i+Y_i)}|\Pi\right]\right] = \mathbb{E}\left[\prod_i \int_{\mathbb{R}^d} e^{-uf(X_i+y)}\,\nu(dy)\right] \\
&= \mathbb{E}\left[\exp\left(\log \prod_i \int_{\mathbb{R}^d} e^{-uf(X_i+y)}\,\nu(dy)\right)\right] \\
&= \mathbb{E}\left[\exp\left(-\sum_i \left(-\log \int_{\mathbb{R}^d} e^{-uf(X_i+y)}\,\nu(dy)\right)\right)\right] \\
&= \mathbb{E}\left[\exp\left(-\Pi(g)\right)\right],
\end{aligned}$$

where $g(x) = -\log \int_{\mathbb{R}^d} e^{-uf(x+y)}\,\nu(dy)$. By Theorem 7.6 we have

$$\begin{aligned}
\mathbb{E}\left[\exp\left(-\Pi(g)\right)\right] &= \exp\left(\lambda \int_{\mathbb{R}^d} \left(\exp\left(\log \int e^{-uf(x+y)}\,\nu(dy)\right) - 1\right)\,dx\right) \\
&= \exp\left(\lambda \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} e^{-uf(x+y)}\,\nu(dy) - 1\right)\,dx\right) \\
&= \exp\left(\lambda \int_{\mathbb{R}^d} \left(e^{-uf(x)} - 1\right)\,dx\right),
\end{aligned}$$

where in the last step we used Fubini's theorem and the fact that ν is a probability measure on $\mathbb{R}^d$. $\square$

For the rest of this section we are going to consider the following model: let $\Phi(0)$ be a Poisson point process in $\mathbb{R}^d$ of intensity λ, let $\Phi(0) = \{X_i\}$. We now let each point of the Poisson process move independently according to a standard Brownian motion in d dimensions. Namely the point X_i moves according to the Brownian motion $(\xi_i(t))_{t\geq 0}$. This way at every time t we obtain a new process $\Phi(t) = \{X_i + \xi_i(t)\}$, which by Proposition 7.8 is again a Poisson point process of intensity λ.

We can think of the points of the Poisson process $\Phi(0)$ as the users of a wireless network that can communicate with each other when they are at distance at most r from each other. So it is natural to introduce mobility to the model and this is why we let them evolve in space.

We now fix a target particle which is at the origin of $\mathbb{R}^d$ and we are interested in the first time that one of the points of the Poisson process is within distance r from it, i.e. we define

$$T_{\text{det}} = \inf\left\{t \geq 0 : 0 \in \cup_i \mathcal{B}(X_i + \xi_i(t), r)\right\},$$

where $\mathcal{B}(x, r)$ stands for the ball centred at x of radius r.

Theorem 7.9. [Stochastic geometry formula] *Let ξ be a standard Brownian motion in d dimensions and let $W(t) = \cup_{s\leq t}\mathcal{B}(\xi(s), r)$ be the so-called "Wiener sausage" up to time t. Then, for any dimension $d \geq 1$, the detection probability satisfies*

$$\mathbb{P}(T_{\text{det}} > t) = \exp(-\lambda \mathbb{E}[\text{vol}(W(t))]).$$

Proof. Let Π be the set of points of $\Phi(0)$ that have detected 0 by time t, that is

$$\Pi = \{X_i \in \Phi(0) : \exists s \leq t \text{ s.t. } 0 \in \mathcal{B}(X_i + \xi_i(s), r)\}.$$

Since the ξ_i's are independent we have by Proposition 7.7 that Π is a thinned Poisson point process with intensity $\Lambda(x)dx$ where Λ is given by

$$\Lambda(x) = \lambda \mathbb{P}(x \in \cup_{s\leq t} \mathcal{B}(-\xi(s), r)),$$

for ξ is a standard Brownian motion.

So for the probability that the detection time is greater than t we have that

$$\begin{aligned}\mathbb{P}(T_{\text{det}} > t) = \mathbb{P}(\Pi(\mathbb{R}^d) = 0) &= \exp\left(-\lambda \int_{\mathbb{R}^d} \mathbb{P}(x \in \cup_{s\leq t}\mathcal{B}(-\xi(s), r))\, dx\right) \\ &= \exp(-\lambda \mathbb{E}\left[\text{vol}(\cup_{s\leq t}\mathcal{B}(\xi(s), r))\right]) = \exp(-\lambda\mathbb{E}[\text{vol}W(t)]),\end{aligned}$$

where the third equaliy follows by Fubini. $\square$

Theorem 7.10. *The expected volume of the Wiener sausage $W(t) = \cup_{s\leq t}\mathcal{B}(\xi(s), r)$ satisfies as $t \to \infty$*

$$\mathbb{E}\left[\text{vol}(W(t))\right] = \begin{cases} \sqrt{\frac{8t}{\pi}} + 2r & \text{for } d = 1 \\ \frac{2\pi t}{\log t}(1 + o(1)) & \text{for } d = 2 \\ \frac{2\pi^{d/2} r^{d-2} t}{\Gamma\left(\frac{d-2}{2}\right)}(1 + o(1)) & \text{for } d \geq 3. \end{cases}$$

Proof. Dimension $d = 1$ is left as an exercise.

For all d we have that

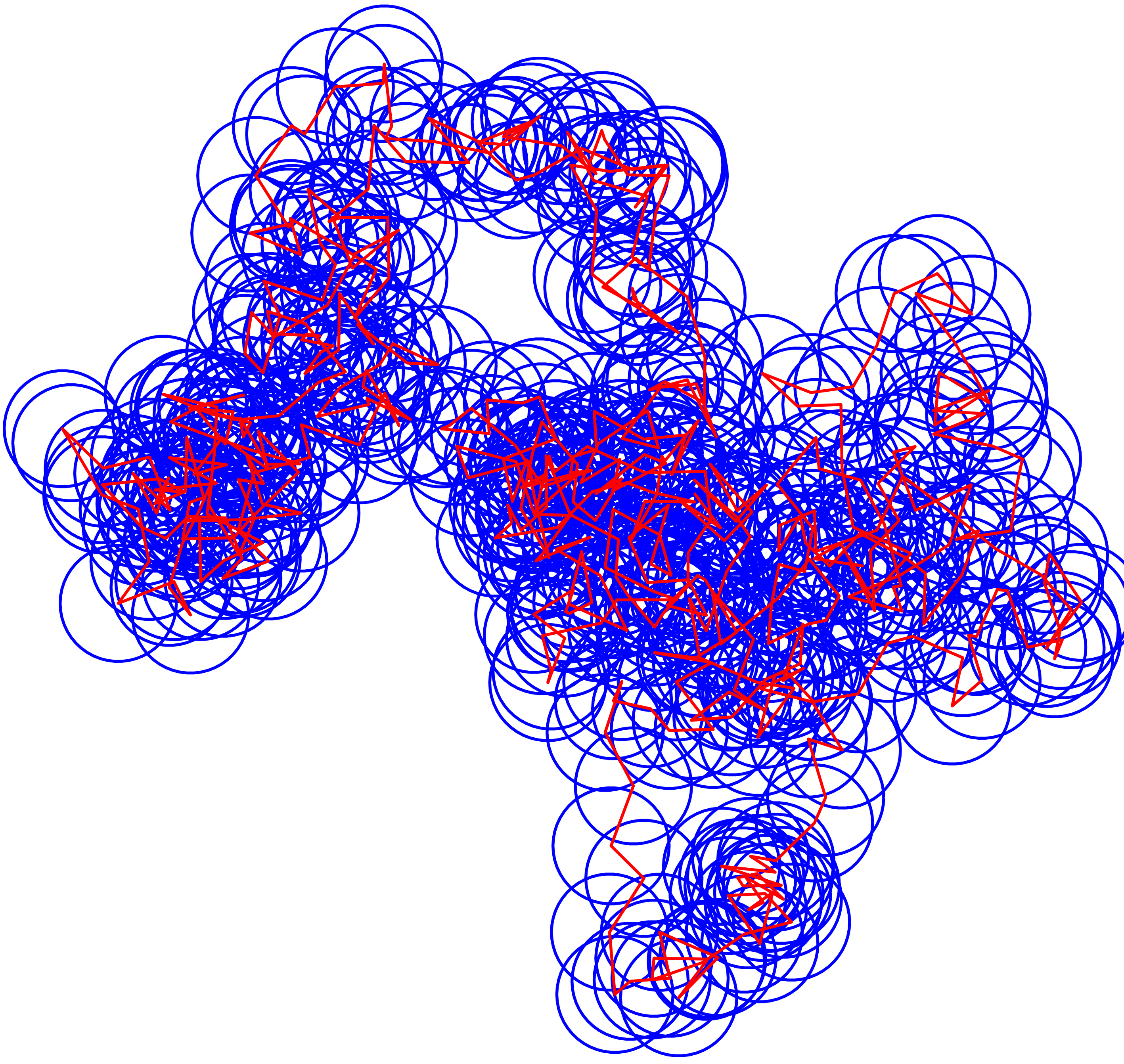
$$\begin{aligned}\mathbb{E}[\text{vol}(W_t)] &= \int_{\mathbb{R}^d} \mathbb{P}(y \in \cup_{s\leq t}\mathcal{B}(\xi(s), r))dy = \int_{\mathbb{R}^d} \mathbb{P}(\tau_{\mathcal{B}(y,r)} \leq t)dy \\ &= \text{vol}(\mathcal{B}(0, r)) + \int_{\mathbb{R}^d\setminus\mathcal{B}(0,r)} \mathbb{P}(\tau_{\mathcal{B}(y,r)} \leq t)dy,\end{aligned}$$

where τ_A is the first hitting time of the set A by the Brownian motion. Define

$$Z_t^y = \int_0^t \mathbf{1}(\xi(s) \in \mathcal{B}(y, r))\, ds, \tag{7.2}$$

i.e. the time that the Brownian motion spends in the ball $\mathcal{B}(y, r)$ before time t. It is clear by the continuity of the Brownian paths that $\{Z_t^y > 0\} = \{\tau_{\mathcal{B}(y,r)} \leq t\}$. We now have $\mathbb{P}(Z_t^y > 0) = \frac{\mathbb{E}[Z_t^y]}{E[Z_t^y | Z_t^y > 0]}$ and for the first moment we have

$$\mathbb{E}[Z_t^y] = \int_0^t \mathbb{P}_0(\xi(s) \in \mathcal{B}(y, r))\, ds = \int_0^t \int_{\mathcal{B}(y,r)} \frac{1}{(2\pi s)^{d/2}} e^{-\frac{|z|^2}{2s}}\, dz\, ds = \int_0^t \int_{\mathcal{B}(0,r)} \frac{1}{(2\pi s)^{d/2}} e^{-\frac{|z+y|^2}{2s}}\, dz\, ds$$



Random walk sausage

and for the conditional expectation $\mathbb{E}[Z_t^y | Z_t^y > 0]$, if we write T for the first time that the Brownian motion hits the boundary of the ball $\mathcal{B}(y,r)$, then we get that in 2 dimensions for all $y \notin \mathcal{B}(0,r)$

$$\mathbb{E}[Z_t^y | Z_t^y > 0] = \mathbb{E}\left[\int_T^t \mathbf{1}(\xi(s) \in \mathcal{B}(y,r))\, ds\right] \leq \int_0^t \mathbb{P}_0(\xi(s) \in \mathcal{B}(0,r))\, ds$$
$$\leq 1 + \int_1^t \int_{\mathcal{B}(0,r)} \frac{1}{2\pi s} e^{-\frac{|z|^2}{2s}}\, dz\, ds \leq 1 + \frac{r^2}{2} \log t.$$

In dimensions $d \geq 3$ we have for all $y \notin \mathcal{B}(0,r)$

$$\mathbb{E}[Z_t^y | Z_t^y > 0] = \mathbb{E}\left[\int_T^t \mathbf{1}(\xi(s) \in \mathcal{B}(y,r))\, ds\right] \leq \int_0^t \mathbb{P}_0(\xi(s) \in \mathcal{B}(0,r))\, ds$$
$$= \int_0^t \int_{\mathcal{B}((0,r),r)} \frac{1}{(2\pi s)^{d/2}} e^{-\frac{|z|^2}{2s}}\, dz\, ds = \frac{1}{(2\pi)^{d/2}} \int_{\mathcal{B}((0,r),r)} \frac{1}{|z|^{d-2}} \int_{\frac{|z|^2}{2t}}^{\infty} s^{d/2-2} e^{-s}\, ds\, dz,$$

where $\mathcal{B}((0,r),r)$ stands for the ball centred at $(0,\ldots,0,r)$ and of radius r and the last step follows by a change of variable. Now notice that

$$\int_{\frac{|z|^2}{2t}}^{\infty} s^{d/2-2} e^{-s}\, ds \to \Gamma\left(\frac{d-2}{2}\right) \text{ as } t \to \infty$$

and by the mean value property for the harmonic function $1/|z|^{d-2}$ we get that

$$\int_{\mathcal{B}((0,r),r)} \frac{dz}{|z|^{d-2}} = \mathrm{vol}(\mathcal{B}(0,1)) r^2.$$

So, putting all things together we obtain that in 2 dimensions

$$\mathbb{E}[\mathrm{vol}(W_t)] = \mathrm{vol}(\mathcal{B}(0,r)) + \int_{\mathbb{R}^2 \setminus \mathcal{B}(0,r)} \frac{\mathbb{E}[Z_t^y]}{\mathbb{E}[Z_t^y | Z_t^y > 0]}\, dy$$
$$\geq \mathrm{vol}(\mathcal{B}(0,r)) + \frac{\int_0^t \int_{\mathcal{B}(0,r)} \left(\int_{\mathbb{R}^2} \frac{1}{2\pi s} e^{-\frac{|z+y|^2}{2s}}\, dy\right) dz\, ds - \int_{\mathcal{B}(0,r)} \mathbb{E}[Z_t^y]\, dy}{1 + \frac{r^2}{2}\log t}$$
$$= \mathrm{vol}(\mathcal{B}(0,r)) + \frac{2\pi t r^2}{2 + r^2 \log t} - \frac{2\int_{\mathcal{B}(0,r)} \mathbb{E}[Z_t^y]\, dy}{2 + r^2 \log t}.$$

It is easy to see that $\int_{\mathcal{B}(0,r)} \mathbb{E}[Z_t^y]\, dy = O(\log t)$ and hence in 2 dimensions we get

$$\liminf_{t\to\infty} \mathbb{E}[\mathrm{vol}(W_t)] \frac{\log t}{2\pi t} \geq 1.$$

In $d \geq 3$ we obtain in the same way as above

$$\liminf_{t\to\infty} \mathbb{E}[\mathrm{vol}(W_t)] \frac{\Gamma\left(\frac{d-2}{2}\right)}{2\pi^{d/2} r^{d-2} t} \geq 1,$$

since $\int_{\mathcal{B}(0,r)} \mathbb{E}[Z_t^y]\, dy = O(1)$. It remains to show that in 2 dimensions

$$\limsup_{t\to\infty} \mathbb{E}[\mathrm{vol}(W_t)] \frac{\log t}{2\pi t} \le 1 \tag{7.3}$$

and in $d \ge 3$ that

$$\limsup_{t\to\infty} \mathbb{E}[\mathrm{vol}(W_t)] \frac{\Gamma\left(\frac{d-2}{2}\right)}{2\pi^{d/2} r^{d-2} t} \le 1. \tag{7.4}$$

Let $\varepsilon > 0$. We define $\widetilde{Z}_t^y = \int_0^{t(1+\varepsilon)} \mathbf{1}(\xi(s) \in \mathcal{B}(y,r))\, ds$ and use the obvious inequality

$$\mathbb{P}(Z_t^y > 0) \le \frac{\mathbb{E}[\widetilde{Z}_t^y]}{E[\widetilde{Z}_t^y | Z_t^y > 0]}.$$

We can now lower bound the conditional expectation appearing in the denominator above as follows. In $d = 2$ we have

$$\begin{aligned}\mathbb{E}[\widetilde{Z}_t^y | Z_t^y > 0] &\ge \int_0^{t\varepsilon} \int_{\mathcal{B}((0,r),r)} \frac{1}{2\pi s} e^{-\frac{|z|^2}{2s}}\, dz\, ds \ge \int_{\log(t\varepsilon)}^{t\varepsilon} \int_{\mathcal{B}((0,r),r)} \frac{1}{2\pi s} e^{-\frac{|z|^2}{2s}}\, dz\, ds \\ &\ge \frac{r^2 e^{-\frac{2r^2}{\log(t\varepsilon)}}}{2} (\log(t\varepsilon) - \log\log(t\varepsilon)).\end{aligned}$$

For $d \ge 3$ we have

$$\mathbb{E}[\widetilde{Z}_t^y | Z_t^y > 0] \ge \int_0^{t\varepsilon} \int_{\mathcal{B}((0,r),r)} \frac{1}{(2\pi s)^{d/2}} e^{-\frac{|z|^2}{2s}}\, dz\, ds = \frac{1}{(2\pi)^{d/2}} \int_{\mathcal{B}((0,r),r)} \frac{1}{|z|^{d-2}} \int_{\frac{|z|^2}{2t\varepsilon}}^{\infty} s^{d/2-2} e^{-s}\, ds\, dz.$$

Similarly to the calculations leading to the lower bound we get that in $d = 2$

$$\mathbb{E}[\mathrm{vol}(W_t)] \le \frac{2\pi t(1+\varepsilon) e^{\frac{2r^2}{\log(t\varepsilon)}}}{\log t + \log \varepsilon - \log\log(t\varepsilon)}$$

and hence for $d = 2$

$$\limsup_{t\to\infty} \mathbb{E}[\mathrm{vol}(W_t)] \frac{\log t}{2\pi t} \le 1 + \varepsilon,$$

for all $\varepsilon > 0$, and thus letting ε go to 0 proves (7.3).

For $d \ge 3$ in the same way we obtain

$$\limsup_{t\to\infty} \mathbb{E}[\mathrm{vol}(W_t)] \frac{\Gamma\left(\frac{d-2}{2}\right)}{2\pi^{d/2} r^{d-2} t} \le 1 + \varepsilon,$$

for all $\varepsilon > 0$, and thus letting ε go to 0 proves (7.4). $\square$

Now suppose that the target particle is moving according to a deterministic function $f : \mathbb{R}_+ \to \mathbb{R}^d$. We define the detection time

$$T_{\mathrm{det}}^f = \inf\{t \ge 0 : f(t) \in \cup_i \mathcal{B}(X_i + \xi_i(t), r)\}.$$

Then we have the following theorem which is non-examinable:

Theorem 7.11 (Peres-Sousi)**.** *For all times t and all dimensions d we have*

$$\mathbb{P}(T_{\mathrm{det}}^f > t) \leq \mathbb{P}(T_{\mathrm{det}} > t).$$

Using a straightforward generalization of Theorem 7.9 we get the equivalent statement

Theorem 7.12 (Peres-Sousi)**.** *For all times t and all dimensions d we have*

$$\mathbb{E}[\mathrm{vol}(W_f(t))] \geq \mathbb{E}[\mathrm{vol}(W(t))],$$

where $W_f(t) = \cup_{s \leq t} \mathcal{B}(\xi(s) + f(s), r)$.

We now conclude this course by stating an open question.

Question 7.13. Does the stochastic domination inequality

$$\mathbb{P}(\mathrm{vol}(\cup_{s \leq t} \mathcal{B}(\xi(s) + f(s), r)) \geq \alpha) \geq \mathbb{P}(\mathrm{vol}(\cup_{s \leq t} \mathcal{B}(\xi(s), r)) \geq \alpha) \quad \forall \alpha$$

also hold?

Acknowledgements
These notes are based in part on the lecture notes by James Norris and Grégory Miermont for the same course and [1].

References

[1] Peter Mörters and Yuval Peres. *Brownian motion.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2010. With an appendix by Oded Schramm and Wendelin Werner.

[2] David Williams. *Probability with martingales.* Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.