

# Part III Advanced Probability

## Based on lectures by P. Sousi

Notes taken by Pantelis Tassopoulos

Michaelmas 2023

## Contents

<b>1</b>	<b>Conditional Expectation</b>	<b>3</b>
1.1	Basic definitions . . . . .	3
1.2	Expectation . . . . .	3
1.3	Conditional expectation with respect to countably generated sigma algebras . . . .	4
1.4	General case . . . . .	4
<b>2</b>	<b>Discrete Time Martingales</b>	<b>9</b>
<b>3</b>	<b>Martingale Convergence Theorem</b>	<b>13</b>
<b>4</b>	<b>Doob's inequalities</b>	<b>15</b>
<b>5</b>	<b>Applications of martingales</b>	<b>19</b>
<b>6</b>	<b>Continuous Time processes</b>	<b>21</b>
<b>7</b>	<b>Weak Convergence</b>	<b>28</b>
<b>8</b>	<b>Large deviations</b>	<b>32</b>
<b>9</b>	<b>Brownian Motion</b>	<b>37</b>
9.1	Properties of Brownian Motion . . . . .	38
9.2	Martingales for Brownian motion . . . . .	44
9.3	Transience and recurrence . . . . .	46
9.4	Dirichlet Problem . . . . .	48
<b>10</b>	<b>Donsker's invariance principle</b>	<b>53</b>
<b>11</b>	<b>Poisson random measures</b>	<b>55</b>

# 1 Conditional Expectation

## Lecture 1 1.1 Basic definitions

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space. Remember the following definitions

**Definition 1.1 (Sigma algebra).**  $\mathcal{F}$  is a sigma algebra if and only if: ( $\mathcal{F} \in \mathcal{P}\Omega$ )

1.  $\Omega \in \mathcal{F}$
2.  $A \in \mathcal{F} \implies A^c \in \mathcal{F}$
3.  $(A_n)_{n \in \mathbb{N}} \subseteq \mathcal{F} \implies \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{F}$

**Definition 1.2 (Probability measure).**  $\mathbb{P}$  is a probability measure if

1.  $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$  (i.e. a set function)
2.  $\mathbb{P}(\Omega) = 1$ , and  $\mathbb{P}(\emptyset) = 0$
3.  $(A_n)_{n \in \mathbb{N}}$  pairwise disjoint  $\implies \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$ .

**Definition 1.3 (Random Variable).**  $X : \Omega \rightarrow \mathbb{R}$  is a random variable if for all  $B$  open in  $\mathbb{R}$ ,  $X^{-1}(B) \in \mathcal{F}$ .

**Remark.** Observe that the sigma algebra  $\mathcal{G} = \{B \subseteq \mathbb{R} : X(B) \in \mathcal{F}\} \supseteq \mathcal{O} \implies \mathcal{G} \supseteq \mathcal{B}(\mathbb{R})$ , the former being the collection of open sets in  $\mathbb{R}$  and the latter the Borel sigma algebra on  $\mathbb{R}$  with the usual topology, namely,  $\sigma(\mathcal{O})$  (see below for the notation).

Let  $\mathcal{A}$  be a collection of subsets of  $\Omega$ . We define

$$\begin{aligned} \sigma(\mathcal{A}) &= \text{smallest sigma algebra containing } \mathcal{A} \\ &= \bigcap \{ \mathcal{T} : \mathcal{T} \text{ sigma algebra containing } \mathcal{A} \}. \end{aligned}$$

**Definition 1.4 (Borel sigma algebra on  $\mathbb{R}$ ).** Let  $\mathcal{O} = \{\text{open sets in } \mathbb{R}\}$ . Then, the Borel sigma algebra  $\mathcal{B}(\mathbb{R})$  ( $:= \mathcal{B}$ ) is defined as above, namely,

$$\mathcal{B}(\mathbb{R}) := \sigma(\mathcal{O}).$$

Let  $(X_i)_{i \in I}$  be a family of random variables, then  $\sigma(X_i : i \in I)$  = the smallest sigma algebra that makes them all measurable. We also have the characterisation  $\sigma(X_i : i \in I) = \sigma(\{\underbrace{\{\omega \in \Omega : X_i(\omega) \in B\}}_{X_i^{-1}(B)}, i \in I, B \in \mathcal{B}(\mathbb{R})\})$ .

## 1.2 Expectation

Note we use the following for the indicator function on some event  $A$

$$\mathbf{1}(A)(x) = \mathbf{1}(x \in A) := \left. \begin{array}{l} 1, \quad x \in A \\ 0, \quad x \notin A \end{array} \right\}, \quad A \in \mathcal{F}.$$

We now begin the construction of the expectation of generic random variables.

Positive simple random variables:  $X = \sum_{i=1}^n \mathbf{1}(A_i), c_i \geq 0, A_i \in \mathcal{F}$ .

$$\mathbb{E}[X] := \sum_{i=1}^n c_i \mathbb{P}(A_i).$$

**Non-negative random variables:** ( $X \geq 0$ ). We proceed by approximation. Namely, let  $X_n(\omega) := 2^{-n} \lfloor 2^n \cdot X(\omega) \rfloor \wedge n \uparrow X(\omega), n \rightarrow \infty$ . Now, by monotone convergence,

$$\mathbb{E}[X] := \uparrow \lim_{n \rightarrow \infty} \mathbb{E}[X_n] = \sup \mathbb{E}[X].$$

**General random variables:** Have the decomposition  $X = X^+ - X^-$ , where  $X^+ = X \vee 0$ ,  $X^- = -X \wedge 0$ . If one of  $\mathbb{E}[X^+], \mathbb{E}[X^-] < \infty$  then set

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

**Definition 1.5.**  $X$  is called integrable if  $\mathbb{E}[|X|] < \infty$ .

**Definition 1.6.** Let  $B \in \mathcal{F}$  with  $\mathbb{P}(B) > 0$ . Then for all  $A \in \mathcal{F}$ , set

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

Now for an integer-valued random variable  $X$ , we set:

$$\mathbb{E}[X|B] := \frac{\mathbb{E}[X \cdot \mathbf{1}_B]}{\mathbb{P}(B)}$$

### 1.3 Conditional expectation with respect to countably generated sigma algebras

Lecture 2

We now extend the definition of the conditional expectation for a countably generated sigma algebra. Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space. We call the sigma algebra  $\mathcal{G}$  countably generated if there exists a collection  $(B_n)_{n \in \mathbb{N}}$  of pairwise disjoint events such that  $\bigcup_{n \in I} B_n = \Omega$  with ( $I$  countable) and  $\mathcal{G} = \sigma(B_i : i \in I)$ .

Let  $X$  be an integrable random variable. We want to define  $\mathbb{E}[X|\mathcal{G}]$ .

Define  $X'(\omega) = \mathbb{E}[X|B_i]$ , whenever  $\omega \in B_i$ , i.e.

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

We make the convention that  $\mathbb{E}[X|B_i] = 0$  if  $\mathbb{P}(B_i) = 0$ . It is easy to check that  $X'$  is  $\mathcal{G}$ -measurable. We also have that

$$\mathcal{G} = \left\{ \bigcup_{j \in J} B_j : J \subseteq I \right\}$$

and  $X'$  satisfies for all  $G \in \mathcal{G}: \mathbb{E}[X \cdot \mathbf{1}_G] = \mathbb{E}[X' \cdot \mathbf{1}_G]$  and

$$\begin{aligned} \mathbb{E}[|X'|] &\leq \mathbb{E} \left[ \sum_{i \in I} |\mathbb{E}[X|B_i]| \mathbf{1}(B_i) \right] \\ &= \sum_{i \in I} \mathbb{P}(B_i) \cdot |\mathbb{E}[X|B_i]| \\ &\leq \sum_{i \in I} \mathbb{P}(B_i) \cdot \underbrace{\mathbb{E}[X \cdot \mathbf{1}(B_i)]}_{\mathbb{P}(B_i)} \\ &= \mathbb{E}[|X|] < \infty. \end{aligned}$$

### 1.4 General case

We say  $A \in \mathcal{F}$  happens a.s. if  $\mathbb{P}(A) = 1$ . Recall (from measure theory and basic functional analysis):

**Theorem 1.1 (Monotone Convergence Theorem (MCT)).** Let  $(X_n)_{n \in \mathbb{N}}$  be such that  $X_n \geq 0, X$  be random variables such that  $X_n \uparrow X$  as  $n \rightarrow \infty$ . Then,  $\mathbb{E}[X_n] \uparrow \mathbb{E}[X]$  as  $n \rightarrow \infty$ .

**Theorem 1.2 (Dominated Convergence Theorem (DCT)).** Let  $(X_n)_{n \in \mathbb{N}}$  be random variables such that  $X_n \rightarrow X$  a.s. as  $n \rightarrow \infty$  and  $|X_n| \leq Y$  a.s. for all  $n \in \mathbb{N}$ , where  $Y$  is integrable, then  $\mathbb{E}[X_n] \rightarrow \mathbb{E}[X]$ , as  $n \rightarrow \infty$ .

Let  $1 \leq p < \infty$  and  $f$  a measurable function, then set  $\|f\|_p := (\mathbb{E}[\|f\|^p])^{\frac{1}{p}}$ . If  $p = \infty$ , then set  $\|f\|_\infty := \inf\{\lambda : |f| \leq \lambda \text{ a.s.}\}$ . Recall for all  $p$ , the Lebesgue spaces,  $\mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P}) = \{f : \|f\|_p < \infty\}$ .

**Theorem 1.3.**  $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$  is a Hilbert space, with inner product  $\langle u, v \rangle_2 = \mathbb{E}[u \cdot v]$ . Furthermore, for any closed subspace  $\mathcal{H}$ , if  $f \in \mathcal{L}^2$ , there exists a unique  $g \in \mathcal{H}$  s.t.  $\|f - g\|_{\mathcal{L}^2} = \inf_{h \in \mathcal{H}} \|f - h\|_{\mathcal{L}^2}$  and  $\langle f - g, h \rangle = 0$ , for all  $h \in \mathcal{H}$ . We say that  $g$  is the orthogonal projection of  $f$  in  $\mathcal{H}$ .

We now construct the conditional expectation in the general case, for any integrably random variable with respect to an arbitrary sigma algebras.

**Theorem 1.4 (Conditional Expectation).** Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space,  $\mathcal{G} \subseteq \mathcal{F}$  a sub-sigma algebra,  $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ . Then there exists an integrable random variable  $Y$  satisfying:

1.  $Y$  is  $\mathcal{G}$ -measurable
2. for all  $G \in \mathcal{G}, \mathbb{E}[X \cdot \mathbf{1}(G)] = \mathbb{E}[Y \cdot \mathbf{1}(G)]$ .

Moreover,  $Y$  unique in the sense that if  $Y'$  also satisfies the above 1), 2), then  $Y = Y'$  a.s.. We call  $Y$  a version of the conditional expectation of  $X$  given  $\mathcal{G}$ . We write  $Y = \mathbb{E}[X|\mathcal{G}]$  a.s. If  $\mathcal{G} = \sigma(Z)$ , where  $Z$  is a random variable, then we write  $\mathbb{E}[Z] = \mathbb{E}[X|\mathcal{G}]$ .

**Remark.** 2) could be replaced by  $\mathbb{E}[X \cdot Z] = \mathbb{E}[Y \cdot Z]$  for all  $Z$  bounded  $\mathcal{G}$ -measurable random variables.

We now state and prove the main theorem of this section:

*Proof.* (Theorem 1.4) Uniqueness: Let  $Y, Y'$  satisfy 1), 2). Let  $A = \{Y > Y'\} \in \mathcal{G}$ . Then 2)

$$\begin{aligned} \implies \mathbb{E}[Y \cdot \mathbf{1}(A)] &= \mathbb{E}[Y' \cdot \mathbf{1}(A)] = \mathbb{E}[X \cdot \mathbf{1}(A)] \\ \implies \mathbb{E}[(Y - Y') \cdot \mathbf{1}(A)] &= 0 \\ \implies \mathbb{P}(A) = \mathbb{P}(Y > Y') &= 0 \\ \implies Y \leq Y' \text{ a.s.} \end{aligned}$$

We similarly obtain  $Y \geq Y'$  a.s., hence we deduce that  $Y = Y'$  a.s.

Existence: three steps.

1. Assume that  $X \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ . Observe that  $\mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$  is a closed subspace of  $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ . Hence, Theorem 1.3, we have the decomposition  $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P}) = \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P}) \oplus \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})^\perp$ . Then, we have the corresponding decomposition  $X = Y + Z$ , where  $Y \in \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$  and  $Z \in \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})^\perp$  respectively. Define  $\mathbb{E}[X|\mathcal{G}] := Y$ ,  $Y$  is  $\mathcal{G}$ -measurable and for all  $A \in \mathcal{G}, \mathbb{E}[X \cdot \mathbf{1}(A)] = \mathbb{E}[Y \cdot \mathbf{1}(A)] + \mathbb{E}[Z \cdot \mathbf{1}(A)]$  since  $Z \in \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})^\perp$ .

Claim: If  $X \geq 0$ , a.s. then  $Y \geq 0$  a.s. Indeed, let  $A = \{Y < 0\} \in \mathcal{G}$ . Then observe that  $0 \leq \mathbb{E}[X \cdot \mathbf{1}(A)] = \mathbb{E}[Y \cdot \mathbf{1}(A)] \leq 0$ . Hence  $\mathbb{E}[Y \cdot \mathbf{1}(A)] = 0$  and so  $\mathbb{P}(A) = 0$ , giving  $Y = 0$  a.s.

2. Assume  $X \geq 0$ .

Define  $X_n = X \wedge n \leq n$ , meaning  $X_n$  is bounded for all  $n \in \mathbb{N}$ . So  $X_n \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ . Let  $Y_n = \mathbb{E}[X_n]$  a.s..  $(X_n)_{n \in \mathbb{N}}$  is an increasing sequence. By the claim above, so is  $(Y_n)_{n \in \mathbb{N}}$  a.s. Define  $Y = \limsup_n Y_n$  meaning  $Y$  is  $\mathcal{G}$ -measurable and  $Y = \uparrow \lim_{n \rightarrow \infty} Y_n$  a.s. Now, we have that for all  $A \in \mathcal{G}, \mathbb{E}[X_n \cdot \mathbf{1}(A)] = \mathbb{E}[Y_n \cdot \mathbf{1}(A)]$ . Thus, by theorem 1.1 (MCT),  $\mathbb{E}[X \cdot \mathbf{1}(A)] = \lim_{n \rightarrow \infty} \mathbb{E}[X_n \cdot \mathbf{1}(A)] = \lim_{n \rightarrow \infty} \mathbb{E}[Y_n \cdot \mathbf{1}(A)] = \mathbb{E}[Y \cdot \mathbf{1}(A)]$ .

3.  $X$  general in  $\mathcal{L}^1$ .

Decompose as before  $X = X^+ - X^-$ . Define,  $\mathbb{E}[X\mathcal{G}] = \mathbb{E}[X^+|\mathcal{G}] - \mathbb{E}[X^-|\mathcal{G}]$ .

□

Lecture 3

**Remark.** From the second step of the proof of Theorem 1.4 we see that we can define  $\mathbb{E}[X|\mathcal{G}]$  for all  $X \geq 0$ , not necessarily integrable. It satisfies all conditions 1), 2) except for the integrability one.

**Definition 1.7.**  $\underbrace{\mathcal{G}_1, \mathcal{G}_2, \dots}_{\text{sigma algebras}} \subset \mathcal{F}$ . We call them independent if whenever  $G_i \in \mathcal{G}_i$  and

$$i_1 < \dots < i_k \text{ for some } k \in \mathbb{N}, \text{ then } \mathbb{P}(G_{i_1} \cap \dots \cap G_{i_k}) = \prod_{j=1}^k \mathbb{P}(G_{i_j}).$$

Moreover, let  $X$  be a random variable and  $\mathcal{G}$  a sigma algebra, then they are said to be int if  $\sigma(X)$  is independent of  $\mathcal{G}$ .

Properties of conditional expectations: Fix  $X, y \in \mathcal{L}^1, G \in \mathcal{F}$ .

1.  $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$  (take  $A = \Omega$ )
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  $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$
4. If  $X \geq 0$  a.s., then  $\mathbb{E}[X|\mathcal{G}] \geq 0$  a.s.
5. For  $\alpha, \beta \in \mathbb{R}$   $\mathbb{E}[\alpha X + \beta Y|\mathcal{G}] = \alpha \mathbb{E}[X|\mathcal{G}] + \beta \mathbb{E}[Y|\mathcal{G}]$
6.  $\mathbb{E}[|X|\mathcal{G}] \geq \mathbb{E}[X|\mathcal{G}]$  a.s.

Below we proved:we expansions of useful measure theoretic results for the expectation to their corresponding conditional counterparts. First recall:

**Lemma 1.1 (Fatou's Lemma).** Let  $X_n \geq 0$  for all  $n \in \mathbb{N}$ . Then

$$\mathbb{E}[\liminf_n X_n] \leq \liminf_n \mathbb{E}[X_n] \quad \text{a.s.}$$

**Theorem 1.5 (Jensen's Inequality).** If  $X$  is integrable and  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  is a convex function, then

$$\phi(\mathbb{E}[X]) \leq \mathbb{E}[\phi(X)] \quad \text{a.s.}$$

Now the results themselves:

**Theorem 1.6 (Conditional Monotone Convergence theorem (MCT)).** Let  $\mathcal{G} \subset \mathcal{F}$  be sigma algebras,  $X_n \geq 0$  a.a. and  $X_n \uparrow X$ , as  $n \rightarrow \infty$  a.s. Then

$$\mathbb{E}[X_n|\mathcal{G}] \uparrow \mathbb{E}[X|\mathcal{G}] \quad \text{a.s.}$$

*Proof.* Theorem 1.6 Set  $Y_n = \mathbb{E}[X_n|\mathcal{G}]$  a.s. Observe that  $Y_n$  is a.s. increasing. Set  $Y = \limsup_n Y_n$ .  $Y_n$  is  $\mathcal{G}$ -measurable, hence, so is  $Y$  (as a limsup of  $\mathcal{G}$ -measurable random variables) is also  $\mathcal{G}$ -measurable. Also,  $Y = \lim_{n \rightarrow \infty} Y_n$  a.s.

Need to show:  $\mathbb{E}[Y \cdot \mathbf{1}(A)] = \mathbb{E}[X \cdot \mathbf{1}(A)]$  for all  $A \in \mathcal{G}$ . Indeed,

$$\begin{aligned} \mathbb{E}[Y \cdot \mathbf{1}(A)] &= \mathbb{E}[\lim_{n \rightarrow \infty} Y_n \cdot \mathbf{1}(A)] \stackrel{\text{MCT}}{=} \lim_{n \rightarrow \infty} \mathbb{E}[Y_n \cdot \mathbf{1}(A)] \\ &= \lim_{n \rightarrow \infty} \mathbb{E}[X_n \cdot \mathbf{1}(A)] = \mathbb{E}[X \cdot \mathbf{1}(A)]. \end{aligned}$$

□

*Proof.* Theorem 1.1  $\liminf_n X_n = \lim_{n \rightarrow \infty} \left( \inf_{k \geq n} X_k \right)$ , the limit of an increasing sequence. By Theorem 1.1, we have

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

and

$$\mathbb{E}[\inf_{k \geq n} X_k | \mathcal{G}] \stackrel{\text{a.s.}}{\leq} \inf_{k \geq n} \mathbb{E}[X_k | \mathcal{G}]^1$$

which gives the result

$$\mathbb{E}[\liminf_n X_n] \leq \liminf_n \mathbb{E}[X_n] \quad \text{a.s.}$$

□

**Theorem 1.7 (Conditional Dominated Convergence Theorem).** Suppose  $X_n \rightarrow X$  a.s.  $n \rightarrow \infty$  and  $|X_n| \leq Y$  a.s. for all  $n \in \mathbb{N}$  with  $Y$  integrable. Then  $\mathbb{E}[X_n | \mathcal{G}] \rightarrow \mathbb{E}[X | \mathcal{G}]$  a.s. as  $n \rightarrow \infty$ .

*Proof.* From  $-Y \leq X_n \leq Y$ , we have  $X_n + Y \geq 0$  for all  $n \in \mathbb{N}$  and  $Y - X_n \geq 0$  a.s. By Theorem 1.1,

$$\begin{aligned} \mathbb{E}[X + Y | \mathcal{G}] &= \mathbb{E}[\liminf_n (X_n + Y) | \mathcal{G}] \\ &\leq \liminf_n \mathbb{E}[X_n + Y | \mathcal{G}] = \liminf_n \mathbb{E}[X_n | \mathcal{G}] + \mathbb{E}[Y] \end{aligned}$$

Thus,

$$\begin{aligned} \mathbb{E}[|X - Y| | \mathcal{G}] &= \mathbb{E}[Y - \liminf_n X_n | \mathcal{G}] \\ &\leq \mathbb{E}[Y] + \liminf_n \mathbb{E}[X_n | \mathcal{G}] \end{aligned}$$

Hence,

$$\limsup_n \mathbb{E}[X_n | \mathcal{G}] \leq \mathbb{E}[X | \mathcal{G}]$$

and

$$\liminf_n \mathbb{E}[X_n | \mathcal{G}] \geq \mathbb{E}[X | \mathcal{G}]$$

a.s., concluding the proof.

□

**Theorem 1.8 (Conditional Jensen).** Let  $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ ,  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  be a convex function s.t.  $\phi(X)$  is integrable or  $\phi(X) \geq 0$

$$\phi(\mathbb{E}[X | \mathcal{G}]) \leq \mathbb{E}[\phi(X) | \mathcal{G}] \quad \text{a.s.}$$

*Proof.* Claim: (true for any convex function, no proof given)  $\phi(x) = \sup_{i \in \mathbb{N}} (a_i x + b_i)$ ,  $a_i b_i \in \mathbb{R}$ . Thus,

$$\mathbb{E}[\phi(X) | \mathcal{G}] \geq a_i \mathbb{E}[X | \mathcal{G}] + b_i \quad \text{for all } i \in \mathbb{N}.$$

Taking the supremum gives <sup>2</sup>

$$\begin{aligned} \mathbb{E}[\phi(X) | \mathcal{G}] &\geq \sup_{i \in \mathbb{N}} (a_i \mathbb{E}[X | \mathcal{G}] + b_i) \\ &= \phi(\mathbb{E}[X | \mathcal{G}]) \quad \text{a.s.} \end{aligned}$$

□

**Corollary 1.8.1.** For all  $1 \leq p < \infty$   $\|\mathbb{E}[X | \mathcal{G}]\|_p \leq \|X\|_p$ .

*Proof.* Apply conditional Jensen.

□

<sup>1</sup>can take the infimum due to countability that preserves a.s.

<sup>2</sup>can take the supremum due to countability which again preserves a.s.

**Proposition 1.1 (Tower Property).** Let  $X$  be integrable and  $\mathcal{H} \subseteq \mathcal{G}$  sigma algebras. Then

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

*Proof.* (a)  $\mathbb{E}[X|\mathcal{H}]$  is  $\mathcal{H}$ -measurable.

(b) For all  $A \in \mathcal{H}$  NTS:

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

Indeed, both terms above are equal to  $\mathbb{E}[X \cdot \mathbf{1}(A)]$  since  $A \in \mathcal{G} \subseteq \mathcal{H}$ . □

**Proposition 1.2.** Let  $X \in \mathcal{L}^1$ ,  $\mathcal{G} \subseteq \mathcal{F}$ ,  $Y$  bounded  $\mathcal{G}$ -measurable. Then

$$\mathbb{E}[X \cdot Y|\mathcal{G}] = Y \cdot \mathbb{E}[X|\mathcal{G}].$$

*Proof.* (a) RHS is clearly  $\mathcal{G}$ -measurable.

(b) For all  $A \in \mathcal{G}$ :

$$\begin{aligned} \mathbb{E}[X \cdot Y \cdot \mathbf{1}(A)] &= \mathbb{E}[Y \cdot \mathbb{E}[X|\mathcal{G}] \cdot \mathbf{1}(A)] \\ \mathbb{E}[X \cdot \underbrace{(Y \cdot \mathbf{1}(A))}_{\mathcal{G}\text{-meas. and bounded}}] &= \mathbb{E}[\mathbb{E}[X|\mathcal{G}] \cdot Y \cdot \mathbf{1}(A)] = RHS. \end{aligned}$$

(Also observe that by a monotone class argument, we have for any integrable function  $f : \Omega \rightarrow \mathbb{R}$ ,  $\mathbb{E}[X \cdot f] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}] \cdot f]$ ) □

Lecture 4 We are building towards the Theorem

**Theorem 1.9.**  $X \in \mathcal{L}^1$ ,  $\mathcal{G}, \mathcal{H} \subseteq \mathcal{F}$ . Assume  $\sigma(\mathcal{G}, \mathcal{H}) \perp \mathcal{H}$ , Then

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

We begin with a definition

**Definition 1.8.** Let  $\mathcal{A}$  be a collection of sts. It is called a  $\pi$ -system if for all  $A, B \in \mathcal{A}$ , we also have  $A \cap B \in \mathcal{A}$ .

**Theorem 1.10 (Uniqueness of extension).** Let  $(E, \xi)$  be a measurable space and let  $\mathcal{A}$  be a  $\pi$ -system generating the sigma algebra  $\xi$ . Let  $\mu, \nu$  be two measures on  $(E, \xi)$  with  $\mu(E) = \nu(E) < \infty$ . If  $\mu = \nu$  on  $\mathcal{A}$ , then  $\mu = \nu$  on  $\xi$ .

*Proof.* (Theorem 1.9) NTS: for all  $F \in \sigma(\mathcal{G}, \mathcal{H})$

$$\mathbb{E}[X \cdot \mathbf{1}_F] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}] \cdot \mathbf{1}_F]$$

Now, set  $\mathcal{A} = \{A \cap B : A \in \mathcal{G}, B \in \mathcal{H}\}$ . It is easy to check that  $\mathcal{A}$  is a  $\pi$ -system generating  $\sigma(\mathcal{G}, \mathcal{H})$ . If  $F = A \cap B$  for some  $A \in \mathcal{G}$  and  $B \in \mathcal{H}$ , Then

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

Now assume  $X \geq 0$ ; in the general case, decompose  $X = X^+ - X^-$  and apply same argument to both  $X^\pm$ . Define the measures  $\mu(F) = \mathbb{E}[X \cdot \mathbf{1}(F)]$  and  $\nu(F) = \mathbb{E}[X \cdot \mathbf{1}(F)]$  for all  $F \in \sigma(\mathcal{G}, \mathcal{H})$ . Observe that  $\mu(\Omega) = \nu(\Omega) = \mathbb{E}[X] < \infty$  and we have shown that  $\mu = \nu$  on  $\mathcal{A}$ . Thus,  $\mu = \nu$  on  $\sigma(\mathcal{G}, \mathcal{H})$  which finally implies the result

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

□



**Examples:**

**Definition 1.9 (Gaussian).**  $(X_1, X_2, \dots, X_n) \in \mathbb{R}^n$  has the Gaussian distribution if and only if for all scalars  $a_1, a_2, \dots, a_n \in \mathbb{R}$ ,  $a_1 X_1 + \dots + a_n X_n$  has the Gaussian distribution in  $\mathbb{R}$ .

A stochastic process (more on that later)  $(X_t)_{t \geq 0}$  is a Gaussian process if for all  $t_1 < t_2 < \dots < t_n$  the vector  $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$  is Gaussian.

Let  $(X, Y)$  be a Gaussian vector in  $\mathbb{R}^2$ . We compute  $\mathbb{E}[X|Y]$ .

Let  $X' = \mathbb{E}[X|Y]$ . Looking for  $f$  a Borel measurable function s.t.  $\mathbb{E}[X|Y] = f(Y)$  a.s. Let  $f(y) = ay + b$  for some  $a, b \in \mathbb{R}$  to be determined. Now,  $X' = aY + b$ ,  $\mathbb{E}[X'] = \mathbb{E}[X] = a\mathbb{E}[Y] + b$  and  $\mathbb{E}[X' \cdot Y] = \mathbb{E}[X \cdot Y] \implies \mathbb{E}[(X - X') \cdot Y] = 0$ . Thus  $\text{Cov}(X - X', Y) = 0 \implies \text{Cov}(X, Y) = a^2 \text{Var}(Y)$ .

Need to check: that for all  $Z$  bounded  $\sigma(Y)$ -measurable,  $\mathbb{E}[(X - X') \cdot Z] = 0$ .

Indeed, observe that  $(X - X', Y)$  is a Gaussian vector and since  $\text{Cov}(X - X', Y) = 0 \implies X - X' \perp Y \implies (X - X') \perp Z$ .

2. Let  $(X, Y)$  be a random vector with density in  $\mathbb{R}^2$  with joint density function  $f_{X,Y} : \mathbb{R}^2 \rightarrow \mathbb{R}$ . Let  $h : \mathbb{R} \rightarrow \mathbb{R}$  be a Borel function such that  $h(X)$  is integrable. We now compute  $\mathbb{E}[h(X)|Y]$ .

We have for all  $g$  bounded  $\sigma Y$ -measurable functions.

$$\begin{aligned} \int_{\mathbb{R}^2} h(x)g(y)f_{X,Y}(x,y) dx dy &= \mathbb{E}[h(X)g(Y)] \\ &= \mathbb{E}[\mathbb{E}[h(X)|Y]g(Y)] = \mathbb{E}[\phi(Y)g(Y)] \\ &= \int_{\mathbb{R}^2} \phi(y)g(y)f_Y(y) dy \end{aligned}$$

where  $f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x,y) dx$  and  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  is some Borel measurable function. Hence,

$$\phi(y) = \begin{cases} \int_{\mathbb{R}} h(x) \frac{f_{X,Y}(x,y)}{f_Y(y)} dx, & f_Y(y) > 0 \\ 0, & \text{otherwise} \end{cases}$$

can be seen to work. Thus, we obtain

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

## 2 Discrete Time Martingales

**Definition 2.1 (Filtration).** Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space. A filtration is a sequences of increasing sigma sub-algebras of  $\mathcal{F}$ ,  $(\mathcal{F}_n)_{n \in \mathbb{N}}$ ,  $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$  for all  $n \in \mathbb{N}$ . We call  $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \in \mathbb{N}})$  a filtered probability space.

Let  $X = (X_n)_{n \in \mathbb{N}}$  be a sequence of random variables/a stochastic process. Then it induces  $(\mathcal{F}_n^X)_{n \in \mathbb{N}}$ , where  $\mathcal{F}_n^X := \sigma(X_{k \leq n})$  for all  $n \in \mathbb{N}$ : the canonical filtration associated to  $X$ . We call  $X$  adapted to a filtration  $(\mathcal{F}_n)_{n \in \mathbb{N}}$  if  $X$  is  $\mathcal{F}_n$ -measurable for all  $n \in \mathbb{N}$ .  $X$  is called integrable if  $X_n$  is integrable for all  $n \in \mathbb{N}$ .

**Definition 2.2 (Martingale discrete time).** Let  $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \in \mathbb{N}}, \mathbb{P})$  be a filtered probability space. Let  $X = (X_n)_{n \in \mathbb{N}}$  be an integrable and adapted process.

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

- $X$  is called a super-martingale if  $\mathbb{E}[X_n|\mathcal{F}_m] \leq X_m$  a.s. for all  $n \geq m$ .
- $X$  is called a sub-martingale if  $\mathbb{E}[X_n|\mathcal{F}_m] \geq X_m$  a.s. for all  $n \geq m$ .

**Remark.** If  $X$  is a (super/sub)martingale with respect to  $(\mathcal{F}_n)_{n \in \mathbb{N}}$ , then it is also a martingale with respect to  $(\mathcal{F}_n^X)_{n \in \mathbb{N}}$ . To see this, one has to use the tower property 1.1:  $\mathcal{F}_n^X \subseteq \mathcal{F}_n$  for all  $n \in \mathbb{N}$  implies  $\mathbb{E}[X_n|\mathcal{F}_m^X] = \mathbb{E}[\mathbb{E}[X_n|\mathcal{F}_m]|\mathcal{F}_m^X]$  (since  $\mathbb{E}[X_n|\mathcal{F}_m]$  a.s.).

**Examples:**

1. Let  $(\xi_i)_{i \in \mathbb{N}}$  be iid. s.t.  $\mathbb{E}[\xi_i] = 0$  for all  $i \in \mathbb{N}$  and define  $X = (X_n)_{n \in \mathbb{N}}$  by  $X_n = \xi_1 + \dots + \xi_n$  for all  $n \in \mathbb{N}$ ,  $X_0 = 0$ .  $X$  is a martingales with respect to  $(\mathcal{F}_n^\xi)_{n \in \mathbb{N}}$ .
2. Let  $(\xi_i)_{i \in \mathbb{N}}$  be iid. s.t.  $\mathbb{E}[\xi_i] = 1$  for all  $i \in \mathbb{N}$  and define  $X = (X_n)_{n \in \mathbb{N}}$  by  $X_n = \prod_{i=1}^n \xi_i$  for all  $n \in \mathbb{N}$ ,  $X_0 = 1$ .  $X$  is again a martingales with respect to  $(\mathcal{F}_n^\xi)_{n \in \mathbb{N}}$ .

Lecture 5

Let  $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \in \mathbb{N}}, \mathbb{P})$  be a filtered probability space.

**Definition 2.3 (Stopping time discrete time).** A stopping time  $T$  is a random variable  $T : \Omega \rightarrow \mathbb{Z}_+ \cup \{\infty\}$  s.t.  $\{T \leq n\} \in \mathcal{F}_n$  for all  $n \in \mathbb{N}$ . Equivalently, if  $\{f = n\} \in \mathcal{F}_n$  for all  $n \in \mathbb{N}$  since

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

and

$$\{T \leq n\} = \bigcup_{k=1}^n \{T = k\} \in \mathcal{F}_k \subset \mathcal{F}_n.$$

**Examples:**

1. Constant time are trivially stopping times.
2. Let  $X = (X_n)_{n \in \mathbb{N}}$  be a stochastic process taking values in  $\mathbb{R}$  and  $A \in \mathcal{B}(\mathbb{R})$  ( $X$  adapted). Define

$$T_A = \{n \geq 0 : X_n \in A\}.$$

Then  $\{T_A \leq n\} = \bigcup_{k=0}^n \{X_k \in A\} \in \mathcal{F}_n$  for all  $n \in \mathbb{N}$  (with convention  $\inf \emptyset = \infty$ ).

3.  $L_A = \sup\{n \geq 0 : X_n \in A\}$  is NOT a stopping time.

Properties:  $S, T, (T_n)_{n \in \mathbb{N}}$  stopping times. Then  $S \wedge T, S \vee T, \inf_n T_n, \sup_n T_n, \liminf_n T_n, \limsup_n T_n$  are also stopping times.

**Definition 2.4 (Stopping time sigma algebra).** If  $T$  is a stopping time, define

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

Let  $(X_n)_{n \geq 0}$  be a process. Write  $X_T(\omega) = X_{T(\omega)}(\omega)$  for  $\omega \in \Omega$  whenever  $T(\omega) < \infty$ . Define the stopped process:  $X_t^T := X_{T \wedge t}$ .

**Proposition 2.1.** Let  $S$  and  $T$  be stopping times, and let  $X$  be an adapted process, then:

1. If  $S \leq T$ , then  $\mathcal{F}_S \subseteq \mathcal{F}_T$ .
2.  $X_T \cdot$  is  $\mathcal{F}_T$ -measurable.
3.  $X^T$  is adapted.

4. If  $X$  is integrable, then the stopped process is integrable.

*Proof.* 1. Immediate from definition.

2. Let  $A \in \mathcal{B}(\mathbb{R})$ . Need to show:

$$\{X_T \mathbf{1}(T < \infty)\} \cap \{T \leq t\} \in A, \quad \text{for all } t \geq 0.$$

Indeed, we have that

$$\{X_T \mathbf{1}(T < \infty)\} = \bigcup_{s=0}^t \underbrace{\{X_s \in A\}}_{\mathcal{F}_s \subseteq \mathcal{F}_t} \cap \underbrace{\{T = s\}}_{\mathcal{F}_s} \in \mathcal{F}_t.$$

3.  $X_t^T = X_{T \wedge t}$ , this being  $\mathcal{F}_{T \wedge t}$ -measurable  $\subseteq \mathcal{F}_t$ -measurable by 1), so we conclude it is  $\mathcal{F}_t$ -measurable.

4.

$$\begin{aligned} \mathbb{E}[|X_t^T|] &= \mathbb{E}[|X_{T \wedge t}|] \\ &= \sum_{s=0}^{t-1} \mathbb{E}[|X_s| \cdot \mathbf{1}(T = s)] + \mathbb{E}[|X_t| \cdot \mathbf{1}(T \geq t)] \\ &\leq \sum_{s=0}^t \mathbb{E}[|X_s|] \underbrace{\leq \infty}_{X_t \text{ is integrable}}. \end{aligned}$$

□

We now state and prove a fundamental theorem in Martingale theory:

**Theorem 2.1 (Optional Stopping Theorem discrete time).** Let  $(X_n)$  be a martingale.

1. If  $T$  is a stopping time, then the stopped process  $X^T$  is also a martingale. In particular for all  $t \geq 0$ :

$$\mathbb{E}[X_{T \wedge t}] = \mathbb{E}[X_0].$$

2. If  $S \leq T$  are bounded stopping times, then

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

and hence  $\mathbb{E}[X_T] = \mathbb{E}[X_S]$ .

3. If there exists an integrable random variable  $Y$  such that  $|X_n| \leq Y$  for all  $n \geq 0$  and  $T$  is finite, then  $\mathbb{E}[X_T] = \mathbb{E}[X_0]$ .

4. If there exists  $M \geq 0$  such that  $|X_{n+1} - X_n| \leq M$  for all  $n \in \mathbb{N}$  and  $T$  is a stopping time with  $\mathbb{E}[T] < \infty$ , then  $\mathbb{E}[X_T] = \mathbb{E}[X_0]$ .

*Proof.* 1. NTS: for all  $t \geq 0$ ,  $\mathbb{E}[X_{T \wedge t} | \mathcal{F}_{t-1}] = X_{T \wedge t}$  a.s. Indeed,

$$\begin{aligned} \mathbb{E}[X_{T \wedge t} | \mathcal{F}_{t-1}] &= \sum_{s=0}^{t-1} \mathbb{E}[X_s \cdot \mathbf{1}(T = s) | \mathcal{F}_{t-1}] \mathbb{E}[X_{t-s} \cdot \mathbf{1}(T \geq t) | \mathcal{F}_{t-1}] \\ &= \sum_{s=0}^{t-1} \mathbf{1}(T = s) \cdot X_s + X_{t-1} \cdot \mathbf{1}(T \geq t) \quad \text{a.s.} \\ &= \sum_{s=0}^{t-2} \mathbf{1}(T = s) \cdot X_s + X_{t-1} \cdot \mathbf{1}(T \geq t-1) \quad \text{a.s.} \\ &= X_{T \wedge t-1} \quad \text{a.s.} \end{aligned}$$

2.  $S \leq T \leq n, n \in \mathbb{N}$  fixed. Let  $A \in \mathcal{F}_S$ . NTS:  $\mathbb{E}[X_T \cdot \mathbf{1}(A)] = \mathbb{E}[X_S \cdot \mathbf{1}(A)]$ . We compute

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

Thus,

$$\mathbb{E}[X_T \cdot \mathbf{1}(A)] \stackrel{(A \in \mathcal{F}_S)}{=} \mathbb{E}[X_S \cdot \mathbf{1}(A)] + \sum_{k=0}^{n-1} \mathbb{E}[(X_{k+1} - X_k) \cdot \mathbf{1}(S \leq k < T) \cdot \mathbf{1}(A)]$$

Have,  $A \cap \{S \leq k\} \in \mathcal{F}_k$  and  $A \cap \{T > k\} \in \mathcal{F}_k$ . Thus,  $\mathbf{1}(S \leq k < T) \cdot \mathbf{1}(A)$  is  $\mathcal{F}_k$ -measurable. Using  $\mathbb{E}[X_{k+1} | \mathcal{F}_k] = X_k$  a.s. we deduce that

$$\mathbb{E}[(X_{k+1} - X_k) \cdot \mathbf{1}(S \leq k < T) \cdot \mathbf{1}(A)] = \mathbb{E}[\mathbb{E}[(X_{k+1} - X_k) | \mathcal{F}_k] \cdot \mathbf{1}(S \leq k < T) \cdot \mathbf{1}(A)] = 0$$

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

3. By the Optional Stopping Theorem applied to  $(X_{T \wedge n})_{n \geq 0}$ , we have

$$\mathbb{E}[X_{T \wedge n}] = \mathbb{E}[X_0] \quad \text{for all } n \geq 0.$$

Now,  $T$  being finite a.s. implies that  $X_T = \lim_{n \rightarrow \infty} X_{T \wedge n}$  a.s. By assumption, have  $|X_{T \wedge n}| \leq Y$  a.s. for all  $n \in \mathbb{N}$  and so can apply DCT to conclude.

4. Observe that for all  $n \geq 1$

$$X_{T \wedge n} - X_0 = \sum_{k=0}^{n-1} (X_k - X_0) \cdot \mathbf{1}(T = k) + (X_n - X_0) \mathbf{1}(T \geq n)$$

Hence, we have the bound (using that  $|X_{k+1} - X_k| \leq M$  a.s. for all  $k \geq 0$ )

$$\begin{aligned} |X_{T \wedge n} - X_0| &\leq M \sum_{k=0}^{n-1} k \mathbf{1}(T = k) + n \mathbf{1}(T \geq n) \\ &\leq \mathbb{E}[T] < \infty \quad \text{a.s.} \end{aligned}$$

Now,  $\mathbb{E}[T] < \infty$  gives  $T < \infty$  a.s. and so can deduce as before that  $X_T = \lim_{n \rightarrow \infty} X_{T \wedge n}$  and use the DCT to conclude the argument. □

**Corollary 2.1.1.** Let  $X$  be a positive supermartingale,  $T$  a stopping time such that  $T < \infty$  a.s., then

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

*Proof.* Use Fatou 1.1:  $\mathbb{E}[\liminf_{t \uparrow \infty} X_{T \wedge t}] \leq \liminf_{t \uparrow \infty} \mathbb{E}[X_{T \wedge t}] \leq \mathbb{E}[X_0]$ . □

**Simple random walk on  $\mathbb{Z}$**

Let  $(\xi_i)_{i \geq 0}$  be iid Bernoulli random variables with success probability  $1/2$ . Define the process  $(X_n)_{n \geq 0}$  by setting  $X_n = \xi_1 + \dots + \xi_n$  for all  $n \geq 1$  and  $X_0 = 0$ . Furthermore, let  $T = \inf\{n \geq 0 : X_n = 1\}$ . Using the analysis below, we will see that  $\mathbb{P}(T < \infty) = 1$ . The Optional Stopping Theorem gives  $\mathbb{E}[X_{T \wedge t}] = 0$  for all  $t \geq 0$ . Yet,  $1 = \mathbb{E}[X]_T \neq 0$ . We thus see that the condition  $\mathbb{E}[T] < \infty$  in 4) is necessary, since  $\mathbb{E}[T] = \infty$ .

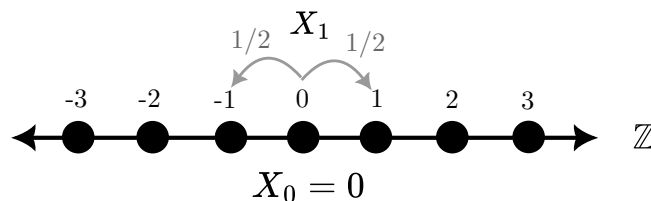


Figure 1: Illustration of simple random walk (first step) on  $\mathbb{Z}$ .

We consider again the example of the simple random walk 2  $(X_n)_{n \in \mathbb{N}}$  and define the stopping times

$$T_c = \inf n \geq 0 : X_n = c, \quad c \in \mathbb{Z}$$

Set  $T = T_{-a} \wedge T_b$  for  $ab \in \mathbb{Z}$ . We now ask what is  $\mathbb{P}(T_{-a} \wedge T_b)$ ?

To answer this, note first that  $X_n^T = X_{T \wedge n}$  is a martingale by the Optional Stopping Theorem and we also have the bounded differences  $|X_{n+1} - X_n| \leq 1$  for all  $n \geq 1$ .

Claim:  $\mathbb{E}[T] < \infty$ .

To show this, we will *stochastically dominate*  $T$  by a geometric random variable, which automatically has a finite expectation and then conclude using the non-negativity of  $T$ . Now we have that for the sequence  $\xi_1, \xi_2, \dots, \xi_{a+b}$  the probability that they all are either  $+1$  or  $-1$  is  $2 \cdot 2^{-(a+b)}$  by independence, call this event  $A_1$ . The same is true for the shifted sequence  $\xi_{k(a+b)+1} \dots \xi_{(k+1)(a+b)}$  for all  $k \in \mathbb{N}$ , where we call the corresponding event  $A_k$ .

Thus, we can bound  $T$  by the random variable

$$Z(\omega) = \inf\{n \geq 0 : \omega \in A_n\}$$

which has the distribution  $Z \sim \text{Geom}(2 \cdot 2^{-(a+b)})$ . Thus,  $\mathbb{E}[T] < \mathbb{E}[Z] \leq (a+b) \cdot 2^{a+b-1} < \infty$ . Thus, we conclude by the OST that  $\mathbb{E}[X_T] = \mathbb{E}[X_0] = 0$ . Hence,  $-a\mathbb{P}(T_a < T_b) + b\mathbb{P}(T_b < T_{-a}) = 0$  and so a quick computation yields that  $\mathbb{P}(T_{-a} < T_b) = \frac{b}{a+b}$ .

### 3 Martingale Convergence Theorem

**Theorem 3.1 (Almost sure martingale convergence theorem).** Let  $X$  be a supermartingale bounded in  $\mathcal{L}^1$ , i.e. satisfying  $\sup \mathbb{E}[|X_n|] < \infty$ . Then, there exists  $X_\infty \in \mathcal{L}^1(\mathcal{F}_\infty)$ ,  $\mathcal{F}_\infty = \sigma(\mathcal{F}_n : n \geq 0)$  such that  $X_n \xrightarrow{n \rightarrow \infty} X_\infty$ , a.s.

Before we embark on the proof of this theorem, we need some supporting results. First we have a result from analysis and we set up some notation. Let  $x = (x_n)_{n \in \mathbb{N}}$  be a real sequence and let  $a < b$  be reals. We proceed to define the *number of upcrossings of the sequence* before time  $n \in \mathbb{N}$ . We construct recursively the sequence of times:

$$\begin{aligned} T_0(x) &= 0 \\ S_{k+1}(x) &= \inf\{n \geq T_k(x) : x_n \leq a\} \\ T_{k+1}(x) &= \inf\{n \geq S_{k+1}(x) : x_n \geq b\} \end{aligned}$$

and

$$N_n([a, b], X) = \sup\{k \geq 0 : T_k(x) \leq n\}$$

Observe that as  $n \rightarrow \infty$ ,  $N_n([a, b], X) \uparrow N([a, b], X) = \sup\{k \geq 0 : T_k(x) < \infty\}$  (see figure 2 for an illustration).

**Lemma 3.1.** Let  $X = (X_n)$  be a real sequence. Then  $X$  converges in  $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$  if and only if for all  $a < b$ ,  $a, b \in \mathbb{Q}$ ,  $N([a, b], X) < \infty$ .

*Proof.*  $\implies$ : Suppose  $x$  converges, if  $a < b$  such that  $N([a, b], x) = \infty$ , then  $\liminf_n x_n \leq a < b \leq \limsup_n x_n$ , a contradiction.

$\impliedby$ : if not, then  $\liminf_n x_n < \limsup_n x_n$  which implies that there exists  $a < b$  in  $\mathbb{Q}$  with  $\liminf_n x_n < a < b < \limsup_n x_n$ , and hence  $N([a, b], x) = \infty$ , a contradiction.  $\square$

Now we state it Doob's upcrossing Inequality

**Lemma 3.2 (Doob's upcrossing inequality).** Let  $X$  be a supermartingale, then for all  $n \in \mathbb{N}$ :

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

*Proof.* It is not hard to check that the sequences of times in 3 are stopping times. Now we have:

$$\begin{aligned} & \sum_{k=1}^n (X_{T_k \wedge n} - X_{S_k \wedge n}) \\ &= \underbrace{\sum_{k=1}^{N_n} (X_{T_k} - X_{S_k})}_{\geq N_n \cdot (b-a)} + (X_n - X_{S_{N_n+1}}) \mathbf{1}(S_{N_n+1} \leq n) \end{aligned}$$

Since  $T_k \wedge n \geq S_k \wedge n$ , the OST gives  $\mathbb{E}[X_{T_k \wedge n}] \leq \mathbb{E}[X_{S_k \wedge n}]$ . Note:

$$\begin{aligned} & \underbrace{X_n - X_{S_{N_n+1}}}_{\geq (X_n - a) \wedge 0 = -(X_n - a)^-} \mathbf{1}(S_{N_n+1} \leq n). \end{aligned}$$

Thus, taking expectations on both sides gives:

$$0 \geq (b - a) \cdot \mathbb{E}[N_n] - \mathbb{E}[(X_n - a)^-].$$

thus concluding the proof.  $\square$

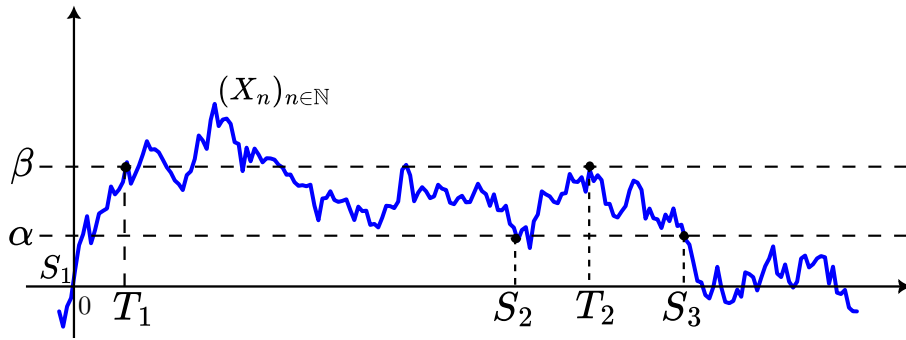


Figure 2: Illustration of upcrossings for the process  $(X_n)_{n \in \mathbb{N}}$ .

Now we proceed to the proof of the martingale convergence theorem:

*Proof.* (Theorem 3.1) Fix  $a < b$ , in  $\mathbb{Q}$ . Have

$$\begin{aligned} \mathbb{E}[N_n([a, b], X)] &\leq (b - a)^- \underbrace{\mathbb{E}[(X_n - a)^-]}_{\leq \mathbb{E}[|X_n| + a]} \\ &\leq (b - a)^- \left( \sup_{n \geq 0} \underbrace{\mathbb{E}[|X_n|]}_{< \infty} + a \right) \end{aligned}$$

Also have  $N_n([a, b], X) \uparrow N([a, b], X)$  as  $n \rightarrow \infty$ . By monotone convergence:  $\mathbb{E}[N([a, b], X)] < \infty$ . Set

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

and  $\mathbb{P}(\Omega_0) = 1$ . On  $\Omega_0$ ,  $X$  converges. set

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

So,  $X_\infty$  is  $\mathcal{F}_\infty$ -measurable,  $X_n \xrightarrow{n \rightarrow \infty} X_\infty$  a.s. and

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

$\square$

**Corollary 3.1.1.** Let  $B$  be a upermaartingale. Then,  $X$  converges a.s.

*Proof.*  $\mathbb{E}[|X_n|] = \mathbb{E}[X_n] \leq \mathbb{E}[X_0]$ . Apply the martingale convergence theorem to conclude.  $\square$

Lecture 7

## 4 Doob's inequalities

**Theorem 4.1 (Doob's maximal inequality).** Let  $X$  be a non-negative submartingale and set  $X_n^* = \sup_{0 \leq k \leq n} X_k$ . Then for all  $\lambda \geq 0$ ,

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

*Proof.* Let  $T = \inf\{k \geq 0 : X_k \geq \lambda\}$  (it is a stopping time). Then  $\{X_n^* \geq \lambda\} = \{T \leq n\}$ . Also have that  $X_{T \wedge n}$  is a submartingale by the OST. Then  $\mathbb{E}[X_{T \wedge n}] \leq \mathbb{E}[X_n]$ . Now,

$$\begin{aligned} \mathbb{E}[X_{T \wedge n}] &= \mathbb{E}[X_T \cdot \mathbf{1}(T \leq n)] \\ &\quad + \mathbb{E}[X_n \cdot \mathbf{1}(T > n)] \\ &\geq \lambda \cdot \mathbb{P}(T \leq n) + \mathbb{E}[X_n \cdot \mathbf{1}(T > n)] \\ &\implies \lambda \cdot \mathbb{P}(T \leq n) \leq \mathbb{E} \left[ \underbrace{X_n \cdot \mathbf{1}(T \leq n)}_{=\{X_n^* \geq \lambda\}} \right] \\ &\leq \mathbb{E}[X_n] \end{aligned}$$

$\square$

**Theorem 4.2 (Doob's  $\mathcal{L}^1$  inequality).** Let  $p > 1$  and let  $X$  be a martingale or a non-negative submartingale. Set  $X_n^* = \sup_{0 \leq k \leq n} |X_k|$ . Then

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

*Proof.* By Jensen, it is enough to prove 4.2 for a non-negative submartingale. Now, observe that

$$\begin{aligned} &= b \\ (y \wedge k)^p &= \int_k^0 p x^{p-1} \mathbf{1}(y \geq x) dx = \mathbb{E} \left[ \int_0^k [x^{p-1} \mathbf{1}(X_n^* \geq x)] dx \right] \\ &\stackrel{\text{Fubini}}{=} \int_0^k p x^{p-1} \underbrace{\mathbb{P}(X_n^* \geq x)}_{\leq \frac{1}{x} \mathbb{E}[X_n \cdot \mathbf{1}(X_n^* \geq x)]} dx \\ &\leq \mathbb{E} \left[ \int_0^k p x^{p-2} \cdot \mathbf{1}(X_n^* \geq x) dx \cdot X_n \right] \\ &= \mathbb{E} \left[ \frac{p}{p-1} (X_n^* \wedge k)^{p-1} \cdot X_n \right] \\ &\stackrel{\text{H\"older}}{\leq} \frac{p}{p-1} \cdot \|X_n\|_p \cdot \|X_n^* \wedge k\|_p^{p-1}. \end{aligned}$$

So we proved  $\|X_n^* \wedge k\|_p^p \leq \frac{p}{p-1} \|X_n\|_p \cdot \|X_n^* \wedge k\|_p^{p-1}$ , which implies  $\|X_n^* \wedge k\|_p \leq \frac{p}{p-1} \cdot \|X_n\|_p$ . Now take  $k \rightarrow \infty$  and use monotone convergence to conclude the argument.  $\square$

**Theorem 4.3 ( $\mathcal{L}^p$ -convergence theorem).** Let  $X$  be a martingale and  $1 < p < \infty$ , then the following are equivalent:

1.  $X$  is bounded in  $\mathcal{L}^p$ , i.e.  $\sup_{n \geq 0} \|X_n\|_p < \infty$ .
2.  $X$  converges underlinealmost surely and in  $\mathcal{L}^p$  to a limit  $X_\infty \in \mathcal{L}^p$ .

3. There exists  $Z \in \mathcal{L}^p$  s.t.  $X_n = \mathbb{E}[Z|\mathcal{F}_n]$  a.s.

*Proof.* 1)  $\implies$  2):  $X$  bounded in  $\mathcal{L}^p$  implies  $X$  is bounded in  $\mathcal{L}^1$ . So there exists  $X_\infty$  such that  $X_n \xrightarrow{n \rightarrow \infty} X_\infty$  a.s.

Also,  $\mathbb{E}[|X_\infty|^p] = \mathbb{E}\left[\liminf_n |X_n|^p\right] \stackrel{\text{Fatou}}{\leq} \liminf_n \mathbb{E}[|X_n|^p] < \infty$ . Thus,  $X_\infty \in \mathcal{L}^p$ .

Now, let  $X_n^* = \sup_{0 \leq k \leq n} |X_k|$ ,  $X_\infty^* = \sup_{k \in \mathbb{N}} |X_k|$ . Thus,

$$|X_n - X_\infty| \leq 2X_\infty^*$$

for all  $n \in \mathbb{N}$ . Thus, it is enough to show by DCT that  $X_\infty^* \in \mathcal{L}^p$ . By Doob's  $\mathcal{L}^p$ -inequality,  $\|X_n^*\|_p = \frac{p}{p-1} \cdot \sup_{n \in \mathbb{N}} \|X_n\|_p < \infty$ . By MCT ( $X_n^* \uparrow X_\infty^*$ ):  $\|X_\infty^*\|_p \leq \frac{p}{p-1} \sup_{n \in \mathbb{N}} \|X_n\|_p < \infty$ . Thus,  $X_\infty^* \in \mathcal{L}^p$ .

2)  $\implies$  3):  $X_n \xrightarrow{n \rightarrow \infty} X_\infty$  a.s. and in  $\mathcal{L}^p$ . Set  $Z = X_\infty$ . Need to show:  $X_n = \mathbb{E}[X_\infty|\mathcal{F}_n]$  for all  $n \in \mathbb{N}$ .

$$\|X_n - \mathbb{E}[X_\infty|\mathcal{F}_n]\|_p \stackrel{m \geq n}{\leq} \|\mathbb{E}[X_m - X_\infty|\mathcal{F}_n]\|_p \stackrel{\text{contraction (Jensen)}}{\leq} \|X_m - X_\infty\|_p \rightarrow 0, \quad m \rightarrow \infty.$$

3)  $\implies$  1): By conditional Jensen, we can conclude.  $\square$

**Definition 4.1.** A martingale of the form  $X_n = \mathbb{E}[Z|\mathcal{F}_n]$ ,  $Z \in \mathcal{L}^p$  is called a martingale closed in  $\mathcal{L}^p$ .

**Corollary 4.3.1.** Let  $Z \in \mathcal{L}^p$ ,  $X_n = \mathbb{E}[Z|\mathcal{F}_n]$  a.s. Then  $X_n \xrightarrow{n \rightarrow \infty} \mathbb{E}[Z|\mathcal{F}_\infty]$  a.s. and in  $\mathcal{L}^p$  where  $\mathcal{F}_\infty = \sigma(X_n, n \geq 0)$ .

*Proof.* By theorem 4.3, we have  $X_n \xrightarrow{n \rightarrow \infty} X_\infty$  a.s. And in  $\mathcal{L}^p$ . Now, we need to show:

$$X_\infty = \mathbb{E}[Z|\mathcal{F}_\infty] \quad \text{a.s.}$$

Now, we have that  $X_\infty$  is  $\mathcal{F}_\infty$ -measurable (being the point wise limit of  $X_n, n \geq 0$ ) and for all  $A \in \mathcal{F}_\infty$ ,  $\mathbb{E}[Z \cdot \mathbf{1}(A)] = \mathbb{E}[X_\infty \cdot \mathbf{1}(A)]$ . Fix  $A \in \bigcup_{n \geq 0} \mathcal{F}_n$ , a  $\pi$ -system generating  $\mathcal{F}_\infty$ . There exists  $N \in \mathbb{N}$  such that  $A \in \mathcal{F}_N$ . Let  $n \geq N$ , now

$$\mathbb{E}[Z \cdot \mathbf{1}(A)] = \mathbb{E}[X_n \cdot \mathbf{1}(A)] \xrightarrow{n \rightarrow \infty} \mathbb{E}[X_\infty \cdot \mathbf{1}(A)].$$

$\square$

**Definition 4.2 (Uniform integrability).** A collection of variables  $(X_i)_{i \in I}$  is called uniformly integrable (UI) if

$$\sup_{i \in I} \mathbb{E}[|X_i| \cdot \mathbf{1}(|X_i| > M)] \xrightarrow{M \rightarrow \infty} 0.$$

Equivalently,  $(X_i)_{i \in I}$  is UI if  $(X_i)$  is bounded in  $\mathcal{L}^1$  and for all  $\epsilon > 0$ , there exists  $\delta > 0$  such that for all  $A \in \mathcal{F}$  with  $\mathbb{P}(A) < \delta$ ,

$$\sup_{i \in I} \mathbb{E}[|X_i| \cdot \mathbf{1}(A_i)] < \epsilon.$$

- A UI family is bounded in  $\mathcal{L}^1$ .
- If a family  $(X_i)$  is bounded in  $\mathcal{L}^p$ ,  $p > 1$ , then it is also UI.



**Lemma 4.1.** Let  $(X_n)_{n \in \mathbb{N}}$ ,  $X$  be in  $\mathcal{L}^1$  and  $X_n \xrightarrow{n \rightarrow \infty} X$  a.s. Then  $X_n \xrightarrow{n \rightarrow \infty}$  in  $\mathcal{L}^1$  if and only if  $(X_n)_{n \in \mathbb{N}}$  is UI.

**Theorem 4.4.** Let  $X \in \mathcal{L}^1$ . The family  $\{\mathbb{E}[X|\mathcal{G} : \mathcal{G} \subset \mathcal{F}]\}$  is uniformly integrable (UI).

*Proof.* Need to show for all  $\epsilon > 0$ , there exists  $\lambda$  large enough such that for all  $\mathcal{G} \subset \mathcal{F}$

$$\begin{aligned} & \mathbb{E}[|\mathbb{E}[X|\mathcal{G}] \cdot \mathbf{1}(|\mathbb{E}[X|\mathcal{G}]| > \lambda)] < \epsilon \\ & \leq \mathbb{E}[\mathbb{E}[|X||\mathcal{G}] \cdot \mathbf{1}(\underbrace{|\mathbb{E}[X|\mathcal{G}]|}_{\leq \mathbb{E}[|X||\mathcal{G}]} > \lambda)]. \end{aligned}$$

Since  $X \in \mathcal{L}^1$ , for all  $\epsilon > 0$ , there exists  $\delta > 0$  such that if  $A \in \mathcal{F}$ ,  $\mathbb{P}(A) < \delta$ , then  $\mathbb{E}[|X| \cdot \mathbf{1}(A)] < \epsilon$ . Now,

$$\begin{aligned} \mathbb{P}(|\mathbb{E}[X|\mathcal{G}]| > \lambda) & \stackrel{\text{Markov}}{\leq} \frac{\mathbb{E}[|\mathbb{E}[X|\mathcal{G}]|]}{\lambda} \\ & \leq \frac{\mathbb{E}[\mathbb{E}[|X||\mathcal{G}]]}{\lambda} = \frac{\mathbb{E}[|X|]}{\lambda}. \end{aligned}$$

Take  $\lambda = \frac{\mathbb{E}[|X|]}{\delta}$ , then we are done. □

**Definition 4.3.**  $X = (X_n)_{n \in \mathbb{N}}$  is called UI (super/sub) martingale if it is a (super/sub) martingale and  $(X_n)_{n \geq 0}$  is UI.

**Examples:**

Let  $X_1, X_2, \dots$  be an iid sequence with  $\mathbb{P}(X_1 = 0) = \mathbb{P}(X_1 = 2) = 1/2$ . Set  $Y_n = X_{12} \cdots X_n$ , which can be seen to be a martingale. Also have  $\mathbb{E}[Y_n] = 1$  for all  $n \in \mathbb{N}$  and  $Y_n \xrightarrow{n \in \mathbb{N}} Y_\infty = 0$  a.s. by the martingale convergence theorem, not not in  $\mathcal{L}^1$  (because it is not UI).

**Theorem 4.5.** Let  $X$  be a martingale. Then the following are equivalent:

1.  $X$  is UI.
2.  $X$  converges a.s. and in  $\mathcal{L}^1$  to  $X_\infty$  as  $n \rightarrow \infty$ .
3. There exists  $Z \in \mathcal{L}^1$  such that  $X_n = \mathbb{E}[Z|\mathcal{F}_n]$  for all  $n \geq 0$ .

*Proof.* 1)  $\implies$  2):  $X$  is bounded in  $\mathcal{L}^1$  implies (by the martingale convergence theorem),  $X_n \rightarrow$  a.s. Since  $X_n$  is UI, then  $X_n \rightarrow X_\infty$  in  $\mathcal{L}^1$ .

2)  $\implies$  3): Set  $Z = X_\infty$ . Need to show:  $X_n = \mathbb{E}[X_\infty|\mathcal{F}_n]$  a.s. Indeed,

$$\begin{aligned} \|X_n - \mathbb{E}[X_\infty|\mathcal{F}_n]\|_1 & \stackrel{m \geq n}{\leq} \|\mathbb{E}[X_m - X_\infty|\mathcal{F}_n]\|_1 \\ & \leq \|X_m - X_\infty\|_1 \xrightarrow{m \rightarrow \infty} 0. \end{aligned}$$

3)  $\implies$  1): The tower property implies  $(X_n)_{n \in \mathbb{N}}$  is a martingale and the previous theorem implies that  $(X_n)_{n \in \mathbb{N}}$  is UI. □

**Remark.** 1. We get as before,  $X_\infty = \mathbb{E}[Z|\mathcal{F}]$  a.s., where  $\mathcal{F}_\infty = \sigma(X_n : n \geq 0)$ .

2. If  $X$  were a UI super/sub martingale, then we would get  $\mathbb{E}[X_\infty|\mathcal{F}_n] \stackrel{\geq \text{sub}}{\leq} X_n$  (check!).

$X$  is UI implies  $X_n \rightarrow X_\infty$  in  $\mathcal{L}^1$  and a.s. Now let  $T$  be a stopping time. We can then define

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

**Theorem 4.6 (Optional stopping theorem for UI martingales).** Let  $X$  be a UI martingale and let  $S, T$  be stopping times with  $S \leq T$ . Then

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

*Proof.* We know that  $X_n = \mathbb{E}[X_\infty | \mathcal{F}_n]$  a.s. since  $X$  is UI. It suffices to prove that for any stopping times  $T$ ,  $\mathbb{E}[X_\infty | \mathcal{F}_T] = X_T$  a.s. Indeed,  $\mathbb{E}[X_T | \mathcal{F}_S] = \mathbb{E}[\mathbb{E}[X_\infty | \mathcal{F}_T] | \mathcal{F}_S]$  and since  $S \leq T$  we have  $\mathcal{F}_S \subseteq \mathcal{F}_T$  and hence the tower property would give:

$$\mathbb{E}[X_T | \mathcal{F}_S] = \mathbb{E}[X_\infty | \mathcal{F}_S] = X_S$$

a.s. Thus, we need to show: for all  $T$  stopping times,  $\mathbb{E}[X_\infty | \mathcal{F}_T] = X_T$  a.s.

1. NTS:  $X_T \in \mathcal{L}^1$ :

$$\begin{aligned} \mathbb{E}[|X_T|] &= \sum_{n=0}^{\infty} \mathbb{E}[|X_n \cdot \mathbf{1}(T = n)|] + \mathbb{E}[|X_\infty| \cdot \mathbf{1}(T = \infty)] \\ &\stackrel{\text{have } X_n = \mathbb{E}[X_\infty | \mathcal{F}_n]}{\leq} \sum_{n=0}^{\infty} \mathbb{E} \left[ \mathbb{E}[|X_\infty| \cdot \underbrace{\mathbf{1}(T = n)}_{\in \mathcal{F}_n}] \right] \\ &\quad + \mathbb{E}[|X_\infty| \cdot \mathbf{1}(T = \infty)] \\ &= \sum_{n=0}^{\infty} \mathbb{E}[|X_\infty| \cdot \mathbf{1}(T = n)] + \mathbb{E}[|X_\infty| \cdot \mathbf{1}(T = \infty)] \\ &= \mathbb{E}[|X_\infty|] < \infty \end{aligned}$$

as  $X_\infty \in \mathcal{L}^1$ . It is also not hard to check that  $X_T$  is  $\mathcal{F}_T$ -measurable.

2. NTS: for all  $B \in \mathcal{F}_T$ :  $\mathbb{E}[X_\infty \cdot \mathbf{1}(B)] = \mathbb{E}[X_T \cdot \mathbf{1}(B)]$

$$\begin{aligned} \mathbb{E}[X_T \cdot \mathbf{1}(B)] &= \sum_{n=0}^{\infty} \mathbb{E} \left[ X_n \cdot \underbrace{\mathbf{1}(T = n) \cdot \mathbf{1}(B)}_{\in \mathcal{F}_n} \right] \\ &\quad + \mathbb{E}[X_\infty \cdot \mathbf{1}(T = \infty) \cdot \mathbf{1}(B)] \\ &= \sum_{n=0}^{\infty} \mathbb{E}[X_\infty \cdot \mathbf{1}(T = n) \cdot \mathbf{1}(B)] \\ &= \mathbb{E}[X_\infty \cdot \mathbf{1}(B)] \end{aligned}$$

□

**Definition 4.4 (Backwards martingales).** Let  $\dots \subseteq \mathcal{G}_{-2} \subseteq \mathcal{G}_{-1} \subseteq \mathcal{G}_0$  be a decreasing family of sub sigma algebras of  $\mathcal{F}$ . We call  $X = (X_n)_{n \leq 0}$  a backwards martingale if  $X_0 \in \mathcal{L}^1$  and for all  $n \leq -1$   $\mathbb{E}[X_{n+1} | \mathcal{G}_n] = X_n$  a.s. By the tower property,  $\mathbb{E}[X_0 | \mathcal{G}_n] = X_n$  for all  $n \leq 0$ . Since  $X_0 \in \mathcal{L}^1$ , a backwards martingale is automatically UI.

**Theorem 4.7 ( $\mathcal{L}^p$ /a.s. backwards martingale convergence theorem).** Let  $X$  be a backwards martingale with  $X_0 \in \mathcal{L}^p$ ,  $1 \leq p < \infty$ . Then  $X_n \rightarrow X_{-\infty}$  as  $m \rightarrow -\infty$  a.s. and in  $\mathcal{L}^p$  and  $X_{-\infty} = \mathbb{E}[X_0 | \mathcal{G}_{-\infty}]$  a.s., where  $\mathcal{G}_{-\infty} = \bigcap_{n \leq 0} \mathcal{G}_n$ .

*Proof.* Set  $\mathcal{F}_k = \mathcal{G}_{-n+k}$ ,  $0 \leq k \leq n$ . This is an increasing filtration and  $(X_{-n+k})_{0 \leq k \leq n}$  is  $\mathcal{F}_k$ -martingale. Let  $N_{-n}([a, b], X)$  be the number of upcrossings of the interval  $[a, b]$  between  $-n$  and 0. Doob's upcrossing inequality gives:

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

As before, we get that  $X_n \rightarrow X_{-\infty}$  as  $n \rightarrow -\infty$  a.s. We also have  $X_{-\infty}$  is  $\mathcal{G}_{-\infty}$ -measurable. Also observe that  $nX_o \in \mathcal{L}^p$  implies  $X_n \in \mathcal{L}^p$  for all  $n \leq 0$ .

Lecture 9  $X_n = \mathbb{E}[X_n|\mathcal{G}_n]$  a.s. (backwards martingale). If  $X_n \in \mathcal{L}^p, p \in [1, \infty) X_{n \rightarrow X_{-\infty}}$  a.s.  $n \rightarrow -\infty$  a.s. and  $X_{-\infty}$  is  $\mathcal{G}_{-\infty} = \bigcap_{n \leq 0} \mathcal{G}_n$ -measurable. □

Observe we have that  $X_n \in \mathcal{L}^p$  by conditional Jensen and using Fatou, we obtain  $X_{-\infty} \in \mathcal{L}^p$ . Now we need to show that  $X_n \rightarrow X_{-\infty}$  in  $\mathcal{L}^p$ . Indeed,

$$\begin{aligned} |X_n - X_{-\infty}|^p &= |\mathbb{E}[X_0|\mathcal{G}_n] - \mathbb{E}[X_{-\infty}|\mathcal{G}_n]|^p \\ &= |\mathbb{E}[X_0 - X_{-\infty}|\mathcal{G}_n]|^p \\ &\stackrel{\text{Jensen}}{\leq} \underbrace{\mathbb{E}[|X_0 - X_{-\infty}|^p|\mathcal{G}_n]}_{\text{UI family}}. \end{aligned}$$

Hence,  $(|X_n - X_{-\infty}|^p)_{n \leq 0}$  is UI, hence giving  $\mathcal{L}^1$  convergence.

NTS:  $X_{-\infty} = \mathbb{E}[X_o|\mathcal{G}_{-\infty}]$  a.s.

Let  $A \in \mathcal{G}_{-\infty} = \bigcap_{n \geq 0} \mathcal{G}_n$  implies that  $A \in \mathcal{G}_n$  for all  $n \leq 0$ . Hence,  $\mathbb{E}[X_n \cdot \mathbf{1}(A)] = \mathbb{E}[X_0 \cdot \mathbf{1}(A)]$ , for all  $n \leq 0$ . Take  $n \rightarrow -\infty$  and use  $\mathcal{L}^1$  convergence to get  $\mathbb{E}[X_{-\infty} \cdot \mathbf{1}(A)] = \mathbb{E}[X_0 \cdot \mathbf{1}(A)]$  so conclude.

## 5 Applications of martingales

sec: applications of mgs

**Theorem 5.1 (Kolmogorov's 0 - 1 law).** Let  $(X_i)$  be iid and for all  $n \in \mathbb{N}, \mathcal{F}_n = \sigma(X_k : k \geq n), \mathcal{F}_\infty = \bigcap_{n \geq 0} \mathcal{F}_n$ . Then,  $\mathcal{F}_\infty$  is trivial, i.e. for all  $A \in \mathcal{F}_\infty, \mathbb{P}(A) \in \{0, 1\}$ .

*Proof.* Let  $A \in \mathcal{F}_\infty$ . Define  $\mathcal{G}_\setminus = \sigma(\mathcal{X}_\setminus : \setminus \leq \setminus)$  and  $\mathcal{G}_\infty = \sigma(\mathcal{G}_n, n \geq 0)$ . Now, we have that  $\mathbb{E}[\mathbf{1}(A)|\mathcal{G}_n]$  is a martingale and

$$\mathbb{E}[\mathbf{1}(A)|\mathcal{G}_n] \xrightarrow{n \rightarrow \infty} \mathbb{E}[\mathbf{1}(A)|\mathcal{G}_\infty] \quad \text{a.s.}$$

Now,  $A \in \mathcal{F}_\infty$  implies that  $A \in \mathcal{F}_{n+1}$  and also have  $\mathcal{G}_n \perp \mathcal{F}_{n+1}$  and  $\mathbb{E}[\mathbf{1}(A)|\mathcal{G}_n] = \mathbb{P}(A)$  a.s.,  $\mathbb{E}[\mathbf{1}(A)|\mathcal{G}_\infty] = \mathbf{1}(A)$  a.s. since  $\mathcal{F}_\infty \subseteq \mathcal{G}_\infty$  implies that  $A \in \mathcal{G}_\infty$ . So  $\mathbb{P}(A) = \mathbf{1}(A)$  a.s. finally giving  $\mathbb{P}(A) \in \{0, 1\}$ . □

**Theorem 5.2 (Strong law of large numbers).** Let  $(X_i)_{i \in I}$  be an iid sequence in  $\mathcal{L}^1$  with  $\mathbb{E}[X_1]$ . Define  $S_n = X_1 + \dots + X_n$ . Then  $\frac{S_n}{n}$  converges a.s. and in  $\mathcal{L}^1$  to  $\mu$  as  $n \rightarrow \infty$  a.s.

*Proof.* Define  $\mathcal{G} = \sigma(S_n, S_{n+1}, \dots) = \sigma(S_n, X_{n+1}, \dots)$ . For  $n \leq -1, M^n = \frac{S_{-n}}{-n}$ . We will show that  $(M_n)_{n \leq -1}$  is a backwards martingale with respect to  $(\mathcal{G}_{-n, n \leq -1})$ . Indeed,

$$\begin{aligned} \mathbb{E}[M_{m+1}|\mathcal{G}_{-m}] &= M_{-m} \quad \text{a.s. for } m \leq -1 \\ &= \mathbb{E}\left[\frac{S_{-m-1}}{-m-1} \middle| \mathcal{G}_{-m}\right] \stackrel{\text{set } n = -m}{=} \mathbb{E}\left[\frac{S_{n-1}}{n-1} \middle| \mathcal{G}_n\right] \\ &= \mathbb{E}\left[\frac{S_{n-1}}{n-1} \middle| S_{n-1}, X_{n+1}, \dots\right] \\ &= \mathbb{E}\left[\frac{S_n - X_n}{n-1} \middle| S_n\right] \\ &= \frac{S_n}{n-1} - \mathbb{E}\left[\frac{X_n}{n-1} \middle| S_n\right]. \end{aligned}$$

Now since  $S_n = X_1 + \dots + X_n$ , we have that  $\mathbb{E}[X_k|S_n] = \mathbb{E}[X_1|S_n]$  and so  $\frac{S_n}{n-1} - \frac{1}{n-1} \left(\frac{S_n}{n}\right) = \frac{S_n}{n-1} \left(\frac{n-1}{n}\right) = \frac{S_n}{n}$ . Hence  $\frac{S_n}{n} \xrightarrow{n \rightarrow \infty} Y$  a.s. and in  $\mathcal{L}^1$  measurable for all  $k \geq 0$ . Thus  $Y$  is

$\underbrace{\bigcap_k \sigma(X_{k+1}, \dots)}_{\text{Kolmogorov 0-1 law} \implies \text{trivial}}$  –measurable. So there exists  $c \in \mathbb{R}$  such that  $\mathbb{P}(Y = c) = 1$ . So  $\frac{S_n}{n} \xrightarrow{n \rightarrow \infty} c$  in  $\mathcal{L}^1$  and hence  $c = \mathbb{E}[Y] = \lim_{i \rightarrow \infty} \mathbb{E}[\frac{S_n}{n}] = \mu$  and so finally  $c = \mu$ .  $\square$

**Theorem 5.3 (Radon-Nikodym Theorem).** Let  $P$  and  $Q$  be two probability measures on the space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Suppose that  $\mathcal{F}$  is countable generated, i.e. there exists a sequence  $(F_n)_{n \in \mathbb{N}}$  such that  $\mathcal{F} = \sigma(F_n : n \in \mathbb{N})$ . Then the following are equivalent:

1. For all  $A \in \mathcal{F}$ ,  $\mathbb{P}(A) = 0$  implies  $Q(A) = 0$ . ( $Q \ll P$ ).
2. For all  $\epsilon > 0$ , there exists  $\delta > 0$  such that if  $A \in \mathcal{F}$  with  $\mathbb{P}(A) < \delta$ , then  $Q(A) < \epsilon$ .
3. There exists a non-negative random variable  $X$  such that  $Q(A) = \mathbb{E}[X \cdot \mathbf{1}(A)]$ , for all  $A \in \mathcal{F}$ .

**Remark.**  $X$  is called a version of the Radon-Nikodym derivative of  $Q$  with respect to  $P$ , or  $X = \frac{dQ}{dP}$  on  $\mathcal{F}$  a.s.

*Proof.* 1)  $\implies$  2): Suppose 2) does not hold, then there exists an  $\epsilon > 0$  such that for all  $n \in \mathbb{N}$ , there exist  $A_n$  with  $P(A_n) \leq \frac{1}{n^2}$  and  $Q(A_n) \geq \epsilon$ . Now, since  $\sum_{n=1}^{\infty} P(A_n) < \infty$  Borel-Cantelli implies  $P(A_n \text{ i.o.}) = 0$  and so  $Q(A_n) = 0$ . However,

$$\begin{aligned} \{A_n \text{ i.o.}\} &= \bigcap_n \bigcup_{k \geq n} A_k \implies Q(A_n \text{ i.o.}) \\ &= \lim_{n \rightarrow \infty} Q\left(\bigcup_{k \geq n} A_k\right) \\ &\geq \lim_{n \rightarrow \infty} Q(A_n) \geq \epsilon \end{aligned}$$

a contradiction.

3)  $\implies$  1): trivial.

2)  $\implies$  3): Let  $\mathcal{A}_n = \{H_1 \cap \dots \cap H_n : H_i = F_i \text{ or } F_i^c \text{ for all } i\}$ . In other words  $\mathcal{A}_n = \{F_1, F_2, \dots, F_n, \bigcup_{k \geq n} F_k\}$ . Let  $\mathcal{F}_N = \sigma(\mathcal{A}_N)$ , so  $\mathcal{F}_N$  is a filtration.

Now defined

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

Thus, for all  $A \in \mathcal{F}_n$ ,  $\mathbb{E}[X_n \cdot \mathbf{1}(A)] = Q(A) = \mathbb{E}[X_{n+1} \cdot \mathbf{1}(A)]$ . So  $(X_n)_{n \in \mathbb{N}}$  is indeed a martingale. Furthermore  $\mathbb{E}[X_n] = Q(\Omega) = 1$  (and since  $X_n \geq 0$  for all  $n \geq 0$ ), we have that  $X_n$  is an  $\mathcal{L}^1$  bounded martingale. Thus,  $X_n \xrightarrow{n \rightarrow \infty} X_\infty$  a.s.

Now we show that  $(X_n)_{n \in \mathbb{N}}$  is UI:

$$\begin{aligned} \mathbb{P}(X_n \geq \lambda) &\leq 1/\lambda < \infty \\ &\leq \delta \end{aligned}$$

using Markov's inequality and taking  $\lambda = 1/\delta$ . Thus,  $\mathbb{E}[X_n \cdot \mathbf{1}(X_n \geq \lambda)] = Q(X_n \geq \lambda) < \epsilon$ . Thus  $(X_n)_{n \in \mathbb{N}}$  is UI and so  $X_n \rightarrow X_\infty$  in  $\mathcal{L}^1$ .

Now define  $\tilde{Q}(A) = \mathbb{E}[X_\infty \cdot \mathbf{1}(A)]$ . Want to show:  $\tilde{Q}(A) = Q(A)$  for all  $A \in \mathcal{F}$ . Indeed, we have  $X_n = X_\infty | \mathcal{F}_n$ . Now if we let for a moment  $A \in \bigcup_{n \geq 0} \mathcal{F}_n$ , there exists some  $N \in \mathbb{N}$  such that  $A \in \mathcal{F}_N$ . Thus,

$$\underbrace{\mathbb{E}[X_N \cdot \mathbf{1}(A)]}_{=Q(A)} = \underbrace{\mathbb{E}[X_\infty \cdot \mathbf{1}(A)]}_{=\tilde{Q}(A)}.$$

Hence,  $Q = \tilde{Q}$  on a  $\pi$ -system,  $(\bigcup_n \mathcal{F}_n)$ , that generates  $\mathcal{F}$ , and by the extension theorem we have that  $Q \equiv \tilde{Q}$  everywhere.  $\square$

## 6 Continuous Time processes

Let  $X = (X_n)_{n \in \mathbb{N}}$  be a process, that is for all  $n \in \mathbb{N}$   $X_n$  is a random variable on some underlying probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ .  $X$  can also be viewed as the map

$$X : (\omega, n) \mapsto X_n(\omega).$$

and observe that this map is  $\mathcal{F} \otimes \mathcal{P}(\mathbb{N}) = \sigma(\{A \times \{k\} : A \in \mathcal{F}, k \in \mathbb{N}\})$  as long as  $X_n$  is  $\mathcal{F}$ -measurable for all  $n \in \mathbb{N}$ . Now we consider random variables taking values in the spaces  $\mathbb{R}^d$ ,  $d \geq 1$ .

**Definition 6.1 (Stochastic process).** The family  $(X_t)_{t \in \mathbb{R}_+}$  is called a stochastic process if for all  $t$  positive  $X_t$  is a random variable.

**Remark.** The map  $X : (\omega, t) \mapsto X_t(\omega)$  need not be  $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+)$ -measurable.

Claim: If for all  $\omega \in \Omega$ ,  $t \mapsto X_t(\omega)$  is a continuous function for  $t \in (0, 1]$ , then the map  $X : (\omega, t) \mapsto X_t(\omega)$  is  $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+)$ -measurable.

Indeed, by continuity we can write

$$X_t(\omega) = \lim_{n \rightarrow \infty} \overbrace{\sum_{k=0}^{2^n-1} \mathbf{1}(t \in (k \cdot 2^{-n}, (k+1) \cdot 2^{-n}]) X_{k \cdot 2^{-n}}(\omega)}^{\text{for all } n \text{ this sum is } \mathcal{F} \otimes \mathcal{B}((0,1]) \text{-meas.}}$$

Thus  $X$  is measurable with as a limit of measurable functions.

From now onwards, we will always (unless otherwise stated) assume that  $X$  is right-continuous and admits left limits, almost everywhere. We call such processes cadlag.

We now revisit some of the earlier definition we have made in the discrete setting and extend the to the continuous case. A filtration is an increasing family of sigma algebras  $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$  whenever  $t \leq t'$ . We say  $X$  is adapted to the filtration above if  $X_t$  is  $\mathcal{F}_t$ -measurable for all  $t \in \mathbb{R}_+$ . A random variable  $T : \Omega \rightarrow [0, \infty]$  is called a stopping time if for all  $t$ ,  $\{T \leq t\} \in \mathcal{F}_t$ . Define  $\mathcal{F}_T = \{A \mid \text{in } \mathcal{F} : A \cap \{T \leq t\} \in \mathcal{F}_t \text{ for all } t\}$  and  $A \mid \text{in } \mathcal{B}(\mathbb{R})$ . Furthermore,  $T_A = \inf_{t \geq 0: X_t \in A}$  is not always a stopping time.

$$\{T_A \leq t\} = \bigcup_{s \leq t} \{X_s \in A\}$$

an uncountable union so not immediately clear whether it in  $\mathcal{F}_t$ .

**Examples:**

Let  $J = \begin{cases} 1, & \text{with probability } \frac{1}{2} \\ -1, & \text{with probability } \frac{1}{2} \end{cases}$  and

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

Let  $(\mathcal{F}_t)_{t \geq 0} = (\mathcal{F}_t^X)_{t \geq 0}$  and fix  $A \in (1, 2)$ . Then  $\{T_A \leq 1\} \mid \text{in } \mathcal{F}_1 = \{\emptyset, \Omega\}$ , since  $\{T_A \leq 1\} = \{J = 1\}$ .

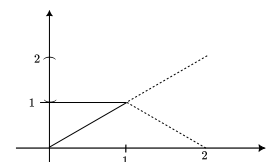


Figure 3: Illustration of  $X$ .

Again, we say  $X_t^T = X_{T \wedge t}$ ,  $X_T(\omega) = X_{T(\omega)}(\omega)$  whenever  $T(\omega) < \infty$ .

**Proposition 6.1.** Let  $S, T$  be stopping times and  $X$  a cadlag adapted process. Then

1. If  $S \leq T$ , then  $\mathcal{F}_S \subseteq \mathcal{F}_T$ .
2.  $S \wedge T$  is a stopping time.
3.  $X_T \cdot \mathbf{1}(T < \infty)$  is  $\mathcal{F}_T$ -measurable.

4.  $X^T$  is adapted.

*Proof.* 1), 2) are clear (check!) and 4) is immediate from 3), since  $X_{T \wedge t}$  is  $\mathcal{F}_{T \wedge t}$ -measurable and  $\mathcal{F}_{T \wedge t} \subseteq \mathcal{F}_t$ .

proof of 3): Claim: a random variable  $Z$  is  $\mathcal{F}_{T \wedge t}$ -measurable if and only if  $Z \cdot \mathbf{1}(T \leq t)$  is  $\mathcal{F}_t$ -measurable for all  $t \geq 0$ . Indeed,

$\Leftarrow$ ): is true by definition.

$\Rightarrow$ ): if  $Z = c \cdot \mathbf{1}(A)$ ,  $A \in \mathcal{F}$ , then  $A \in \mathcal{F}_T$  which means that  $Z$  is  $\mathcal{F}_T$ -measurable. Now, if  $Z = \sum_i c_i \cdot \mathbf{1}(A_i)$ , a finite sum with  $c_i > 0$ ,  $A_i \in \mathcal{F}$ , then  $Z$  is  $\mathcal{F}_T$ -measurable.

$Z$  general ( $\geq 0$ ): let  $Z_n \uparrow Z$ , where

$$Z_n = 2^{-n} \lfloor 2^n Z \rfloor, \quad \text{for all } n \in \mathbb{N}.$$

Observe that  $Z_n$  are simple for all  $n$  and so by the previous steps  $Z_n$  is  $\mathcal{F}_T$ -measurable and hence so is  $Z$ , being an a.s. pointwise limit of measurable functions.

The case for completely general  $Z$  follows by decomposing  $Z = Z^+ - Z^-$ ,  $Z^+ = Z \vee 0$ ,  $Z^- = (-Z) \vee 0$  and apply the previous case to  $Z^+$ ,  $Z^-$ .

Now, by the above claim, it suffices to show:  $X_T \cdot \mathbf{1}(T \leq t)$  is  $\mathcal{F}_t$ -measurable for all  $t$ . We have  $X_T \mathbf{1}(T \leq t) = X_T \cdot \mathbf{1}(T < t) + X_t \cdot \mathbf{1}(T = t)$ . Hence, it suffices to show that  $X_T \cdot \mathbf{1}(T < t)$  is  $\mathcal{F}_t$ -measurable for all  $t$ .

Define  $T_n = 2^{-n} \lfloor 2^n T \rfloor$ , stopping times since

$$\begin{aligned} \{T_n \leq t\} &= \{\lfloor 2^n T \rfloor \leq 2^n t\} \\ &= \{2^n T \leq \lfloor 2^n t \rfloor + 1\} = \{T \leq 2^{-n} \lfloor 2^n t \rfloor + 2^{-n}\} \\ &\in \mathcal{F}_{2^{-n} \lfloor 2^n t \rfloor + 2^{-n}} \subseteq \mathcal{F}_t. \end{aligned}$$

Also,  $T_n \downarrow T$ , as  $n \rightarrow \infty$ . Now by the cadlag property of  $X$ ,  $X_T \cdot \mathbf{1}(T < t) = \lim_{n \rightarrow \infty} X_{T_n \wedge t} \cdot \mathbf{1}(T < t)$ .

Furthermore,  $T_n$  takes values in  $\mathcal{D}_n = \{k \cdot 2^{-n}, k \in \mathbb{N}\}$ . Now,

$$\begin{aligned} X_{T_n \wedge t} \cdot \mathbf{1}(T < t) &= \sum_{d \in \mathcal{D}_n, d \leq t} \overbrace{X_d \cdot \mathbf{1}(T_n = d) \cdot \mathbf{1}(T < t)}^{\mathcal{F}_t\text{-meas.}} \\ &\quad + X_t \cdot \mathbf{1}(T_n = t) \cdot \mathbf{1}(T < t). \end{aligned}$$

Hence,  $X_T \cdot \mathbf{1}(T < \infty)$  is  $\mathcal{F}_t$ -measurable as a limit of  $\mathcal{F}_t$ -measurable functions. □

**Proposition 6.2.** Let  $X$  be a continuous and adapted process and let  $A$  be a closed set. Then  $T_A = \{t \geq 0 : X_t \in A\}$  is a stopping time.

*Proof.* Need to show:  $\{T_A \leq t\} = \left\{ \inf_{s \in \mathbb{Q}, s \leq t} d(X_s, A) = 0 \right\}$ .

$(\subseteq)$ :  $d(x, A)$  = distance of  $x$  from  $A$ . Let  $T_A = s \leq t$ , then there exists a sequence  $s_n \downarrow s$ , such that  $X_{s_n} \in A$ . Since  $A$  is closed, we have  $d(X_{s_n}, A) = 0$  and  $X_{s_n} \rightarrow X_s$ , as  $n \rightarrow \infty$ . Again  $A$  being closed implies that  $d(X_s, A) = 0$ . The continuity of  $X$  and  $d(\cdot, A)$  means that there exists another sequence  $(q_n)_{n \in \mathbb{N}} \subseteq \mathbb{Q}$  such that  $q_n \uparrow s$  such that  $d(X_{q_n}, A) \rightarrow 0$  hence  $\inf_{s \in \mathbb{Q}, s \leq t} d(X_s, A) = 0$ .

$(\supseteq)$ : If  $\inf_{s \in \mathbb{Q}, s \leq t} d(X_s, A) = 0$ , then there exists a sequence  $(s_n)_{n \in \mathbb{N}}$  such that  $s_n \leq t$  for all  $n$  and  $d(X_{s_n}, A) \rightarrow 0$  as  $n \rightarrow \infty$ . Then by compactness, there exists a convergent subsequence of  $s_n \rightarrow s$  (without relabelling), such that  $s \leq t$  and  $d(X_{s_n}, A) \rightarrow 0$  as  $n \rightarrow \infty$  and by continuity we obtain  $d(X_s, A) = 0$ , hence  $X_s \in A$  and so  $T_A \leq t$ . □

**Definition 6.2.** Given a filtration  $(\mathcal{F}_t)_{t \geq 0}$ , we define  $\mathcal{F}_{t+} = \bigcap_{s > t} \mathcal{F}_s$ , for all  $t \geq 0$ . Observe that  $(\mathcal{F}_{t+})_{t \geq 0}$  is a filtration. If for all  $t \geq 0$ ,  $\mathcal{F}_{t+}$ , we say  $(\mathcal{F}_t)_{t \geq 0}$  is right-continuous.

Lecture 11

**Proposition 6.3.** Let  $X$  be a continuous process, and  $A$  be an open set. Then

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

is a stopping time with respect to the filtration  $(\mathcal{F}_{t+})_{t \geq 0}$ .

*Proof.* Need to show: for all  $t \geq 0$ ,  $\{T_A \leq t\} \in \mathcal{F}_{t+}$ . Have,

$$\begin{aligned} \{T_A < s\} &= \bigcup_{q \in \mathbb{Q}, q < s} \underbrace{X_q \in A}_{\in \mathcal{F}_s} \in \mathcal{F}_s \\ \{T_A \leq t\} &= \bigcap_n \underbrace{\{T_A < t + \frac{1}{n}\}}_{\in \mathcal{F}_{t+\frac{1}{n}}} \in \mathcal{F}_{t+}. \end{aligned}$$

□

Let  $(X_t)_{t \geq 0}$  be a stochastic process. It can be viewed, as a random element in the space of functions  $\{f : \mathbb{R}_+ \rightarrow E\}$  endowed with the product sigma-algebra making all projections measurable. Further, let  $\mathcal{C}(\mathbb{R}_+, E)$  be the space of all continuous functions and  $\mathcal{D}(\mathbb{R}_+, E)$  the space of all cad lag functions. Endow the spaces  $\mathcal{C}, \mathcal{D}$  with the sigma algebra that makes all projections  $\pi_t : f \mapsto f_t$  measurable for all  $t \geq 0$ . This sigma algebra is generated by the cylinder sets

$$\left\{ \bigcap_{s \in J} \{f_s \in A_s : \text{for all } T \subseteq \mathbb{R}_+, \text{ finite, } A_s \in \mathcal{B}(E)\} \right\}.$$

For  $A$  in the product sigma algebra, we write  $\mu(A) = \mathbb{P}(X \in A)$  and we call  $\mu$  the law of  $X$ . (“ $X_*\mathbb{P} = \mu$ ”). For every  $J$  finite subset of  $\mathbb{R}_+$ , write  $\mu_J$  for the law of  $(X_t)_{t \in J}$ . The measures  $(\mu_J)$  are called the finite dimensional marginals of  $X$ . The  $\mu_J$  completely characterise the law of  $\mu$ . This follows because the sets above form a  $\pi$ -system that generates the sigma fields previously mentioned.

**Examples:**

Let  $X = 0$  for all  $t \in [0, 1]$  and  $U \sim [0, 1]$  (uniform) and  $X_{t'} = \mathbf{1}(U = t)$  for  $t \in [0, 1]$ . Both of them have the same finite dimensional distributions which are Dirac masses at zero, but the processes are not equal.

$$\begin{aligned} \mathbb{P}(X_t = 0 \text{ for all } t \in [0, 1]) &= 1 \\ \mathbb{P}(X'_t = 0 \text{ for all } t \leq 1) &= 0. \quad \text{But,} \\ \mathbb{P}(X_t = X'_t) &= 1 \quad \text{for all } t \in [0, 1]. \end{aligned}$$

**Definition 6.3.** Let  $X$  and  $X'$  be two processes on  $(\Omega, \mathcal{F}, \mathbb{P})$ , we say  $X'$  is a version of  $X$  if  $(X_t = X'_t \text{ a.s.})$  for all  $t$ . That is

$$\text{For all } t \geq 0 : \mathbb{P}(X_t = X'_t) = 1.$$

**Definition 6.4.** Fix a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ . Set  $\mathcal{N}$  to be the collection of sets of measure zero. Furthermore, set

$$\tilde{\mathcal{F}}_t = \sigma(\mathcal{F}_t, \mathcal{N})$$

for all  $t \geq 0$ . If for all  $t$ ,  $\mathcal{F}_t = \tilde{\mathcal{F}}_t$ , we say that  $(\mathcal{F}_t)_{t \geq 0}$  satisfies the usual conditions.

**Theorem 6.1 (Martingale regularisation theorem).** Let  $(X_t)_{t \geq 0}$  be a martingale wrt  $(\mathcal{F}_t)_{t \geq 0}$ . Then, there exists a cadlag process  $(\tilde{X}_t)_{t \geq 0}$  satisfying for all  $t \geq 0$ :

$$X_t = \mathbb{E} \left[ \tilde{X}_t | \mathcal{F}_t \right] \quad \text{a.s.}$$

and  $X$  is a martingale with respect to the augmented filtration  $(\tilde{\mathcal{F}}_t)_{t \geq 0}$ . If  $(\mathcal{F}_t)_{t \geq 0}$  satisfies the usual conditions, then  $\tilde{X}$  is a version of  $X$ .

We start with a Lemma

**Lemma 6.1.** Let  $f : \mathbb{Q}_+ \rightarrow \mathbb{R}$  such that for all  $I \subseteq \mathbb{Q}_+$  bounded,  $f$  is bounded on  $I$  and for any  $a < b, a, b \in \mathbb{Q}_+$ , for all  $I$  bounded and suppose

$$\mathcal{N}([a, b], I, f) = \sup \{n \geq 0 : \text{there exists } 0 < s_1 < t_1 < \dots < s_n < t_n, \\ s_i, t_i \in I \text{ s.t. } f(s_i) < a, f(t_i) > b\} < \infty.$$

Then, for all  $t \geq 0$ , the limits

$$\lim_{s \uparrow t, s \in \mathbb{Q}_+} f(s), \quad \lim_{s \downarrow t, s \in \mathbb{Q}_+} f(s)$$

exist and are finite.

*Proof.* Let  $s_n \downarrow t$ , the sequence  $(f(s_n))$  will converge by the finite upcrossing property (see lemma 3.1). Now suppose  $t_n \downarrow t$  is another such sequence, then combining them (by selecting elements from each sequence in an alternating fashion exploiting convergence) we get a decreasing sequence converging to  $t$  to conclude  $\lim_{n \rightarrow \infty} f(s_n) = \lim_{n \rightarrow \infty} f(t_n)$ . Finally,  $f$  being bounded gives that both limits are equal and finite.  $\square$

Goal: To define  $\tilde{X}_t = \lim_{s \downarrow t, s \in \mathbb{Q}_+} X_s$  on a set of measure 1, and zero otherwise. We now outline below the main steps in the proof of Theorem 6.1.

Steps:

1. Show that the limit exists and is finite on a set of measure one.
2. Show that  $\tilde{X}$  is  $\tilde{\mathcal{F}}_t$ -measurable and satisfies  $\mathbb{E} \left[ \tilde{X}_t | \mathcal{F}_t \right]$  a.s. for all  $t \geq 0$ .
3.  $\tilde{X}$  is a  $(\tilde{\mathcal{F}}_t)_{t \geq 0}$  martingale.
4.  $\tilde{X}$  is cadlag.

*Proof.* (Theorem 6.1)

1. Let  $I$  be a bounded subset of  $\mathbb{Q}_+$ . Need to show that  $\mathbb{P} \left( \sup_{t \in I} |X_t| < \infty \right) = 1$ . Observe that

$$\sup_{t \in I} |X_t| = \sup_{J \subseteq I, J \text{ finite}} \sup_{t \in J} |X_t|.$$

Now, let  $J = \{j_1, \dots, j_n\} \subseteq I$  with  $j_1 < \dots < j_n$  and  $k > \sup I$ . Then  $(X_t)_{t \in J}$  is a discrete time martingale. Hence the maximal inequality in 4.1 gives

$$\lambda \cdot \mathbb{P}(\sup_{t \in J} |X_t| \geq \lambda) \leq \mathbb{E} [|X_{j_n}|] \leq \mathbb{E} [|X_k|]$$

by the martingale property and Jensen. Now taking the limit as  $J \uparrow I$ ,

$$\lambda \cdot \mathbb{P} \left( \sup_{t \in I} |X_t| \geq \lambda \right) \leq \mathbb{E} [|X_{j_n}|] \leq \mathbb{E} [|X_k|]$$

So,  $\mathbb{P} \left( \sup_{t \in I} |X_t| \geq \lambda \right) = 1$ . Now for  $M \in \mathbb{N}$  define  $I_M = \mathbb{Q}_+ \cap [0, M]$ , then by the above,

$$\mathbb{P} \left( \bigcap_{M \in \mathbb{N}} \left\{ \sup_{t \in I_M} |X_t| < \infty \right\} \right) = 1$$



Lecture 12

and on the above event,  $X_t$  is bounded on bounded intervals of  $\mathbb{Q}_+$ .

Let  $a < b$ ,  $a, b \in \mathbb{Q}_+$ ,  $I \subseteq \mathbb{Q}_+$ , bounded. Observe that

$$\mathcal{N}([a, b], I, X) = \sup_{I \subseteq J, J \text{ finite}} \mathcal{N}([a, b], J, X).$$

Now, let  $J = \{j_1, \dots, j_n\} \subseteq I$  with  $j_1 < \dots < j_n$  and  $k > \sup I$ . Then  $(X_t)_{t \in J}$  is a discrete time martingale. Now, Doob's upcrossing inequality from 3.2 gives

$$(b - a) \cdot \mathbb{E}[\mathcal{N}([a, b], J, X)] \leq \mathbb{E}[(X_{j_n} - a)^-] \leq \mathbb{E}[(X_k - a)^-].$$

By monotone convergence, we get

$$(b - a) \cdot \mathbb{E}[\mathcal{N}([a, b], I, X)] < \infty.$$

Let  $M \in \mathbb{N}$ ,  $I_M = \mathbb{Q}_+ \cap [0, M]$  and

$$\Omega_0 = \bigcap_{m \in \mathbb{N}} \left( \bigcap_{a < b, a, b \in \mathbb{Q}} \{\mathcal{N}([a, b], I_M, X) < \infty\} \cup \left\{ \sup_{t \in I_m} |X_t| < \infty \right\} \right).$$

On  $\Omega_0$ , from lemma 6.1,  $\lim_{s \downarrow t, X_s}$  exists and we have  $\mathbb{P}(\Omega_0) = 1$ . Now, define

$$\tilde{X}_t = \begin{cases} \lim_{s \downarrow t, s \in \mathbb{Q}_+} X_s, & \text{on } \Omega_0 \\ 0, & \text{otherwise.} \end{cases}$$

Recall  $\tilde{\mathcal{F}}_t = \sigma(\mathcal{F}_t, \mathcal{N})$  for all  $t \geq 0$ . From the definition definition, we see that  $\tilde{X}$  is  $\tilde{\mathcal{F}}$ -adapted.

It remains to show that  $X_t = \mathbb{E}[\tilde{X}_t | \mathcal{F}_t]$  a.s. and  $\tilde{X}$  is cadlag and a martingale.

2. Let  $t_n \downarrow t$ ,  $t_n \in \mathbb{Q}_+$ , then

$$\tilde{X}_t = \lim_{n \rightarrow \infty} X_{t_n}$$

a.s. Observe that  $(X_{t_n})$  is a backwards martingale with respect to the filtration  $(\mathcal{F}_{t_n})_{n \in \mathbb{N}}$ . So  $(X_{t_n})$  converges a.s. and in  $\mathcal{L}^1$ . In other words,  $X_t = \lim_{n \rightarrow \infty} X_{t_n}$  a.s. So  $X_t = \mathbb{E}[\tilde{X}_t | \mathcal{F}_t]$  a.s.

3. We now prove that  $\tilde{X}$  is a martingale. Let  $s < t$ , we need to show that  $\mathbb{E}[\tilde{X}_t | \tilde{\mathcal{F}}_s] = \tilde{X}_s$  a.s.

Claim:  $\mathbb{E}[X_t | \mathcal{F}_{t+}] = \tilde{X}_s$  a.s. Indeed, first observe that for  $Y$  any random variable and  $\mathcal{G}$  a sigma algebra it follows that

$$\mathbb{E}[Y | \sigma(\mathcal{G}, \mathcal{N})] = \mathbb{E}[X | \mathcal{G}]$$

which is clear because the conditional expectation is defined almost surely and  $\mathcal{N}$  only contains sets of measure zero.

Now, fix  $s < t$  and let  $s_n \downarrow s$ ,  $s_n \in \mathbb{Q}_+$ ,  $s_0 < t$ . We have by the tower property that  $(\mathbb{E}[X_t | \mathcal{F}_{s_n}])_{n \in \mathbb{N}}$  is a backwards martingale and so it converges a.s. and in  $\mathcal{L}^1$  to  $\mathbb{E}[X_t | \mathcal{F}_{t+}]$ . But  $\mathbb{E}[X_t | \mathcal{F}_{s_n}] = X_{s_n}$  a.s. and  $X_{s_n} \rightarrow \tilde{X}_s$  a.s. as  $n \rightarrow \infty$ . So  $\tilde{X}_s = \mathbb{E}[X_t | \mathcal{F}_{s+}]$ .

4. Finally, we show that  $\tilde{X}$  is a cadlag. First we show that  $\tilde{X}$  is right continuous. Suppose not. Then, there exists  $\omega \in \Omega_0$  and some  $t \geq 0$  such that  $\tilde{X}(\omega)$  is not right continuous at  $t$ . That is there exists a sequence  $s_n \downarrow t$  such that  $|\tilde{X}_{s_n} - \tilde{X}_t| \geq \epsilon > 0$  (for some positive  $\epsilon$ ). By the definition of  $\tilde{X}$ , there exists another sequence  $s'_n > s_n$ , for all  $n \in \mathbb{N}$  and  $s'_n \downarrow t$ ,  $s'_n \in \mathbb{Q}_+$  such that  $|\tilde{X}_{s_n} - X_{s'_n}| \leq \frac{\epsilon}{2}$ . So  $|X_{s'_n} - \tilde{X}_t| \geq \frac{\epsilon}{2}$ , a contradiction since  $s'_n \downarrow t$ ,  $s'_n \in \mathbb{Q}_+$ . The argument for left continuity is entirely analogous.

□

**Examples:**

Let  $\xi, \eta$  be independent iid symmetric Bernoulli with success probability  $1/2$ . Define

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

and let  $\mathcal{F}_t = \sigma(X_s, s \leq t)$  for all  $t \geq 0$ . Observe that  $X$  is an  $(\mathcal{F}_t)_{t \geq 0}$  martingale. Also,  $\tilde{X}$  satisfies  $X_t = \mathbb{E}[\tilde{X}_t | \mathcal{F}_t]$  where

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

Furthermore,  $\mathcal{F}_1 = \sigma(\xi)$  and  $\mathcal{F}_t = \sigma(\xi, \eta)$  for all  $t > 1$ ,  $\tilde{X}$  is cadlag with respect to  $\tilde{F}$ . Observe finally that  $\mathcal{F}_{1+} = \sigma(\xi, \eta)$  and so the filtration  $\mathcal{F}$  is not right continuous and  $\tilde{X}$  is not a version of  $X$ . We thus see that the right-continuity of  $(\mathcal{F}_t)_{t \geq 0}$  is necessary in Theorem 6.1.

**Theorem 6.2 (Almost sure martingale convergence theorem).** Let  $X$  be a cadlag martingale bounded in  $\mathcal{L}^1$ . Then  $X_t \rightarrow X_\infty$  a.s. with  $X_\infty \in \mathcal{L}^1(\mathcal{F}_\infty)$ .

*Proof.* Let  $I_M = \mathbb{Q}_+ \cap [0, M]$ . Then Doob's upcrossing inequality 3.2 from the discrete setting and a monotone convergence argument give for  $a < b, a, b \in \mathbb{Q}_+$

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

Taking  $M \rightarrow \infty$  gives  $\mathcal{N}([a, b], \mathbb{Q}_+, X) < \infty$  a.s. Hence, for the event

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

we have  $\mathbb{P}(\Omega_0) = 1$  and on  $\Omega_0$ ,  $\lim_{q \rightarrow \infty, q \in \mathbb{Q}_+} X_q$  exists and is finite. We thus have  $X_\infty = \lim_{q \rightarrow \infty, q \in \mathbb{Q}_+} X_q$  on  $\Omega_0$ . Now for all  $\epsilon > 0$ , there exists  $q_0$  such that  $|X_{q_0} - X_\infty| \leq \frac{\epsilon}{2}$  for all  $q > q_0, q \in \mathbb{Q}_+$ . Now let  $t > q_0$ . Then there exists some  $q > t, q \in \mathbb{Q}_+$  such that  $|X_t - X_q| \leq \frac{\epsilon}{2}$  by right continuity of  $X$ . So  $|X_t - X_\infty| \leq \epsilon$ . □

**Theorem 6.3 (Doob's maximal inequality).** Let  $X$  be a cadlag martingale,  $X_t^* = \sup_{s \leq t} |X_s|$ . Then for all  $\lambda > 0$ ,

$$\lambda \cdot \mathbb{P}(X_t^* \geq \lambda) \leq \mathbb{E}[|X_t| \cdot \mathbf{1}(X_t^* \geq \lambda)] \leq \mathbb{E}[|X_t|].$$

*Proof.* Have

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

and use the beginning of the proof of theorem 6.1. □

**Theorem 6.4 (Optional stopping theorem for cadlag UI martingales).** Let  $X$  be a cadlag UI martingale, then for all  $S \leq T$  stopping times

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

*Proof.* Let  $T_n = 2^{-n} \lfloor 2^n T \rfloor$  and  $S_n = 2^{-n} \lfloor 2^n S \rfloor$ . Both are stopping times and  $T_n \downarrow T, S_n \downarrow S$  as  $n \rightarrow \infty$ . need to show: for  $A \in \mathcal{F}_S$ , then  $\mathbb{E}[X_T \cdot \mathbf{1}(A)] = \mathbb{E}[X_S \cdot \mathbf{1}(A)]$ . Indeed,  $X_{T_n} \rightarrow X_T$  and  $X_{S_n} \rightarrow X_S$  a.s. as  $n \rightarrow \infty$  ( $X$  is right continuous).

Now, by the discrete optional stopping theorem applied to the martingale  $(X_{k \cdot 2^{-n}})_{k \in \mathbb{N}}$  with respect to the filtration  $(\mathcal{F}_{K \cdot 2^{-n}})_{k \in \mathbb{N}}$ ,  $X_{T_n} = \mathbb{E}[X_\infty | \mathcal{F}_{T_n}]$ , so  $X_{T_n}$  is UI (since  $T_n$  take values in

$2^{-n} \cdot \mathbb{N}$ ). Thus,  $X_{T_n} \rightarrow X_T$  in  $\mathcal{L}^1$ , and the same holds for  $X_{S_n} \rightarrow X_S$  using the exact same argument. By the discrete optional stopping theorem, we have that  $\mathbb{E}[X_{T_n} | \mathcal{F}_{S_n}] = X_{S_n}$  a.s. Now for  $A \in \mathcal{F}_S$ , we have that  $A \in \mathcal{F}_{S_n}$  for all  $n \in \mathbb{N}$  since  $S_n \geq S$ . So  $\mathbb{E}[X_{T_n} \cdot \mathbf{1}(A)] = \mathbb{E}[X_{S_n} \cdot \mathbf{1}(A)]$ .  $\square$

Lecture 13

**Theorem 6.5 (Kolmogorov's continuity criterion).** Let  $\mathcal{D}_n = \{K \cdot 2^{-n} : 0 \leq k \leq 2^n\}$  and  $\mathcal{D} = \bigcup_{n \geq 0} \mathcal{D}_n$ . Let  $(X_t)_{t \in \mathcal{D}}$  be a stochastic process taking real values. Suppose there exists some  $\epsilon > 0$   $p > 0$ , such that

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

where  $c$  is a positive constant. Then for all  $\alpha \in (0, \epsilon/p)$ , the process is  $\alpha$ -Hölder continuous, that is there exists a random variable  $K_\alpha < \infty$  such that

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

*Proof.*

$$\mathbb{P}(|X_{k \cdot 2^{-n}} - X_{(k+1) \cdot 2^{-n}}| \geq 2^{-n\alpha}) \stackrel{\text{Markov + assumption}}{\leq} c \cdot 2^{-n\alpha} p \cdot 2^{-n(1+\epsilon)}.$$

Thus,

$$\mathbb{P}\left(\max_{0 \leq k \leq 2^n} |X_{k \cdot 2^{-n}} - X_{(k+1) \cdot 2^{-n}}| \geq 2^{-n\alpha}\right) \stackrel{\text{union bound}}{\leq} c \cdot 2^{n\alpha p n \epsilon}, \quad \alpha \in (0, \frac{\epsilon}{p}).$$

By Borel-Cantelli,

$$\max_{0 \leq k \leq 2^n} |X_{k \cdot 2^{-n}} - X_{(k+1) \cdot 2^{-n}}| \leq 2^{-n\alpha}$$

for all  $n \in \mathbb{N}$  sufficiently large. Thus,

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

a.s. For some random variable  $M$ .

Need to show: there exists some  $M'$  such that  $|X_t - X_s| \leq M' \cdot |t - s|^\alpha$  for all  $s, t \in \mathcal{D}$ .

Let  $s < t$ ,  $s, t \in \mathcal{D}$  and let  $r$  be the unique integer such that  $2^{-(r+1)} < t - s \leq 2^{-r}$ . Then there exists some  $k \in \mathbb{N}$  such that  $s < k \cdot 2^{-(r+1)} < t$ . Now, observe that  $t - s \leq 2^{-r}$  so

$$t - s = \sum_{j=r+1}^{\infty} \frac{x_j}{2^j}, \quad x_j \in \{0, 1\}$$

and

$$\alpha - s = \sum_{j=r+1}^{\infty} \frac{y_j}{2^j}, \quad y_j \in \{0, 1\}.$$

Observe that  $[s, t)$  is a disjoint union of dyadic intervals each of them having length  $2^{-n}$  with  $n \geq r + 1$  and each interval of length will appear at most twice. Thus, we get the bound

$$\begin{aligned} |X_t - X_s| &\leq \sum_{d,n} \overbrace{|X_d - X_{d+2^{-n}}|}^{\leq 2^{-n\alpha} M} \\ &\leq 2 \cdot M \cdot \sum_{n=r+1}^{\infty} 2^{-n\alpha} = \frac{2M \cdot 2^{-(r+1)\alpha}}{1 - 2^{-\alpha}} < \frac{2M}{1 - 2^{-\alpha}} |t - s|^\alpha. \end{aligned}$$

$\square$

## 7 Weak Convergence

We fix  $(\mathcal{M}, d)$  a metric space endowed with its Borel sigma algebra.

**Definition 7.1.** Let  $(\mu_n)_{n \in \mathbb{N}}$  be a sequence of probability measures on  $\mathcal{M}$ . We say  $(\mu_n)_{n \in \mathbb{N}}$  converges weakly to  $\mu$  and write  $\mu_n \rightrightarrows \mu$  as  $n \rightarrow \infty$  if

$$\mu_n(f) := \int_{\mathcal{M}} f(x) \mu_n(dx) \xrightarrow{n \rightarrow \infty} \int_{\mathcal{M}} f(x) \mu(dx) := \mu(f)$$

for any  $f$  continuous and bounded.

### Examples:

1. Let  $x_n \rightarrow x$  as  $n \rightarrow \infty$  in  $(\mathcal{M}, d)$  then  $\delta_{x_n} \xrightarrow{n \rightarrow \infty} \delta_x$ , since  $\delta_{x_n}(f) = f(x_n) \xrightarrow{n \rightarrow \infty} f(x) = \delta_x(f)$ .
2. Let  $\mathcal{M} = [0, 1]$ , with the Euclidean metric and its Borel sigma algebra. Let  $\mu_n = \frac{1}{n} \sum_{0 \leq k \leq n} \delta_{k/n}$ . Then  $\mu_n$  converges weakly to the Lebesgue measure. Indeed,  $\mu_n(f) = \frac{1}{n} \sum_{0 \leq k \leq n} f(k/n) \xrightarrow{n \rightarrow \infty} \int_0^1 f(x) dx$ , being Riemann sums.
3.  $\mu_n = \delta_{\frac{1}{n}} \rightrightarrows \delta_0$ , as  $n \rightarrow \infty$ . Notice however that for  $A = (0, 1)$ ,  $\mu_n(A) = 1$  for all  $n \geq 1$  and so  $\mu_n(A) \not\rightrightarrows \delta_0(A) = 0$ .

**Theorem 7.1.** Let  $(\mu_n)_{n \in \mathbb{N}}$  be a sequence of probability measures on  $(\mathcal{M}, d)$ . Then the following are equivalent:

1.  $\mu_n \rightrightarrows \mu$ .
2. For all  $G$  open,  $\liminf_n \mu_n(G) \geq \mu(G)$ .
3. For all  $A$  closed,  $\limsup_n \mu_n(A) \leq \mu(A)$ .
4. For all  $A$  with  $\mu(\partial A) = 0$ , then  $\mu_n(A) \rightarrow \mu(A)$ .

*Proof.* 1  $\implies$  2: Let  $G$  be open with  $G^c \neq \emptyset$ . Let  $M > 0$  and set  $f_M(x) = \mathbf{1}(Md(x, G^c)) \leq \mathbf{1}(x \in G)$ . Observe that  $f_M(x) \uparrow \mathbf{1}(x \in G)$  as  $M \rightarrow \infty$ ,  $f_M$  is bounded and continuous for all  $M$ . So  $\mu_n(f_M) \rightarrow \mu(f_M)$  as  $n \rightarrow \infty$  for all  $M$ . Thus,

$$\liminf_n \mu_n(G) \geq \liminf_n \mu_n(f_M) = \mu(f_M) \xrightarrow{\text{monotone convergence}} \mu(G).$$

2  $\implies$  3: follows from the previous case by taking complements. 2, 3  $\implies$  4:  $0 = \mu(\partial A) = \mu(A \setminus \text{int } A)$ , hence  $\mu(\overline{A}) = \mu(A) = \mu(\text{int } A)$ . 2:  $\liminf_n \mu_n(\int A) \geq \mu(\text{int } A) = \mu(A)$ . 3:  $\limsup_n \mu_n(\overline{A}) \leq \mu(\overline{A}) = \mu(A)$ .

4  $\implies$  1: Need to show for any  $f$  continuous and bounded,  $\mu_n(f) \rightarrow \mu(f)$ . We can assume further that  $f \geq 0$ . Fix  $K > \sup f$ . Have,

$$\int_{\mathcal{M}} f(x) \mu_n(dx) = \int_{\mathcal{M}} \left( \int_0^K \mathbf{1}(t \leq f(x)) dt \right) \mu_n(dx) \\ \stackrel{\text{Fubini}}{=} \int_0^K \mu_n(f \geq t) dt.$$

It suffices to show  $\mu_n(f \geq t) \rightarrow \mu(f \geq t)$  as  $n \rightarrow \infty$ . Since then we can conclude using dominated convergence. Thus it suffices to show that  $\mu(\partial\{f \geq t\}) = 0$ . Indeed,

$$\partial\{f \geq t\} \subset \{f = t\}.$$

since  $f$  is continuous and  $\{f > t\}$  is open and  $\subset \{f \geq t\}$ . Also observe that there exists an at most countable number of  $t$  such that  $\mu(f = t) > 0$ . Thus,

$$\{t : \mu(f = t) > 0\} = \bigcup_n \underbrace{\{t : \mu(\{f = t\}) \geq \frac{1}{n}\}}_{\# \leq n}.$$

Thus,  $\partial\{f \geq t\}$  is countable and has Lebesgue measure zero. □

Now, let  $\mathcal{M} = \mathbb{R}$ . Let  $\mu$  be a probability measure on  $\mathbb{R}$ . We define the distribution function of  $\mu$  to be the function  $F_\mu : x \mapsto \mu((-\infty, x])$ ,  $F_\mu : \mathbb{R} \rightarrow [0, 1]$ .

**Proposition 7.1.** Let  $(\mu_n)_{n \in \mathbb{N}}$  be a sequence of probability measures on  $\mathbb{R}$ . Then the following are equivalent:

$$\mu_n \implies \mu, \text{ as } n \rightarrow \infty. F_{\mu_n}(x) \xrightarrow{n \rightarrow \infty} F_\mu(x) \text{ for all } x \in \mathbb{R} \text{ continuity points of } F_\mu.$$

1. *Proof.* 1  $\implies$  2: Let  $x$  be a continuity point of  $F_\mu$ . Have  $F_{\mu_n}(x) = \mu_n((-\infty, x])$  and

$$\begin{aligned} \mu(\partial(-\infty, x]) &= \mu(\{x\}) \\ &= \mu((-\infty, x]) - \lim_{n \rightarrow \infty} \mu((-\infty, x - \frac{1}{n}]) \\ &= F_\mu(x) - \lim_{n \rightarrow \infty} F_\mu(x - \frac{1}{n}) = 0 \end{aligned}$$

since  $x$  is a continuity point of  $F_\mu$ . □

2  $\implies$  1: Let  $G$  be an open set in  $\mathbb{R}$ . Then  $G = \bigcup_n (a_k, b_k)$ , a union of disjoint open intervals.

Now,

$$\begin{aligned} \liminf_n \mu_n(G) &= \liminf_n \sum_k \mu_n(a_k, b_k) \\ &\stackrel{\text{Fat}}{\geq} \sum_k \liminf_n \mu_n(a_k, b_k). \\ &= +d \end{aligned}$$

So it suffices to show that  $\liminf_n \mu_n(a, b) \geq \mu(a, b)$  for all  $a < b \in \mathbb{R}$ .

Indeed, We have  $\mu_n((a, b)) = F_{\mu_n}(b-) - F_{\mu_n}(a)$  and since  $F_\mu$  is non-decreasing and has at most countably many discontinuities, there exist  $a', b'$  continuity points of  $F_\mu$ . Hence,  $F_{\mu_n}(a') \xrightarrow{n \rightarrow \infty} F_\mu(a')$  and  $F_{\mu_n}(b') \xrightarrow{n \rightarrow \infty} F_\mu(b')$ . This means that

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

By the density of continuity points, there exist  $(b'_m)_{m \in \mathbb{N}}$ , such that  $b'_m \uparrow b'$  and  $(a'_m)_{m \in \mathbb{N}}$ ,  $a'_m \downarrow a'$  all continuity points. Thus,

$$\begin{aligned} \liminf_n \mu_n((a, b)) &\geq \sup_n F_{\mu_n}(b'_m) - F_\mu(a'_m) \\ &= F_\mu(b-) - F_\mu(a) = \mu((a, b)). \end{aligned}$$

**Definition 7.2.** Let  $(X_n)_{n \in \mathbb{N}}$  be a sequence of random variables taking values in  $(\mathcal{M}, d)$ , defined on probability spaces  $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$ . We say that  $(X_n)_{n \in \mathbb{N}}$  converges weakly (or in distribution) to a random variable  $X$  defined on  $(\Omega, \mathcal{F}, \mathbb{P})$  if  $\mathcal{L}(X_n) \implies \mathcal{L}(X)$  (i.e. the laws converge weakly).

**Remark.** Equivalently,  $X_n \xrightarrow{w/d} X$  if for all  $F$  continuous and bounded,  $\mathbb{E}_{\mathbb{P}_n} [f(X_n)] \rightarrow \mathbb{E}_{\mathbb{P}} [f(X)]$ , as  $n \rightarrow \infty$ .

**Proposition 7.2.** 1. If  $X_n \xrightarrow{\mathbb{P}} X$  as  $n \rightarrow \infty$ , then  $X_n \xrightarrow{d} X$  as  $n \rightarrow \infty$ .  
 2. If  $X_n \xrightarrow{d} c$ ,  $c$  a constant, then  $X_n \xrightarrow{\mathbb{P}} c$

**Examples: (CLT)**

Let  $(X_n)_{n \in \mathbb{N}}$  be iid and  $\mathbb{E}[X_1] = m$  and  $\sigma^2 = \text{Var}(X_1)$ . Then with  $S_n = \sum_{i=1}^n X_i$

$$\frac{S_n - n \cdot m}{\sqrt{n\sigma^2}} \xrightarrow{d} \mathcal{N}(0, 1)$$

as  $n \rightarrow \infty$ .

**Definition 7.3 (Tightness).** Let  $(\mathcal{M}, d)$  be a metric space. A sequence of probability measures  $(\mu_n)_{n \in \mathbb{N}}$  on  $\mathcal{M}$  is called tight if for all  $\epsilon > 0$ , there exists a compact set  $K \subseteq \mathcal{M}$  such that

$$\sup_{n \geq 0} \mu(\mathcal{M} \setminus K) \leq \epsilon.$$

**Remark.** If  $\mathcal{M}$  is compact, then all sequences of probability measures are tight.

**Theorem 7.2 (Prohorov).** Let  $(\mu_n)_{n \in \mathbb{N}}$  be a tight sequence of probability measures, then there exists a subsequence  $(\mu_{n_k})_{k \in \mathbb{N}}$  and a probability measure  $\mu$  such that

$$\mu_{n_k} \xrightarrow{d} \mu, \quad \text{as } k \rightarrow \infty.$$

*Proof.* We focus on the case  $\mathcal{M} = \mathbb{R}$ . Let  $\mathbb{Q} = \{x_1, x_2, \dots\}$  be an enumeration of  $\mathbb{Q}$  and  $F_n = F_{\mu_n}$ . Then, the sequence  $(F_n(x_1))_{n \in \mathbb{N}}$  in  $[0, 1]$  has a convergent subsequence  $F_{n_k^{(1)}}(x_1) \xrightarrow{k \rightarrow \infty} F(x_1)$  by compactness. So does  $(F_{n_k^{(1)}}(x_2))_{k \in \mathbb{N}}$ . Thus, continuing so inductively, we obtain for all  $i \in \mathbb{N}$  that there exist sequences  $(n_k^{(i)})_{k \in \mathbb{N}}$  such that

$$F_{n_k^{(i)}}(x_j) \xrightarrow{k \rightarrow \infty} F(x_j), \quad \text{for all } 1 \leq j \leq i.$$

Thus, we can extract a diagonal sequence  $(m_k)_{k \in \mathbb{N}}$ , where  $m_k = n_k^{(k)}$  for all  $k \in \mathbb{N}$  and Have

$$F_{m_k}(x) \xrightarrow{k \rightarrow \infty} F(x), \quad \text{for all } x \in \mathbb{Q}.$$

Observe now that the functions  $F_{m_k}$  are non-decreasing, and so  $F$  is non-decreasing, so for  $x \in \mathbb{R}$  define  $F(x) = \lim_{q \downarrow x, q \in \mathbb{Q}} F(q)$ . Thus,  $F$  is right continuous, non-decreasing and so  $F$  has left-limits.

Let  $x \in \mathbb{R}$  be a continuity point of  $F$ . We need to show that  $F_{m_k}(x) \xrightarrow{k \rightarrow \infty} F(x)$ . Indeed, for any  $\epsilon > 0$ , there exist  $s_1 < x < s_2$ ,  $s_i \in \mathbb{Q}$  such that  $F(s_i) - F(x) < \epsilon/2$  (since  $F$  is continuous at  $x$ ). We now have the chain of inequalities

$$F(x) - \epsilon \leq F(s_1) - \frac{\epsilon}{2} \leq F_{m_k}(s_1) \leq F_{m_k}(x) \leq F_{m_k}(s_2) \stackrel{\text{conv. in } \mathbb{Q}}{\leq} F(s_2) + \frac{\epsilon}{2} \leq F(x) + \epsilon$$

for all  $k \in \mathbb{N}$  sufficiently large.

Finally, it remains to show that there exists some probability measure  $\mu$  such that  $F = F_\mu$ . Indeed, by tightness, we have that for all  $\epsilon > 0$ , there exists  $N \in \mathbb{R}$  large enough so that (with  $\pm N$  being continuity points of  $F$ )

$$\sup_{n \geq 0} \mu_n([-N, N]^c) \leq \epsilon.$$

Thus,  $F(-N) \leq \epsilon$  and  $1 - F(N) \leq \epsilon$ . This guarantees that

$$\lim_{x \rightarrow -\infty} F(x) = 0, \quad \lim_{x \rightarrow \infty} F(x) = 1.$$

Finally, define  $\mu((a, b]) = F(b) - F(a)$ . Then,  $\mu$  can be extended to the Borel sigma algebra by Calathea dory's extension theorem.  $\square$

**Definition 7.4.** Let  $X$  be a random variables with values in  $\mathbb{R}^d$ . The characteristic function of  $X$  is defined as

$$\phi_X(u) = \mathbb{E} \left[ e^{i\langle u, X \rangle} \right], \quad u \in \mathbb{R}^d.$$

Properties of  $\phi_X$ :

1.  $\phi_X$  is continuous on  $\mathbb{R}^d$  and  $\phi_X(0) = 1$ .
2.  $\phi_X$  completely determines the law of  $X$ , that is if  $\phi_X(u) = \phi_Y(u)$  for all  $u \in \mathbb{R}^d$ , then  $\mathcal{L}(X) = \mathcal{L}(Y)$ .

Lecture 15

**Theorem 7.3 (Lévy's convergence theorem).** Let  $(X_n)_{n \in \mathbb{N}}$ ,  $X$  be random variables taking values in  $\mathbb{R}^d$ . Then

1.  $\mathcal{L}(X_n) \implies \mathcal{L}(X)$  as  $k \rightarrow \infty$ , then  $\phi_{X_n}(u) \xrightarrow{n \rightarrow \infty} \phi_X(u)$  for all  $u \in \mathbb{R}^d$ .
2. Suppose there exists  $\psi : \mathbb{R}^d \rightarrow \mathbb{C}$  such that  $\psi(0) = 1$ ,  $\psi$  is continuous at zero and  $\phi_{X_n}(u) \xrightarrow{n \rightarrow \infty} \psi(u)$  for all  $u \in \mathbb{R}^d$ . Then there exists a random variable  $X$  with characteristic function  $\psi = \phi_X$  and  $\mathcal{L}(X_n) \implies \mathcal{L}(X)$ .

Before we proceed with the proof of the theorem, we state a Lemma

**Lemma 7.1.** Let  $X$  be a random variable in  $\mathbb{R}^d$ . Then, for all  $K > 0$ ,

$$\mathbb{P}(\|X\|_\infty \leq K) \leq C \cdot \left(\frac{K}{2}\right)^d \int_{[-\frac{1}{K}, \frac{1}{K}]^d} (1 - \phi_X(u)) du,$$

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

*Proof.* Fix  $\lambda > 0$  and let  $\mu = \mathcal{L}(X)$ . Then,

$$\begin{aligned} \int_{[-\lambda, \lambda]^d} \phi_X(u) du &= \int_{[-\lambda, \lambda]^d} \left( \int_{\mathbb{R}^d} \prod_{j=1}^d e^{iu_j \cdot x_j} \mu(dx) \right) du \\ &\stackrel{\text{Fubini}}{=} \int_{\mathbb{R}^d} \mu(dx) \prod_{j=1}^d \left( \int_{[-\lambda, \lambda]} e^{iu_j \cdot x_j} du_j \right) \\ &= \int_{\mathbb{R}^d} \mu(dx) \prod_{j=1}^d \left( \frac{e^{ix_j \lambda} - e^{-ix_j \lambda}}{ix_j} \right) \\ &= \int_{\mathbb{R}^d} \prod_{j=1}^d \frac{2 \cdot \sin(\lambda x_j)}{x_j} \mu(dx) \\ &= (2\lambda)^d \int_{\mathbb{R}^d} \prod_{j=1}^d \left( \frac{2 \cdot \sin(\lambda x_j)}{\lambda x_j} \right) \mu(dx). \end{aligned}$$

Thus,

$$\int_{[-\lambda, \lambda]^d} (1 - \phi_X(u)) du = (2\lambda)^d \int_{\mathbb{R}^d} \prod_{j=1}^d \left( 1 - \frac{2 \cdot \sin(\lambda x_j)}{\lambda x_j} \right) \mu(dx)$$

Now, let  $f(u) = \prod_{j=1}^d \left( \frac{2 \cdot \sin(u_j)}{u_j} \right)$ ,  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ .

Claim: not hard to see that if  $x \geq 1$ , then  $\left| \frac{\sin(x)}{x} \right| \leq \sin(1)$ . Hence, if  $\|u\|_\infty \geq 1$ , then  $|f(u)| \leq \sin(1)$ . So  $\mathbf{1}(\|u\|_\infty \geq 1) \leq C \cdot (1 - f(u))$ , where  $C = (1 - \sin(1))^{-1}$ . Hence,

$$\mathbb{P}(\|X\|_\infty \geq k) \leq C \cdot \mathbb{E} \left[ 1 - f\left(\frac{X}{k}\right) \right]$$

and by simple scaling, one can conclude for the general case.  $\square$

*Proof.* (Theorem 7.3)

$f(x) = e^{i\langle u, x \rangle}$  is continuous and bounded so by bounded convergence, have

$$\phi_{X_n}(u) = \mathbb{E}[f(X_n)] \rightarrow \mathbb{E}[f(X)]$$

as  $n \rightarrow \infty$ .

1. First we prove that  $\mathcal{L}(X_n)_{n \in \mathbb{N}}$  is tight. By Lemma 7.1, have that

$$\mathbb{P}(\|X_n\|_\infty \leq K) \leq C \cdot \left(\frac{K}{2}\right)^d \int_{[-\frac{1}{K}, \frac{1}{K}]^d} (1 - \phi_{X_n}(u)) \, du$$

and  $|1 - \phi_{X_n}(u)| \leq 2$  for all  $u \in \mathbb{R}^d, n \in \mathbb{N}$ . Thus, by dominated convergence,

$$\int_{[-\frac{1}{K}, \frac{1}{K}]^d} (1 - \phi_{X_n}(u)) \, du \xrightarrow{n \rightarrow \infty} \int_{[-\frac{1}{K}, \frac{1}{K}]^d} (1 - \psi(u)) \, du.$$

Since  $\psi$  is continuous at zero and  $\psi(0) = 1$ , taking  $K$  large enough we get

$$\int_{[-\frac{1}{K}, \frac{1}{K}]^d} (1 - \psi(u)) \, du < \frac{\epsilon}{2^d C d} (2K^{-1})^d.$$

Thus,  $\mathbb{P}(\|X_n\|_\infty \geq K) \leq \epsilon$  for all  $n \in \mathbb{N}$  sufficiently large. Taking  $K$  possibly even larger, we conclude that

$$\sup_{n \geq 0} \mathbb{P}(\|X\|_\infty \geq K) \leq \epsilon,$$

hence showing that  $(\mathcal{L}_n)_{n \in \mathbb{N}}$  is tight. By Pro horror, there exists a subsequence  $(n_k)_{k \in \mathbb{N}}$  such that

$$\mathcal{L}(X_{n_k}) \xrightarrow{n \rightarrow \infty} \mathcal{L}(X)$$

and so  $\phi_{X_{n_k}}(u) \rightarrow \phi_X(u)$  for all  $u \in \mathbb{R}^d$ . Thus,  $\psi \equiv \phi$ .

Suppose for a contradiction that  $\mathcal{L}_{X_n}$  does not converge. Then there exists  $f$  continuous and bounded and a subsequence  $m_k$  such that

$$|\mathbb{E}_{m_k}[f(X_{m_k})] - \mathbb{E}[f(X)]| \geq \epsilon$$

for all  $k \in \mathbb{N}$ . Now, since  $(\mathcal{L}(X_{m_k}))_{k \in \mathbb{N}}$  is tight, there exist a subsequence, without relabelling, such that  $(\mathcal{L}(X_{m_k}))$  converges weakly, a contradiction. Thus, the limit must also be  $X$ .  $\square$

Now, we briefly embark on a discussion of the theory of *large deviations*.

## 8 Large deviations

Let  $X_1, X_2, \dots$  be iid  $\sim \mathcal{N}(0, 1)$  random variables. Let  $\hat{S}_n = \frac{1}{n} \sum_{i=1}^n X_i \sim \mathcal{N}(0, 1/n)$ . Let  $\delta > 0$ , we by the weak law of large numbers that

1.

$$\mathbb{P}(|\hat{S}_n| \geq \delta) \xrightarrow{n \rightarrow \infty} 0.$$

2.

$$\mathbb{P}(\sqrt{n}|\hat{S}_n| \in \overset{\text{interval}}{A}) \xrightarrow{\text{CLT}} \int_A \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dx.$$



3.

$$\mathbb{P}(|\hat{S}_n| \geq \delta) = 1 - \int_{-\delta\sqrt{n}}^{\delta\sqrt{n}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

In other words,

$$\frac{\log \mathbb{P}(|\hat{S}_n| \geq \delta)}{n} \xrightarrow{n \rightarrow \infty} -\frac{\delta^2}{2}.$$

Observe that  $\hat{S}_n$ , the “typical” value is of the order  $\frac{1}{\sqrt{n}}$  and it can take relatively large values ( $\geq \delta > 0$ ) with very small probability  $\tilde{e}^{-\frac{\delta^2 n}{2}}$ . Furthermore, 1, 2 are universal but 3 depends on the distribution. We shall focus on quantifying 3 for an appropriate class of random variables.

Let  $X_1, X_2, \dots$  be an iid family of random variables, such that  $\mathbb{E}[X_1] = \bar{x}$ ,  $S_n = X_1 + X_2 + \dots + X_n$ . Let  $a \in \mathbb{R}$ . Now

$$\mathbb{P}(S_{n+m} \geq a(n+m)) \stackrel{\text{independence}}{\geq} \mathbb{P}(S_n \geq a_n) \cdot \mathbb{P}(S_{m \geq a_m}).$$

Now, with  $b_n = -\log \mathbb{P}(S_n \geq an)$  for all  $n \in \mathbb{N}$ , have that  $b_{n+m} \leq b_n + b_m$ . This is called sub-additive sequence. Actually, for such sequences one has

$$\lim_{n \rightarrow \infty} \frac{b_n}{n} = \inf_n \frac{b_n}{n}.$$

**Sub-additive sequences**

To quickly see this, suppose first that  $\inf_n \frac{b_n}{n} > -\infty$ . Fix any  $\epsilon > 0$ , then there exists some  $m \in \mathbb{N}$  such that  $\frac{b_m}{m} < \inf_n \frac{b_n}{n} + \epsilon$ . Hence, for any  $k \geq m$ , we have by Euclidean division that there exists some  $q \in \mathbb{Z}_+$  and  $r \in [0, m) \cap \mathbb{N}$  such that  $k = qm + r$ . Thus, the sub-additivity of  $(b_n)_{n \in \mathbb{N}}$  implies that

$$\begin{aligned} \frac{b_k}{k} = \frac{b_{qm+r}}{qm+r} &\leq \frac{q \cdot b_m + b_r}{qm+r} \\ &\leq \frac{qm}{qm+r} \inf_n \frac{b_n}{n} + \epsilon \cdot mq + \frac{b_r}{qm+r} \end{aligned}$$

as  $k \rightarrow \infty$ . The case where  $\inf_n \frac{b_n}{n} = -\infty$  can be dealt with similarly.

So, we have that

$$-\frac{1}{n} \log \mathbb{P}(S_n \geq an) \xrightarrow{n \rightarrow \infty} I(a).$$

Also,

$$\begin{aligned} \mathbb{P}(S_n \geq an) &\stackrel{\lambda \geq 0}{\leq} \mathbb{P}(e^{\lambda S_n} \geq e^{\lambda an}) \\ &\stackrel{\text{Markov}}{\leq} \mathbb{E}[e^{\lambda S_n}] \cdot e^{-n\lambda a} = \mathbb{E}[e^{\lambda X_1}]^n \cdot e^{-n\lambda a}. \end{aligned}$$

Define  $M(\lambda) = \mathbb{E}[e^{\lambda \cdot X_1}]$ ,  $\psi(\lambda) = \log M(\lambda)$ ,  $\lambda \in \mathbb{R}$ . In other words, we have

$$\mathbb{P}(S_n \geq an) \leq \exp(-n(\lambda a - \psi(\lambda))).$$

Furthermore, let  $\psi^*(a) = \sup_{\lambda \geq 0} (\lambda a - \psi(\lambda)) \geq 0$ . So  $\mathbb{P}(S_n \geq an) \leq \exp(-n\psi^*(a))$  and so have obtained

$$\frac{-\log \mathbb{P}(S_n \geq an)}{n} \geq \psi^*(a).$$

Lecture 16

**Theorem 8.1 (Cramer’s Theorem).** Let  $X_1, X_2, \dots$  be an iid sequence of random variables with  $\mathbb{E}[X_1] = \bar{x}$ . Let  $S_n = \sum_{i=1}^n X_i$ . Then,

$$-\frac{1}{n} \log \mathbb{P}(S_n \geq an) \xrightarrow{n \rightarrow \infty} \psi^*(a)$$

for all  $a \geq \bar{x}$  where  $\psi^*(a) = \sup_{\lambda \geq 0} (\lambda a - \psi(\lambda))$ ,  $\psi(\lambda) = \log \mathbb{E}[e^{\lambda X_1}]$  ( $\psi^*$  is known as the *Legendre transform*).

We collect some basic facts about the function  $M(\lambda) = \mathbb{E}[e^{\lambda X_1}]$ ,  $\lambda \in \mathbb{R}$ .

**Lemma 8.1.** The functions  $M$  and  $\psi$  are continuous on  $\mathcal{D} = \{\lambda : M(\lambda) < \infty\}$  and differentiable in  $\text{int } \mathcal{D}$  with  $M'(\lambda) = \mathbb{E}[X_1 e^{\lambda X_1}]$  and  $\psi'(\lambda) = \frac{M'(\lambda)}{M(\lambda)}$ ,  $\lambda \in \mathcal{D}$ .

*Proof. Continuity:* Fix a sequence  $\lambda_n \xrightarrow{n \rightarrow \infty} \lambda \in \mathcal{D}$ . Then, pointwise,  $e^{\lambda_n X_1} \xrightarrow{n \rightarrow \infty} e^{\lambda X_1}$  and take  $n \in \mathbb{N}$  such that for all  $n \geq N$ ,  $e^{\lambda_n X_1} \leq e^{\lambda_N X_1} + e^{\lambda X_1} \in \mathcal{L}^1$  (which holds by since  $\lambda_N \leq \lambda_n \leq \lambda$  for  $n$  possible larger). Thus, can conclude by dominated convergence that  $\psi(\lambda_n) \xrightarrow{n \rightarrow \infty} \psi(\lambda)$ .

*Differentiability:* Fix  $\eta \in \text{int } \mathcal{D}$ . We can now bound

$$\left| \frac{M(\eta+\epsilon) - M(\eta)}{\epsilon} \right| = \left| \mathbb{E} \left[ \frac{e^{(\eta+\epsilon) X_1} - e^{\eta X_1}}{\epsilon} \right] \right| \leq e^{\eta X_1} \left| \frac{e^{\epsilon X_1} - 1}{\epsilon} \right|.$$

Now, let  $\delta > 0$  sufficiently small such that  $(\eta - \delta, \eta + \delta) \subseteq \text{int } \mathcal{D}$ . Now, for all  $\epsilon \in (-\delta, \delta)$

$$\left| \frac{e^{\epsilon X_1} - 1}{\epsilon} \right| \stackrel{\text{comparing power series}}{\leq} \frac{e^{\delta |X_1|} - 1}{\delta}.$$

So

$$\left| \frac{e^{(\eta+\epsilon) X_1} - e^{\eta X_1}}{\epsilon} \right| \leq e^{\eta X_1} \cdot \frac{e^{\delta |X_1|} - 1}{\delta}.$$

Now, since  $e^{\eta X_1} \cdot e^{\delta |X_1|} \leq e^{\eta X_1} \cdot (e^{\delta X_1} + e^{-\delta X_1}) \in \mathcal{L}^1$  since  $\eta \in \text{int } \mathcal{D}$  and we can thus conclude by dominated convergence.  $\square$

*Proof.* (Theorem 8.1) From the previously derived Chernoff bound, we have

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(S_n \geq an) \geq \psi^*(a).$$

It suffices to show now that

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(S_n \geq an) \leq \psi^*(a), \quad \text{for all } a \geq \bar{x}.$$

Observe that we can replace each  $X_i$  by  $\tilde{X}_i = X_i - a$  and define  $\tilde{S}_n = \sum_{i=1}^n \tilde{X}_i$  and

$$\tilde{M}(\lambda) = \mathbb{E}[e^{\lambda \tilde{X}}] = e^{-a\lambda} M(\lambda), \quad \text{where } \tilde{\psi}(\lambda) = \psi(\lambda) - a\lambda, \quad \lambda \in \mathbb{R}. \quad \square$$

Thus we can restate the original inequality as follows

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(S_n \geq an) = \lim_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(\tilde{S}_n \geq 0) \leq \tilde{\psi}^*(0)$$

where  $\tilde{\psi}^*(\lambda) = \sup_{\lambda \geq 0} (-\tilde{\psi}(\lambda))$ . Thus, without loss of generality, it suffices to show that

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(S_n \geq 0) \geq \inf_{\lambda \geq 0} \psi(\lambda),$$

when  $\bar{x} \leq 0$ .

For the remainder of the proof, we let  $\mu = \mathcal{L}(X)$  and break the proof into several cases.

Case 1:  $M(\lambda) < \infty$  for all  $\lambda \in \mathbb{R}$ .

Define a new measure  $\mu_\theta$  for all  $\theta \geq 0$ , absolutely continuous with respect to  $\mu$  and radon-Nikodym derivative

$$\frac{d\mu_\theta}{d\mu} = \frac{e^{\theta X_1}}{M(\theta)}.$$

We compute

$$\mathbb{E}_\theta [f(X_1)] = \int_{\mathbb{R}} \frac{e^{\theta x} f(x)}{M(\theta)} \mu(dx).$$

Now, if  $X_1, \dots, X_n$  are iid  $\sim \mu$ . Then

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

Set  $g(\theta) = \mathbb{E}_\theta [X_1] = \int x \frac{e^{\theta x}}{M(\theta)} d\mu = \frac{M'(\theta)}{M(\theta)} = \psi'(\theta)$ .

Seek:  $\theta$  such that  $g(\theta) = \psi'(\theta) = 0$ .

If  $\mathbb{P}(X_1 > 0) = 0$ , then  $\mathbb{P}(S_n \geq 0) = (\mathbb{P}(X_1 = 0))^n$  by independence. Thus,

$$\frac{1}{n} \log \mathbb{P}(S_n \geq 0) = \mathbb{P}(X_1 = 0)$$

and

$$\inf_{\lambda \geq 0} \leq \lim_{\lambda \rightarrow \infty} \psi(\lambda) = \lim_{\lambda \rightarrow \infty} \mathbb{E} [e^{\lambda X_1}] \stackrel{\text{DCT}}{=} \lim_{\lambda \rightarrow \infty} \mathbb{E} [e^{\lambda X_1} \mathbf{1}(X_1 = 0)] = \mathbb{P}(X_1 = 0).$$

We can now focus on the case where  $\mathbb{P}(X_1 > 0) > 0$ . Now, there exists an  $N \in \mathbb{N}$  such that  $\mathbb{P}(X_1 > \frac{1}{N}) > 0$ . We deduce that

$$\lim_{\theta \rightarrow \infty} \psi(\theta) = \lim_{\theta \rightarrow \infty} \mathbb{E} [e^{\theta X_1}] \geq \lim_{\theta \rightarrow \infty} \mathbb{E} \left[ e^{\frac{\theta}{N}} \mathbf{1} \left( X_1 > \frac{1}{N} \right) \right] = \infty.$$

Thus, there exists some  $\eta \geq 0$  such that  $\inf_{\lambda \geq 0} \psi(\lambda) = \psi(\eta)$  and  $\psi'(\eta) = 0$ . Now,

$$\begin{aligned} \mathbb{P}(S_n \geq 0) &\geq \mathbb{P}(S_n \in [0, \epsilon n]) \geq \mathbb{E} [e^{\eta S_n - \eta \epsilon n} \mathbf{1}(S_n \in [0, \epsilon n])] \\ &= e^{-\eta \epsilon n} (M(\eta))^n \cdot \mathbb{P}_\eta(S_n \in [0, \epsilon n]) \end{aligned}$$

where  $\mathbb{P}_\eta(X_1 \in \cdot) = \mu_\eta(\cdot)$ . Now, since  $\mathbb{E}_\eta[X_1] = 0$ , we claim that we can use the CLT on iid copies of  $X_1$  with law  $\mu_\eta$  to deduce

$$\mathbb{P}(S_n \in [0, \epsilon n]) \xrightarrow{n \rightarrow \infty} \frac{1}{2}.$$

**Proof of claim**

This is a little messy, be warned! Fix any  $\epsilon' > 0$ . We have by the triangle inequality

$$\left| \mathbb{P}_\eta(S_n \in [0, \epsilon n]) - \frac{1}{2} \right| \leq \left| \mathbb{P}_\eta(S_n \in [0, \epsilon n]) - \mathbb{P}_\eta(S_n \in [0, \infty)) \right| + \left| \mathbb{P}_\eta(S_n \in [0, \infty)) - \frac{1}{2} \right|.$$

for all  $n \in \mathbb{N}$ . Now, by the CLT and Theorem 7.1 we have that

$$\mathbb{P}(S_n \in [0, \infty)) \xrightarrow{n \rightarrow \infty} \frac{1}{2}.$$

Thus, for all  $n$  sufficiently large, we have that  $|\mathbb{P}(S_n \in [0, \infty)) - 1/2| < \epsilon'/3$ . Furthermore, there exists some  $N \in \mathbb{N}$  such that  $\mathbb{P}(\mathcal{N} \in (\epsilon\sqrt{N}, \infty)) < \epsilon'/3$  where  $\mathcal{N}$  denotes a standard

normal random variable. Thus, for all  $n \in \mathbb{N}$  sufficiently large

$$\begin{aligned}
 \left| \mathbb{P}_\eta(S_n \in [0, \epsilon n]) - \frac{1}{2} \right| &\leq \frac{\epsilon'}{3} + \left| \mathbb{P}_\eta\left(\frac{S_n}{\sqrt{n}} \in (\epsilon\sqrt{n}, \infty)\right) \right| \leq \frac{\epsilon'}{3} + \left| \mathbb{P}_\eta\left(\frac{S_n}{\sqrt{n}} \in (\epsilon\sqrt{n}, \infty)\right) \right| \\
 &\leq \frac{\epsilon'}{3} + \left| \mathbb{P}_\eta\left(\frac{S_n}{\sqrt{n}} \in (\epsilon\sqrt{N}, \infty)\right) \right| \leq \frac{\epsilon'}{3} + \mathbb{P}(\mathcal{N} \in (\epsilon\sqrt{N}, \infty)) \\
 &\quad + \left| \mathbb{P}_\eta\left(\frac{S_n}{\sqrt{n}} \in (\epsilon\sqrt{N}, \infty)\right) - \mathbb{P}_\eta\left(\frac{\mathcal{N}}{\sqrt{n}} \in (\epsilon\sqrt{N}, \infty)\right) \right| \\
 &\leq \frac{\epsilon'}{3} + \frac{\epsilon'}{3} + \left| \mathbb{P}_\eta\left(\frac{S_n}{\sqrt{n}} \in (\epsilon\sqrt{N}, \infty)\right) - \mathbb{P}_\eta\left(\frac{\mathcal{N}}{\sqrt{n}} \in (\epsilon\sqrt{N}, \infty)\right) \right| \\
 &\leq \epsilon' \quad \text{(CLT)} \leq \frac{\epsilon'}{3}
 \end{aligned}$$

as required.

Thus,

$$\frac{\log \mathbb{P}(S_n \geq 0)}{n} \geq -\eta\epsilon + \log M(\eta) + \frac{\log \mathbb{P}_\eta(S_n \in [0, \epsilon n])}{n}.$$

Now, for all  $\epsilon > 0$ ,

$$\liminf_n \frac{1}{n} \log \mathbb{P}(S_n \geq 0) \geq \log M(\eta) - \eta\epsilon = \psi(\eta) \geq \inf_{\lambda \geq 0} \psi(\lambda).$$

Sending  $\epsilon \rightarrow 0$  gives the desired inequality.

#### General Case:

Without loss of generality, (arguing as in the previous case), let  $K > 0$  sufficiently large so that  $\mu([0, K]) > 0$ . Then define the conditional laws  $\nu = \mathcal{L}(X_1 | |X_1| \leq K)$ ,  $\nu_n = \mathcal{L}\left(S_n \mid \bigcap_{i=1}^n \{|X_i| \leq K\}\right)$ .

Have

$$\mu_n([0, \infty)) \geq \nu_n([0, \infty)) \cdot (\mu([-K, K]))^n$$

and

$$\log \mu_n([0, \infty)) \geq \frac{\log \nu_n([0, \infty))}{n} + \log \mu([-K, K]).$$

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

exists again by sub-additivity

$$\overbrace{\lim_{n \rightarrow \infty} \frac{1}{n} \log \mu_n([0, \infty))}^{\text{exists again by sub-additivity}} \geq \inf_{\lambda \geq 0} \left( \log \int_{-\infty}^{\infty} e^{\lambda x} d\nu(x) \right) + \log \mu([-K, K]) = \inf_{\lambda \geq 0} \psi_K(\lambda) := J_K > -\infty.$$

Now, as observe that  $\psi_K$  is a non-decreasing family of continuous functions. Hence, the  $(J_k)_{k \in \mathbb{N}}$  are non-decreasing and so one has  $J_k \uparrow J > -\infty$   $K \rightarrow \infty$ . Furthermore, the sets  $\{\lambda : \psi_K(\lambda) \leq J\}$  are compact by the continuity of the  $\psi_K$  and the fact that  $\mu([0, K]) > 0$  implies  $\lim_{\lambda \rightarrow \infty} \psi_K(\lambda) = \infty$ , as well as nested. Thus, there exists some  $\lambda_0 \in \bigcap_k \{\lambda : \psi_K(\lambda) \leq J\}$ . hence,  $\psi(\lambda_0) = \lim_{k \rightarrow \infty} \psi_K(\lambda) \leq J$

by monotone convergence. So,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mu_n([0, \infty)) \geq J \geq \psi(\lambda_0) \geq \inf_{\lambda \geq 0} \psi(\lambda)$$

as required.

## 9 Brownian Motion

Lecture 17

**Definition 9.1.** A process  $(B_t)_{t \in \mathbb{R}_+}$  is called a Brownian motion in  $\mathbb{R}^d$ ,  $d \geq 1$  starting from  $x \in \mathbb{R}^d$  if  $(B_t)_{t \geq 0}$  is a continuous process and

1.  $B_0 = x$  a.s.
2. For all  $s < t$ ,  $B_t - B_s \sim \mathcal{N}(0, (t-s) \cdot \text{Id}_d)$ .
3.  $(B_t)_{t \geq 0}$  has independent increments independent of  $B_0$ .

If  $x = 0$  we call it a standard Brownian motion. Observe that  $i$  determine uniquely its law.

### Examples:

Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion in  $\mathbb{R}$ ,  $U \sim [0, 1]$  uniformly distributed and independent from  $(B_t)_{t \geq 0}$  and define

$$\tilde{B}_t = \begin{cases} B_t, & t \neq U \\ 0, & t = U \end{cases}$$

Then  $\tilde{B}$  is a.s. discontinuous, so even though  $B, \tilde{B}$  have the same finite dimensional distributions,  $\tilde{B}$  is not a Brownian motion.

**Theorem 9.1 (Wiener).** There exists a Brownian motion on some probability space.

*Proof.* (Lévy and Kolmogorov)

1. We shall proceed to construct a BM on  $[0, 1]$  in  $d = 1$ . Let  $\mathcal{D}_0 = \{0, 1\}$ ,  $\mathcal{D}_n = \{k \cdot 2^{-n} : 0 \leq k \leq 2^n\}$  for  $n \in \mathbb{N}$  and  $\mathcal{D} = \bigcup_{n \geq 0} \mathcal{D}_n$ .

We will now construct  $(B_d, d \in \mathcal{D})$  inductively. First for  $\mathcal{D}_0$ . Let  $(Z_d, d \in \mathcal{D})$  be an iid sequence  $\sim \mathcal{N}(0, 1)$  on some probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Set  $b_0 = 0$ ,  $B_1 = Z_1$  (clearly satisfies properties in 9.1). Suppose now we have constructed  $(B_d, d \in \mathcal{D}_{n-1})$  satisfying properties 2&3. We need to construct  $(B_d, d \in \mathcal{D}_n)$ .

For  $d \in \mathcal{D}_n \setminus \mathcal{D}_{n-1}$ , have  $d_{\pm} = d \pm 2^{-n} \in \mathcal{D}_{n-1}$ . Now, set

$$B_d = \begin{cases} \frac{B_{d-} + B_{d+}}{2} + \frac{Z_d}{2^{\frac{n+1}{2}}}, & d \in \mathcal{D}_n \setminus \mathcal{D}_{n-1} \\ B_d, & d \in \mathcal{D}_{n-1}. \end{cases}$$

We now show that our candidate process  $(B_d)_{d \in \mathcal{D}_n}$  has independent increments. Indeed, we have that for  $d \in \mathcal{D}_n \setminus \mathcal{D}_{n-1}$ ,

$$\begin{aligned} B_d - B_{d-} &= \frac{B_{d+} - B_{d-}}{2} + \frac{Z_d}{2^{\frac{n+1}{2}}} \\ B_{d+} - B_d &= \frac{B_{d+} - B_{d-}}{2} - \frac{Z_d}{2^{\frac{n+1}{2}}} \end{aligned}$$

are independent. To see this, note that by induction we have that  $\frac{B_{d+} - B_{d-}}{2} \sim \mathcal{N}(0, \frac{d+ - d-}{4})$  and the same holds for  $\frac{Z_d}{2^{\frac{n+1}{2}}}$ . Thus,  $d - B_{d-}$ ,  $B_{d+} - B_d$  are two mean-zero uncorrelated Gaussians, hence they are independent.

Now for any two disjoint intervals of length  $2^{-n}$ , the corresponding increments of the process  $(B_d)_{d \in \mathcal{D}_n}$  are independent since one can express every increment as half the increment of the previous scale plus an independent Gaussian and apply the induction step.

Thus, we have been able to construct  $(B_d, d \in \mathcal{D})$  satisfying the conditions 2&3. Furthermore, by Gaussianity we have

$$\mathbb{E}[|B_d - B_q|^p] = |d - q|^{\frac{p}{2}} \cdot \mathbb{E}[|N|^p],$$

where  $N \sim \mathcal{N}(0, 1)$ . Since for all  $p > 0$   $\mathbb{E}[|N|^p] < \infty$ . By Kolmogorov's continuity criterion, for all  $\alpha \in (0, \frac{\epsilon}{p})$  with  $\epsilon = \frac{p}{2} - 1$   $(B_d, d \in \mathcal{D})$  is a.s.  $\alpha$ -Hölder continuous for all  $\alpha < \frac{1}{2}$ .

We now extend to the whole of  $[0, 1]$  by setting  $B_t = \lim_{i \rightarrow \infty} B_{d_i}$ ,  $d \in \mathcal{D}$ ,  $d_i \rightarrow t$ ,  $i \rightarrow \infty$ . It is immediate that  $(B_{t, t \in [0, 1]})$  is a.s.  $\alpha$ -Hölder continuous for all  $\alpha < \frac{1}{2}$ . Now it remains to check conditions 2&3 are satisfied.

Let  $0 = t_0 \leq t_1 \leq \dots \leq t_n \leq 1$ . Then, we claim the increments  $(B_{t_i} - B_{t_{i-1}})_{i=1, \dots, k}$  are independent Gaussian with  $(B_{t_i} - B_{t_{i-1}}) \sim \mathcal{N}(0, t_i - t_{i-1})$  for all  $1 \leq i \leq k$ . Indeed, let

$$\begin{array}{ccccccc} 0 \leq t_0^n \leq t_1^n \leq \dots \leq t_k^n \leq 1 \\ \downarrow \quad \downarrow \quad \dots \quad \downarrow \\ 0 \leq t_0 \leq t_1 \leq \dots \leq t_k \leq 1 \end{array}$$

be dyadic rationals. By continuity, we have a.s.  $B_{t_j^n} - B_{t_{j-1}^n} \xrightarrow{n \rightarrow \infty} B_{t_j} - B_{t_{j-1}}$  for all  $j \leq k$ . Thus, by bounded convergence,

$$\begin{aligned} \mathbb{E} \left[ \exp \left( i \sum_{j=1}^k u_j \overbrace{(B_{t_j^n} - B_{t_{j-1}^n})}^{\text{independent, normal}} \right) \right] &= \prod_{j=1}^k \exp \left( \frac{-u_j^2 (t_j^n - t_{j-1}^n)}{2} \right) \\ &\xrightarrow{n \rightarrow \infty} \prod_{j=1}^k \exp \left( \frac{-u_j^2 (t_j - t_{j-1})}{2} \right) \equiv \phi(u). \end{aligned}$$

By Lévy's convergence theorem, since  $\phi : \mathbb{R}^k \rightarrow \mathbb{R}$  is the characteristic function of independent Gaussians  $\sim \mathcal{N}(0, t_j - t_{j-1})$  and since the characteristic functions of the increments and the independent Gaussians agree, this forces the law of  $(B_{t_j} - B_{t_{j-1}})_{j \leq k}$  to be that of  $k$  independent  $\mathcal{N}(0, t_j - t_{j-1})$  gaussians. Hence,  $(B_t, t \in [0, 1])$  satisfies all the properties.

2. Extending the construction to all of  $\mathbb{R}$ . Let  $(B_t^i, t \in [0, 1])$  be independent brownian motions and define

$$B_t = B_{t - \lfloor t \rfloor}^i + \sum_{i=0}^{\lfloor t \rfloor - 1} B_t^i, \quad t \geq 0.$$

It is not hard to see that the conditions in 9.1 are satisfied.

3. Now for  $d > 1$ , let  $(B_t^1)_{t \geq 0}, (B_t^2)_{t \geq 0}, \dots, (B_t^d)_{t \geq 0}$  be independent one dimensional Brownian motions. Set  $(B_t)_{t \geq 0} = (B_t^1, \dots, B_t^d)_{t \geq 0}$  and it is easy to check that the conditions are met.

□

**Theorem 9.2.** Let  $B$  be a standard Brownian motion in  $\mathbb{R}^d$ . Then

1. If  $U$  is an orthogonal matrix, then  $UB = (UB_t)_{t \geq 0}$  is also a standard Brownian motion. Hence so is  $-B$ .
2. (Scale invariance:) Let  $\lambda > 0$  be given. Then  $\left( \frac{B_{\lambda t}}{\sqrt{\lambda}} \right)_{t \geq 0}$  is also a standard brownian motion.
3. (Simple Markov property:) For all  $c \geq 0$ ,  $(B_{t+s} - B_s)_{t \geq 0}$  is also a standard Brownian motion and is independent of  $\mathcal{F}_s^B$ , where  $\mathcal{F}_s^B = \sigma(B_u : u \leq s)$ .

*Proof.* Easy to check that it follows from the definition of Brownian motion.

□

## 9.1 Properties of Brownian Motion

**Theorem 9.3 (Time inversion).** Let  $B$  be a standard Brownian motion in  $d = 1$ . Let

$$X_t = \begin{cases} tB_{\frac{1}{t}}, & t > 0 \\ 0, & t = 0. \end{cases}$$

Then  $(X_t)_{t \geq 0}$  is a standard brownian motion.

*Proof.* Fix  $t_1, \dots, t_k > 0$ . Then  $(B_{t_1}, \dots, B_{t_k})$  is Gaussian random vector with zero mean and  $\text{Cov}(B_s, B_t) = s \wedge t$ . Need to check that  $(X_{t_1}, \dots, X_{t_k})$  is Gaussian and with the same covariance as above. By inspection, we see that this vector is clearly Gaussian with zero mean. Now for the covariance, we compute

$$\text{Cov}(X_{t_i}, X_{t_j}) = \text{Cov}(t_i B_{\frac{1}{t_i}}, t_j B_{\frac{1}{t_j}}) = t_i t_j \text{Cov}(B_{\frac{1}{t_i}}, B_{\frac{1}{t_j}}) = t_i t_j \left( \frac{1}{t_i} \wedge \frac{1}{t_j} \right) = t_i \wedge t_j.$$

Now it remains to show that  $X$  is continuous. Indeed, for positive  $t$ ,  $X$  is clearly continuous. Now, we also claim that  $\lim_{t \downarrow 0} X_t = 0$  a.s. Observe that

$$(X_t, t \in \mathbb{Q}_+) \stackrel{d}{=} (B_t, t \in \mathbb{Q}_+)$$

and so

$$\begin{aligned} \mathbb{P} \left( \lim_{t \downarrow 0, t \in \mathbb{Q}_+} X_t = 0 \right) &= \mathbb{P} \left( \bigcap_{N \in \mathbb{N}} \bigcup_{r \in \mathbb{Q}_+} \bigcap_{q \in \mathbb{Q}_+, q < r} \left\{ |X_q| \leq \frac{1}{N} \right\} \right) \\ &= \mathbb{P} \left( \bigcap_{N \in \mathbb{N}} \bigcup_{r \in \mathbb{Q}_+} \bigcap_{q \in \mathbb{Q}_+, q < r} \left\{ |B_q| \leq \frac{1}{N} \right\} \right) = \mathbb{P} \left( \lim_{t \downarrow 0, t \in \mathbb{Q}_+} B_t = 0 \right) = 1. \end{aligned}$$

Finally, since  $\mathbb{Q}_+$  is dense in  $\mathbb{R}_+$  and  $X$  is continuous for  $t > 0$ , we have that

$$\lim_{t \downarrow 0} X_t = \lim_{t \downarrow 0, t \in \mathbb{Q}_+} X_t = 0, \quad \text{a.s.}$$

□

**Corollary 9.3.1.** Let  $B$  be a standard brownian motion in  $d = 1$ . Then,

$$\frac{B_t}{t} \xrightarrow{t \rightarrow \infty} 0, \quad \text{a.s.}$$

*Proof.* By theorem 9.3, we have that with  $X$  defined therein,

$$\lim_{t \rightarrow \infty} \frac{B_t}{t} = \lim_{t \rightarrow \infty} X \left( \frac{1}{t} \right) = 0$$

by the continuity of  $X$  at zero. □

**Definition 9.2.** For  $s \geq 0$ , let  $\mathcal{F}_s^+ = \bigcap_{t > s} \mathcal{F}_t^B = \sigma(B_u : u \leq t)$ . Have  $\mathcal{F}_s^B \subseteq \mathcal{F}_s^+$ .

**Remark.** From the simple Markov property, we have that

$$(B_{t+s} - B_s)_{t \geq 0} \perp \mathcal{F}_s^B.$$

In fact we have more, that is

**Theorem 9.4.** For all  $s \geq 0$ ,

$$(B_{t+s} - B_s)_{t \geq 0} \perp \mathcal{F}_s^+.$$

*Proof.* It suffices to show that if  $t_1, \dots, t_k \in \mathbb{R}_+$  and  $F$  is a continuous and bounded, function on  $(\mathbb{R}^d)^k$  and if  $A \in \mathcal{F}_s^+$  then

$$\mathbb{E}[F(B_{t_1+s} - B_s, \dots, B_{t_k+s} - B_s) \cdot \mathbf{1}(A)] = \mathbb{E}[F(B_{t_1+s} - B_s, \dots, B_{t_k+s} - B_s)] \cdot \mathbb{P}(A).$$

Since, for any open set,  $U \subset (\mathbb{R}^d)^k$ , one can approximate  $F_m \uparrow \mathbf{1}(U)$  from below by bounded continuous functions  $F_m(x) = f_m(\text{dist}(x, U^c))$ ,  $x \in (\mathbb{R}^d)^k$  where  $f : \mathbb{R} \rightarrow \mathbb{R}$  is the continuous, bounded function

$$f(r) = \begin{cases} 1, & r \geq \epsilon \\ \frac{1}{\epsilon}r, & r < \epsilon. \end{cases}$$

for  $r \in \mathbb{R}$  and apply monotone convergence. Then one just has to observe that the collection of open sets generates the borel sigma algebra on  $(\mathbb{R}^d)^k$  and apply the uniqueness of extension theorem.

Now, let  $s_n \downarrow s$  be a strictly decreasing sequence. Then, by continuity, have  $B_{t_i s_n} - B_{s_n} \xrightarrow{n \rightarrow \infty} B_{t_i+s} - B_s$  a.s. for all  $i \leq k$ . Thus, we have

$$\mathbb{E}[F(B_{t_1+s} - B_s, \dots, B_{t_k+s} - B_s) \cdot \mathbf{1}(A)] \stackrel{\text{DCT}}{=} \mathbb{E}[F(B_{t_1+s_n} - B_{s_n}, \dots, B_{t_k+s_n} - B_{s_n}) \cdot \mathbf{1}(A)]$$

and observe that  $A \in \mathcal{F}_s^+$  implies  $A \in \mathcal{F}_{s_n}^B$  for all  $n \in \mathbb{N}$ . Thus, we can conclude by the simple Markov Property and another application of Dominated convergence.  $\square$

**Corollary 9.4.1 (Blumenthal's 0-1 Law).** The sigma algebra  $\mathcal{F}_0^+$  is trivial, i.e. if  $A \in \mathcal{F}_0^+$ , then  $\mathbb{P}(A) \in \{0, 1\}$ .

*Proof.* Take  $A \in \mathcal{F}_0^+ \subseteq \sigma(B_t : t \geq 0)$ . But, by the above, we have  $\sigma(B_t : t \geq 0) \perp \mathcal{F}_0^+$  and so  $A \perp A$  which gives

$$\mathbb{P}(A) = \mathbb{P}(A \cap A) = \mathbb{P}(A) \cdot \mathbb{P}(A).$$

$\square$

**Theorem 9.5.** Let  $B$  be a standard Brownian motion in  $d = 1$ . 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  $n \in \mathbb{N}$ , have that  $\{\tau = 0\} = \bigcap_{k \geq n} \underbrace{\{\epsilon \in (0, 1/k) \text{ s.t. } B_\epsilon > 0\}}_{\mathcal{F}_{\frac{1}{n}}^B}$  and so have  $\{\tau = 0\} \in \mathcal{F}_0^+$

which means that  $\mathbb{P}(\tau = 0) \in \{0, 1\}$ . Now,  $\mathbb{P}(\tau \leq t) \geq \mathbb{P}(B_t > 0) = \frac{1}{2}$  for all  $t > 0$ . So,

$$\mathbb{P}(\tau = 0) = \lim_{t \downarrow 0} \mathbb{P}(\tau \leq t) \geq \frac{1}{2}$$

which gives that  $\mathbb{P}(\tau = 0) = 1$ . By symmetry ( $-B$  is a std BM) we also have that

$$\inf\{t > 0 : B_t < 0\} = 0, \quad \text{a.s.}$$

Since  $B$  is continuous, by the intermediate value theorem we get that  $\sigma = 0$  a.s.  $\square$

**Proposition 9.1.** Let  $B$  be a standard brownian motion in  $d = 1$ . For all  $t \geq 0$ , set  $S_t = \sup_{s \leq t} B_s$  and  $I_t = \inf_{s \leq t} B_s$ . Then,

1. For all  $\epsilon > 0$ , have  $S_\epsilon > 0$  and  $I_\epsilon < 0$  a.s. In other words, in every interval  $(0, \epsilon)$  there exists a zero of BM.
2.  $\sup_{t \geq 0} B_t = +\infty$  and  $\inf_{t \geq 0} B_t = -\infty$  a.s.



*Proof.* 1. Let  $t_n \downarrow t$  as  $n \rightarrow \infty$ . Then we have

$$\{B_{t_n} \text{ i.o.}\} \subseteq \{S_\epsilon > 0\}.$$

It is not hard to see that  $\{B_{t_n} \text{ i.o.}\} \in \mathcal{F}_0^+$ . Thus applying Fatou's lemma we deduce

$$\begin{aligned} \mathbb{P}(B_{t_n} \text{ i.o.}) &= \mathbb{P}(\limsup_n \{B_{t_n} \geq 0\}) \\ &\geq \limsup_n \mathbb{P}(\{B_{t_n} \geq 0\}) = \frac{1}{2} \end{aligned}$$

Thus,  $\mathbb{P}(B_{t_n} \text{ i.o.}) = 1$  and so  $\mathbb{P}(S_\epsilon > 0) = 1$ . By symmetry, ( $-B$  is a std BM) we conclude that  $\mathbb{P}(I_\epsilon < 0) = 1$ .

2. Have for all  $\lambda > 0$  that

$$S_\infty = \sup_{t \geq 0} B_t = \sup_{t \geq 0} B_{\lambda t} \stackrel{d}{=} \sqrt{\lambda} \sup_{t \geq 0} \frac{B_{\lambda t}}{\sqrt{\lambda}}.$$

So  $S_\infty \stackrel{d}{=} \alpha S_\infty$  for all  $\alpha > 0$ . We also know now that  $S_\infty$ . Hence it can only be the case that  $S_\infty = +\infty$  a.s. One can show that  $\inf_{t \geq 0} B_t = -\infty$  a.s. □

**Proposition 9.2.** Let  $B$  be a standard Brownian motion and let  $C$  be a cone with origin at zero and non-empty interior, that is  $C = \{tu : t > 0, u \in A\}$  with  $A \subseteq \mathbb{S}^1$  (= unit sphere in  $\mathbb{R}^d$ ). Set  $H_C = \inf\{t > 0 : B_t \in C\}$ . Then,  $\mathbb{P}(H_C = 0) = 1$ .

*Proof.* Observe that  $\{H_C = 0\} \in \mathcal{F}_0^+$  and  $\mathbb{P}(B_t \in C) = \mathbb{P}(B_1 \in C)$  by scale invariance of Brownian motion and  $C$ . Since  $\text{int } C \neq \emptyset$ ,  $\mathbb{P}(B_1 \in C) > 0$ . Thus,  $\mathbb{P}(H_C \leq t) \geq \mathbb{P}(B_t \in C) > 0$ . Taking  $t \downarrow 0$  and applying Blumenthal concludes the argument. □

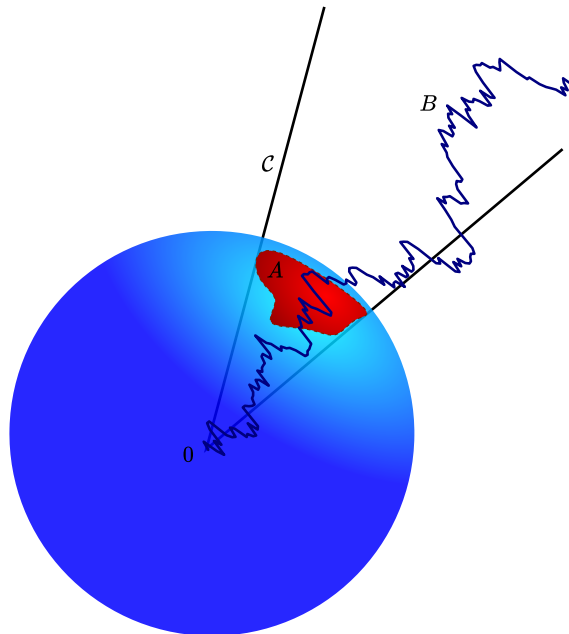


Figure 4: Illustration of cone in proposition 9.2

**Theorem 9.6 (Strong Markov Property).** Let  $B$  be a standard Brownian motion and let  $T$  be an a.s. finite stopping time. Then,  $(B_{t+T} - B_T)_{t \geq 0}$  is a standard Brownian motion and

$$(B_{t+T} - B_T)_{t \geq 0} \perp \mathcal{F}_T^+.$$

*Proof.* Let  $T_n = 2^{-n} \lceil 2^n T \rceil$ ,  $T_n \downarrow T$ ,  $n \rightarrow \infty$ . For  $k \in \mathbb{N}$ , let  $B_t^{(k)} = B_{t+k \cdot 2^{-n}} - B_{k \cdot 2^{-n}}$  and  $B_*^{(n)}(t) = B_{t+T_n} - B_{T_n}$ . Will show that  $B_*$  is a Brownian motion independent of  $\mathcal{F}_{T_n}^+$ .

Clearly,  $B_*^{(n)}$  is continuous. Now, let  $A$  be any event and fix  $E \in \mathcal{F}_{T_n}^+$ . Then, we compute

$$\begin{aligned} \mathbb{P}(B_* \in A, E) &= \sum_{k=0}^{\infty} \mathbb{P}(\overbrace{T_n = k \cdot 2^{-n}}^{\in \mathcal{F}_{k \cdot 2^{-n}}^+}, \overbrace{B^{(k)} \in A, E}^{\perp \mathcal{F}_{k \cdot 2^{-n}}^+}) \\ &= \sum_{k=0}^{\infty} \mathbb{P}(T_n = k \cdot 2^{-n}, E) \cdot \mathbb{P}(B^{(k)} \in A) \\ &= \mathbb{P}(B \in A) \cdot \mathbb{P}(E). \end{aligned}$$

He have thus shown that  $B_* \stackrel{d}{=} B$  and  $\perp \mathcal{F}_{T_n}^+$ . Now, observe that

$$\underbrace{B_{s+t+T} - B_{s+T}}_{\mathcal{N}(0,t)} = \lim_{n \rightarrow \infty} \underbrace{(B_{s+t+T_n} - B_{s+T_n})}_{\mathcal{N}(0,t)}.$$

So,  $(B_{t+T} - B_T)_{t \geq 0}$  is a standard BM.

It remains to show that  $(B_{t+T} - B_T)_{t \geq 0} \perp \mathcal{F}_T^+$ . Indeed, fix  $t_1, \dots, t_k > 0$  and let  $F : (\mathbb{R}^d)^k \rightarrow \mathbb{R}$  be a continuous and bounded function. Fix  $A \in \mathcal{F}_T^+$  and compute

$$\mathbb{E}[F(B_{t_1+T} - B_T, \dots, B_{t_k+T} - B_T) \cdot \mathbf{1}(A)] \stackrel{\text{DCFT}}{=} \lim_{n \rightarrow \infty} \mathbb{E}[F(B_{t_1+T_n} - B_{T_n}, \dots, B_{t_k+T_n} - B_{T_n}) \cdot \mathbf{1}(A)].$$

Since  $A \in \mathcal{F}_T^+$ ,  $A \in \mathcal{F}_{T_n}^+$  for all  $n \in \mathbb{N}$ . Finally, using that  $B_*^{(n)} \perp \mathcal{F}_{T_n}^+$  concludes the proof.  $\square$

**Theorem 9.7 (Reflection principle).** Let  $B$  be a standard Brownian motion in  $d = 1$  and  $T$  an a.s. finite stopping time. Define

$$\tilde{B}_t = \begin{cases} B_t, & 0 \leq t \leq T \\ 2B_T - B_t, & t > T \end{cases}$$

Then  $\tilde{B}$  is a standard Brownian motion.

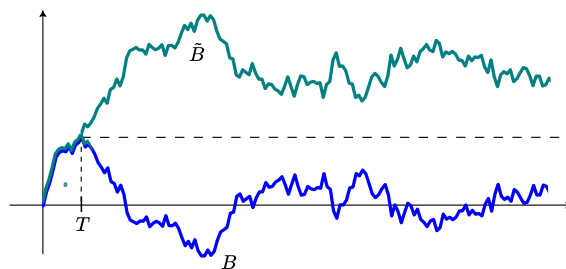


Figure 5: Illustration of reflection of  $B$  at time  $T$ .

*Proof.* We have by the Strong Markov Property that  $B^{(T)} = (B_{t+T} - B_T)_{t \geq 0}$  is a standard Brownian Motion independent of  $\mathcal{F}_T^+$ . Let  $\mathcal{C}_t = \mathcal{C}_0([0, \infty) : \mathbb{R})$  denote the space of continuous functions on the positive reals that vanish at zero, endowed with the topology of local uniform convergence and  $\mathcal{A}$  the induced Borel sigma algebra.

**Metrisability of topology of local uniform convergence**

Recall from Topology that this topology is induced by the metric

$$d : \mathcal{C}_0([0, \infty) : \mathbb{R}) \times \mathcal{C}_0([0, \infty) : \mathbb{R}) \rightarrow \mathbb{R}_+$$

$$(f, g) \mapsto d(f, g) := \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{\sup_{x \in [0, n]} |f(x) - g(x)|}{1 + \sup_{x \in [0, n]} |f(x) - g(x)|}$$

We also have the useful fact that

**Characterisation of  $\mathcal{A}$** 

We have, see Kallenberg's book on the 'Foundations of Modern Probability' for instance, that

$$\mathcal{A} = \sigma(\{\pi_t : t \geq 0\})$$

where for  $t \geq 0$ ,  $\pi_t : \mathcal{C}_0 \rightarrow \mathbb{R}$  denotes the projection onto the  $t$  coordinate.

Now define the function

$$\psi : (\mathcal{C}_0 \times [0, \infty) \times \mathcal{C}_0, \mathcal{A} \otimes \mathcal{B}([0, \infty)) \otimes \mathcal{A}) \rightarrow (\mathcal{C}_0, \mathcal{A})$$

$$(X, T, Y) \mapsto \psi_T(X, Y)(t)$$

$$:= X(t) \cdot \mathbf{1}([0, T])(t) + (X(t) + Y(t - T))\mathbf{1}(T, \infty).$$

is a continuous map in the product topology, therefore measurable. To see that  $\psi$  is continuous,

**Continuity of  $\psi$** 

Fix  $(X, T, Y) \in \mathcal{C}_0 \times [0, \infty) \times \mathcal{C}_0$ . Due to the metrisability of the underlying topologies, it suffices to show that for any sequence  $(X_n, T_n, Y_n)_{n \in \mathbb{N}} \subseteq \mathcal{C}_0 \times [0, \infty) \times \mathcal{C}_0$  such that  $X_n \xrightarrow{d} X$ ,  $Y_n \xrightarrow{d} Y$  and  $T_n \rightarrow T$  as  $n \rightarrow \infty$ ,  $\psi(X_n, T_n, Y_n) \xrightarrow{d} \psi(X, T, Y)$ .

Now, fix  $\epsilon > 0$ , an arbitrary compact set  $K \subseteq \mathbb{R}_+$  and let  $t \in K$  be arbitrary. We estimate

$$\begin{aligned} & |\psi(X_n, T_n, Y_n)(t) - \psi(X, T, Y)(t)| \\ & \leq |(X(t) - X(T)) \cdot \mathbf{1}(t \leq T) - (X_n(t) - X_n(T_n)) \cdot \mathbf{1}(t \leq T_n)| + \cancel{|X(T) - X_n(T_n)|}^0 \\ & \quad + |Y(t - T) \cdot \mathbf{1}((T_n \wedge T, T_n \vee T])| + |Y(t - T) - Y_n(t - T_n)| \\ & \leq |(X(t) - X(T)) \cdot \mathbf{1}((T_n \wedge T, T_n \vee T])| + \cancel{\|X - X_n\|_{\infty, K}}^0 + \cancel{|X(T) - X_n(T_n)|}^0 \\ & \quad + |Y(t - T) \cdot \mathbf{1}((T_n \wedge T, T_n \vee T])| + \cancel{|Y(t - T) - Y_n(t - T_n)|}^0 \\ & \leq |(X(t) - X(T)) \cdot \mathbf{1}((T_n \wedge T, T_n \vee T])| + |Y(t - T) \cdot \mathbf{1}((T_n \wedge T, T_n \vee T])| + \epsilon \end{aligned}$$

(where we make the set harmlessly  $Y(t - T) = 0$  for  $t \leq T$ ) for  $n \in \mathbb{N}$  large enough independent of  $t \in K$ , since the crossed-out terms converge to zero uniformly in  $t \in K$  due to local uniform convergence and uniform continuity on compact sets. The fact that  $Y(t - T)$  and  $X(t) - X(T)$  vanish at  $T$  and that  $T_n \rightarrow T$ ,  $n \rightarrow \infty$  enables us to bound for  $n$  sufficiently large independent of  $t$ :

$$|\psi(X_n, T_n, Y_n)(t) - \psi(X, T, Y)(t)| \leq \sup_{t \in (T_n \wedge T, T_n \vee T]} (|(X(t) - X(T))| + |Y(t - T)|) + \epsilon \leq 2\epsilon$$

and conclude the argument.

Also observe that

$$\begin{aligned} \psi((B_{t \wedge T})_{t \geq 0}, T, B^{(T)}) &= B \\ \psi((B_{t \wedge T})_{t \geq 0}, T, -B^{(T)}) &= \tilde{B} \end{aligned}$$

By observing that  $B_{(T)}$  is independent of the stopped process  $(B_{t \wedge T})_{t \geq 0}$ , we have that

$$((B_t)_{t \geq 0}), T, (B_t^T)_{t \geq 0} \stackrel{d}{=} ((B_t)_{t \geq 0}), T, -(B_t^T)_{t \geq 0}$$

and so it follows that  $B \stackrel{d}{=} \tilde{B}$ . □

**Corollary 9.7.1.** For  $t \geq 0$ , let  $S_t = \sup_{s \leq t} B_s$  and fix  $b > 0$  and  $a \leq b$ . Then

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

*Proof.* Fix  $x > 0$  and define  $T_x = \inf\{t \geq 0 : B_t = x\}$ . Since  $S_\infty < \infty$  a.s., it follows that  $T_x < \infty$  a.s. and  $B_{T_x} = x$ . Observe that  $\{S_t \geq b\} = \{T_b \leq t\}$ . Now we compute,

$$\begin{aligned} \mathbb{P}(S_t \geq b, B_t \leq a) &= \mathbb{P}(\underbrace{T_b \leq t, \tilde{B}_t = 2b - B_t}_{\text{on } T_b \leq t, \tilde{B}_t = 2b - B_t}) \\ &= \mathbb{P}(\tilde{B}_t \geq 2b - a, T_b \leq t) = \mathbb{P}(\underbrace{\tilde{B}_t \geq 2b - a \implies T_b \leq t}_{\tilde{B}_t \geq 2b - a \implies T_b \leq t}) \\ &= \mathbb{P}(B_t \geq 2b - a). \end{aligned}$$

□

**Corollary 9.7.2.**  $S_t \stackrel{d}{=} |B_t|$ .

*Proof.*

$$\begin{aligned} \mathbb{P}(S_t \geq a) &= \mathbb{P}(\overbrace{S_t \geq a, B_t > a}^{=\mathbb{P}(B_t \geq a)}) + \mathbb{P}(\overbrace{S_t \geq a, B_t \leq a}^{=\mathbb{P}(B_t \geq 2b - a) \text{ by the reflection principle}}) \\ &= 2\mathbb{P}(B_t \geq a) \\ &= \mathbb{P}(|B_t| \geq a). \end{aligned}$$

□

**Corollary 9.7.3.** Fix  $x > 0$  and let  $T_x = \inf\{t \geq 0 : B_t = x\}$ . Then

$$T_x \stackrel{d}{=} \left(\frac{x}{B_1}\right)^2.$$

## 9.2 Martingales for Brownian motion

**Theorem 9.8.** Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion in  $d = 1$ . Then

1.  $(B_t)_{t \geq 0}$  is a martingale with respect to the filtration  $(\mathcal{F}_t^+)_{t \geq 0}$
2.  $(B_t^2 - t)_{t \geq 0}$  is a martingale with respect to the filtration  $(\mathcal{F}_t^+)_{t \geq 0}$ .

*Proof.* Fix  $s \leq t$ . Compute

$$\mathbb{E}[B_t | \mathcal{F}_s^+] = \mathbb{E}[\overbrace{B_t - B_s}^{\perp \mathcal{F}_s^+} + B_s | \mathcal{F}_s^+] = B_s, \quad \text{a.s.}$$

and

$$\mathbb{E}[(B_t^2 - t) | \mathcal{F}_s^+] = \mathbb{E}[(B_t - B_s)^2 | \mathcal{F}_s^+] + 2\mathbb{E}[(B_t - B_s)B_s | \mathcal{F}_s^+] + \mathbb{E}[B_s^2 | \mathcal{F}_s^+] - t = B_s^2 - s, \quad \text{a.s.}$$

□

**Corollary 9.8.1.** Let  $B$  be a standard Brownian motion in  $d = 1$  and suppose  $x, y > 0$ . Then

$$\mathbb{P}(T_{-x} < T_y) = \frac{y}{x + y}$$

and

$$\mathbb{E}[T_{-x} \wedge T_y] = x \cdot y$$

with  $T$  defined as in corollary 9.7.3.

**Proposition 9.3.** Let  $B$  be a standard Brownian motion in  $\mathbb{R}^d$ . Set

$$M_t = \exp\left(\langle u, B_t \rangle - \frac{|u|^2 t}{2}\right)$$

is an  $\mathcal{F}_t^+$  martingale for all  $u \in \mathbb{R}^d$ .

*Proof.* Fix  $u \in \mathbb{R}^d$ . Integrability and adaptedness are clear. Now, for the martingale property, we have

$$\begin{aligned} \mathbb{E}[M_t | \mathcal{F}_s^+] &= \mathbb{E}[\exp(\langle u, B_t - B_s \rangle - \langle u, B_s \rangle) | \mathcal{F}_s^+] \cdot e^{-\frac{|u|^2 t}{2}} \\ &= \exp(\langle u, B_s \rangle) \cdot \exp\left(\frac{|u|^2(t-s)}{2}\right) \cdot e^{-\frac{|u|^2 t}{2}} = M_s \end{aligned}$$

□

Lecture 20

**Theorem 9.9.** Let  $f(t, x) : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}$  be continuously differentiable in  $t$  and twice continuously differentiable in  $x$ . Assume  $f$  and all its derivatives are bounded. Then the process

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

is an  $\mathcal{F}_t^+$ -martingale.

*Proof.* By the boundedness assumption,  $M$  is integrable and is clearly adapted. Now it remains to show the martingale property, that is for all  $t, z \geq 0$   $\mathbb{E}[M_{t+s} - M_s | \mathcal{F}_s^+] = 0$ . We have

$$\begin{aligned} M_{t+s} - M_s &= f(t+s, B_{t+s}) - f(s, B_s) - \int_s^{t+s} \left(\frac{\partial}{\partial r} + \frac{1}{2}\Delta\right) f(r, B_r) dr \\ &= f(t+s, B_{t+s}) - f(s, B_s) - \int_0^t \left(\frac{\partial}{\partial r} + \frac{1}{2}\Delta\right) f(r+s, B_{r+s}) dr. \end{aligned}$$

Now, taking conditional expectations, we have

$$\begin{aligned} \mathbb{E}[M_{t+s} - M_s | \mathcal{F}_s^+] &= -f(s, B_s) + \mathbb{E}[f(t+s, B_{t+s} - B_s + B_s) | \mathcal{F}_s^+] \\ &\quad - \mathbb{E}\left[\int_0^t \left(\frac{\partial}{\partial r} + \frac{1}{2}\Delta\right) f(r+s, B_{r+s} - B_s + B_s) dr | \mathcal{F}_s^+\right] \\ &\stackrel{(B_{r+s} - B_s)_{r \geq 0} \perp \mathcal{F}_s^+}{=} -f(s, B_s) - \int_{\mathbb{R}^d} \left(\int_0^t \left(\frac{\partial}{\partial r} + \frac{1}{2}\Delta\right) f(r+s, x + B_s) dr\right) p_r(0, x) dx \\ &\quad + \int_{\mathbb{R}^d} f(t+s, x + B_s) p_t(0, x) dx \end{aligned}$$

where  $p_t(0, x) = \frac{1}{\sqrt{2\pi t^d}} \exp\left(\frac{-|x|^2}{2t}\right)$  for  $x \in \mathbb{R}^d, t > 0$ . Note that  $p_t$  satisfies the heat equation:

$$\frac{\partial p_t}{\partial t} = \frac{1}{2}\Delta p_t.$$

Using dominated convergence, we have that

$$\begin{aligned} &\int_{\mathbb{R}^d} \left(\int_0^t \left(\frac{\partial}{\partial r} + \frac{1}{2}\Delta\right) f(r+s, x + B_s) dr\right) p_r(0, x) dx \\ &= \lim_{\epsilon \downarrow 0} \int_{\mathbb{R}^d} \left(\int_{\epsilon}^t \left(\frac{\partial}{\partial r} + \frac{1}{2}\Delta\right) f(r+s, x + B_s) dr\right) p_r(0, x) dx. \end{aligned}$$

On  $(\epsilon, t)$ , we have enough regularity to integrate by parts and s theorem to obtain

$$\begin{aligned} & \int_{\mathbb{R}^d} \left( \int_0^t \left( \frac{\partial}{\partial r} + \frac{1}{2} \Delta \right) f(r+s, x+B_s) dr \right) p_r(0, x) dx \\ &= \lim_{\epsilon \downarrow 0} \int_{\mathbb{R}^d} \left( \int_\epsilon^t \left( \frac{\partial}{\partial r} + \frac{1}{2} \Delta \right) f(r+s, x+B_s) dr \right) p_r(0, x) dx. \\ &= \int_{\mathbb{R}^d} f(t+s, x+B_s) p_t(0, x) dx - \lim_{\epsilon \downarrow 0} \int_{\mathbb{R}^d} f(\epsilon+s, x+B_s) p_\epsilon(0, x) dx \\ &+ \lim_{\epsilon \downarrow 0} \int_{\mathbb{R}^d} \left( \int_\epsilon^t \left( -\frac{\partial p_r(0, x)}{\partial r} + \frac{1}{2} \Delta p_r(0, x) \right) f(r+s, x+B_s) dr \right) dx \\ &= \mathbb{E}[f(t+s, B_{t+s})] - \lim_{\epsilon \downarrow 0} \mathbb{E}[f(\epsilon+s, B_{\epsilon+s}) | \mathcal{F}_s^+] \\ &\stackrel{\text{DCT}}{=} \mathbb{E} \int_{\mathbb{R}^d} f(t+s, x+B_s) p_t(0, x) dx - f(s, B_s). \end{aligned}$$

Combining all of the above together yields the desired equality  $\mathbb{E}[M_{T+s} - M_s | \mathcal{F}_s^+] = 0$  a.s.  $\square$

### 9.3 Transience and recurrence

Recall that if  $B$  is Brownian motion  $B_0 = 0$  then it is called a standard Brownian motion. More generally, if  $B_0 = x$  then call its law  $\mathbb{P}_x$  and note that  $(B_t - x, t \geq 0)$  is a standard Brownian motion.

**Theorem 9.10.** Let  $B$  be a standard Brownian motion in  $\mathbb{R}^d$ .

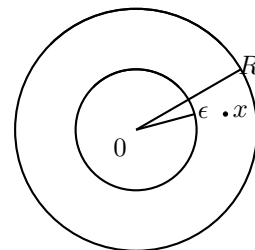
1. If  $d = 1$ , then  $B$  is point-recurrent, i.e. for all  $x, z \{t \geq 0 : B_t = x\}$  is unbounded  $\mathbb{P}_z$ -a.s.
2. If  $d = 2$ , then  $B$  is neighbourhood recurrent, that is for all  $\epsilon > 0$  and  $x, z \in \mathbb{R}^d$  the set of times  $\{t \geq 0 : |B_t - z| \leq \epsilon\}$  is unbounded  $\mathbb{P}_x$ -a.s. However, it does not hit points that is  $\mathbb{P}_x(\exists t \geq 0 : B_t = z) = 0$ .
3. If  $d = 3$ ,  $B$  is transient, that is  $|B_t| \rightarrow \infty$ , as  $t \rightarrow \infty$   $\mathbb{P}_x$ -a.s.

*Proof.* 1.  $d = 1$ : we have almost surely that  $\limsup_{t \rightarrow \infty} B_t = \infty, \liminf_{t \rightarrow \infty} B_t = -\infty$  which gives the result.

2.  $d = 2$ : by translation, it suffices to consider  $z = 0$ . Fix radii  $\epsilon < |x| < R$ . Let  $T_r = \inf\{t \geq 0 : |B_t| = r\}$  for  $r > 0$ . We want to compute  $\mathbb{P}_x(T_\epsilon < T_R)$ . Let  $H = T_\epsilon \wedge T_R$ , an a.s. finite stopping time. Let  $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$  be given by  $\phi(y) = \log |y|$  on the annulus  $\epsilon < |y| < R$  and extended outside that region in a fashion so that  $\phi \in \mathcal{C}_b^2(\mathbb{R}^2)$ . Then,  $\Delta \phi = 0$  in the annulus. By theorem 9.9, the process

$$M_t = \phi(B_t) - \phi(B_0) - \int_0^t \frac{1}{2} \Delta \phi(B_s) ds$$

is a continuous  $(\mathcal{F}_t^+)_{t \geq 0}$ -martingale. An argument similar to that in Theorem 4.6<sup>3</sup> gives  $\mathbb{E}[M_{n \wedge H}] = 0$  for all  $n \in \mathbb{N}$ , in other words,  $\mathbb{E}[\log(|B_{n \wedge H}|)] = \log |x|$ . Taking  $n \uparrow \infty$  and applying DCT gives  $\mathbb{E}[\log(|B_H|)] = \log |x|$ . In other words, expressed in terms of the stopping times  $T_\epsilon, T_R$ , this leads to



$$\mathbb{P}_x(T_\epsilon < T_R) = \frac{\log R - \log |x|}{\log R - \log \epsilon}. \tag{*}$$

Now, taking  $R \rightarrow \infty, T_R \rightarrow \infty$  a.s. and so  $\mathbb{P}_x(T_\epsilon < \infty) = 1$ . We now compute

$$\begin{aligned} \mathbb{P}_x(|B_t| \leq \epsilon \text{ for some } t > n) &= \mathbb{P}_x(|B_{t+n} - B_n + B_n| \leq \epsilon \text{ for some } t > 0) \\ &= \int_{\mathbb{R}} \mathbb{P}_0(|B_t + y| \leq \epsilon \text{ for some } t > 0) p_{n(x,y)} dy = 1. \end{aligned}$$

<sup>3</sup>Approximate  $n \wedge H$  from above by the sequence  $(T_m)_{m \in \mathbb{N}} = (2^{-m} [2^m n \wedge H])_{m \in \mathbb{N}}$ , use the discrete OST on the UI martingale  $(M_{d \wedge T_m})_{d \in \mathcal{D}_m}$  of bdd stopping times and pass to the limit as  $m \rightarrow \infty$  using DCT.

Hence,  $\{t \geq 0 : |B_t| \leq \epsilon\}$  is unbounded  $\mathbb{P}_x$ -a.s. Now, in (\*), letting  $\epsilon \rightarrow 0$ ,  $\mathbb{P}_x(\text{hit } 0 \text{ before } R) = 0$ . Let  $R \rightarrow \infty$  we finally obtain  $\mathbb{P}_x(\exists t > 0 : B_t = 0) = 0$  for all  $x \neq 0$ .

Lecture 21

It remains to show now that  $\mathbb{P}_0(B_t = 0 \text{ for some } t > 0) = 0$ . Indeed, let  $a > 0$ , and observe that

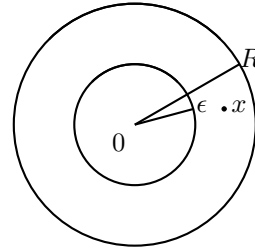
$$\begin{aligned} \mathbb{P}_0(B_{t+a} = 0 \text{ for some } t > 0) &= \mathbb{P}_0(\overbrace{B_{t+a} - B_a}^{\perp \mathcal{F}_a^+} + B_a = 0 \text{ for some } t > 0) \\ &= \int_{\mathbb{R}^2} \mathbb{P}_0(\overbrace{B_{t+a} - B_a}^{\text{std BM}} + y = 0 \text{ for some } t > 0) p_a(y) dy \\ &= \int_{\mathbb{R}^2} \mathbb{P}_y(B_t = 0 \text{ for some } t > 0) p_a(y) dy = 0 \end{aligned}$$

since  $\mathbb{P}_y(\exists t > 0 : B_t = 0) = 0$  for all  $y \neq 0$ . So taking the limit as  $a \downarrow 0$ , we get

$$\mathbb{P}_0(B_t = 0 \text{ for some } t > 0) = \lim_{a \downarrow 0} \mathbb{P}_0(B_t = 0 \text{ for some } t > a) = 0.$$

3. We now show that  $B$  is transient for  $d \geq 3$ , that is  $|B_t| \rightarrow \infty$ , as  $t \rightarrow \infty$ . To this end, it clearly suffices to prove transience for  $d = 3$ .

As in the case  $d = 2$ , start by fixing radii  $\epsilon < |x| < R$ . Let  $T_r = \inf\{t \geq 0 : |B_t| = r\}$  for  $r > 0$ . Let  $H = T_\epsilon \wedge T_R$ , an a.s. finite stopping time. Let  $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$  be given by  $\phi(y) = \left(\frac{1}{|y|}\right)^{2-d}$  on the annulus  $\epsilon < |y| < R$  and extended outside that region in a fashion so that  $\phi \in C_b^2(\mathbb{R}^2)$ . Then,  $\Delta\phi = 0$  in the annulus. By theorem 9.9, the process



$$M_t = \phi(B_t) - \phi(B_0) - \int_0^t \frac{1}{2} \Delta\phi(B_s) ds$$

is a continuous  $(\mathcal{F}_t^+)_{t \geq 0}$ -martingale. Arguing in the same way as above, we obtain  $\mathbb{E} \left[ \left(\frac{1}{|B_H|}\right)^{2-d} \right] = \left(\frac{1}{|x|}\right)^{2-d}$ . In other words, expressed in terms of the stopping times  $T_\epsilon, T_R$ , this leads to

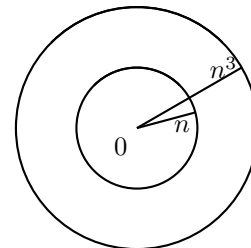
$$\mathbb{P}_x(T_\epsilon < T_R) = \frac{\left(\frac{1}{R}\right)^{2-d} - \left(\frac{1}{|x|}\right)^{2-d}}{\left(\frac{1}{\epsilon}\right)^{2-d} - \left(\frac{1}{R}\right)^{2-d}}. \tag{**}$$

Now, taking  $R \rightarrow \infty, T_R \rightarrow \infty$  a.s. and so  $\mathbb{P}_x(T_\epsilon < \infty) = \left(\frac{\epsilon}{|x|}\right)^{2-d}$ .

For  $n \in \mathbb{N}$ , let  $A_n = \{|B_t| > n \text{ for all } t \geq T_{n^3}\}$ ,  $T_{n^3}$  being almost surely finite for all  $n \in \mathbb{N}$ . To prove  $|B_t| \rightarrow \infty$  a.s. as  $t \rightarrow \infty$ , it suffices to show that the  $A_n$  happen eventually a.s. (recall  $d = 3$ ). We now compute

$$\begin{aligned} \mathbb{P}_0(A_n^c) &= \mathbb{P}_0(|B_t| \leq n \text{ for some } t \geq T_{n^3}) \\ &\stackrel{\text{SMP}}{=} \mathbb{E}_0 \left[ \mathbb{P}_{B_{T_{n^3}}}(|B_t| \leq n \text{ for some } t \geq 0) \right] = \frac{1}{n^2} \end{aligned}$$

so  $\sum_{n=1}^\infty \mathbb{P}(A_n^c) < \infty$  and so we conclude that  $A_n$  occurs eventually in  $n \in \mathbb{N}$  a.s. thereby showing transience. □



### 9.4 Dirichlet Problem

**Definition 9.3 (Poincaré cone condition).**  $D \subseteq \mathbb{R}^d$  is called a domain if it is open, non-empty and connected. We say that  $D$  satisfies the Poincaré cone condition at  $x \in \partial D$  if there exists a non-empty open cone  $\mathcal{C}$  with origin at  $x$  and  $r > 0$  such that  $\mathcal{C} \cap B(x, r) \subseteq D^c$ .

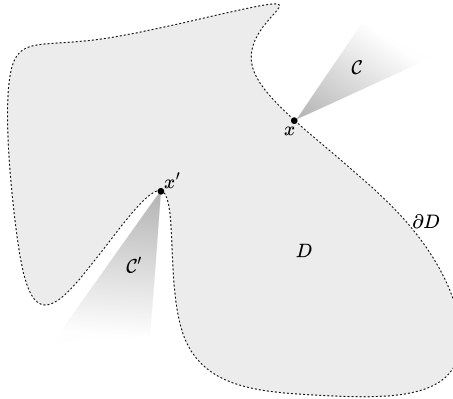


Figure 6: Illustration of Poincaré cone condition for a domain  $D \subseteq \mathbb{R}^d$ .

**Theorem 9.11 (Dirichlet problem).** let  $D$  be a bounded domain in  $\mathbb{R}^d$  such that every boundary point of  $D$  satisfies the Poincaré cone condition, (see figure 6). Let  $\phi$  be continuous on  $\partial D$  and let  $B$  be a Brownian motion,  $\tau_{\partial D} = \inf_{\{t \geq 0: B_t \in \partial D\}}$ . Then the function

$$u(x) = \mathbb{E}_x [\phi(B_{\tau_{\partial D}})], \quad x \in \bar{D}$$

is the unique continuous function satisfying the boundary value problem

$$\begin{cases} \Delta u = 0, & \text{in } D \\ u = \phi, & \text{on } \partial D. \end{cases}$$

Before we proceed with the proof we recall some facts from the theory of PDEs.

**Theorem 9.12.** Let  $D \subseteq \mathbb{R}^d$  be a domain and  $u : D \rightarrow \mathbb{R}$  be measurable and locally bounded. Then the following are equivalent:

1.  $u$  is twice continuously differentiable and  $\Delta u = 0$ .
2. For all balls  $B(x, r) \subseteq D$ ,

$$u(x) = \frac{1}{\mathcal{L}(B(x, r))} \int_{B(x, r)} u(y) \, dy.$$

3. For all balls  $B(x, r) \subseteq D$ ,

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

where  $\sigma_{x,r}$  denotes the surface area measure of  $\partial B(x, r)$ .

**Definition 9.4 (Harmonic).** if  $u$  satisfies any of the above, we call  $u$  harmonic in  $D$ .



**Theorem 9.13 (Maximum principle).** Let  $u : \mathbb{R}^d \rightarrow \mathbb{R}$  be harmonic in  $D$ . Then

1. If  $u$  attains its maximum in  $D$ , then  $u$  is constant in  $D$ .
2. If  $u$  is continuous in  $\bar{D}$  and  $D$  is bounded, then  $\max_{x \in \bar{D}} u(x) = \max_{x \in \partial D} u(x)$ .

*Proof.* 1. Let  $M$  be the maximum, let  $V = \{x \in D : u(x) = M\}$ , then by assumption such that the ball  $B(x, r) \subseteq D$ . Then, by the mean value property

$$M = u(x) = \frac{1}{\mathcal{L}(B(x, r))} \int_{B(x, r)} u(y) dy \leq M.$$

Hence,  $u(y) = M$  for almost all  $y \in B(x, r)$ . By the continuity of  $u$ , we have equality everywhere in  $B(x, r)$ . Thus,  $B(x, r) \subseteq V$  and so  $V$  is now open, closed and also non-empty. Since,  $D$  is connected, we deduce that  $V = D$ .

2.  $u$  is continuous in  $\bar{D}$  and  $D$  is bounded implies that  $u$  attains a maximum in  $\bar{D}$ . By 1,  $\max_{\bar{D}} u = \max_{\partial D} u$ . □

**Corollary 9.13.1.** If  $u_1, u_2 : \mathbb{R}^d$  are harmonic in  $D$ , with  $D$  bounded and  $u_1, u_2$  agree on  $\partial D$ , then  $u_1 = u_2$  in  $D$ .

*Proof.* Have

$$\max_{\bar{D}} (u_1 - u_2) = \max_{\partial D} (u_1 - u_2) = 0$$

by the maximum principle. Thus,  $u_1 \leq u_2$  for all  $x \in \bar{D}$  and similarly we obtain  $u_2 \leq u_1$  in  $\bar{D}$ . Thus, we obtain  $u_1 \equiv u_2$  in  $\bar{D}$ . □

Lecture 22

*Proof.* (Theorem 9.11) With

$$u(x) = \mathbb{E}_x [\phi(B_{\tau_{\partial D}})], \quad \text{on } \bar{D}$$

to show that  $u$  is twice continuously differentiable and harmonic, Theorem 9.12 it suffices to show that it satisfies the mean value property.

Now, we have that  $\tau = \inf\{t \geq 0 : B_t \in (x, \delta)\} < \infty$  a.s. and by the tower property

$$u(x) = \mathbb{E}_x [\phi(B_{\tau_{\partial D}})] = \mathbb{E} [\mathbb{E}_x [\phi(B_{\tau_{\partial D}}) | \mathcal{F}_\tau^+]]$$

Now, define the function

$$F : \mathbb{R}_+ \times \mathcal{C}_0([0, \infty)) \rightarrow \mathbb{R} \\ (z, f) \mapsto F(z, f) := \inf\{t \geq 0 : z + f(t) \in D^c\}$$

which is measurable with respect to  $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{A}$ , where  $\mathcal{A}$  is the Borel sigma algebra induced by the topology of local uniform convergence, as in the proof of Theorem 9.7. Observe now that  $\tau_{B_{\partial D}} = \tau + F(B_\tau, (B_{\tau+t} - B_\tau)_{t \geq 0})$ . a.s. By the strong Markov property,  $(B_{\tau+t} - B_\tau)_{t \geq 0} \perp \mathcal{F}_\tau^+$  and so we can conclude

$$\begin{aligned} u(x) &= \mathbb{E} \left[ \mathbb{E}_x \left[ \phi(B_{\tau + F(B_\tau, (B_{\tau+t} - B_\tau)_{t \geq 0})}^{(\tau)} + B_\tau) | \mathcal{F}_\tau^+ \right] \right] \\ &= \mathbb{E} \left[ \mathbb{E}_x \left[ \phi(B_{\tau + F(B_\tau, (B_{\tau+t} - B_\tau)_{t \geq 0})}^{(\tau)} + B_\tau) | \mathcal{F}_\tau^+ \right] \right] \\ &= \mathbb{E} \left[ \mathbb{E}_x \left[ \phi(g(B_\tau, (B_{\tau+t} - B_\tau)_{t \geq 0}) + B_\tau) | \mathcal{F}_\tau^+ \right] \right] \end{aligned}$$

where  $g$  is the continuous, hence measurable function

$$F : \mathbb{R}_+ \times \mathcal{C}_0([0, \infty)) \rightarrow \mathbb{R} \\ (z, f) \mapsto g(z, f) := f(z)$$

thus another application of the strong Markov property gives

$$\begin{aligned}
 u(x) & \stackrel{\text{SMP \& indep.}}{=} \mathbb{E}_x \left[ \int_{\mathcal{C}_0([0,\infty))} \left[ \phi(g(B_\tau, w) + B_\tau) \right] \underbrace{\mu(dw)}_{\substack{\mu \text{ denotes the Wiener measure of BM started at } x \\ \text{the law of } B_\tau \text{ is invariant under rotations and by uniqueness is the spherical measure } \sigma(x,\delta)}} \right] \\
 & = \frac{1}{\sigma_{x,r}(B(x,\delta))} \int_{\partial B(x,\delta)} \int_{\mathcal{C}_0([0,\infty))} \phi(g(y, w) + y) \mu(dw) d\sigma_{x,\delta} y \\
 & = \frac{1}{\sigma_{x,r}(B(x,\delta))} \int_{\partial B(x,\delta)} \mathbb{E}_y [\phi(g(y, B) + y)] d\sigma_{x,\delta} y \\
 & = \frac{1}{\sigma_{x,r}(B(x,\delta))} \int_{\partial B(x,\delta)} u(y) d\sigma_{x,\delta} y
 \end{aligned}$$

thus, showing that  $u$  is indeed harmonic. Uniqueness follows from the result 9.13.1 established earlier.

It remains to show that  $u$  is continuous up to the boundary,  $\partial D$ . Let  $z \in \partial D$ . Need to show that  $u$  is continuous at  $z$ . Since  $\phi$  is continuous on  $\partial D$ , we have that for all  $\epsilon > 0$ , there exists a  $\delta > 0$  such that if  $|y - z| \leq \delta$ ,  $y \in \partial D$ ,  $|\phi(y) - \phi(z)| < \epsilon$ .

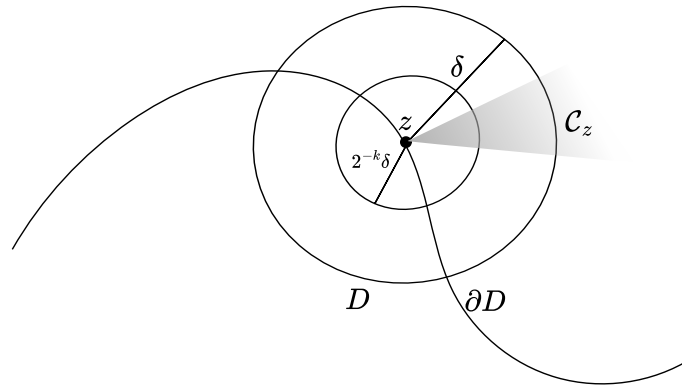


Figure 7: Illustration of situation near the boundary.

Let  $k \in \mathbb{N}$  to be determined and let  $x$  be such that  $|x - z| \leq 2^{-k} \cdot \delta$ , then we estimate

$$\begin{aligned}
 |u(x) - u(z)| & \leq |\mathbb{E}_x [\phi(B_{\tau_{\partial D}}) - \phi(z)]| \\
 & \leq \mathbb{E}_x [|\phi(B_{\tau_{\partial D}}) - \phi(z)|] \\
 & \leq \epsilon \cdot \mathbb{P}_x(\tau_{\partial D} < \tau_{\partial B(z,\delta)}) + 2 \|\phi\|_\infty \cdot \mathbb{P}_x(\tau_{\partial B(z,\delta)} < \tau_{\partial D}).
 \end{aligned}$$

Now, By the Poincaré cone condition, let  $\mathcal{C}_z$  be an open cone at  $z$  that lies in  $D^c$  sufficiently close to  $z$ . Then

$$\mathbb{P}_x(\tau_{\partial B(z,\delta)} < \tau_{\partial D}) \leq \mathbb{P}_x(\tau_{\partial B(z,\delta)} < \tau_{\partial \mathcal{C}_z}).$$

We claim that

$$\sup_{x \in B(0, \frac{1}{2})} \mathbb{P}_x(\tau_{\partial B(0,1)} < \tau_{\mathcal{C}}) \leq \alpha < 1. \tag{***}$$

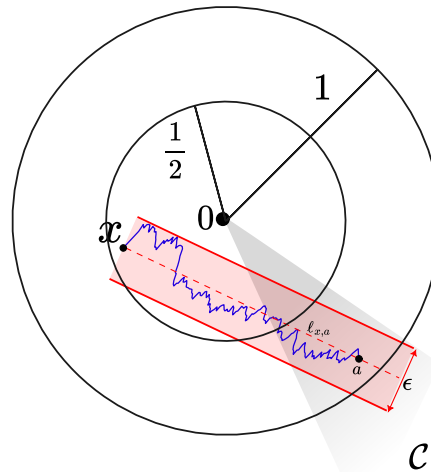
where  $\mathcal{C}$  is a translate of the cone  $\mathcal{C}_z$  to the origin.

**Proof of (\*\*\*)**

To establish (\*\*\*) it suffices to show that Brownian motion stays arbitrarily close to straight bounded segment of lines (and in fact to any continuous function) with positive probability, i.e.

$$p_{\epsilon,x,a} := \mathbb{P}_x(\|B - \ell_{x,a}\|_{\infty,[0,1]} \leq \epsilon) > 0$$

$\epsilon > 0, x \in \mathbb{R}^d$  and lines  $\ell_{x,a}$  connecting points  $x, a \in \mathbb{R}^d$ , i.e.  $\ell_{x,a}(t) = tx + (1-t)a, t \in [0, 1]$  (see figure 9.4). To see this, the bound (\*\*\*) essentially reduces to bounding uniformly from zero the probability that a Brownian path starting from a point  $x \in B(0, 1/2)$  stays within some uniform in  $x$  amount  $\epsilon > 0$  close to a line of length  $\leq 2$  (figure 9.4).



By the geometry of the situation, an  $\epsilon > 0$  exists, so that no matter where the BM starts in  $B(0, 1/2)$ , there is a direction it can stay within epsilon to for times  $t \in [0, 1]$  that would guarantee it never touches the boundary  $\partial B(0, 1)$  before the cone, in  $[0, 1]$ .

We proceed with several reduction steps. By translation, we can let without loss of generality  $x = 0$  and consider only  $p_{\epsilon,a} := p_{\epsilon,0,a}$  for  $\epsilon > 0, a \in \mathbb{R}^d$ . Moreover, by containment of events that  $p_{\epsilon,a}$  is decreasing in  $\epsilon > 0$  and less than one away from zero with  $p_{\epsilon,0} = 1$  for all  $\epsilon > 0$ . By independence of the components of  $B$  and rotational symmetry, we have

$$\mathbb{P}_0(\|B_t - at\|_{\infty,[0,1]} \leq \epsilon) = \left(\mathbb{P}_0(\|B\|_{\infty,[0,1]} \leq \epsilon)\right)^{d-1} \cdot \mathbb{P}_0\left(\|B_t^1 - \|a\| \cdot t\|_{\infty,[0,1]} \leq \epsilon\right)$$

We now show that for fixed  $\epsilon, > 0$  and  $|\lambda|$  bounded, one has a positive uniform lower bound on the probabilities

$$\mathbb{P}_0(|W_t - \lambda t| \leq \epsilon).$$

where  $W$  is a standard Brownian motion in one dimension. To show this it suffices to note that from Lévy's construction of Brownian motion, one constructs  $(W_t, t \in [0, 1])$  as an a.s. uniformly convergent power series (starting from the iid sequence  $(Z_d)_{d \in \mathcal{D}}$ ) as

$$W_t = \sum_{n=1}^{\infty} F_n(t)$$

where the  $F_n$  are independent, piecewise linear functions given by

$$F_n(t) = \begin{cases} 2^{-\frac{n+1}{2}} Z_t, & t \in \mathcal{D}_n \setminus \mathcal{D}_{n-1} \\ 0, & t \in \mathcal{D}_{n-1} \\ \text{linear between consecutive points in } \mathcal{D}_n. \end{cases}$$

for  $n \geq 1$  and define  $F_0$  by interpolating linearly between  $Z_1$  and zero in  $[0, 1]$ . Now, a straightforward computation yields that for  $c > 0, n \in \mathbb{N}$  sufficiently large independent of  $c, \mathbb{P}(|Z_d| \geq c\sqrt{n}) \leq \exp\left(-\frac{c^2 n}{2}\right)$ . This in conjunction with the Borell-Cantelli lemma gives

$$\mathbb{P}(\{\exists d \in \mathcal{D}_n, \text{ s.t. } |Z_d| < c\sqrt{n}\}) = 0, \quad \text{for } c > \sqrt{2 \log 2}.$$

Thus, using the continuity of  $\mathbb{P}$ , there exists some  $M \in \mathbb{N}$  such that

$$\mathbb{P}\left(\bigcap_{n \geq M} \{\|Z_n\|_{\infty} < c\sqrt{n}2^{-\frac{n}{2}}\}\right) > 0.$$

Additionally, observe that for the truncated series  $\sum_{n=1}^N F_n$  corresponds to the piecewise linear extension of the dyadic approximations to  $W$  in Theorem 9.1. Now, for  $|\lambda|$  bounded,

$M$  possibly even larger, by the independence of the Gaussians  $(Z_d, d \in \mathcal{D})$  it is not hard to approximate  $\lambda t$  on  $[0, 1]$  by the truncated series up to stage  $M - 1$  and then use the independence of the  $F_n$  and the above to obtain with a positive probability that the tail is also uniformly close and the approximation at stage  $M - 1$ , hence proving there is a uniform positive lower bound, thus proving (\*\*\*) .

In the final step, essentially we will iterate the bound (\*\*\*) on nested balls surrounding the origin to get exponential decay in the corresponding version of (\*\*\*) . More precisely,

### Exponential decay step

Let  $k \in \mathbb{N}$ . We want to bound

$$\mathbb{P}_x(\tau_{\partial B(0,1)} \subseteq \tau_{\mathcal{C}}), \quad x \in B(0, 1/2^k).$$

Now, we have

$$\begin{aligned} \mathbb{P}_x(\tau_{\partial B(0,1)} \subseteq \tau_{\mathcal{C}}) &= \mathbb{E}_x \left[ \mathbf{1}(\tau_{\partial B(0,1)} < \tau_{\mathcal{C}}) \right] \\ &= \mathbb{E}_x \left[ \mathbf{1}(\tau_{\partial B(0,2^{-k})} < \tau_{\mathcal{C}}) \cdot \mathbb{E} \left[ \mathbf{1}(\tau_{\partial B(0,1)} < \tau_{\mathcal{C}}) \mid \mathcal{F}_{\tau_{\partial B(0,2^{-k})}}^+ \right] \right] \\ &\stackrel{SMP}{=} \mathbb{E}_x \left[ \mathbf{1}(\tau_{\partial B(0,2^{-k})} < \tau_{\mathcal{C}}) \cdot \mathbb{P}_{B_{\tau_{\partial B(0,2^{-k})}}}( \tau_{\partial B(0,1)} < \tau_{\mathcal{C}}) \right] \\ &\leq \mathbb{P}_x(\tau_{\partial B(0,1)} < \tau_{\mathcal{C}}) \cdot \sup_{y \in B(0,2^{-(k-1)})} \mathbb{P}_y(\tau_{\partial B(0,1)} < \tau_{\mathcal{C}}). \end{aligned} \tag{A}$$

Now, for  $x \in B(0, 2^{-k})$  consider  $\mathbb{P}_x(\tau_{\partial B(0,2^{-(k-1)})} < \tau_{\mathcal{C}})$ . By scale invariance of Brownian motion and the cone,  $\mathcal{C}$  we have with  $\lambda = 2^{(k-1)}$ ,

$$\begin{aligned} &\mathbb{P}_x(\tau_{\partial B(0,2^{-(k-1)})} < \tau_{\mathcal{C}}) \\ &= \mathbb{P}_0(\inf\{t \geq 0 : \overbrace{W_t}^{\text{std BM}} \in \overline{B(0, 2^{-(k-1)})}\} < (\inf\{t \geq 0 : W_t \in \overline{\mathcal{C}}\})) \\ &\stackrel{\text{scale invariance}}{=} \mathbb{P}_0(\inf\{t \geq 0 : 1/\lambda W_{\lambda^2 t} \in \overline{B(0, 2^{-(k-1)})}\} < (\inf\{t \geq 0 : 1/\lambda W_{\lambda^2 t} \in \overline{\mathcal{C}}\})) \\ &= \mathbb{P}_0(\inf\{t \geq 0 : W_{\lambda^2 t} + \lambda x \in \overline{B(0, 1)}\} < (\inf\{t \geq 0 : W_{\lambda^2 t} + \lambda x \in \overline{\mathcal{C}}\})) \\ &= \mathbb{P}_0(1/\lambda^2 \inf\{\lambda^2 t \geq 0 : W_{\lambda^2 t} + \lambda x \in \overline{B(0, 1)}\} < 1/\lambda^2 \inf\{\lambda^2 t \geq 0 : W_{\lambda^2 t} + \lambda x \in \overline{\mathcal{C}}\}) \\ &= \mathbb{P}_{\lambda x}(\inf\{t \geq 0 : B_t \in \overline{B(0, 1)}\} < \inf\{t \geq 0 : B_t \in \overline{\mathcal{C}}\}) \\ &= \mathbb{P}_{\lambda x}(\tau_{\partial B(0,1)} < \tau_{\mathcal{C}}) \leq \sup_{|y| \leq 2^{-(k-1)}} \mathbb{P}_y(\tau_{\partial B(0,1)} < \tau_{\mathcal{C}}). \end{aligned}$$

Thus, inducting on  $k \in \mathbb{N}$  and using (\*\*\*) as the base case and (A) as the induction step we deduce that

$$\sup_{x \in B(0, 2^{-k})} \mathbb{P}_x(\tau_{\partial B(z, \delta)} < \tau_{\mathcal{C}_z}) \leq \alpha^k \rightarrow 0, \quad k \rightarrow \infty$$

which allows us to conclude the proof.  $\square$

### Example:

Let  $d = 2$  and let  $\phi : \partial B(0, 1) \rightarrow \mathbb{R}$  continuous. Let  $v : D \rightarrow \mathbb{R}$  where  $D = B(0, 1) \setminus \{0\}$  be the unique solution to the Dirichlet problem on  $B(0, 1)$  with boundary data  $\phi$ . Augment  $\phi$  to  $\partial B(0, 1) \cap \{0\}$  and observe that

$$u(x) = \mathbb{E}_x[\phi(B_{\tau_{\partial D}})]$$

is not a solution if  $v(0) \neq \phi(0)$  since  $u(0) = \phi(0) = v(0)$  because Brownian motion does not hit points, as we proved in Theorem 9.10.

## 10 Donsker's invariance principle

The main theorem of this section is Donsker's invariance principle, which states

**Theorem 10.1 (Donsker's invariance principle).** Let  $X_1, X_2, \dots$  be iid  $\mathbb{R}$ -valued integrable random variables with law  $\mu$ , such that  $\mathbb{E}[X_1] = 0$ , and variance  $\sigma^2 \in (0, \infty)$ . Set  $S_0 = 0$ ,  $S_n = X_1 + \dots + X_n$  for  $n \geq 1$  and  $S_t = (1 - \{t\})S_{[t]} + \{t\}S_{[t]+1}$ , where  $\{t\} = t - [t]$  and  $[t]$  is the integer part of  $t \geq 0$ . Now, define

$$S_t^{[N]} = \frac{S_{tN}}{\sqrt{\sigma^2 N}}$$

for  $0 \leq t \leq 1$ . Then,  $(S_t^{[N]}, 0 \leq t \leq 1)$  converges weakly to  $(B_t, 0 \leq t \leq 1)$ , that is to a standard Brownian motion. More explicitly, we have for all continuous (in the local uniform topology) and bounded functionals  $F : \mathcal{C}([0, 1], \mathbb{R}) \rightarrow \mathbb{R}$

$$\mathbb{E}[F(S^{[N]})] \xrightarrow{N \rightarrow \infty} \mathbb{E}[F(B)].$$

Before we prove it we need a supporting result, the so-called Skorokhod embedding.

**Theorem 10.2 (Skorokhod embedding).** Let  $\mu$  be a probability measure with zero mean and variance  $\sigma^2 \in (0, \infty)$ . Then, there exists a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , a filtration  $(\mathcal{F}_t)_{t \geq 0}$ , a Brownian motion  $(B_t)_{t \geq 0}$  and a sequence of stopping times  $0 = T_0 \leq T_1 \leq \dots$  such that

1. The sequence defined by  $(S_n)_{n \in \mathbb{N}} = (B_{T_n})_{n \in \mathbb{N}}$  for  $n \in \mathbb{N}$  is a random walk with step distribution  $\mu$ .
2. The sequence  $(T_n)_{n \in \mathbb{N}}$  and steps of mean  $\sigma^2$ .

### Lecture 23

*Proof.* Define the Borel measures on  $\mathcal{B}([0, \infty))$ , for  $A \in \mathcal{B}([0, \infty))$

$$\begin{aligned} \mu_+(A) &= \mu(A \cap [0, \infty)) \\ \mu_-(A) &= \mu(-A \cap (-\infty, 0)). \end{aligned}$$

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space on which we define a standard Brownian motion  $(B_t)_{t \geq 0}$  and the iid sequence  $(X_n, Y_n)_{n \in \mathbb{N}}$  with law  $\nu(dx, dy) = C \cdot \mu_-(dx)\mu_+(dy)$  (independent from  $B$ ) and  $C > 0$  a normalising constant. We have

$$\begin{aligned} \int_0^\infty \int_0^\infty \nu(dx, y dy) = 1 &= C\mu([0, \infty)) \int_0^\infty x\mu_-(dx) \\ &+ C\mu((-\infty, 0)) \int_0^\infty y\mu_+(dy). \end{aligned}$$

Since  $\mu$  has mean zero, we also have that  $\int_0^\infty x\mu_-(dx) = \mu((-\infty, 0)) \int_0^\infty y\mu_+(dy)$  which gives

$$C \int_0^\infty x\mu_-(dx) = C\mu((-\infty, 0)) \int_0^\infty y\mu_+(dy) = 1.$$

Now, define the random sequence  $T_0 = 0$ , and for  $n \geq 1$

$$T_{n+1} = \inf\{t \geq T_n : B_t - B_{T_n} \in \{-X_{n+1}, Y_{n+1}\}\}.$$

We claim that the  $(T_n)_{n \in \mathbb{N}}$  are stopping times with respect to the filtration  $\mathcal{F}_t = \sigma(\mathcal{F}_t^B, \mathcal{F}_0)$  for  $t \geq 0$  where  $\mathcal{F}_0 = \sigma((X_n, Y_n) : n \in \mathbb{N})$ .

**$(T_n)_{n \in \mathbb{N}}$  are stopping times**

To see that the  $(T_n)_{n \in \mathbb{N}}$  are stopping times, we proceed by induction. Clearly  $T_0 \equiv 0$  is a stopping time. Now, suppose for  $n \in \mathbb{N}$  that  $T_n$  is a stopping time. Now, for  $s \geq 0$  and observe that since  $T_{n+1} = T_n + \eta$  a.s. where  $\eta = \inf\{t \geq 0 : B^{(T_n)} \in \{-X_{n+1}, Y_{n+1}\}\}$

$$\{T_{n+1} \leq s\} = \overbrace{\{T_n \leq r\}}^{\{T_{n+1} \leq t\} \subseteq \{T_n \leq t\}} \cap \{\eta \leq s\}.$$

Observe, that  $\eta$  is  $\sigma(\mathcal{F}_{T_n}^B, \mathcal{F}_0)$ -measurable. Hence, we have that  $\{T_{n+1} \leq s\} \in \sigma(\mathcal{F}_{T_n}^B, \mathcal{F}_0) \cap \mathcal{F}_s^B$ . Noting that we can express  $\sigma(\mathcal{F}_{T_n}^B, \mathcal{F}_0) = \sigma(\{A \cap B : A \in \mathcal{F}_{T_n}^B, B \in \mathcal{F}_0\}) \subseteq \mathcal{F}_s^B$ , we conclude that  $\{T_{n+1} \leq s\} \in \mathcal{F}_s^B$  for any  $s \geq 0$ , hence proving the statement.

Now, define the measurable (wrt to the usual sigma algebras) function

$$\begin{aligned} \tau : \mathcal{C}([0, \infty)) \times \mathbb{R} &\rightarrow \mathbb{R} \\ (f, x) &\mapsto \inf\{t \geq 0 : f(t) = x\} \end{aligned}$$

and conditioning on  $X_1, Y_1$  we compute

$$\begin{aligned} \mathbb{P}(B_{T_1} = Y_1 | X_1, Y_1) &= \mathbb{E}[\mathbf{1}(\tau(B, Y) < \tau(B, -X)) | X_1, Y_1] \\ &= \frac{X_1}{X_1 + Y_1} \end{aligned}$$

using the well-known Gambler's ruin identity and the independence of  $B$  from  $(X_n, Y_n)_{n \in \mathbb{N}}$ . Similarly, we also obtain  $\mathbb{E}[T_1 | X_1, Y_1] = X_1 \cdot Y_1$ .

Now we determine the law of  $B_{T_1}$ . Fix  $A \in \mathcal{B}([0, \infty))$  and compute

$$\begin{aligned} \mathbb{P}(B_{T_1} \in A) &= \int_A \int_0^\infty \frac{x}{x+y} C(x+y) \mu_-(dx) \mu_+(dy) \\ &= \mu_+(A) = \mu(A). \end{aligned}$$

Similarly, we have for  $A \in \mathcal{B}((-\infty, 0))$   $\mathbb{P}(B_{T_1} \in A) = \mu(A)$  and

$$\mathbb{E}[T_1] = \int_0^\infty \int_0^\infty x \cdot y \cdot C(x+y) \mu_-(dx) \mu_+(dy) = \int_0^\infty x^2 \mu_-(dx) + \int_0^\infty y^2 \mu_+(dy) = \sigma^2.$$

This ties with the case  $n = 1$ , for the general case one proceeds inductively using the strong Markov property, that is  $(B_{t+T_n} - B_{T_n})_{t \geq 0} \perp \mathcal{F}_{T_n}^B$  and essentially reduce the argument to what we have already done.  $\square$

We now return to Theorem 10.1.

*Proof.* (Theorem 10.1) Without loss of generality, let  $\sigma^2 = 1$  (by scaling). Now, let  $B$  be a Brownian motion and a sequence of stopping times  $(T_n)_{n \in \mathbb{N}}$  as in Skorokhod's embedding theorem, on a possibly enlarged probability space such that

$$(B_{T_n})_{n \in \mathbb{N}} \stackrel{d}{=} (S_n)_{n \in \mathbb{N}}.$$

Now, define  $B_t^{(N)} = \sqrt{N} B_{\frac{t}{N}}$  a standard Brownian motion by scale invariance. Let  $(T_n^{(N)})_{n \in \mathbb{N}}$  be stopping times corresponding to  $B^{(N)}$  (again on a possibly enlarged probability space). Set  $S_n^{(N)} = B_{T_n^{(N)}}^{(N)}$  for all  $n \in \mathbb{N}$ . Let  $S_t^{(N)}$  be the linear interpolation of  $(S_n^{(N)})_{n \in \mathbb{N}}$ . Observe that we have

$$\left( (S_t^{(N)})_{t \geq 0}, (T_n^{(N)})_{n \in \mathbb{N}} \right) \stackrel{d}{=} \left( (S_t)_{t \geq 0}, (T_n)_{n \in \mathbb{N}} \right).$$

Now, we need to show

$$\mathbb{E} \left[ F \left( (S_t^{(N)})_{t \leq 1} \right) \right] \xrightarrow{N \rightarrow \infty} \mathbb{E} [F((B_t)_{t \leq 1})]$$

for all continuous and bounded functionals  $F : \mathcal{C}([0, \infty)) \rightarrow \mathbb{R}$ . It suffices to show that  $S^{(N)}$  converges uniformly in probability to  $B$ , that is

$$\mathbb{P} \left( \sup_{0 \leq t \leq 1} |\tilde{S}_t^{(N)} - B_t| > \epsilon \right) \xrightarrow{N \rightarrow \infty} 0$$

for all  $\epsilon > 0$ , by Dominated convergence.

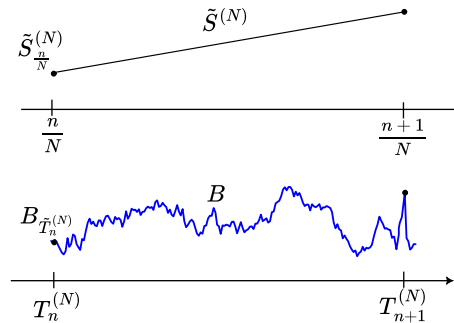
Now, for  $n \leq N$ ,

$$\tilde{S}_{n/N}^{(N)} = \frac{S^{(N)}}{\sqrt{N}} = \frac{B_{T_n^{(N)}}^{(N)}}{\sqrt{N}} = \frac{\sqrt{N} B_{T_n^{(N)}/N}^{(N)}}{\sqrt{N}} = B_{\tilde{T}_n^{(N)}}$$

where  $\tilde{T}_n^{(N)} = \frac{T_n^{(N)}}{N}$ . By the Strong law of large numbers, we have that  $\frac{T_n}{n} \xrightarrow{n \rightarrow \infty} 1$  a.s. (from Skorokhod embedding, the  $T_n$  are a random walk with independent and identically distributed steps). Thus,

$$\frac{1}{N} \cdot \sup_{n \in \mathbb{N}} |T_n - n| \xrightarrow{n \rightarrow \infty} 0, \quad \text{a.s.}$$

Since  $\tilde{S}_{n/N}^{(N)} = B_{\tilde{T}_n^{(N)}}$  for all  $n \leq N - 1$ , we claim that for all  $\frac{n}{N} \leq t \leq \frac{n+1}{N}$ , there exists  $\tilde{T}_n^{(N)} \leq u \leq \tilde{T}_{n+1}^{(N)}$  such that  $\tilde{S}_t^{(N)} = B_u$ . This follows from an application of the Implicit Function Theorem and using the continuity of  $B$ ,  $\tilde{S}$  and that  $\tilde{S}$  is piecewise linear.



So now we have

$$\begin{aligned} A &:= \left\{ |\tilde{S}_t^{(N)} - B_t| > \epsilon \text{ for some } t \in [0, 1] \right\} \\ &\subseteq \left\{ |\tilde{T}_{n/N}^{(N)} - \frac{n}{N}| \geq \delta \text{ for some } n \geq N \right\} := A_1 \\ &\cup \left\{ |B_t - B_n| > \epsilon \text{ for some } t \in [0, 1] \text{ and } |u - t| \leq \delta + \frac{1}{N} \right\} := A_2. \end{aligned}$$

Hence we have the bound  $\mathbb{P}(A) \leq \mathbb{P}(A_1) + \mathbb{P}(A_2)$ . Take  $N \geq 1/\delta$  and  $\delta > 0$  sufficiently small so that  $\mathbb{P}(A_2) < \epsilon/2$  since Brownian motion is uniformly continuous on  $[0, 1]$ .  $\square$

Lecture 24

## 11 Poisson random measures

Recall that  $X \sim Po(\lambda)$ ,  $\lambda > 0$  if  $\mathbb{P}(X = n) = e^{-\lambda} \frac{\lambda^n}{n!}$  for all  $n \in \mathbb{N}$ . If  $\lambda = 0$  set  $X \equiv 0$  and if  $\lambda = \infty$ , set  $X \equiv \infty$ . Also recall the following basic facts about Poisson random variables.

**Proposition 11.1 (Addition property).** Let  $(N_k)_{k \in \mathbb{N}}$  be independent Poisson  $N_k \sim Po(\lambda_k)$ ,  $\lambda_k > 0$  for all  $k \in \mathbb{N}$ . Then

$$\sum_{k=0}^{\infty} N_k \sim Po\left(\sum_{k=0}^{\infty} \lambda_k\right).$$

**Proposition 11.2 (Splitting property).** Let  $N \sim Po(\lambda)$ ,  $\lambda > 0$  and let  $(Y_n)_{n \in \mathbb{N}}$  be an iid sequence and independent of  $N$  with  $\mathbb{P}(Y_1 = j) = p_j$ ,  $j = 1 \dots, k$ . Set  $N_j = \sum_{n=1}^N \mathbf{1}(Y_n = j)$ . Then  $N_1, \dots, N_k$  are independent and  $N_j \sim Po(\lambda p_j)$ .

**Definition 11.1.** Let  $(E, \mathcal{E}, \mu)$  be a  $\sigma$ -finite measure space. A Poisson random measure with intensity  $\mu$   $M$  is a random map  $M : \Omega \times \mathcal{E} \rightarrow \mathbb{Z} \cup \{\infty\}$  such that if  $(A_k)_{k \in \mathbb{N}}$  is a disjoint collection in  $\mathcal{E}$ , then

1.  $M\left(\bigcup_{k \in \mathbb{N}} A_k\right) = \sum_{k=0}^{\infty} M(A_k)(\omega)$ , for all  $\omega \in \Omega$
2.  $(M(A_k))_{k \in \mathbb{N}}$  are independent random variables.

3. For all  $k \in \mathbb{N}$ ,  $M(A_k) \sim Po(\mu(A_k))$ .

Let  $E^* = \{\mathbb{Z}_+ \cup \{\infty\}\}$ -valued measures on  $(E, \mathcal{E})$ . Now for  $A \in \mathcal{E}$  define the maps

$$\begin{aligned} X &: E^* \times \mathcal{E} \rightarrow \mathbb{Z}_+ \cup \{\infty\} \\ X_A &: E^* \rightarrow \mathbb{Z}_+ \cup \{\infty\} \\ (m, A) &\mapsto X_A(m) := m(A). \end{aligned}$$

Furthermore, set  $\mathcal{E}^* = \sigma(X_A : A \in \mathcal{E})$ . We now can state the following existence (and uniqueness) theorem for Poisson random measures.

**Theorem 11.1.** There exists a unique probability measure  $\mu^*$  on  $(E^*, \mathcal{E}^*)$  such that under  $\mu^*$ ,  $X$  is a Poisson random measure of intensity  $\mu$ .

*Proof. Uniqueness:* Let  $A_1, \dots, A_k$  be disjoint in  $\mathcal{E}$  and  $n_1, \dots, n_k \in \mathbb{Z}_+$ . Set

$$A^* = \{m \in E^* : m(A_1) = n_1, \dots, m(A_k) = n_k\}.$$

Let  $\mu^*$  be as in the statement. Then compute

$$\mu^*(A^*) = \prod_{j=1}^k e^{-\mu(A_j)} \frac{(\mu(A_j))^{n_j}}{n_j!}.$$

But,  $A^*$  of the above form is a  $\pi$ -system that generates  $\mathcal{E}^*$ , so  $\mu^*$  is uniquely determined.

*Existence:* First assume  $\lambda = \mu(E) < \infty$ . Let  $N \sim Po(\lambda)$  and  $(Y_n)_{n \in \mathbb{N}}$  be an iid sequence independent of  $N$  with law  $\mu/\mu(E)$ . Set

$$M(A) = \sum_{n=1}^N \mathbf{1}(Y_n \in A), \quad A \in \mathcal{E}^*.$$

Let  $A_1, \dots, A_k$  be disjoint in  $\mathcal{E}$ . Need to show that  $M(A_i)_{i \leq k}$  are independent  $\sim Po(\mu(A_i))$  random variables. Consider  $X_n = j$  whenever  $Y_n \in A_j$ . The  $(X_n)_{n \leq N}$  are iid and  $M(A_j) = \sum_{n=1}^N \mathbf{1}(X_n = j)$ . By the splitting property 11.2, we get that  $M(A_1), \dots, M(A_k)$  are independent and  $M(A_j) \sim Po\left(\mu(E) \cdot \frac{\mu(A_j)}{\mu(E)}\right)$ .

If  $\mu(E) = \infty$ , let  $(E_k)_{k \in \mathbb{N}}$  be a partition of  $E$  into sets with  $\mu(E_k) < \infty$  for all  $k \in \mathbb{N}$ . Then on some probability space we can construct independent Poisson random measures  $M_k$  with intensity  $\mu|_{E_k}$  (on some suitable product space). Then for  $A \in \mathcal{E}$ , set

$$M(A) = \sum_{k=0}^{\infty} \overbrace{M_k(A \cap E_k)}^{\sim Po(\mu(A \cap E_k))}.$$

By the addition property 11.1,  $M(A) \sim Po\left(\sum_{k=0}^{\infty} \mu(A \cap E_k) = \mu(A)\right)$ . Independence follows since the  $(M_k)_{k \in \mathbb{N}}$  are PRM.

We have now constructed Poisson random measures on some probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Now simply observe that  $\mu^* = \mathbb{P}_M$  (the pushforward under of  $\mathbb{P}$  under  $M$ ) is the probability measure on  $(E^*, \mathcal{E}^*)$ . □

**Proposition 11.3.** Let  $M$  be a Poisson random measure with intensity  $\mu$ . Let  $A \in \mathcal{E}$  be such that  $\mu(A) < \infty$ . Then  $M(A) \sim Po(\mu(A))$  and conditional on  $M(A) = k$ , then we can express  $M = \sum_{i=1}^k \delta_{X_i}$ , where  $(X_1, \dots, X_k)$  are independent and identically distributed, with



law  $\frac{\mu(\cdot \cap A)}{\mu(A)}$ . Moreover, if  $A \cap B = \emptyset$ ,  $\mu|_A$  is independent of  $\mu|_B$ .

We leave the following as an exercise: let  $E = \mathbb{R}_+$ ,  $\theta > 0$ ,  $\mu = \theta \cdot \mathbf{1}(t \geq 0) dt$ . Let  $M$  be a PRM( $\mu$ ), let  $T_0 = 0$ ,  $(T_n - T_{n-1})_{n \geq 1}$  be iid  $\sim \text{Exp}(\theta)$ . Set

$$N_t = \sum_{n=1}^{\infty} \mathbf{1}(T_n \leq t).$$

Then,  $(N_t, t \geq 0) \stackrel{d}{=} (M([0, t]), t \geq 0)$ .

**Theorem 11.2.** Let  $M$  be a Poisson random measure with intensity  $\mu$ . Let  $f \in \mathcal{L}^1(\mu)$  and define  $M(f) = \int f(y)M(dy)$ . Then  $M(f) \in \mathcal{L}^1(\mu)$  and

$$\mathbb{E}[M(f)] = \int f(y)\mu(dy).$$

Fix  $f : E \rightarrow \mathbb{R}_+$  measurable. Then for all  $u > 0$ ,

$$\mathbb{E}\left[e^{-uM(f)}\right] = \exp\left(\int_E (e^{-uf(y)} - 1)\mu(dy)\right) \quad (\text{Campbell's formula})$$

*Proof.* The first part follows from a standard approximation by simple functions argument and Dominated Convergence. Let  $(E_n)_{n \in \mathbb{N}} \subseteq \mathcal{E}^*$  be such that  $\mu(E_n) < \infty$ . Have

$$\mathbb{E}\left[e^{-yM(f \cdot \mathbf{1}(E_n))}\right] = \sum_{k=0}^{\infty} \mathbb{E}\left[e^{-yM(f \cdot \mathbf{1}(E_n))} | M(E_n) = k\right] \cdot e^{-\mu(E_n)} \frac{\mu(E_n)^k}{k!}.$$

Now, given  $M(E_n) = k$ ,  $M = \sum_{i=1}^k \delta_{X_i}$  with  $(X_1, \dots, X_k)$  independent and each  $\sim \mu|_{E_n}$ , hence

$$\mathbb{E}\left[e^{-yM(f \cdot \mathbf{1}(E_n))}\right] = \sum_{k=0}^{\infty} \left(\int_{E_n} e^{-yf(x)} \frac{\mu(dx)}{\mu(E_n)}\right)^k \cdot e^{-\mu(E_n)} \frac{\mu(E_n)^k}{k!}$$

□

by independence and conclude with monotone convergence.