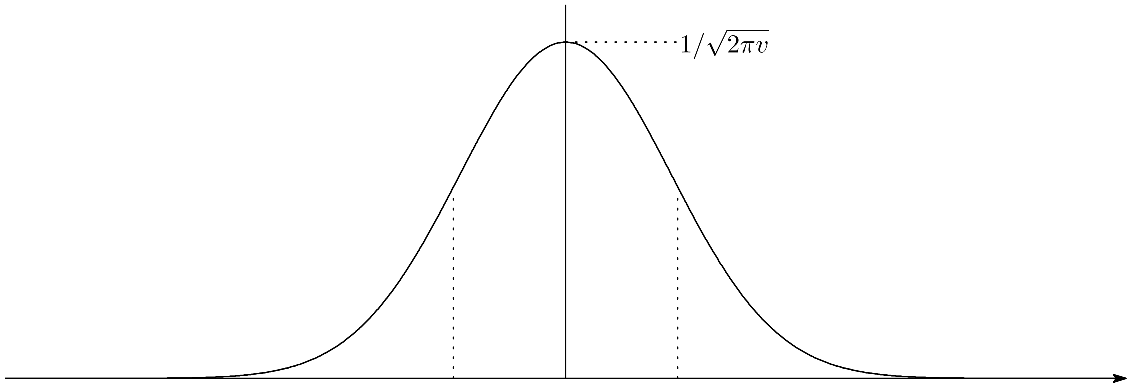
A Course in Probability[1]

Nobuo Yoshida[2]



[1]Notes for a course April 8, 2024.

[2][e-mail] noby@math.nagoya-u.ac.jp, [URL] http://www.math.nagoya-u.ac.jp/~noby

0 Introduction

The purpose of this course is to provide a quick and self-contained exposition of some basic notions and theorems in the probability theory. We try to get the feeling of "real world" probabilistic phenomena, rather than to learn a rigorous framework of "measure theoretical probability theory" (though we do use the measure theory as a convenient tool to describe the "real world").

We start by introducing the notion of independent random variables. Then, without too much preparations, we proceed to random walks, which will be the central topic of this course. Some interesting properties of random walks will be explained and proved. Classical theorems in the probability theory, like the law of large numbers and the central limit theorem, are presented in the context of random walks. We first show as an application of the law of large numbers, that the random walk travels along a constant velocity motion (including the case of zero velocity). We then see from the central limit theorem that the fluctuation around the constant velocity motion, if properly scaled in space and time, looks like a normally distributed random variable. Finally, we investigate a question whether or not the random walk comes back to its starting point with probability one, the answer to which depends on the dimension of the space.

If we have enough time, then we will also discuss Brownian motion.

Contents

0.1 Overview

To start with, we outline the content of this course.

• **Random variables**

Imagine a game such that its outcome is determined by chance, e.g., tossing a coin and seeing if it falls head or tail. Suppose that you play the game and that you record the outcome as follows;

$$X = \begin{cases} +1 & \text{if the coin falls head,} \\ -1 & \text{if the coin falls tail.} \end{cases} \tag{0.1}$$

The value X is not always the same (may be -1 for the first toss and $+1$ for the second) and hence is considered as a function $X : \Omega \to \{-1, +1\}$ on a suitable set Ω. Since one cannot predict the value X for sure, you may be interested in how large is the "probability" $P(X = \pm 1)$ of the "event" $\{\omega \in \Omega; X(\omega) = \pm 1\}$. In this overview, we temporarily adopt the following convention[3]:

- There is a set Ω and number $P(A) \in [0, 1]$ for each "measurable" $A \subset \Omega$. $P(A)$ is called the *probability* of the *event* A.
- A random quantity is described by a function

$$X : \Omega \to \mathbb{R}^d \ \ (\omega \mapsto X(\omega)) \tag{0.2}$$

such that

$$\{\omega \in \Omega \ ; \ X(\omega) \in I\} \text{ is measurable for all interval } I \subset \mathbb{R}^d. \tag{0.3}$$

A function with the above property is called a *random variable* (abbreviated as "r.v."). The above set $\{\omega \in \Omega \ ; \ X(\omega) \in I\}$ and its probability $P(\{\omega \in \Omega \ ; \ X(\omega) \in I\})$ are often denoted simply as $\{X \in I\}$ and $P(X \in I)$, respectively.

- For a r.v. $X : \Omega \to S$, where S is a finite subset and a function $f : S \to \mathbb{R}$, we define the *expectation* of $f(X)$ as:

$$Ef(X) = E[f(X)] = \sum_{s \in S} f(s) P(X = s). \tag{0.4}$$

• **Random walk**

Imagine that you walk "randomly" on $\mathbb{Z}^d$, the d-dimensional integer lattice. Let:

$$\begin{aligned} X_n &= \text{the displacement made at } n\text{-th step,} \\ S_n &= X_1 + ... + X_n = \text{the position at the } n\text{-th step} \end{aligned} \tag{0.5}$$

[3]Here, to keep the presentation as elementary as possible, we leave the notion of measurability ambiguous. We will discuss it in section 1.

We now describe how the random vectors $X_1, X_2, ...$ is determined. Let $e_1, ..., e_d$ be the canonical basis of $\mathbb{R}^d$, that is, $e_\alpha = (\delta_{\alpha,\beta})_{\beta=1}^d$. We introduce

$$\mathcal{E} = \bigcup_{\alpha=1}^{d} \{e_\alpha, -e_\alpha\} \subset \mathbb{Z}^d,$$
$$p : \mathcal{E} \to [0, 1), \ \sum_{e \in \mathcal{E}} p(e) = 1.$$

That is, $\mathcal{E}$ is the set of all nearest neighbors of the origin, and p is a probability distribution on $\mathcal{E}$. A typical example $p : \mathcal{E} \to [0, 1)$ is given by:

$$p(e) \equiv \frac{1}{2d}, \quad \forall e \in \mathcal{E}. \tag{0.6}$$

We suppose that $X_1, X_2, ...$ are determined by the following rule:

$$P\left(\bigcap_{j=1}^{n} \{X_j = x_j\}\right) = \prod_{j=1}^{n} p(x_j) \quad \text{for any } n \geq 1 \text{ and } x_1, ..., x_n \in \mathcal{E}. \tag{0.7}$$

In particular,

$$P(X_n = x) = p(x), \qquad \text{for any } n \geq 1 \text{ and } x \in \mathcal{E}. \tag{0.8}$$

Recall the standard notation of conditional probability:

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

Then, it follows from (0.7) that:

$$P\left(X_n = x_n \,\middle|\, \bigcap_{j=1}^{n-1} \{X_j = x_j\}\right) = P(X_n = x_n) \quad \text{for any } n \geq 1 \text{ and } x_1, ..., x_n \in \mathcal{E}. \tag{0.9}$$

We see from (0.9) that the values of $X_1, .., X_{n-1}$ have no influence on how X_n is determined. For this reason, $X_1, ..., X_n$ are said to be *independent*. For the moment, we call the sequence $(S_n)_{n \geq 1}$ defined by (0.5) a *random walk* (More general definition will be given later, cf. Definition 3.1.1.). In particular, the special case (0.6) will be called the *simple random walk*.

• **The law of large numbers**

We are interested in the behavior of the random walk S_n when $n \nearrow \infty$. Here is the first question we ask:

$$\text{Is there a particular direction in which the random walk prefers to travel?} \tag{0.10}$$

To investigate it, we introduce the following vector:

$$m = (m_\alpha)_{\alpha=1}^d \in \mathbb{R}^d, \quad m_\alpha = p(e_\alpha) - p(-e_\alpha). \tag{0.11}$$

If we write $X_n = (X_{n,\alpha})_{\alpha=1}^d$, we have

$$E[X_{n,\alpha}] = m_\alpha. \tag{0.12}$$

To see this, note that

$$X_{n,\alpha} = \begin{cases} \pm 1 & \text{if } X_n = \pm e_\alpha, \\ 0 & \text{if otherwise.} \end{cases} \tag{0.13}$$

Therefore,

$$E[X_{n,\alpha}] \stackrel{(0.4)}{=} 1 \cdot P(X_n = e_\alpha) + (-1)P(X_n = -e_\alpha) \stackrel{(0.8)}{=} p(e_\alpha) - p(-e_\alpha).$$

An answer to the question (0.10) is provided by:

Theorem 0.1.1 (Law of Large Numbers)

$$P\left(\frac{S_n}{n} \stackrel{n\to\infty}{\longrightarrow} m\right) = 1.$$

Let us decompose S_n in a silly expression:

$$S_n = nm + (S_n - nm).$$

Then Theorem 0.1.1 says that $S_n - nm$ is of order $o(n)$. In this sense, one can conclude that, if $n \to \infty$, then

$$S_n \text{ is close to } nm \text{ up to the random correction: } S_n - nm = o(n). \tag{0.14}$$

- **The central limit theorem**

Having understood (0.14), we proceed to address a further question.

$$\text{How does the correction term } S_n - nm \text{ look like?} \tag{0.15}$$

To investigate this question, we introduce the following $d \times d$ matrix:

$$V = (v_{\alpha,\beta})_{\alpha,\beta=1}^d, \quad v_{\alpha,\beta} = \delta_{\alpha,\beta}(p(e_\alpha) + p(-e_\alpha)) - m_\alpha m_\beta. \tag{0.16}$$

The component $v_{\alpha,\beta}$ stands for the *covariance* of $X_{n,\alpha}$ and $X_{n,\beta}$. Indeed, it follows from (0.13) that

$$X_{n,\alpha}X_{n,\beta} = \begin{cases} 1 & \text{if } \alpha = \beta \text{ and } X_n = \pm e_\alpha, \\ 0 & \text{if otherwise.} \end{cases}$$

This implies that

$$E[X_{n,\alpha}X_{n,\beta}] = \delta_{\alpha,\beta}(p(e_\alpha) + p(-e_\alpha)). \tag{0.17}$$

Therefore,

$$\begin{aligned} \text{cov}(X_{n,\alpha}X_{n,\beta}) &\stackrel{\text{def}}{=} E[X_{n,\alpha}X_{n,\beta}] - E[X_{n,\alpha}]E[X_{n,\beta}] \\ &\stackrel{(0.12),(0.17)}{=} \delta_{\alpha,\beta}(p(e_\alpha) + p(-e_\alpha)) - m_\alpha m_\beta. \end{aligned} \tag{0.18}$$

From here on, we will assume for simplicity that

$$p(\pm e_\alpha) > 0, \quad \forall \alpha = 1, ..., d. \tag{0.19}$$

Now, we list two facts, whose proofs are omitted here[4]:

$$\det V > 0, \tag{0.20}$$

$$\int_{\mathbb{R}^d} \rho_V = 1, \quad \text{where} \quad \rho_V(x) = \frac{1}{\sqrt{\det(2\pi V)}} \exp\left(-\tfrac{1}{2} x \cdot V^{-1} x\right). \tag{0.21}$$

The function ρ_V is the density of mean-zero *Gaussian distrubution* with the covariance matrix V (See Example 1.2.4 for more details).

We are now in position to state:

Theorem 0.1.2 (Central Limit Theorem)
For every interval $I \subset \mathbb{R}^d$,

$$P\left(\frac{S_n - nm}{\sqrt{n}} \in I\right) \stackrel{n\to\infty}{\longrightarrow} \int_I \rho_V. \tag{0.22}$$

To answer the question (0.15), we introduce a random variable Y with values in $\mathbb{R}^d$ such that

$$P(Y \in I) = \int_I \rho_V \quad \text{for every interval } I \subset \mathbb{R}^d.$$

By (0.22), for every interval $I \subset \mathbb{R}^d$,

$$P\left(\frac{S_n - nm}{\sqrt{n}} \in I\right) \text{ is close to } P(Y \in I) \text{ if } n \text{ is large enough.} \tag{0.23}$$

If we are allowed to replace I above by $I/\sqrt{n}$, we would be able to answer the question (0.15) in the following form:

$$P(S_n - nm \in I) \text{ is close to } P(\sqrt{n}Y \in I) \text{ if } n \text{ is large enough.} \tag{0.24}$$

Although, the replacement of I by $I/\sqrt{n}$ suggested above is not rigorous, the approximation (0.24) is known to be good enough for some applications, and is used in statistics.

- **Transience and recurrence**

Here, we take up a question whether a simple random walk $(S_n)_{n\geq 1}$ (cf. (3.3)) comes back to its starting point with probability one. Note that the simple random walk satisfies $m = 0$ (cf. (0.11)) and (0.19). We will prove the following

Theorem 0.1.3 *Suppose that $m = 0$ (cf. (0.11)) and (0.19), then,*

$$P\left(S_n = 0 \text{ for some } n \geq 1\right) \begin{cases} = 1 & d \leq 2, \\ < 1 & d \geq 3. \end{cases}$$

Theorem 0.1.3 is often explained with a joke:

" A drunk man will find his way home but a drunk bird may get lost forever".

[4] See Example 3.2.3 for a proof of (0.20).

0.2 Notations

For a set S,

2^S: the colloection of all subsets of S,
$\sigma(\mathcal{A})$: the σ-algebra generated by $\mathcal{A} \subset 2^S$, i.e., the smallest σ-algebra which contains $\mathcal{A}$.

For x and y in $\mathbb{R}$,

$x \vee y = \max\{x, y\}$,
$x \wedge y = \min\{x, y\}$.

For $x = (x_i)_{i=1}^d$ and $y = (y_i)_{i=1}^d$ in $\mathbb{R}^d$,

$x \cdot y = \sum_{i=1}^d x_i y_i$,
$|x| = (x \cdot x)^{1/2}$,
$\mathbf{e}_x(y) = \mathbf{e}_y(x) = \exp\left(\sqrt{-1} x \cdot y\right)$,

For a topological space S,

$C(S)$: the set of continuous functions on S
$C_{\mathrm{b}}(S)$: the set of bounded continuous functions on S
$C_{\mathrm{c}}(S)$: the set of continuous functions on S, which vanish outside a compact subset.
$\mathcal{B}(S)$: the Borel σ-algebra of S, i.e., the σ-algebra generated by all open subsets of S.

1 Independent Random Variables

1.1 Random Variables

The reader is supposed to be familiar with basics of the measure theory such as Lebesgue's monotone convergence theorem, Fatou's lemma, Lebesgue's dominated convergence theorem and Fubini's theorem. Nevertheless, we start by reviewing some basic terminology.

Definition 1.1.1 (Measurability)

▶ A couple $(S, \mathcal{B})$ is called a *measurable space* when S is a set and $\mathcal{B} \subset 2^S$ is a *σ-algebra*, i.e.,

S1) $S \in \mathcal{B}$.

S2) If $B \in \mathcal{B}$, then $B^{\mathsf{c}} \in \mathcal{B}$, where B^{c} denotes the complement of the set B.

S3) If $B_1, B_2, \ldots \in \mathcal{B}$, then $\cup_{n \geq 1} B_n \in \mathcal{B}$.

Let $(\Omega, \mathcal{F})$ and $(S, \mathcal{B})$ be measurable spaces.

▶ A map $X : \Omega \to S$ is said to be *measurable* if

$$\sigma[X] \stackrel{\text{def}}{=} \{X^{-1}(B) \ ; \ B \in \mathcal{B}\} \subset \mathcal{F}. \tag{1.1}$$

The σ-algebra $\sigma[X]$ is called *the σ-algebra generated by X*.

Example 1.1.2 (**The Borel σ-algebra**) When S is a topological space, we let $\mathcal{B}(S)$ denote the Borel σ-algebra of S, i.e., the smallest σ-algebra that contains all open subsets of S. In this course, S will usually be $\mathbb{R}^d$ or its Borel subset.

Definition 1.1.3 (Probability) Let $(S, \mathcal{B})$ a mesurable space and $\mu : \mathcal{B} \to [0, \infty]$ be a function.

▶ The function μ is called a *measure* when it satisfies

M1) $0 = \mu(\emptyset) \leq \mu(B)$ for all $B \in \mathcal{B}$,

M2) If $B_1, B_2, \ldots \in \mathcal{B}$ are disjoint, then $\mu\left(\cup_{n \geq 1} B_n\right) = \sum_{n \geq 1} \mu(B_n)$.

▶ A measure μ is called a *probability measure* when it satisfies

M3) $\mu(S) = 1$.

We introduce the following notation:

$$\mathcal{P}(S, \mathcal{B}) = \{\mu \ ; \ \mu \text{ is a probability measure on } (S, \mathcal{B})\}. \tag{1.2}$$

We abbreviate $\mathcal{P}(S, \mathcal{B})$ by $\mathcal{P}(S)$ when the choice of the σ-algebra $\mathcal{B}$ is obvious from the context.

▶ A triple $(S, \mathcal{B}, \mu)$ is called a *a measure space* if $(S, \mathcal{B})$ is a measurable space and μ is a measure on $(S, \mathcal{B})$.

▶ A measure space $(S, \mathcal{B}, \mu)$ is called a *probability space* if $\mu \in \mathcal{P}(S, \mathcal{B})$.

We already have a rough idea of the notion of random variable cf. (0.2)–(0.3). We now put it in more solid mathematical framework.

- For the rest of this subsection, let $(\Omega, \mathcal{F}, P)$ be a probability space, $(S, \mathcal{B})$ be a measurable space (cf. Definition 1.1.1, Definition 1.1.3), and $X : \Omega \to S$ be a map.

Definition 1.1.4 (Events and random variables)

▶ A set $A \subset \Omega$ is called an *event* if $A \in \mathcal{F}$.

▶ $X : \Omega \to S$ is called a *random variable* ("r.v." for short) if it is measurable (cf. Definition 1.1.1). The set S in this case is called the *state space* for the r.v. X.

▶ The *law* (or the *distribution*) of the r.v. X is a measure $\mu \in \mathcal{P}(S, \mathcal{B})$ defined by

$$\mu(B) = P(\{\omega \in \Omega \; ; \; X(\omega) \in B\}), \quad B \in \mathcal{B}. \tag{1.3}$$

We abbreviate the above relation of X and μ by

$$X \stackrel{\text{law}}{=} \mu \quad \text{or} \quad X \approx \mu \tag{1.4}$$

For another r.v. $X' : \Omega' \to S$ defined on a probability space (Ω', P'), we write

$$X \stackrel{\text{law}}{=} X' \quad \text{or} \quad X \approx X', \tag{1.5}$$

when X and X' share the same law.

Remark: Here are some remarks on the use of notation.

1) The set $\{\omega \in \Omega \; ; \; X(\omega) \in B\}$ will often be abbreviated by $\{X \in B\}$, and the right-hand side of (1.3) by $P(X \in B)$.

2) The law of a r.v. X, i.e., the measure defined by the right-hand side of (1.3) will often be denoted by $P(X \in \cdot)$.

Let a measurable space $(S, \mathcal{B})$ and a $\mu \in \mathcal{P}(S, \mathcal{B})$ be given. We look at a couple of examples in which a probability space $(\Omega, \mathcal{F}, P)$ and a r.v. $X \to S$ with $X \approx \mu$ are given.

Example 1.1.5 (**Identity map on the state space**) Let:

- $(\Omega, \mathcal{F}, P) = (S, \mathcal{B}, \mu)$, $X(\omega) = \omega$.

Then $\sigma[X] = \mathcal{F}$, and hence X is measurable. Moreover, $X \approx \mu$, since

$$P(X \in B) = \mu(\omega \; ; \; \omega \in B) = \mu(B) \text{ for any } B \in \mathcal{B}.$$

Example 1.1.6 (**Unit interval as a probability space**) Let:

- S=an at most countable set, $\mathcal{B} = 2^S$, $\mu \in \mathcal{P}(S, \mathcal{B})$.

We split $(0, 1]$ into disjoint intervals $\{I_s\}_{s \in S}$ with length $|I_s| = \mu(s)$ for each $s \in S$.

- $\Omega = (0, 1]$, $\mathcal{F} = \mathcal{B}((0, 1])$, P=the Lebesgue measure on $(0, 1]$,
- $X(\omega) = s$ if $\omega \in I_s$.

Then, X is measurable. In fact, for any $B \in \mathcal{B}$,

$$X^{-1}(B) = \bigcup_{s \in B} I_s \in \mathcal{F}.$$

Moreover, we see that $X \approx \mu$ as follows. First, for for any $s \in S$,

$$P(X = s) = P(\omega \in I_s) = |I_s| = \mu(s).$$

Then, for any $B \in \mathcal{B}$,

$$P(X \in B) = \sum_{s \in B} P(X = s) = \sum_{s \in B} \mu(s) = \mu(B).$$

Definition 1.1.7 (Expectation and (co)variance)

► For an $\mathbb{R}$-valued r.v. X, the integral $\int X dP$ is called the *expectation* or *mean* and is usually denoted by

$$EX,\ E(X) \text{ or } E[X]. \tag{1.6}$$

► For $X, Y \in L^1(P)$ such that $XY \in L^1(P)$, we define their *covariance* or *correlation* by

$$\begin{aligned} \text{cov}(X,Y) &= E\left((X - EX)(Y - EY)\right) \\ &= E(XY) - E(X)E(Y). \end{aligned} \tag{1.7}$$

In particular, $\text{cov}(X,X)$ is called the *variance* of X and is denoted by

$$\text{var}\ X \ \text{ or } \ \text{var}\,(X). \tag{1.8}$$

Remark: Notations (1.6) are also used to denote the expectations for complex or vector valued r.v.

Proposition 1.1.8 *Suppose that $X : \Omega \to S$ is a r.v. and that $\mu \in \mathcal{P}(S, \mathcal{B})$. Then, the following are equivalent:*

a) $X \approx \mu$.

b) *For a measurable function $f : S \to [0, \infty]$,*

$$Ef(X) = \int_S f d\mu. \tag{1.9}$$

Proof: a)⇒ b): By (1.3), the equality (1.9) is true for $f = \mathbf{1}_B$ with $B \in \mathcal{B}$. Thus, (1.9) is also true when f is a simple function [5]. Finally, for a measurable function $f : S \to [0, \infty]$, there is a sequence of simple functions f_n such that $f_n \nearrow f$. Thus, by the monotone convergence theorem,

$$Ef(X) = \lim_{n\to\infty} Ef_n(X) = \lim_{n\to\infty} \int_S f_n d\mu = \int_S f d\mu.$$

b)⇒ a): By setting $f = \mathbf{1}_B$ with $B \in \mathcal{B}$ in (1.9), we get (1.3). \(^□^)/

Remark: Suppose that $X \approx \mu$ in the setting of Proposition 1.1.8. Then, it follows from (1.9) that

$$f(X) \in L^1(P) \iff f \in L^1(\mu)$$

and that (1.9) holds true for $f \in L^1(\mu)$.

[5] A function of the form $\sum_{i=1}^n c_i \mathbf{1}_{B_i}$ ($c_i \in \mathbb{R}$, $B_i \in \mathcal{B}$) is called a simple function.

Proposition 1.1.9 (Chebyshev's inequality)

$$P(X \geq a) \leq \frac{EX}{a} \quad \textit{for a r.v. } X : \Omega \to [0,\infty) \textit{ and } a > 0. \tag{1.10}$$

Proof: It is obvious that

$$\mathbf{1}_{\{X \geq a\}} \leq \frac{X}{a}.$$

By taking the expectation of the both hand sides, we get the desired inequality. \\(^□^)/

Exercise 1.1.1 Prove that $\sigma[X]$ (cf (1.1)) is indeed a σ-algebra.

Exercise 1.1.2 Let $-\infty < a < b < \infty$ and suppose that $X \in L^1(P)$ satisfies $X \leq b$ a.s. Prove then that

$$P(X \leq a) \leq \frac{b - EX}{b - a}.$$

Exercise 1.1.3 Suppose that $f \in C^1([0,\infty) \to \mathbb{R})$ is non-decreasing. Use $f(x) - f(0) = \int_0^x f'(t)dt$ and Fubini's theorem to prove;

$$\int (f(x) - f(0))\mu(dx) = \int_0^\infty f'(t)\mu(x : x \geq t)dt$$

for a Borel measure μ on $[0,\infty)$. In particular, for a non-negative r.v. X,

$$Ef(X) = f(0) + \int_0^\infty f'(t)P(X \geq t)dt. \tag{1.11}$$

Exercise 1.1.4 Suppose that $f : \mathbb{N} \to \mathbb{R}$ is non-decreasing. Use $f(n) - f(0) = \sum_{j=1}^n (f(j) - f(j-1))$ and Fubini's theorem to prove that

$$\sum_{n\geq 1}(f(n) - f(0))\mu(n) = \sum_{n \geq 1}(f(n) - f(n-1))\mu(x : x \geq n)$$

for a measure μ on $\mathbb{N}$. In particular, for an $\mathbb{N}$-valued r.v. X,

$$Ef(X) = f(0) + \sum_{n \geq 1}(f(n) - f(n-1))P(X \geq n). \tag{1.12}$$

Exercise 1.1.5 Suppose that $X \in L^1(P)$. Prove then that for any $\varepsilon > 0$, there exists $\delta > 0$ such that $|E[X : A]| < \varepsilon$ for all $A \in \mathcal{F}$ with $P(A) < \delta$. **Hint**: Suppose the contrary. Then, for some $\varepsilon > 0$, there exist $A_n \in \mathcal{F}$, $n \in \mathbb{N}\backslash\{0\}$ such that $P(A_n) < 1/n$ and $|E[X : A_n]| \geq \varepsilon$.]

Exercise 1.1.6 Suppose that X is a r.v. with values in $\mathbb{N} \cup \{\infty\}$ and set $f(s) = E[s^X : X < \infty]$ for $s \in (0,1)$. **(i)** Show that $f'(s) \stackrel{s \to 1}{\longrightarrow} E[X : X < \infty]$, including the possibility that the limit diverges. **Hint**: The monotone convergence theorem. **(ii)** Generalize (i) to the case where X takes values in $(\mathbb{Z} \cap [-m,\infty)) \cup \{\infty\}$ for some $m \in \mathbb{N}$.

Exercise 1.1.7 (**Inclusion-exclusion formula**) Let $A_1, ..., A_n$ be events. Prove then that

$$P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n}(-1)^{k-1} \sum_{1 \le i_1 < ... < i_k \le n} P\left(A_{i_1} \cap ... \cap A_{i_k}\right). \tag{1.13}$$

Hint: Let $A_0 = \bigcup_{k=1}^{n} A_k$, $\chi_j = \mathbf{1}_{A_j}$ $(j = 0, 1, ..., n)$ and $\sigma_{n,k} = \sum_{1 \le i_1 < ... < i_k \le n} \chi_{i_1} \cdots \chi_{i_k}$ $(k = 1, ..., n)$. Then,

$$\chi_0 = 1 - \prod_{k=1}^{n}(1 - \chi_k) = \sum_{k=1}^{n}(-1)^{k-1}\sigma_{n,k}. \tag{1.14}$$

Exercise 1.1.8 (**Bonferroni inequalities**) Let $1 \le m \le n-1$. Then, the following variants of (1.13) hold:

$$P\left(\bigcup_{k=1}^{n} A_k\right) \left\{ \begin{array}{c} \le \\ \ge \end{array} \right\} \sum_{k=1}^{m}(-1)^{k-1} \sum_{1 \le i_1 < ... < i_k \le n} P\left(A_{i_1} \cap ... \cap A_{i_k}\right) \left\{ \begin{array}{l} \text{if } m \text{ is odd,} \\ \text{if } m \text{ is even.} \end{array} \right.$$

Prove these inequalities by going through the following steps (i)–(ii).

(i) $\sum_{k=m+1}^{n}(-1)^{k-1}\binom{n}{k} = \sum_{k=0}^{m}(-1)^{k}\binom{n}{k} \left\{ \begin{array}{ll} \le 0 & \text{if } m \text{ is odd,} \\ \ge 0 & \text{if } m \text{ is even.} \end{array} \right.$ **Hint**: The first equality is nothing but $(1-1)^n = 0$. To prove the inequality for $m \le n/2$, note that $k \mapsto \binom{n}{k}$ is increasing for $k \le n/2$. Then, use $(1-1)^n = 0$ again and the symmetry $\binom{n}{n-k} = \binom{n}{k}$ to take care of the case $m \ge n/2$.

(ii) The following variant of (1.14) holds: $\chi_0 \left\{ \begin{array}{c} \le \\ \ge \end{array} \right\} \sum_{k=1}^{m}(-1)^{k-1}\sigma_{n,k} \left\{ \begin{array}{l} \text{if } m \text{ is odd,} \\ \text{if } m \text{ is even.} \end{array} \right.$ **Hint**: Let $\ell = \sum_{k=1}^{n} \chi_k$. Then, $\sigma_{n,k} = \binom{\ell}{k}\mathbf{1}_{k \le \ell}$. Combine (1.14) with this observation and (i) to see that $\chi_0 - \sum_{k=1}^{m}(-1)^{k-1}\sigma_{n,k} = \sum_{k=m+1}^{\ell}(-1)^{k-1}\binom{\ell}{k} \left\{ \begin{array}{ll} \le 0 & \text{if } m \text{ is odd,} \\ \ge 0 & \text{if } m \text{ is even.} \end{array} \right.$.

Exercise 1.1.9 (**Payley-Zygumund inequality**) Let $X \in L^2(P)$, $m \stackrel{\text{def}}{=} EX > 0$. Prove then that $P(X > cm) \ge \frac{(1-c)^2 m^2}{\operatorname{var} X + (1-c)^2 m^2}$ for $c \in [0, 1)$. **Hint**: Let $Y = X/EX$. Then, $1 - c = E[Y - c] \le E[(Y - c)\mathbf{1}_{\{Y > c\}}]$, and hence $(1-c)^2 \le E[(Y-c)^2]P(Y > c)$.

Exercise 1.1.10 Let S be a real d-dimenisonal vector space equipped with an inner product $x \cdot y$, $(x, y \in S)$, and let $\{u_\alpha\}_{\alpha=1}^{d} \subset S$ and $\{v_\alpha\}_{\alpha=1}^{d} \subset S$ be respectively orthonormal systems. Prove then the following. **(i)** $(u_\alpha \cdot v_\beta)_{\alpha,\beta=1}^{d} \in \mathcal{O}_d$, where $\mathcal{O}_d$ denotes the totality of $d \times d$ real orthogonal matrices. **(ii)** Let $X = (X_\alpha)_{\alpha=1}^{d}$ be an $\mathbb{R}^d$ valued r.v. such that $UX \approx X$ for all $U \in \mathcal{O}_d$. Then, $\sum_{\alpha=1}^{d} X_\alpha u_\alpha \approx \sum_{\alpha=1}^{d} X_\alpha v_\alpha$.

Definition 1.1.10 (Conditional probability) Let $(\Omega, \mathcal{F}, P)$ be a probability space. If $B \in \mathcal{F}$ and $P(B) > 0$, then the *conditional probability given* B is defined by

$$P(A|B) = P(A \cap B)/P(B), \quad A \in \mathcal{F}. \tag{1.15}$$

Exercise 1.1.11 Suppose that $B = \sum_{i=1}^{n} B_i$, where $B_i \in \mathcal{F}$ and $P(B_i) > 0$. Prove then that $P(A|B) = \sum_{i=1}^{n} P(A|B_i)P(B_i|B)$ for any $A \in \mathcal{F}$.

1.2 Examples

Example 1.2.1 (**Uniform distribution**) Let $-\infty < a < b < \infty$ and $I = (a, b) \subset \mathbb{R}$.
▶ A r.v. $U : \Omega \to I$ is said to be a *uniform r.v.* on I if

$$P(U \in B) = \tfrac{1}{b-a} \int_B dt \text{ for all } B \in \mathcal{B}(I). \tag{1.16}$$

The law of U is called the *uniform distribution* on I. One can easily verify (Exercise 1.2.1) that

$$EU = (a+b)/2, \quad \text{var } U = (b-a)^2/12. \tag{1.17}$$

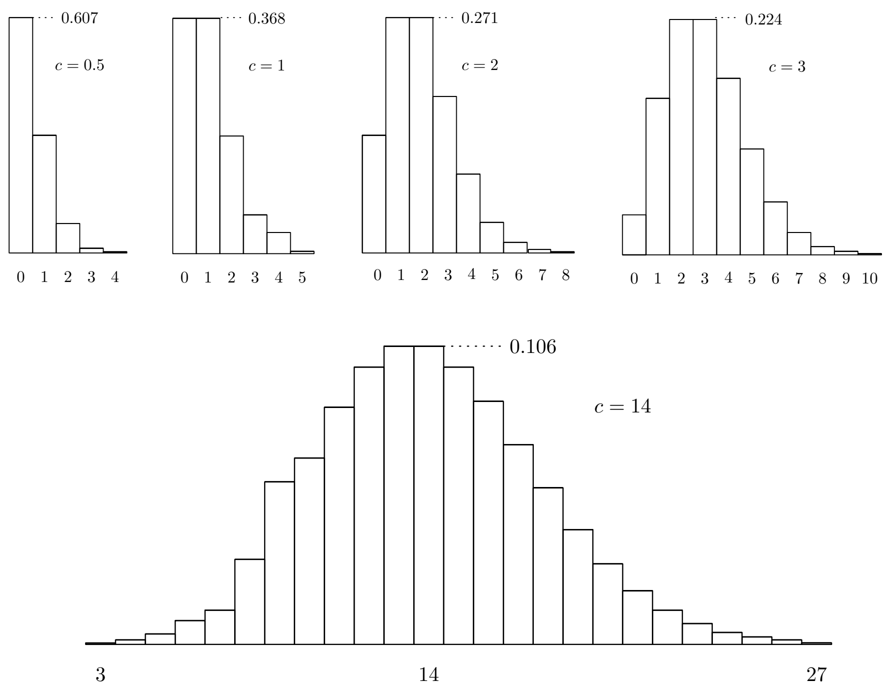
Example 1.2.2 (**Poisson distribution**) Let $c \geq 0$.
▶ A r.v. $N : \Omega \to \mathbb{N}$ is called a *c-Poisson r.v.* if

$$P(N \in B) = \pi_c(B) \stackrel{\text{def.}}{=} \sum_{n \in B} \frac{e^{-c} c^n}{n!}, \quad B \subset \mathbb{N}. \tag{1.18}$$

A probability measure π_c defined above is called *c-Poisson distribution.* It is not hard to see (Exercise 1.2.2) that

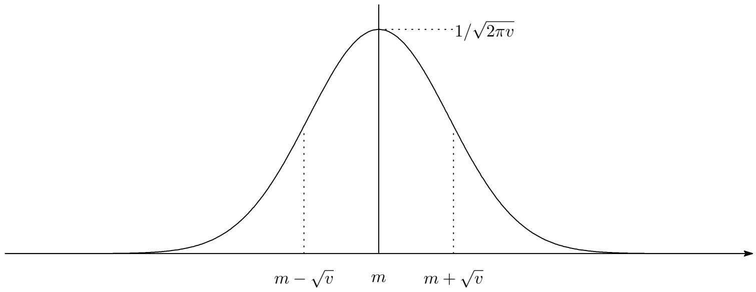
$$EN = \text{var } N = c. \tag{1.19}$$

Here are some pictures of how the function $\frac{e^{-c}c^n}{n!}$ $(n = 0, 1, 2, ...)$ looks like.



Example 1.2.3 (**Gaussian distribution; one dimension**) Let $m \in \mathbb{R}$ and $v > 0$.
▶ A r.v. $X : \Omega \to \mathbb{R}$ is called a (m, v)-*Gaussian* (or *normal*) r.v. if

$$P(X \in B) = \frac{1}{\sqrt{2\pi v}} \int_B \exp\left(-\frac{(x-m)^2}{2v} \right) dx \quad \text{for } B \in \mathcal{B}(\mathbb{R}). \tag{1.20}$$



The law of an (m, v)-Gaussian r.v. is denoted by $N(m, v)$. In particular, $N(0, 1)$ is called the *standard Gaussian* (or *standard normal*) distribution. $N(m, v)$ and $N(0, 1)$ is related as follows.

$$Y \approx N(0,1) \iff X \stackrel{\text{def}}{=} m + \sqrt{v}Y \approx N(m, v). \tag{1.21}$$

To prove $(\Rightarrow)$, we take a measurable $f : \mathbb{R} \to [0, \infty)$ and compute:

$$\begin{aligned} Ef(X) &\stackrel{(1.9)}{=} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(m + \sqrt{v}y) \exp\left(-\tfrac{1}{2}y^2\right) dy \\ &\stackrel{x=m+\sqrt{v}y}{=} \frac{1}{\sqrt{2\pi v}} \int_{\mathbb{R}} f(x) \exp\left(-\frac{(x-m)^2}{2v}\right) dx. \end{aligned}$$

This proves $(\Rightarrow)$ of (1.21). The converse can be proved similarly.
Next, let us verify that

$$X \approx N(m, v) \implies EX = m, \quad \operatorname{var} X = v. \tag{1.22}$$

By (1.21), this boils down to the case of $(m, v) = (0, 1)$, where we have that:

$$\begin{aligned} EX &\stackrel{(1.9)}{=} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x \exp\left(-\tfrac{1}{2}x^2\right) dx = 0, \\ \operatorname{var} X &\stackrel{(1.9)}{=} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 \exp\left(-\tfrac{1}{2}x^2\right) dx = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} x^2 \exp\left(-\tfrac{1}{2}x^2\right) dx \\ &\stackrel{x=\sqrt{2y}}{=} \frac{2}{\sqrt{2\pi}} \int_0^{\infty} 2y \exp(-y) \frac{\sqrt{2}}{2} y^{-1/2} dy = \frac{2}{\sqrt{\pi}} \Gamma(3/2). \end{aligned}$$

Here, we have introdued the Gamma function as usual:

$$\Gamma(a) = \int_0^{\infty} x^{a-1} e^{-x} dx, \quad a \in \mathbb{C}, \ \mathrm{Re}(a) > 0. \tag{1.23}$$

Recall that $\Gamma(a+1) = a\Gamma(a)$ and that $\Gamma(1/2) = \sqrt{\pi}$. Hence,

$$\operatorname{var} X = \frac{2}{\sqrt{\pi}} \Gamma(3/2) = \frac{1}{\sqrt{\pi}} \Gamma(1/2) = 1.$$

\(^□^)/

Example 1.2.4 (**Gaussian distribution; higher dimensions**) Let $m \in \mathbb{R}^d$, and V be a symmetric, strictly positive definite $d \times d$-matrix.

► A r.v. $X : \Omega \to \mathbb{R}^d$ is called a (m, V)-*Gaussian* (or *normal*) r.v. if

$$P(X \in B) = \frac{1}{\sqrt{\det(2\pi V)}} \int_B \exp\left(-\tfrac{1}{2}(x-m)\cdot V^{-1}(x-m)\right) dx \quad \text{for } B \in \mathcal{B}(\mathbb{R}^d). \tag{1.24}$$

The law of an (m, V)-Gaussian r.v. is denoted by $N(m, V)$. (See also Example 2.2.4 for the case where the matrix V may degenerate.) When $m = 0$ and V is the identity matrix I_d, $N(0, I_d)$ is called the *standard normal* (or *standard Gaussian*) distribution.

Let A be a $d \times d$ matrix, not necessarily symmetric, such that $V = AA^*$. See Proposition 8.2.4 for a characterization of such A for a given V. Now, $N(m, V)$ and $N(0, I_d)$ is related as:

$$Y \approx N(0, I_d) \iff X \stackrel{\text{def}}{=} m + AY \approx N(m, V). \tag{1.25}$$

The proof goes similarly as that of (1.21). To prove $(\Rightarrow)$, we take a measurable $f : \mathbb{R}^d \to [0, \infty)$ and write

1) $$Ef(X) = Ef(m + AY) \stackrel{(1.9)}{=} \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(m + Ay) \exp\left(-\tfrac{1}{2}|y|^2\right) dy.$$

We rewrite the integral on the right-hand side of 1) in terms of the new variable $x \stackrel{\text{def}}{=} m + Ay$. We first compute the Jacobian of the transformation $x \to y$. Since

$$y = A^{-1}(x - m) \text{ and } (\det A)^2 = \det A \det A^* = \det V,$$

we have

2) $$\left|\det\left(\frac{\partial y_\alpha}{\partial x_\beta}\right)_{\alpha,\beta=1}^d\right| = |\det(A^{-1})| = \frac{1}{|\det A|} = \frac{1}{\sqrt{\det V}}.$$

Next, we express $|y|^2$ in terms of the variable x. We have

$$|y|^2 = |A^{-1}(x-m)|^2 = A^{-1}(x-m)\cdot A^{-1}(x-m) = (x-m)\cdot (A^{-1})^* A^{-1}(x-m)$$

and

$$(A^{-1})^* A^{-1} = (A^*)^{-1}A^{-1} = (AA^*)^{-1} = V^{-1}.$$

Therefore,

3) $$|y|^2 = (x-m)\cdot V^{-1}(x-m).$$

By 1),2) and 3), we obtain

$$Ef(X) = \frac{1}{\sqrt{\det(2\pi V)}} \int_{\mathbb{R}^d} f(x) \exp\left(-\tfrac{1}{2}(x-m)\cdot V^{-1}(x-m)\right) dx.$$

This proves $(\Rightarrow)$ of (1.25). The converse can be proved similarly.

The relation (1.25) can be used to verify (Exercise 1.2.5) that

$$m = (EX_\alpha)_{\alpha=1}^d, \quad V = (\text{cov}(X_\alpha, X_\beta))_{\alpha,\beta=1}^d. \tag{1.26}$$

Example 1.2.5 (**Gamma, exponential, and χ^2 distributions**) Let $a, c > 0$.

▶ We define *(c, a)-gamma distribution* $\gamma_{c,a} \in \mathcal{P}((0,\infty))$ by

$$\gamma_{c,a}(B) = \frac{c^a}{\Gamma(a)} \int_B x^{a-1} e^{-cx} dx, \quad \text{for } B \in \mathcal{B}((0,\infty)), \tag{1.27}$$

$\gamma_{c,a}$ is also denoted by $\gamma(c, a)$. There are two important special cases of $\gamma_{c,a}$:

▶ $\gamma_{c,1}$ is called the *c-exponential distribution.*

▶ $\gamma_{1/2,d/2}$ ($d \in \mathbb{N}\backslash\{0\}$) is called the *$\chi_d^2$-distribution.*

For a r.v. $X \approx \gamma_{c,a}$, we easily see that

$$E[X^p] = c^{-p} \frac{\Gamma(p+a)}{\Gamma(a)}, \quad p > -a. \tag{1.28}$$

Indeed, since

$$\frac{c^{p+a}}{\Gamma(p+a)} \int_0^\infty x^{p+a-1} e^{-cx} dx = 1,$$

we have

$$E[X^p] = \frac{c^a}{\Gamma(a)} \int_0^\infty x^{p+a-1} e^{-cx} dx = \frac{c^a}{\Gamma(a)} \frac{\Gamma(p+a)}{c^{p+a}} = c^{-p} \frac{\Gamma(p+a)}{\Gamma(a)}.$$

It follows from (1.28) that

$$EX = a/c, \quad \text{var}\, X = a/c^2. \tag{1.29}$$

Example 1.2.6 (**Square of a Gaussian r.v.**) Let $v > 0$. Then,

$$X \approx N(0, vI_d) \implies |X|^2 \approx \gamma\left(\tfrac{1}{2v}, \tfrac{d}{2}\right). \tag{1.30}$$

In particular,

- $v = 1 \;\Rightarrow\; |X|^2 \approx \chi_d^2$;
- $d = 2 \;\Rightarrow\; |X|^2 \approx \frac{1}{2v}$-exponentail distribution.

To prove (1.30), let $f : [0,\infty) \to [0,\infty)$ be measurable. We compute by the polar coordinate transformation. Let $A_d = 2\pi^{\frac{d}{2}}/\Gamma(\frac{d}{2})$ (the area of the unit sphere in $\mathbb{R}^d$). Then,

$$\begin{aligned} Ef(|X|^2) &\overset{(1.9)}{=} \frac{1}{(2\pi v)^{d/2}} \int_{\mathbb{R}^d} f(|x|^2) \exp\left(-\tfrac{|x|^2}{2v}\right) dx \\ &= \frac{A_d}{(2\pi v)^{d/2}} \int_0^\infty f(r^2) r^{d-1} \exp\left(-\tfrac{r^2}{2v}\right) dr \\ &\overset{s=r^2}{=} \left(\tfrac{1}{2v}\right)^{\frac{d}{2}} \frac{1}{\Gamma(\frac{d}{2})} \int_0^\infty f(s) s^{\frac{d}{2}-1} \exp\left(-\tfrac{s}{2v}\right) ds = \int_0^\infty f d\gamma_{\frac{1}{2v},\frac{d}{2}}. \end{aligned}$$

This proves the relation (1.30). This relation can be combined with (1.28) to verify that

$$E[|X|^p] = (2v)^{p/2} \frac{\Gamma(\frac{p+d}{2})}{\Gamma(\frac{d}{2})}. \tag{1.31}$$

Indeed,

$$E[|X|^p] = E[(|X|^2)^{p/2}] \overset{(1.28),\,(1.30)}{=} \left(\frac{1}{2v}\right)^{-p/2} \frac{\Gamma(\frac{p+d}{2})}{\Gamma(\frac{d}{2})} = (2v)^{p/2} \frac{\Gamma(\frac{p+d}{2})}{\Gamma(\frac{d}{2})}.$$

Example 1.2.7 (**Beta distribution**) We define the Beta function as usual:

$$B(a,b) = \int_{(0,1)} x^{a-1}(1-x)^{b-1}dx, \quad a,b > 0. \tag{1.32}$$

We define (a,b)-*beta distribution* $\beta_{a,b} \in \mathcal{P}((0,1))$ by

$$\beta_{a,b}(B) = \frac{1}{B(a,b)} \int_B x^{a-1}(1-x)^{b-1}dx \quad \text{for } B \in \mathcal{B}((0,1)) \tag{1.33}$$

$\beta_{a,b}$ are also denoted by $\beta(a,b)$. For a r.v.$Y \approx \beta_{a,b}$, we have that

$$EY = \frac{a}{a+b}, \quad \operatorname{var} Y = \frac{ab}{(a+b)^2(a+b+1)}, \quad \text{cf. Exercise 1.2.11.} \tag{1.34}$$

There are two important special cases:

► $\beta_{1,1}$ is the uniform distribution on $(0,1)$.
► $\beta_{1/2,1/2}$ is called the *arcsin law.* Since $B(\frac{1}{2},\frac{1}{2}) = \pi$, the arcsin law has the density $\frac{1}{\pi}\frac{1}{\sqrt{x(1-x)}}$ on $(0,1)$. To explain why $\beta_{1/2,1/2}$ is called the arcsin law, let Y be a r.v. with values in $(-1,1)$ such that for $-1 \le a \le b \le 1$,

$$P(a < Y \le b) = \frac{2}{\pi}\int_a^b \frac{dx}{\sqrt{1-x^2}} = \frac{2}{\pi}(\text{Arcsin } b - \text{Arcsin } a).$$

Then, $Y^2 \approx \beta(\frac{1}{2},\frac{1}{2})$ as is easily verified. In this respect, it would be more correct to call $\beta(\frac{1}{2},\frac{1}{2})$ the "squared arcsin law" rather than the arcsin law.

Example 1.2.8 (**Cauchy distribution, $\mathbf{T_n}$-distribution**)
► Let $a, c > 0$. We define *the generalized Cauchy distribution* $\mu_{c,a} \in \mathcal{P}(\mathbb{R}^d)$ by:

$$\mu_{c,a}(B) = \frac{c^{2a}\Gamma(\frac{d}{2}+a)}{\pi^{d/2}\Gamma(a)} \int_B \frac{dx}{(c^2+|x|^2)^{\frac{d}{2}+a}}, \quad B \in \mathcal{B}(\mathbb{R}^d). \tag{1.35}$$

We will see in Exercise 1.2.13 below that

$$\frac{c^{2a}\Gamma(\frac{d}{2}+a)}{\pi^{d/2}\Gamma(a)} \int_{\mathbb{R}^d} \frac{dx}{(c^2+|x|^2)^{\frac{d}{2}+a}} = 1. \tag{1.36}$$

There are two important special cases:

► $\mu_{c,1/2}$ is called the (c)-*Cauchy distribution.* For $d = 1$ and $B = [a,b]$, one can compute:

$$\mu_{c,1/2}([a,b]) = \frac{c}{\pi}\int_a^b \frac{dx}{c^2+x^2} = \frac{1}{\pi}\left(\text{Arctan}\frac{b}{c} - \text{Arctan}\frac{a}{c}\right).$$

► For $d = 1$ and $n \in \mathbb{N}$, $\mu_{n/2,n/2}$ is called the T_n-*distribution* and used in statistics.

Exercise 1.2.1 Verify (1.17).

Exercise 1.2.2 Verify (1.19).

Exercise 1.2.3 Let $X : \Omega \to \mathbb{N}$ be a (c)-Poisson r.v. Prove then that for $n \in \mathbb{N}$,

$$P(X = 2n|X \text{ is even}) = \frac{1}{\cosh c}\frac{c^{2n}}{(2n)!}, \quad P(X = 2n+1|X \text{ is odd}) = \frac{1}{\sinh c}\frac{c^{2n+1}}{(2n+1)!}.$$

Exercise 1.2.4 Let X be a r.v. $\approx N(0,1)$ and $x > 0$. Then, prove that

$$\frac{1}{\sqrt{2\pi}}\left(\frac{1}{x} - \frac{1}{x^3}\right)\exp(-x^2/2) \le P(X > x) \le \frac{1}{\sqrt{2\pi}}\frac{1}{x}\exp(-x^2/2). \tag{1.37}$$

Hint: $\int_x^\infty \exp(-y^2/2)dy = x^{-1}\exp(-x^2/2) - \int_x^\infty y^{-2}\exp(-y^2/2)dy$.

Exercise 1.2.5 Verify (1.26). Hint: First, consider the case of for $N(0, I_d)$, where (1.22) and Fubini's theorem can be used. Then, use (1.25) to settle the general case.

Exercise 1.2.6 Let X be a positive r.v. Prove then that the following conditions are equivalent. **(a)** $\exists c \in (0,\infty)$, $X \approx \gamma_{c,1}$. **(b)** $P(X > t+s|X > s) = P(X > t) > 0$ for any $t, s \ge 0$. (The property (b) is referred to as the "*memoryless property*".)

Exercise 1.2.7 Suppose that two positive r.v's X, U are related as $U = \exp(-cX)$ $(c > 0)$. Prove then that U is uniformly distributed on $(0,1)$ if and only if $X \approx \gamma(c,1)$.

Exercise 1.2.8 Let $X \approx \gamma_{c,a}$. Prove then that **(i)** $X/r \approx \gamma_{rc,a}$ for $r > 0$.
(ii) $X^p \approx \dfrac{c^a}{|p|\Gamma(a)} x^{\frac{a}{p}-1}\exp(-cx^{\frac{1}{p}})dx$ for $p \in \mathbb{R}\backslash\{0\}$.

Exercise 1.2.9 $(\star)$ (Preparation for Exercise 1.2.10) Let $h_2(r) = \log r$ $(r > 0)$, $h_d(r) = r^{2-d}$ $(d \ge 3,\ r > 0)$, and $e_1 = (1,0,\ldots,0) \in \mathbb{R}^d$. Let also σ_d be the surface measure on S^{d-1}, so that $A_d \stackrel{\text{def}}{=} \sigma(S^{d-1}) = 2\pi^{d/2}/\Gamma(d/2)$. Prove the following. **(i)** The fnction, $u \mapsto \sup_{r>0} h_d(|e_1 + ru|)$ is integrable on S^{d-1} with respect to σ_d. **Hint**: $|e_1 + ru|^2 \ge (1 \wedge r^2)|e_1 + u|^2 + (r-1)^2$ for $u \in S^{d-1}$. **(ii)** $\int_{S^{d-1}} h_d(|e_1 + ru|)d\sigma_d(u) = A_d h_d(r \vee 1)$. **Hint**: Start with the case of $r \in (0,1)$, noting that h_d is harmonic on $\mathbb{R}^d\backslash\{0\}$.

Exercise 1.2.10 $(\star)$ Let $d \ge 2$ be an integer and $g : [0,\infty) \to [0,\infty)$ be locally bounded, measurable, such that $\gamma_d \stackrel{\text{def}}{=} \int_0^\infty r^{d-1}g(r)dr < \infty$. We consider an $\mathbb{R}^d$-valued r.v. $X \approx \frac{1}{\gamma_d A_d}g(|x|)dx$, where $A_d = 2\pi^{d/2}/\Gamma(d/2)$, the area of the unit sphere in $\mathbb{R}^d$. Using the polar coordinate transform and Exercise 1.2.9, prove the following identities for $m \in \mathbb{R}^d$, $c > 0$. For $d = 2$,

$$E\log|m + cX| = \begin{cases} \log c + \frac{1}{\gamma_2}\int_0^\infty r\log r\exp\left(-\frac{r^2}{2}\right)dr, & (m = 0), \\ \log|m| + \frac{1}{\gamma_2}\int_{|m|/c}^\infty \gamma_2(r)r^{-1}dr, & (m \ne 0). \end{cases}$$

where $\gamma_2(r) = \int_r^\infty ug(u)du$ $(r \ge 0)$. For $d \ge 3$,

$$E[|m + cX|^{2-d}] = \begin{cases} c^{2-d}\gamma_2/\gamma_d, & (m = 0), \\ c^{2-d}(d-2)\frac{1}{\gamma_d}\int_0^{|m|/c}\gamma_2(r)r^{d-3}dr, & (m \ne 0). \end{cases}$$

Remark: The special case $g(r) = \exp(-r^2/2)$ is of particular interest, where $X \approx N(0, I_d)$, $\gamma_d = 2^{\frac{d}{2}-1}\Gamma(\frac{d}{2})$ and $\gamma_2(r) = g(r) = \exp(-r^2/2)$.

Exercise 1.2.11 Verify (1.34).

Exercise 1.2.12 ($\star$) Prove that $\beta_{k,n-k+1}((0,p]) = \sum_{r=k}^{n} \binom{n}{r} p^r(1-p)^{n-r}$ for $p \in [0,1]$ and $1 \le k \le n$. **Hint**: Induction on k.

Exercise 1.2.13 Prove that $\int_0^\infty \frac{r^{a-1}dr}{(1+r^c)^b} = \frac{\Gamma(b-\frac{a}{c})\Gamma(\frac{a}{c})}{c\Gamma(b)}$ for $a,b,c > 0$ such that $bc > a$. Then, use this to see (1.36).

Exercise 1.2.14 Let X be a r.v. with (c)-Cauchy distribution. Then, prove that $\frac{c^2}{c^2+X^2} \approx \beta(\frac{1}{2},\frac{1}{2})$, the arcsin law.

Exercise 1.2.15 Let U be a r.v. with uniform distribution on $(-\frac{\pi}{2},\frac{\pi}{2})$. Then, prove the following. **(i)** $P(\sin U \in B) = \frac{2}{\pi}\int_B \frac{dx}{\sqrt{1-x^2}}$ for $B \in \mathcal{B}((-1,1))$. **(ii)** $\sin^2 U \approx \cos^2 U \approx \beta(\frac{1}{2},\frac{1}{2})$, the arcsin law. **(iii)** $c\tan U \approx (c)$-Cauchy distribution on $\mathbb{R}$ $(c>0)$.

Exercise 1.2.16 Suppose that Y is a r.v. with (1)-Cauchy distribution. Prove the following. **(i)** For $c > 0$, $X \stackrel{\text{def}}{=} c\log|Y| \approx \frac{2}{c\pi}\cosh(x/c)^{-1}dx$. **(ii)** $E[|X|^{s-1}] = \frac{4c^{s-1}}{\pi}\Gamma(s)\sum_{n=0}^\infty \frac{(-1)^n}{(2n+1)^s}$ $(\forall s \in (1,\infty))$.

1.3 When Do Two Measures Coincide?

In this subsection, we take up a question as follows; Let μ_1 and μ_2 be probability measures on a measurable space $(S,\mathcal{B})$, $\mathcal{A} \subset \mathcal{B}$ and

$$\sigma[\mathcal{A}] = \text{the smallest } \sigma\text{-algebra that contains } \mathcal{A}. \tag{1.38}$$

Then, is the following true?

$$\mu_1(A) = \mu_2(A) \text{ for all } A \in \mathcal{A} \implies \mu_1(A) = \mu_2(A) \text{ for all } A \in \sigma[\mathcal{A}]. \tag{1.39}$$

Unfortunately, this is not true in general, see e.g. Example 1.5.3 below. On the other hand, a positive answer is provided by the following:

Lemma 1.3.1 (Dynkin's lemma) *Let μ be a signed measures on a measurable space $(S,\mathcal{B})$ and that $\mu(S)=0$. Suppose that $\mathcal{A} \subset \mathcal{B}$ is a π-system (i.e., $A_1,A_2 \in \mathcal{A} \Rightarrow A_1 \cap A_2 \in \mathcal{A}$). Then,*

$$\mu(A) = 0 \text{ for all } A \in \mathcal{A} \implies \mu(A) = 0 \text{ for all } A \in \sigma[\mathcal{A}]. \tag{1.40}$$

In particular, (1.39) is true for $\mu_1,\mu_2 \in \mathcal{P}(S,\mathcal{B})$, as can be seen by applying (1.40) to $\mu = \mu_1 - \mu_2$.

The proof of this lemma is presented in Section 1.4. It is more important to know how to apply Lemma 1.3.1 than to know how to prove it. Here is an example of such application.

Lemma 1.3.2 *Let S be a metric space with the metric ρ, and $\mathcal{B}$ the Borel σ-algebra. Then, the following conditions for a signed measure μ on $(S,\mathcal{B})$ are equivalent:*

a) $\mu = 0$

b) $\int f d\mu = 0$ *for all bounded, Lipschiz continuous* $f : S \to [0,\infty)$.

c) $\mu(G) = 0$ *for any open subset* $G \subset S$.

Remark $f : S \to \mathbb{R}$ is said to be *Lipschiz continuous*, if there is a constant L such that $|f(x) - f(y)| \le L\rho(x, y)$ for all $x, y \in S$.

Proof: a) $\Rightarrow$ b): Obvious.
b) $\Rightarrow$ c): It is enough to prove that $\mu(F) = 0$ for any closed subset $F \subset S$. For $x \in S$ and a closed set F, let

$$f_n(x) = (1 - n\rho(x, F))^+ \in [0, 1]. \tag{1.41}$$

Then,

$$|f_n(x) - f_n(y)| \le n\rho(x, y) \text{ for all } x, y \in S \tag{1.42}$$

(cf. Exercise 1.3.1) and hence f_n is bounded, Lipschiz continuous. Moreover, $f_n \searrow 1_F$, as $n \nearrow \infty$. Thus, by the bounded convergence theorem,

$$\mu(F) = \lim_{n\to\infty} \int f_n d\mu = 0.$$

c) $\Rightarrow$ a): Let $\mathcal{O}$ be the totality of open subsets in S. Then, $\mathcal{O}$ is a π-system and $\mathcal{B} = \sigma[\mathcal{O}]$. Moreover, $\mu(S) = 0$, since $S \in \mathcal{O}$. Thus, a) follows from c) by Lemma 1.3.1. \(^□^)/

Exercise 1.3.1 Prove (1.42).

Exercise 1.3.2 Suppose that μ is a signed measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Use Lemma 1.3.1 to prove that $\mu = 0$ if and only if

$$\mu\left(\prod_{j=1}^d (-\infty, b_j]\right) = 0 \quad \text{for any } (b_j)_{j=1}^d \in \mathbb{R}^d. \tag{1.43}$$

1.4 ($\star$) Proof of Lemma 1.3.1

Let μ be a signed measures on a measurable space $(S, \mathcal{B})$ and that $\mu(S) = 0$. Let us consider

$$\mathcal{D}_\mu \stackrel{\text{def.}}{=} \{B \in \mathcal{B}\ ;\ \mu(B) = 0\}. \tag{1.44}$$

If the class $\mathcal{D}_\mu$ defined by (1.44) happens to be a π-system, it is then not difficult to prove that $\mathcal{D}_\mu$ is a σ-algebra[6] and hence that $\sigma[\mathcal{A}] \subset \mathcal{D}_\mu$. Unfortunately, $\mathcal{D}_\mu$ is *not* a π-system in general. In fact, we see in Exercise 8.7.2 an example where

- the family $\mathcal{D}_\mu$ in (1.44) is not a σ-algebra and hence is not a π-system (Exercise 1.4.1).
- "$\mu(A) = 0$ for all $A \in \mathcal{A}$" does not imply "$\mu(A) = 0$ for all $A \in \sigma(\mathcal{A})$".

This difficulty can be circumvented as follows. We begin by introducing the abstract terminology.

Definition 1.4.1 Suppose that S is a set.

► A subset $\mathcal{D}$ of 2^S is called a *δ-system* or a *Dynkin class* if the following conditions are satisfied;

D1) $S \in \mathcal{D}$.

[6]Use inclusion and exclusion formula to prove that $\mathcal{D}_\mu$ is closed under finite union.

D2) $\{A_n\}_{n\ge1} \subset \mathcal{D}$, $A_n \subset A_{n+1}$ $(n \ge 1) \Rightarrow A_{n+1}\backslash A_n \in \mathcal{D}$ $(n \ge 1)$, $\cup_{n\ge1} A_n \in \mathcal{D}$.

► For, $\mathcal{A} \subset 2^S$ we define:

$$\delta[\mathcal{A}] = \bigcap \mathcal{D}, \tag{1.45}$$

where the intersection is taken over all δ-system $\mathcal{D}$ that contains $\mathcal{A}$.

Lemma 1.4.2 *Suppose that* S *is a set and that* $\mathcal{A} \subset 2^S$. *Then, the following are equivalent:*

a) $\delta[\mathcal{A}] = \sigma[\mathcal{A}]$.

b) $A \cap B \in \delta[\mathcal{A}]$ *for all* $A, B \in \mathcal{A}$.

c) $\delta[\mathcal{A}]$ *is a* π*-system.*

Before proving Lemma 1.4.2, we first finish the proof of Lemma 1.3.1.
Proof of Lemma 1.3.1: It is easy so see that $\mathcal{D}_\mu$ defined by (1.44) is a δ-system (Here, we use the assumption $\mu(S) = 0$). Since $\mathcal{A} \subset \mathcal{D}_\mu$ and $\mathcal{A}$ is a π-system (and thus, satisfies condidtion b) of Lemma 1.4.2), we see by Lemma 1.4.2 that $\sigma[\mathcal{A}] = \delta[\mathcal{A}] \subset \mathcal{D}_\mu$. \(^□^)/

Proof of Lemma 1.4.2: a) ⇒ b): Obvious.
b) ⇒ c): *Step1*: We first show that $A \in \mathcal{A}$, $B \in \delta[\mathcal{A}] \Rightarrow A \cap B \in \delta[\mathcal{A}]$. To do so, we introduce

$$\mathcal{D}_1 = \bigcap_{A \in \mathcal{A}} \{B \in 2^S \; ; \; A \cap B \in \delta[\mathcal{A}]\}.$$

Then, the claim of Step1 can be paraphrased as $\delta[\mathcal{A}] \subset \mathcal{D}_1$. We have $\mathcal{A} \subset \mathcal{D}_1$ by b). On the other hand, it is easy to verify that $\mathcal{D}_1$ is a δ-system (Exercise 1.4.2). Since $\delta[\mathcal{A}]$ is the smallest δ-system that contains $\mathcal{A}$, we have $\delta[\mathcal{A}] \subset \mathcal{D}_1$.

Step2: We now show that $A, B \in \delta[\mathcal{A}] \Rightarrow A \cap B \in \delta[\mathcal{A}]$, which implies c). To do so, we introduce

$$\mathcal{D}_2 = \bigcap_{A \in \delta[\mathcal{A}]} \{B \in 2^S \; ; \; A \cap B \in \delta[\mathcal{A}]\}.$$

Then, the claim of Step2 can be paraphrased as $\delta[\mathcal{A}] \subset \mathcal{D}_2$. We have $\mathcal{A} \subset \mathcal{D}_2$ by Step1. On the other hand, it is easy to verify that $\mathcal{D}_2$ is a δ-system (Exercise 1.4.2). Since $\delta[\mathcal{A}]$ is the smallest δ-system that contains $\mathcal{A}$, we have $\delta[\mathcal{A}] \subset \mathcal{D}_2$.
c) ⇒ a): $\delta[\mathcal{A}] \subset \sigma[\mathcal{A}]$: $\sigma[\mathcal{A}]$ is one of the δ-system which contains $\mathcal{A}$, while $\delta[\mathcal{A}]$ is the smallest among them.
$\delta[\mathcal{A}] \supset \sigma[\mathcal{A}]$: By b), $\delta[\mathcal{A}]$ is a π-system, which implies that $\delta[\mathcal{A}]$ is a σ-algebra which contains $\mathcal{A}$ (Exercise 1.4.1). Since $\sigma[\mathcal{A}]$ is the smallest σ-algebra that contains $\mathcal{A}$, we have $\delta[\mathcal{A}] \supset \sigma[\mathcal{A}]$.
\(^□^)/

Exercise 1.4.1 Prove that a δ-system $\mathcal{D}$ is a σ-algebra if and only if $\mathcal{D}$ is a π-system.

Exercise 1.4.2 Prove the following: (i) Let $\mathcal{D}_\lambda$ $(\lambda \in \Lambda)$ be δ-systems on a set S. Then $\bigcap_{\lambda\in\Lambda} \mathcal{D}_\lambda$ is a δ-system. (ii) Let $\mathcal{D}$ be a δ-system on a set S and $A \in \mathcal{D}$ be fixed. Then, $\mathcal{D}(A) \stackrel{\text{def}}{=} \{B \in 2^S \; ; \; A \cap B \in \mathcal{D}\}$ is a δ-system. (iii) Conclude from (i) and (ii) that $\mathcal{D}_1$ and $\mathcal{D}_2$ in the proof of Lemma 1.4.2 are δ-systems.

1.5 Product Measures

Thoughout sections 1.5 and 1.6, we will use the following notation. Let Λ be a set and $\{(S_\lambda, \mathcal{B}_\lambda)\}_{\lambda \in \Lambda}$ be measurable spaces. For $\Gamma \subset \Lambda$, let $S_\Gamma = \prod_{\lambda \in \Gamma} S_\lambda$ be the direct product, $S = S_\Lambda$, and $\pi_\Gamma : S \to S_\Gamma$ be the canonical projection, $\pi_\lambda = \pi_{\{\lambda\}}$ for $\lambda \in \Lambda$. Recall that for $\mathcal{A} \subset 2^S$, $\sigma[\mathcal{A}]$ denotes the smallest σ-algebra that contains $\mathcal{A}$, cf. (1.38).

Definition 1.5.1 (The direct product of measurable spaces)

► A subset of S of the form $\pi_\lambda^{-1}(B_\lambda)$ for some $\lambda \in \Lambda$ and $B_\lambda \in \mathcal{B}_\lambda$ is called a *simple cylinder set*. We define $\mathcal{C}_0(S) \subset 2^S$ by

$$\mathcal{C}_0(S) = \text{all the simple cylinder sets of } S. \tag{1.46}$$

► The following σ-algebra is called the *product σ-algebra* on S:

$$\mathcal{B}(S) = \bigotimes_{\lambda \in \Lambda} \mathcal{B}_\lambda \stackrel{\text{def.}}{=} \sigma[\mathcal{C}_0(S)]. \tag{1.47}$$

► The measurable space $(S, \mathcal{B}(S))$ is called the *direct product* of $\{(S_\lambda, \mathcal{B}_\lambda)\}_{\lambda \in \Lambda}$.

Remark: The σ-algebra $\mathcal{B}(S)$ can also be characterized as follows.

$$\mathcal{B}(S) = \{\pi_\Gamma^{-1}(A)\ ;\ \Gamma \subset \Lambda \text{ is at most countable}, A \in \mathcal{B}(S_\Gamma)\}.$$

See Proposition 1.5.6 below.

The following lemma characterizes the measurable maps with values in $(S, \mathcal{B}(S))$ in Definition 1.5.1.

Lemma 1.5.2 *Let $(\Omega, \mathcal{F})$ be a measurable space, $(S, \mathcal{B}(S))$ be as in Definition 1.5.1 and $X(\omega) = (X_\lambda(\omega))_{\lambda \in \Lambda}$ be a map from Ω to S. Then, the following are equivalent.*

a) *$X : (\Omega, \mathcal{F}) \to (S, \mathcal{B}(S))$ is measurable.*

b) *$X_\lambda : (\Omega, \mathcal{F}) \to (S_\lambda, \mathcal{B}_\lambda)$ is measurable for all $\lambda \in \Lambda$.*

Proof: a) $\Rightarrow$ b): $\pi_\lambda : (S, \mathcal{B}(S)) \to (S_\lambda, \mathcal{B}_\lambda)$ is measurable for all $\lambda \in \Lambda$. Thus, by assumption, $X_\lambda = \pi_\lambda \circ X$is measurable for all $\lambda \in \Lambda$.
a) $\Leftarrow$ b): We have to prove that

1) $\forall B \in \mathcal{B}(S)$, $X^{-1}(B) \in \mathcal{F}$.

But this amounts to saying that

2) $\mathcal{B}(S) \subset X(\mathcal{F}) \stackrel{\text{def}}{=} \{B \in 2^S\ ;\ X^{-1}(B) \in \mathcal{F}\}$.

To prove 2), it is enough to verify that $X(\mathcal{F})$ is a σ-algebra which contains $\mathcal{C}_0(S)$, since $\mathcal{B}(S) = \sigma[\mathcal{C}_0(S)]$. It is obvious that $X(\mathcal{F})$ is a σ-algebra. On the other hand, we have for any $\lambda \in \Lambda$ and $B_\lambda \in \mathcal{B}_\lambda$ that

$$X^{-1}(\pi_\lambda^{-1}(B_\lambda)) = (\pi_\lambda \circ X)^{-1}(B_\lambda) = X_\lambda^{-1}(B_\lambda) \in \mathcal{F}.$$

This implies that $\mathcal{C}_0(S) \subset \mathcal{F}$. \(^□^)/

Let $(S, \mathcal{B}(S))$ be as in Definition 1.5.1, and $\mu, \nu \in \mathcal{P}(S, \mathcal{B}(S))$. Although $\mathcal{B}(S) = \sigma[\mathcal{C}_0(S)]$, it is *not* true that

$$\mu(B) = \nu(B) \text{ for all } B \in \mathcal{C}_0(S) \implies \mu = \nu. \tag{1.48}$$

Let us now look at a simple, but enlightening example.

Example 1.5.3 (**Simple cylinder sets do not determine the measure**) Let $S_1 = S_2 = \{0, 1\}$, $S = S_1 \times S_2$ and $\mu_\lambda \in \mathcal{P}(S_\lambda)$, $\lambda = 1, 2$.

We define $\nu_\theta \in \mathcal{P}(S)$ by

1) $$\begin{pmatrix} \nu_\theta(0,0) & \nu_\theta(0,1) \\ \nu_\theta(1,0) & \nu_\theta(1,1) \end{pmatrix} = \begin{pmatrix} \theta & \mu_1(0) - \theta \\ \mu_2(0) - \theta & 1 + \theta - \mu_1(0) - \mu_2(0) \end{pmatrix},$$

where, for ν_θ to be a probability measure, we suppose that $\theta \in [\theta_0, \theta_1]$ with

$$\theta_0 = (\mu_1(0) + \mu_2(0) - 1)^+ \text{ and } \theta_1 = \mu_1(0) \wedge \mu_2(0).$$

(We easily see that $\theta_0 \leq \theta_1$, with equality iff $\mu_1(0) \in \{0, 1\}$ or $\mu_2(0) \in \{0, 1\}$.) We will show that for $\mu \in \mathcal{P}(S)$ and $\theta \in [\theta_0, \theta_1]$,

2) $$\mu = \nu_\theta \iff \mu \circ \pi_\lambda^{-1} = \mu_\lambda \ (\lambda = 1, 2) \text{ and } \mu(0, 0) = \theta.$$

Note that $\theta_0 < \theta_1$ iff $0 < \mu_\lambda(0) < 1$ ($\lambda = 1, 2$). Thus, the above simple example already shows that (infinitely) many different probability measures on a product space can take the same values on $\mathcal{C}_0(S) = \{\pi_\lambda^{-1}(B_\lambda) \ ; \ \lambda = 1, 2, \ B_\lambda \subset S_\lambda\}$.

To prove $\Rightarrow$ of 2), we check that $\nu_\theta \circ \pi_\lambda^{-1} = \mu_\lambda$, $\lambda = 1, 2$ for all $\theta \in [\theta_0, \theta_1]$. Since

$$\pi_1^{-1}(0) = \{(0, 0), (0, 1)\}, \quad \pi_1^{-1}(1) = \{(1, 0), (1, 1)\},$$

we have that

$$\begin{aligned} \nu_\theta \circ \pi_1^{-1}(0) &= \nu_\theta(0,0) + \nu_\theta(0,1) \stackrel{1)}{=} \mu_1(0), \\ \nu_\theta \circ \pi_1^{-1}(1) &= \nu_\theta(1,0) + \nu_\theta(1,1) \stackrel{1)}{=} 1 - \mu_1(0) = \mu_1(1). \end{aligned}$$

Similarly,

$$\begin{aligned} \nu_\theta \circ \pi_2^{-1}(0) &= \nu_\theta(0,0) + \nu_\theta(1,0) \stackrel{1)}{=} \mu_2(0), \\ \nu_\theta \circ \pi_2^{-1}(1) &= \nu_\theta(0,1) + \nu_\theta(1,1) \stackrel{1)}{=} 1 - \mu_2(0) = \mu_2(1). \end{aligned}$$

To prove $\Leftarrow$ of 2), let $\mu \in \mathcal{P}(S)$ be such that $\mu \circ \pi_\lambda^{-1} = \mu_\lambda$ ($\lambda = 1, 2$) and $\theta = \mu(0, 0)$. Then, it is clear from the above computation that $\mu(s_1, s_2) = \nu_\theta(s_1, s_2)$ for all $(s_1, s_2) \in S$. \(^□^)/

Instead of (1.48) which is not true, we have the following

Lemma 1.5.4 (Cylinder sets determin the measure) *Let everything be as in Definition 1.5.1.*

▶ *A finite intersection of simple cylinder sets is called a* **cylinder set**. *We define* $\mathcal{C}(S) \subset 2^S$ *by :*

$$\mathcal{C}(S) = \text{all the cylinder sets of } S. \tag{1.49}$$

a) $\mathcal{B}(S) = \sigma[\mathcal{C}(S)]$.

b) *The set* $\mathcal{C}(S)$ *is a* π*-system.*

c) $\mu, \nu \in \mathcal{P}(S, \mathcal{B}(S))$, $\mu(B) = \nu(B)$ *for all* $B \in \mathcal{C}(S) \implies \mu = \nu$.

Proof: a) It is clear that $\mathcal{C}_0(S) \subset \mathcal{C}(S) \subset \sigma[\mathcal{C}_0(S)]$, and hence $\sigma[\mathcal{C}_0(S)] = \sigma[\mathcal{C}(S)]$.
b) Let $B_1, B_2 \in \mathcal{C}(S)$. Then, there exist finite sets $\mathcal{C}_1, \mathcal{C}_2 \subset \mathcal{C}_0(S)$ such that

$$B_i = \bigcap_{B \in \mathcal{C}_i} B, \quad i = 1, 2.$$

Thus,

$$B_1 \cap B_2 = \bigcap_{B \in \mathcal{C}_1 \cup \mathcal{C}_2} B \in \mathcal{C}(S).$$

c) This follows from a), b), and Lemma 1.3.1. \\(^□^)/

Theorem 1.5.5 (Product measures) *Let everything be as in Definition 1.5.1. Suppose that* $\mu_\lambda \in \mathcal{P}(S_\lambda, \mathcal{B}_\lambda)$ *for each* $\lambda \in \Lambda$. *Then, there exists a unique* $\mu \in \mathcal{P}(S, \mathcal{B}(S))$ *such that*

$$\mu\left(\bigcap_{\lambda \in \Lambda_0} \pi_\lambda^{-1}(B_\lambda)\right) = \prod_{\lambda \in \Lambda_0} \mu_\lambda(B_\lambda) \quad \text{for any finite } \Lambda_0 \subset \Lambda \text{ and } B_\lambda \in \mathcal{B}_\lambda \ (\lambda \in \Lambda_0). \tag{1.50}$$

▶ *The measure* μ *defined by (1.50) is called the* **product measure** *of* $\{\mu_\lambda\}_{\lambda \in \Lambda}$ *and is denoted by* $\otimes_{\lambda \in \Lambda} \mu_\lambda$.

Proof: The uniqueness follows from Lemma 1.5.4. For the existence[7], we refer the reader to [Dud89, page 201, Theorem 8.2.2]. A self-contained exposition is given by Proposition 8.3.1 in a special case that Λ is a countable set and each $(S_\lambda, \mathcal{B}_\lambda)$ is a complete separable metric space with the Borel σ-algebra. \\(^□^)/

Remark: Concerning Theorem 1.5.5, note that:

$$\mu = \otimes_{\lambda \in \Lambda} \mu_\lambda \quad \implies \quad \mu \circ \pi_\lambda^{-1} = \mu_\lambda, \quad \text{for all } \lambda \in \Lambda. \tag{1.51}$$

This can be seen from (1.50) by taking $\Lambda_0 = \{\lambda\}$. Note also that the converse is not true. A counterexample is provided by Example 1.5.3, where $\nu_\theta \circ \pi_j^{-1} = \mu_j$ $(j = 1, 2)$ for all $\theta \in [\theta_0, \theta_1]$, but $\nu_\theta = \mu_1 \otimes \mu_2$ only when $\theta = \mu_1(0)\mu_2(0)$.

[7] If each $(S_\lambda, \mathcal{B}_\lambda)$ is a complete separable metric space with the Borel σ-algebra, then one can also apply Kolmogorov's extension theorem [Dur95, page 26 (4.9)].

Exercise 1.5.1 Let everything be as in Lemma 1.5.4. Prove then that the following conditions for a set $B \subset S$ are equivalent:
a) $B \in \mathcal{C}(S)$.
b) $B = \prod_{\lambda \in \Lambda} \pi_\lambda(B)$ with $\pi_\lambda(B) \in \mathcal{B}_\lambda$ for all λ, and $\pi_\lambda(B) = S_\lambda$ except for finitely many λ.
c) $B = \prod_{\lambda \in \Lambda} B_\lambda$, with $B_\lambda \in \mathcal{B}_\lambda$ for all λ, and $B_\lambda = S_\lambda$ except for finitely many λ.

Exercise 1.5.2 Let $S_1 = S_2 = \{0, 1\}$. Find cylinder sets $A, B \subset S_1 \times S_2$ such that $A \cup B$ is *not* a cylinder set. This in particular shows that the set $\mathcal{C}$ is not closed under union in general.

Exercise 1.5.3 Let everything be as in Theorem 1.5.5.
(i) Suppose that $\Lambda = \{1, 2, ...\}$. Prove then that (1.50) is equivalent to that

$$\mu\left(\bigcap_{j=1}^{n} \pi_j^{-1}(B_j)\right) = \prod_{j=1}^{n} \mu_j\,(B_j)$$
for any $n \geq 1$ and $B_j \in \mathcal{B}_j$ $(1 \leq j \leq n)$.

(ii) Suppose that each S_λ is at most countable. Prove then that (1.50) is equivalent to that

$$\mu\left(\bigcap_{\lambda \in \Lambda_0} \pi_\lambda^{-1}(x_\lambda)\right) = \prod_{\lambda \in \Lambda_0} \mu_\lambda\,(x_\lambda)$$
for any finite $\Lambda_0 \subset \Lambda$ and $x_\lambda \in S_\lambda$ $(\lambda \in \Lambda_0)$.

($\star$) Complement to section 1.5

Proposition 1.5.6 *Referring to Definition 1.5.1,*

$$\mathcal{B}(S) = \{\pi_\Gamma^{-1}(A)\ ;\ \Gamma \subset \Lambda \textit{ is at most countable}, A \in \mathcal{B}(S_\Gamma)\}.$$

Proof: We first show that

1) $\mathcal{B}(S) \supset \mathcal{D}(S) \stackrel{\text{def}}{=} \{\pi_\Gamma^{-1}(A)\ ;\ \Gamma \subset \Lambda \text{ is at most countable}, A \in \mathcal{B}(S_\Gamma)\}$.

To this end, we fix a $\Gamma \subset \Lambda$, at most countable, and verify that

2) $\mathcal{B}(S_\Gamma) \subset \mathcal{A}(S_\Gamma) \stackrel{\text{def}}{=} \{A \subset S_\Gamma\ ;\ \pi_\Gamma^{-1}(A) \in \mathcal{B}(S)\}$.

It is clear that $\mathcal{A}(S_\Gamma)$ is a σ-algebra on S_Γ. We will check that $A \stackrel{\text{def}}{=} \pi_{\lambda,\Gamma}^{-1}(B_\lambda) \subset \mathcal{A}(S_\Gamma)$ for any $B_\lambda \in \mathcal{B}_\lambda$, where $\pi_{\lambda,\Gamma}$ denotes the canonical projection from S_Γ to S_λ. Indeed,

$$\pi_\Gamma^{-1}(A) = (\pi_{\lambda,\Gamma} \circ \pi_\Gamma)^{-1}(B_\lambda) = \pi_\lambda^{-1}(B_\lambda) \in \mathcal{B}(S).$$

Hence, we have 2). Next, we show that

3) $\mathcal{B}(S) \subset \mathcal{D}(S)$.

It is clear that $\pi_\lambda^{-1}(B_\lambda) \subset \mathcal{D}(S)$ for any $B_\lambda \in \mathcal{B}_\lambda$. Thus, it is enough to verify that $\mathcal{D}(S)$ is a σ-algebra. It is easy to see that $S \in \mathcal{D}(S)$ and that $D \in \mathcal{D}(S) \Rightarrow D^c \in \mathcal{D}(S)$. To check that $\mathcal{D}(S)$ is closed under countable union, let $\Gamma_n \subset \Lambda$ be at most countable, $A_n \in \mathcal{B}(S_{\Gamma_n})$, $\Gamma = \bigcup_{n \ge 1} \Gamma_n$. Also, let $\pi_{\Gamma_n,\Gamma}$ denotes the canonical projection from S_Γ to S_{Γ_n} and $A = \bigcup_{n\ge 1} \pi_{\Gamma_n,\Gamma}^{-1}(A_n)$. By 1) (appplied to S_Γ, instead of S), we see that $\pi_{\Gamma_n,\Gamma}^{-1}(A_n) \in \mathcal{B}(S_\Gamma)$ for all $n \ge 1$, so that $A \in \mathcal{B}(S_\Gamma)$. Note that

$$\pi_{\Gamma_n}^{-1}(A_n) = (\pi_{\Gamma_n,\Gamma} \circ \pi_\Gamma)^{-1}(A_n) = \pi_\Gamma^{-1}(\pi_{\Gamma_n,\Gamma}^{-1}(A_n)).$$

Therefore,

$$\bigcup_{n\ge 1} \pi_{\Gamma_n}^{-1}(A_n) = \pi_\Gamma^{-1}(A) \in \mathcal{D}(S),$$

which concludes the proof of 3). \(^□^)/

Corollary 1.5.7 *Referring to Definition 1.5.1, suppose that $U \in \mathcal{B}(S)$. Then, there exists an at most countable set $\Gamma \subset \Lambda$ with the following propoerty.*

$$x \in S,\ y \in U,\ \pi_\Gamma(x) = \pi_\Gamma(y) \implies x \in U.$$

Proof: By Proposition 1.5.6, there exist an at most countable set $\Gamma \subset \Lambda$ and $A \in \mathcal{B}(S_\Gamma)$ such that $U = \pi_\Gamma^{-1}(A)$. Since $y \in U$, we have that $\pi_\Gamma(y) \in A$ and hence that $\pi_\Gamma(x) = \pi_\Gamma(y) \in A$. This implies that $x \in U$. \(^□^)/

We present a following variant of Lemma 1.5.2, which applies to a subset U of S, rather than S itself. The proof is almost the same as that of Lemma 1.5.2, hence is omitted.

Lemma 1.5.8 *Let $(\Omega, \mathcal{F})$ be a measurable space, $(S, \mathcal{B}(S))$ be as in Definition 1.5.1, $U \subset S$ and $\mathcal{B}(U) \stackrel{\rm def}{=} \{B \cap U \ ; \ B \in \mathcal{B}(S)\}$. Let also $X(\omega) = (X_\lambda(\omega))_{\lambda \in \Lambda}$ be a map from Ω to U. Then, the following are equivalent.*

a) $X : (\Omega, \mathcal{F}) \to (U, \mathcal{B}(U))$ *is measurable.*

b) $X_\lambda : (\Omega, \mathcal{F}) \to (S_\lambda, \mathcal{B}_\lambda)$ *is measurable for all $\lambda \in \Lambda$.*

We present a following variant of Lemma 1.5.4, which applies to a subset U of S, rather than S itself.

Lemma 1.5.9 *Let everything be as in Lemma 1.5.4, $U \subset S$ and*

$$\mathcal{C}(U) \stackrel{\rm def}{=} \{C \cap U \ ; \ C \in \mathcal{C}(S)\}.$$

A set in $\mathcal{C}(U)$ is called a cylinder set *in U.*

a) $\mathcal{B}(U) \stackrel{\rm def}{=} \{B \cap U \ ; \ B \in \mathcal{B}(S)\} = \sigma[\mathcal{C}(U)]$.

b) *The set $\mathcal{C}(U)$ is a π-system.*

c) $\mu, \nu \in \mathcal{P}(U, \mathcal{B}(U))$, $\mu(C) = \nu(C)$ *for all $C \in \mathcal{C}(U)$* $\implies \mu = \nu$.

Proof: a) Obviously, $\mathcal{B}(U) \supset \mathcal{C}(U)$, and hence $\mathcal{B}(U) \supset \sigma[\mathcal{C}(U)]$. On the other hand, let $\mathcal{A} = \{B \subset S\ ;\ B \cap U \in \sigma[\mathcal{C}(U)]\}$. Then, $\mathcal{A}$ is a σ-algebra on S, which contains $\mathcal{C}(S)$, and hence $\mathcal{A} \supset \sigma[\mathcal{C}(S)] = \mathcal{B}(S)$ (Lemma 1.5.4). This implies that $\mathcal{B}(U) \subset \mathcal{C}(U)$.
b) Let $C_1, C_2 \in \mathcal{C}(S)$. Then $C_1 \cap C_2 \in \mathcal{C}(S)$ (Lemma 1.5.4), and hence $(C_1 \cap U) \cap (C_2 \cap U) = (C_1 \cap C_2) \cap U \in \mathcal{C}(U)$.
c) This follows from a), b) and Lemma 1.3.1. \(^□^)/

1.6 Independent Random Variables

Let us now come back to our informal description (0.1) of playing a game. If you play two games with outcomes $X_i : \Omega \to \{-1, +1\}$ $(i = 1, 2)$ in such a way that the outcome of one game does not affect that of the other, e.g., tossing two coins on different tables. We then should have

$$P(X_2 = \varepsilon_2 | X_1 = \varepsilon_1) = P(X_2 = \varepsilon_2) \quad \text{for all } \varepsilon_k = \pm 1.$$

The above expression of "independence" is equivalent to that

$$P(X_1 = \varepsilon_1, X_2 = \varepsilon_2) = P(X_1 = \varepsilon_1)P(X_2 = \varepsilon_2) \quad \text{for all } \varepsilon_k = \pm 1.$$

We now come to the definition of independent r.v.'s. In what follows, $(\Omega, \mathcal{F}, P)$ denotes a probability space.

Proposition 1.6.1 (Independent r.v.'s) *Suppose that* $\{(S_\lambda, \mathcal{B}_\lambda, \mu_\lambda)\}_{\lambda \in \Lambda}$ *are probability spaces indexed by a set* Λ *and that* $X_\lambda : \Omega \to S_\lambda$ *is a r.v. such that* $X_\lambda \approx \mu_\lambda$ *for each* $\lambda \in \Lambda$. *Then the following conditions a)–c) are equivalent:*

a) *For any finite* $\Lambda_0 \subset \Lambda$ *and for any* $B_\lambda \in \mathcal{B}_\lambda$ $(\lambda \in \Lambda_0)$,

$$P\left(\bigcap_{\lambda \in \Lambda_0} \{X_\lambda \in B_\lambda\}\right) = \prod_{\lambda \in \Lambda_0} P(X_\lambda \in B_\lambda). \tag{1.52}$$

b1) $(X_\lambda)_{\lambda \in \Lambda} \approx \bigotimes_{\lambda \in \Lambda} \mu_\lambda$.

b2) $(X_\lambda)_{\lambda \in \Lambda_1} \approx \bigotimes_{\lambda \in \Lambda_1} \mu_\lambda$ *for any* $\Lambda_1 \subset \Lambda$.

b3) $(X_\lambda)_{\lambda \in \Lambda_0} \approx \bigotimes_{\lambda \in \Lambda_0} \mu_\lambda$ *for any finite* $\Lambda_0 \subset \Lambda$.

c) *For any finite* $\Lambda_0 \subset \Lambda$ *and for any* $f_\lambda \in L^1(\mu_\lambda)$ $(\lambda \in \Lambda_0)$,

$$E\left[\prod_{\lambda \in \Lambda_0} f_\lambda(X_\lambda)\right] = \prod_{\lambda \in \Lambda_0} E\left[f_\lambda(X_\lambda)\right]. \tag{1.53}$$

▶ *R.v.'s* $\{X_\lambda\}_{\lambda \in \Lambda}$ *are said to be* **independent** *if they satisfy one of (therefore all of) conditions in the proposition.*

▶ *Let* $\{X_\lambda\}_{\lambda \in \Lambda}$ *be independent. If* $(S_\lambda, \mathcal{B}_\lambda, \mu_\lambda)$ *are identical for all* $\lambda \in \Lambda$, *then the r.v.'s are called* **iid** *(independent and identically distributed) r.v.'s.*

Proof: Let μ be the law of $X \stackrel{\text{def}}{=} (X_\lambda)_{\lambda \in \Lambda} : \Omega \to \prod_{\lambda \in \Lambda} S_\lambda$. For any finite $\Lambda_0 \subset \Lambda$ and for any $B_\lambda \in \mathcal{B}_\lambda$ $(\lambda \in \Lambda_0)$,

1) $$P\left(\bigcap_{\lambda \in \Lambda_0} \{X_\lambda \in B_\lambda\}\right) = P\left(X \in \bigcap_{\lambda \in \Lambda_0} \pi_\lambda^{-1}(B_\lambda)\right) \stackrel{(1.3)}{=} \mu\left(\bigcap_{\lambda \in \Lambda_0} \pi_\lambda^{-1}(B_\lambda)\right),$$

2) $$\prod_{\lambda \in \Lambda_0} P(X_\lambda \in B_\lambda) \stackrel{(1.3)}{=} \prod_{\lambda \in \Lambda_0} \mu_\lambda(B_\lambda).$$

a) $\Leftrightarrow$ b1):

$$\begin{aligned} \text{a)} \quad &\stackrel{(1.52)}{\Longleftrightarrow} \quad \text{LHS 1) = LHS 2)}, \ \forall \text{ finite } \Lambda_0 \subset \Lambda \\ &\Longleftrightarrow \quad \text{RHS 1) = RHS 2)}, \ \forall \text{ finite } \Lambda_0 \subset \Lambda \\ &\stackrel{(1.50)}{\Longleftrightarrow} \quad \mu = \bigotimes_{\lambda \in \Lambda} \mu_\lambda \Longleftrightarrow \text{b1)}. \end{aligned}$$

a) $\Rightarrow$ b2): a) implies that (1.52) holds in particular for all finite $\Lambda_0 \subset \Lambda_1$. Then, by letting Λ_1 play the role of Λ in the proof of "a) $\Rightarrow$ b1)" above, we get b2).

b2) $\Rightarrow$ b3): Obvious.
b3) $\Rightarrow$ c): Let $S_{\Lambda_0} = \prod_{\lambda \in \Lambda_0} S_\lambda$ and $\mu_{\Lambda_0} = \otimes_{\lambda \in \Lambda_0} \mu_\lambda$. Then,

$$\text{LHS of (1.53)} \stackrel{\text{b3), (1.9)}}{=} \int_{S_{\Lambda_0}} \prod_{\lambda \in \Lambda_0} f_\lambda d\mu_{\Lambda_0} \stackrel{\text{Fubini}}{=} \prod_{\lambda \in \Lambda_0} \int_{S_\lambda} f_\lambda d\mu_\lambda \stackrel{(1.9)}{=} \text{RHS of (1.53)}.$$

c) $\Rightarrow$ a): This can be seen by plugging $f_\lambda = 1_{B_\lambda}$ into (1.53). \(^□^)/

Remarks:
1) The condition a) in Proposition 1.6.1 amounts to saying that the σ-algebras $\{\sigma(X_\lambda)\}_{\lambda \in \Lambda}$ (cf. (1.1)) are independent in the sense of Definition 8.7.1 b).

2) Let $\mu_\lambda \in \mathcal{P}(S_\lambda, \mathcal{B}_\lambda)$ for each $\lambda \in \Lambda$ be given. Then, of course, there can be r.v.'s $\{X_\lambda\}_{\lambda \in \Lambda}$ with

$$X_\lambda \approx \mu_\lambda \quad \text{for all } \lambda \in \Lambda,$$

which are *not* independent. For example, consider the measure ν_θ in Example 1.5.3 and $\{0,1\}$-valued r.v.'s X_1, X_2 such that $(X_1, X_2) \approx \nu_\theta$ with $\theta \neq \mu_1(0)\mu_2(0)$.

Corollary 1.6.2 *Let $\{(S_\lambda, \mathcal{B}_\lambda)\}_{\lambda \in \Lambda}$ be measurable spaces indexed by a set Λ. Suppose that*

a) *$X_\lambda, Y_\lambda : \Omega \to S_\lambda$ are a r.v.'s such that $X_\lambda \approx Y_\lambda$ for each $\lambda \in \Lambda$,*

b) *$\{X_\lambda\}_{\lambda \in \Lambda}$ are independent,*

c) *$\{Y_\lambda\}_{\lambda \in \Lambda}$ are independent.*

Then, $(X_\lambda)_{\lambda \in \Lambda} \approx (Y_\lambda)_{\lambda \in \Lambda}$.

Proof: Let $\mu_\lambda \in \mathcal{P}(S_\lambda, \mathcal{B}_\lambda)$ be the common law of X_λ and Y_λ. Then, by Proposition 1.6.1, $(X_\lambda)_{\lambda \in \Lambda} \approx \otimes_{\lambda \in \Lambda} \mu_\lambda$ and $(Y_\lambda)_{\lambda \in \Lambda} \approx \otimes_{\lambda \in \Lambda} \mu_\lambda$. \(^□^)/

Proposition 1.6.3 *Suppose that $X_i, Y_i, X_i Y_i \in L^1(P)$ for all $i \geq 1$. Then, conditions a)–c) listed below are related as a) $\Rightarrow$ b) $\Rightarrow$ c);*

a) *X_i and Y_j for $i \neq j$ are independent.*

b) *X_i and Y_j for $i \neq j$ are uncorrelated , i.e., $\text{cov}(X_i, Y_j) = 0$ if $i \neq j$.*

c)

$$\text{cov}(\sum_{i=1}^m X_i, \sum_{j=1}^n Y_j) = \sum_{i=1}^m \text{cov}(X_i, Y_i) \quad \text{if } m \leq n. \tag{1.54}$$

Remark: (1.54) is most commonly applied to the special case: $X_i \equiv Y_i$, where it becomes:

$$\text{var}\,(\sum_{i=1}^m X_i) = \sum_{i=1}^m \text{var}\, X_i. \tag{1.55}$$

Proof: a) $\Rightarrow$ b): Since X_i and Y_j for $i \neq j$ are independent,

$$\text{cov}(X_i, Y_j) = E[X_i Y_j] - EX_i EY_j \stackrel{(1.53)}{=} 0$$

b) $\Rightarrow$ c):

$$\text{cov}(\sum_{i=1}^m X_i, \sum_{j=1}^n Y_j) = \sum_{i=1}^m \sum_{j=1}^n \text{cov}(X_i, Y_j) \stackrel{\text{b)}}{=} \sum_{i=1}^m \text{cov}(X_i, Y_i).$$

\(^□^)/

($\star$) Complement to section 1.6

Lemma 1.6.4 (**Kolmogorov's 0-1 law**) *Referring to Proposition 1.6.1, suppose that $\{X_\lambda\}_{\lambda \in \Lambda}$ are independent. Then, $P(B) = 0$ or 1 for all $B \in \mathcal{T}$, where $\mathcal{T}$ is the* **tail σ-algebra** *defined by*

$$\mathcal{T} = \bigcap_{\substack{\Gamma \subset \Lambda \\ \Gamma \text{ is finite}}} \sigma\left[(X_\lambda)_{\lambda \in \Lambda \setminus \Gamma}\right]. \tag{1.56}$$

Proof: Let $\mathcal{G} = \sigma\left[(X_\lambda)_{\lambda \in \Lambda}\right]$, $\mathcal{G}_\Gamma = \sigma\left[(X_\lambda)_{\lambda \in \Gamma}\right]$ for $\Gamma \subset \Lambda$, and $\mathcal{A} = \bigcup_{\substack{\Gamma \subset \Lambda \\ \Gamma \text{ is finite}}} \mathcal{G}_\Gamma$. Fix $B \in \mathcal{T}$ and consider the following two measures on $(\Omega, \mathcal{G})$,

$$\mu_1(A) = P(A \cap B), \ \ \mu_2(A) = P(A)P(B), \ \ (A \in \mathcal{G}).$$

Then,

1) $\mu_1 = \mu_2$ on $\mathcal{A} \cup \{\Omega\}$.

Indeed, it is clear that $\mu_1(\Omega) = \mu_2(\Omega) = P(B)$. Moreover, If $A \in \mathcal{A}$, then $A \in \mathcal{G}_\Gamma$ for some finite set $\Gamma \in \Lambda$. Since $\mathcal{T} = \bigcap_{\substack{\Gamma \subset \Lambda \\ \Gamma \text{ is finite}}} \mathcal{G}_{\Lambda \setminus \Gamma}$, we have $B \in \mathcal{G}_{\Lambda \setminus \Gamma}$. Therefore, A and B are independent, which implies that $\mu_1(A) = \mu_2(A)$.

Since $\mathcal{A}$ is a π-system and $\mathcal{G} = \sigma[\mathcal{A}]$, it follows from 1) and Lemma 1.3.1 that $\mu_1 = \mu_2$ on $\mathcal{G}$. In particular, we have $P(B) = \mu_1(B) = \mu_2(B) = P(B)^2$, which implies that $P(B) = 0$ or 1. \(^□^)/

Let $(\Omega, \mathcal{F}, P)$ be a probability space and A be a finite set. We consider the following setting.

- For each $\alpha \in A$, $(S_{\alpha,\lambda}, \mathcal{B}_{\alpha,\lambda})$, $\lambda \in \Lambda_\alpha$ are measurable spaces indexed by a set Λ_α and

$$(S_\alpha, \mathcal{B}(S_\alpha)) = \left(\prod_{\lambda \in \Lambda_\alpha} S_{\alpha,\lambda}, \bigotimes_{\lambda \in \Lambda_\alpha} \mathcal{B}_{\alpha,\lambda} \right).$$

  (cf. Definition 1.5.1).

- For each $\alpha \in A$, $X_\alpha : \Omega \to S_\alpha$ is a r.v.

Lemma 1.6.5 *Referring to the above setting, the following conditions are equivalent.*

a) *X_α, $\alpha \in A$ are independent.*

b) *$\pi_{\Gamma_\alpha}(X_\alpha)$, $\alpha \in A$ are independent for arbitrarily choosen finite subset $\Gamma_\alpha \subset \Lambda_\alpha$, where $\pi_{\Gamma_\alpha} : S_\alpha \to \prod_{\lambda \in \Gamma_\alpha} S_{\alpha,\lambda}$ denotes the canonical projection.*

Proof: It is enough to prove that b) implies a). The r.v. $(X_\alpha)_{\alpha\in A}$ takes values in the product space

$$S \stackrel{\rm def}{=} \prod_{\alpha\in A} S_\alpha = \prod_{\alpha\in A}\prod_{\lambda\in\Lambda_\alpha} S_{\alpha,\lambda}.$$

Let $\mu_\alpha \in \mathcal{P}(S_\alpha, \mathcal{B}(S_\alpha))$. We prove that

1) $(X_\alpha)_{\alpha\in A} \approx \bigotimes_{\alpha\in A} \mu_\alpha$.

If C is a cylinder set in S, then,

$$C = \bigcap_{\alpha\in A} \pi_{\Gamma_\alpha}^{-1}\left(\prod_{\lambda\in\Gamma_\alpha} B_{\alpha,\lambda}\right)$$

for some finite set $\Gamma_\alpha \subset \Lambda_\alpha$ and $B_{\alpha,\lambda} \in \mathcal{B}(S_{\alpha,\lambda})$. Therefore, by setting $B_\alpha = \prod_{\lambda\in\Gamma_\alpha} B_{\alpha,\lambda}$, we have that

$$\begin{aligned}
P((X_\alpha)_{\alpha\in A} \in C) &= P\left(\bigcap_{\alpha\in A}\{X_\alpha \in \pi_{\Gamma_\alpha}^{-1}(B_\alpha)\}\right) = P\left(\bigcap_{\alpha\in A}\{\pi_{\Gamma_\alpha}(X_\alpha) \in B_\alpha\}\right) \\
&\stackrel{\rm b)}{=} \prod_{\alpha\in A} P\left(\pi_{\Gamma_\alpha}(X_\alpha) \in B_\alpha\right) = \prod_{\alpha\in A} P\left(X_\alpha \in \pi_{\Gamma_\alpha}^{-1}(B_\alpha)\right) \\
&= \left(\bigotimes_{\alpha\in A} \mu_\alpha\right)(C),
\end{aligned}$$

which proves 1) by Lemma 1.5.4. \(^□^)/

Exercise 1.6.1 Let a r.v. U be uniformly distributed on $(0, 2\pi)$. Prove then that $X = \cos U$ and $Y = \sin U$ are not independent and that $\mathrm{cov}(X, Y) = 0$.

Exercise 1.6.2 [8] Let X, Y be r.v.'s with values in $\{0, 1\}$. Prove then that X, Y are independent if and only if $\mathrm{cov}(X, Y) = 0$. Hint:Example 1.5.3.

Exercise 1.6.3 Suppose that a r.v.X is independent of itself. Prove then that there exists $c \in \mathbb{R}$ such that $X = c$, a.s.

Exercise 1.6.4 Suppose that X_j $j = 1, ..., n$ are independent r.v.'s and that $X_1 + ... + X_n = C$ a.s., where C is a constant. Prove then that there are $c_1, ..., c_n \in \mathbb{R}$ such that $X_j = c_j$, a.s. $(j = 1, .., n)$. Hint: $X_n = C - \sum_{j=1}^{n-1} X_j$. Therefore, X_n is independent of itself.

Exercise 1.6.5 Let $S_n = U_1 + ... + U_n$, where $U_1, U_2, \ldots,$ are i.i.d. with uniform distribution on $(0, T)$. For a measurable function $\varphi : \mathbb{R} \to \mathbb{R}$ with period T, prove that $(\varphi(S_j))_{j=1}^n$ and $(\varphi(U_j))_{j=1}^n$ have the same law for any $n \in \mathbb{N}\backslash\{0\}$.

[8] cf. T. Ohira:"On Statistical Independence and No-Correlation for a Pair of Random Variables Taking Two Values: Classical and Quantum" Progress of Theoretical and Experimental Physics, Volume 2018, Issue 8, 1 August 2018, 083A02

Exercise 1.6.6 Let $(X_k)_{k\geq 1}$ be i.i.d. with values in a measurable space $(S, \mathcal{B})$, and let $(N_k)_{k\geq 1}$ be $\mathbb{N}\backslash\{0\}$ valued r.v.'s such that $N_1 < N_2 < ...$ a.s. Assuming that $(X_k)_{k\geq 1}$ and $(N_k)_{k\geq 1}$ are independent, prove that $(X_k)_{k\geq 1}$ and $(X_{N_k})_{k=1}$ have the same law.

Exercise 1.6.7 ($\star$) Let $(X_k)_{k=0,1}$ be indpendent r.v.'s with values in a measurable space $(S, \mathcal{B})$, and let N be $\{0,1\}$-valued r.v. independent of $(X_k)_{k=0,1}$. Then prove that X_N and X_{1-N} are independent if and only if (i): $(X_k)_{k=0,1}$ is i.i.d., or (ii): N is constant a.s. Hint: Take bounded measurable $f_k : S \to \mathbb{R}$ $(k = 0,1)$ and compute $\mathrm{cov}(f_0(X_N), f_1(X_{1-N}))$.

Exercise 1.6.8 ($\star$) Let $(S, \mathcal{A})$ and $(T, \mathcal{B})$ are measurable spaces. Let also $X_1, .., X_n$ be independent r.v.'s with values in S, and $\varphi_j : S^j \to T$ $(j = 1, .., n)$ be measurable functions such that $\varphi_j(s_1, ..., s_{j-1}, X_j)$ has the same law as $\varphi_1(X_j)$ for all $j = 1, .., n$ and $s_1, ..., s_{j-1} \in S$. Prove then that

$$(\varphi_j(X_1, ..., X_{j-1}, X_j))_{j=1}^n \quad \text{and} \quad (\varphi_1(X_j))_{j=1}^n$$

have the same law. This generalizes Exercise 1.6.5.

Exercise 1.6.9 ($\star$) Let everything be as in Proposition 1.6.1. For a disjoint decomposition $\Lambda = \cup_{\gamma\in\Gamma}\Lambda(\gamma)$ of the index set Λ, consider r.v.'s $\{\widetilde{X}\}_{\gamma\in\Gamma}$ defined by

$$\widetilde{X}_\gamma : \omega \mapsto (X_\lambda(\omega))_{\lambda\in\Lambda(\gamma)} \in \prod_{\lambda\in\Lambda(\gamma)} S_\lambda, \quad \gamma \in \Gamma.$$

Prove that r.v.'s $\{\widetilde{X}_\gamma\}_{\gamma\in\Gamma}$ are independent if $\{X_\lambda\}_{\lambda\in\Lambda}$ are. Hint: Condition b) of Proposition 1.6.1.

1.7 Some Functions of Independent Random Variables

Let $X_1, X_2, ...$ be independent r.v.'s for which the distributions are known. Then, one can compute the distribution of a r.v. of the form $f(X_1, X_2, ...)$. Let us look at some examples.

Definition 1.7.1 For a r.v. $X : \Omega \to \mathbb{N}$ with $X \approx \mu \in \mathcal{P}(\mathbb{N})$, we define its *generating function* by the following expectation, or the absolutely convergent power series:

$$G(\mu; s) \stackrel{\text{def}}{=} Es^X = \sum_{n=0}^{\infty} \mu(n)s^n, \quad s \in \mathbb{C}, \ |s| \leq 1, \tag{1.57}$$

where $\mu(n) = \mu(\{n\})$.

Lemma 1.7.2 *For $j = 1,2$, let $\mu_j \in \mathcal{P}(\mathbb{N})$ and let $X_j : \Omega \to \mathbb{N}$ be independent r.v.'s with $X_j \approx \mu_j$. Then, for $\mu \in \mathcal{P}(\mathbb{N})$, the following conditions are equivalent.*

a) $X_1 + X_2 \approx \mu$.

b) $\mu(n) = \sum_{\substack{k,\ell\in\mathbb{N} \\ k+\ell=n}} \mu_1(k)\mu_2(\ell)$.

c) $G(\mu; s) = G(\mu_1; s)G(\mu_2; s), \quad \forall s \in \mathbb{C}, \ |s| \leq 1.$

Proof: a) $\Leftrightarrow$ b): The equivalence can be seen by the following identity. For any $n \in \mathbb{N}$,

$$P(X_1 + X_2 = n) = \sum_{\substack{k,\ell \in \mathbb{N} \\ k+\ell=n}} P(X_1 = k)P(X_2 = \ell) = \sum_{\substack{k,\ell \in \mathbb{N} \\ k+\ell=n}} \mu_1(k)\mu_2(\ell).$$

b) $\Leftrightarrow$ c): The equivalence can be seen by comparing the following two identities.

$$\begin{aligned} G(\mu; s) &= \sum_{n=0}^{\infty} \mu(n)s^n, \\ G(\mu_1; s)G(\mu_2; s) &= \left(\sum_{k=0}^{\infty} \mu_1(k)s^k\right)\left(\sum_{\ell=0}^{\infty} \mu_2(\ell)s^\ell\right) = \sum_{n=0}^{\infty}\left(\sum_{\substack{k,\ell \in \mathbb{N} \\ k+\ell=n}} \mu_1(k)\mu_2(\ell)\right)s^n. \end{aligned}$$

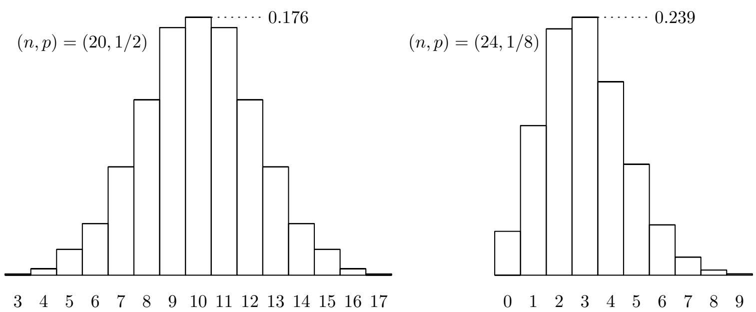
\(^□^)/

Remark Let μ be a complex measure on $\mathbb{N}$. Then, the series $\sum_{n=0}^{\infty} |\mu(\{n\})|$ converges (and equals to the total variation of μ). Thus, we can define its generating function $G(\mu; s)$ ($s \in \mathbb{C}$, $|s| \leq 1$) by the right-hand side of (1.57). Moreover, the equivalence between b) and c) of Lemma 1.7.2 remains valid in the case where μ, μ_1 and μ_2 are complex measures on $\mathbb{N}$.

Example 1.7.3 (**Bin**$(\mathbf{n}, \mathbf{p})$ **and its independent summation**) Let $p \in [0, 1]$ and $n = 1, 2, ...$ A probability measure $\mu_{n,p}$ on $\{0, 1, .., n\}$ defined as follows is called the (n, p)-*binomial distribution*, and will henceforth be denoted by $\mathrm{Bin}(n, p)$:

$$\mu_{n,p}(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, ..., n. \tag{1.58}$$

Here are histograms of $k \mapsto \mu_{n,p}(k)$ for $(n, p) = (20, 1/2)$ and $(n, p) = (24, 1/8)$.



Note in particular that $\mathrm{Bin}(1, p)$ is given by:

$$\mu_{1,p}(k) = \begin{cases} p & \text{if } k = 1, \\ 1-p & \text{if } k = 0. \end{cases} \tag{1.59}$$

Suppose that Z_1, Z_2 are independent r.v.'s, and that $n, n(1), n(2) \in \mathbb{N}$. We show that

$$Z_j \approx \mathrm{Bin}(n(j), p) \ (j = 1, 2), \implies Z_1 + Z_2 \approx \mathrm{Bin}(n, p). \tag{1.60}$$

where $n \stackrel{\text{def}}{=} n(1) + n(2)$. Since the generating function (1.57) for $\mu_{n,p}$ is given by:

$$G(\mu_{n,p}; s) = \sum_{k=0}^{n} \binom{n}{k} (ps)^k (1-p)^{n-k} = (ps + 1 - p)^n, \tag{1.61}$$

we have

$$G(\mu_{n,p}; s) = G(\mu_{n(1),p}; s)G(\mu_{n(2),p}; s),$$

which implies (1.60) via Lemma 1.7.2.
Let $\{X_j\}_{j=1}^n$ be i.i.d. with $X_j \approx \mathrm{Bin}(1,p)$ Then, by applying (1.60) repeatedly, we have

$$S_n \stackrel{\text{def}}{=} X_1 + \ldots + X_n \approx \mathrm{Bin}(n,p). \tag{1.62}$$

The relation (1.62) can also be used to compute the expectation and variance of $\mathrm{Bin}(n,p)$. Note that $X_j^2 = X_j = \mathbf{1}\{X_j = 1\}$. Thus,

$$\begin{aligned} E[X_j^2] &= EX_j = P(X_j = 1) = p, \\ \mathrm{var}\, X_j &= E[X_j^2] - (EX_j)^2 = p(1-p), \\ ES_n &= \sum_{j=1}^n EX_j = np. \end{aligned} \tag{1.63}$$

Since $X_1, ..., X_n$ are independent,

$$\mathrm{var}\, S_n \stackrel{(1.55)}{=} \sum_{j=1}^n \mathrm{var}\, X_j = np(1-p). \tag{1.64}$$

\(^□^)/

Example 1.7.4 (**Summation of independent Poisson r.v.'s**) Suppose that N_1 and N_2 are independent r.v.'s. and that $c(1), c(2) > 0$. We prove that

$$N_j \approx \pi_{c(j)} \ (j = 1,2) \implies N_1 + N_2 \approx \pi_c, \tag{1.65}$$

where $c = c(1) + c(2)$. Since the generating function (1.57) for π_c is given by:

$$G(\pi_c; s) = \exp(-c) \sum_{n=0}^{\infty} \frac{(cs)^n}{n!} = \exp(c(s-1))), \tag{1.66}$$

we have

$$G(\pi_c; s) = G(\pi_{c(1)}; s)G(\pi_{c(2)}; s),$$

which implies (1.65) by Lemma 1.7.2. \(^□^)/

Example 1.7.5 (**Relation between gamma and beta distributions**) Let $a, b, c > 0$ and suppose that X, Y, S, T are r.v.'s such that $X, Y, S \in (0, \infty)$, $T \in (0,1)$ and

$$(S,T) = \left(X + Y, \tfrac{X}{X+Y}\right), \text{ i.e., } (X,Y) = (ST, S(1-T)).$$

Then, the following are equivalent:

a) X and Y are independent, $X \approx \gamma_{c,a}$ and $Y \approx \gamma_{c,b}$;

b) S and T are independent, $S \approx \gamma_{c,a+b}$ and $T \approx \beta_{a,b}$.

Remark: The following well-known formula will also be reproduced in the course of the proof of a) $\Rightarrow$ b):

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \tag{1.67}$$

a) $\Rightarrow$ b): It is enough to show that

1) $P\left((S,T) \in I \times J\right) = \gamma_{c,a+b}(I)\beta_{a,b}(J)$ for all intervals $I \subset (0,\infty)$, $J \subset (0,1)$.

We first show that

2) $$P\left((S,T) \in I \times J\right) = \frac{B(a,b)\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\gamma_{c,a+b}(I)\beta_{a,b}(J).$$

Note the following simple equality for $s > 0$:

3) $$\int_{sJ} x^{a-1}(s-x)^{b-1}dx \overset{x=st}{=} s^{a+b-1}\int_J t^{a-1}(1-t)^{b-1}dt = s^{a+b-1}B(a,b)\beta_{a,b}(J),$$

where $sJ = \{sx, \ ; \ x \in J\}$. Let us write $D = \{(x,y) \in (0,\infty)^2 \ ; \ (x+y, \frac{x}{x+y}) \in I \times J\}$. Then,

$$\begin{aligned}
P\left((S,T) \in I \times J\right) &= P((X,Y) \in D) = (\gamma_{c,a} \otimes \gamma_{c,b})(D) \\
&\overset{(1.27)}{=} \frac{c^{a+b}}{\Gamma(a)\Gamma(b)}\int_D x^{a-1}y^{b-1}e^{-c(x+y)}dxdy \\
&\overset{s=x+y}{=} \frac{c^{a+b}}{\Gamma(a)\Gamma(b)}\int_I e^{-cs}ds\int_{sJ} x^{a-1}(s-x)^{b-1}dx \\
&\overset{3)}{=} \frac{B(a,b)c^{a+b}}{\Gamma(a)\Gamma(b)}\int_I s^{a+b-1}e^{-cs}ds\beta_{a,b}(J) \\
&\overset{(1.27)}{=} \frac{B(a,b)\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\gamma_{c,a+b}(I)\beta_{a,b}(J).
\end{aligned}$$

This proves 2). Letting $I = (0,\infty)$ and $J = (0,1)$ in 2), we get

$$1 = \frac{B(a,b)\Gamma(a+b)}{\Gamma(a)\Gamma(b)}, \quad \text{i.e., (1.67).}$$

Finally, plugging this back in 2), we arrive at 1).
a) $\Leftarrow$ b): Let X' and Y' be independent r.v.'s such that $X' \approx \gamma_{c,a}$ and $Y' \approx \gamma_{c,b}$. Then, we know that

$$S' \overset{\text{def}}{=} X' + Y' \text{ and } T' \overset{\text{def}}{=} \tfrac{X'}{X'+Y'} \text{ are independent, } S' \approx \gamma_{c,a+b} \text{ and } T' \approx \beta_{a,b}.$$

This implies that $(S,T) \approx (S',T')$. Therefore,

$$(X,Y) = (ST, S(1-T)) \approx (S'T', S'(1-T')) = (X',Y'),$$

which implies a). \\(^▫^)/

Example 1.7.6 (**Poisson process**) Let X_j ($j \geq 1$) be iid $\approx \gamma_{c,1}$ (cf. (1.27)) and $S_n = X_1 + ... + X_n$. Then, for $t \geq 0$,

$$N_t \overset{\text{def}}{=} \sup\{n \in \mathbb{N} \ ; \ S_n \leq t\} \approx \pi_{ct}, \quad \text{(cf. (1.18))}. \tag{1.68}$$

$(N_t)_{t\geq 0}$ is called the *Poisson process* with the parameter c. N_t has, for example, the following interpretation; S_n is the time when the n-th customer arrives at the COOP cafeteria in a day and N_t is the number of customers who visited the cafeteria up to time t.

Proof: It is enough to prove that

1) $$P(N_t \geq n) = e^{-ct} \sum_{m=n}^{\infty} \frac{(ct)^m}{m!}.$$

We start by computing:

$$\begin{aligned} P(N_t \geq n) &\stackrel{(1.68)}{=} P(S_n \leq t) \stackrel{\text{Example 1.7.5}}{=} \gamma_{c,n}((0,t]) \\ &\stackrel{(1.27)}{=} \frac{c^n}{(n-1)!} \int_0^t x^{n-1} e^{-xc} dx \stackrel{x=y/c}{=} \frac{1}{(n-1)!} \int_0^{ct} y^{n-1} e^{-y} dy. \end{aligned}$$

Thus, we can conclude 1) from:

2) $$\frac{1}{(n-1)!} \int_0^s y^{n-1} e^{-y} dy = e^{-s} \sum_{m=n}^{\infty} \frac{s^m}{m!}, \quad s \geq 0.$$

We prove 2) in the following generalized form:

3) $$\frac{1}{\Gamma(a)} \int_0^s y^{a-1} e^{-y} dy = e^{-s} \sum_{m=0}^{\infty} \frac{s^{a+m}}{\Gamma(a+m+1)}, \quad a > 0, \quad s \geq 0.$$

In fact,

$$\begin{aligned} \text{LHS 3)} &\stackrel{y=s-x}{=} \frac{e^{-s}}{\Gamma(a)} \int_0^s (s-x)^{a-1} e^x dx = \frac{e^{-s}}{\Gamma(a)} \sum_{m=0}^{\infty} \frac{1}{m!} \int_0^s (s-x)^{a-1} x^m dx \\ &\stackrel{x=sz}{=} \frac{e^{-s}}{\Gamma(a)} \sum_{m=0}^{\infty} \frac{s^{a+m}}{m!} B(a, m+1) \stackrel{(1.67)}{=} \text{RHS 3)} \end{aligned}$$

\(^□^)/

Example 1.7.7 ($\star$) Let $X \approx N(0, vI_d)$ ($d \geq 1$, $v > 0$) and $Y \approx \gamma_{c,a}$ ($c, a > 0$) be independent. Then,

$$X/\sqrt{Y} \approx \frac{(2cv)^a \Gamma(a + \frac{d}{2})}{\pi^{d/2} \Gamma(a)} \frac{dx}{(2cv + |x|^2)^{a+\frac{d}{2}}}. \tag{1.69}$$

The right-hand side is the generalized Cauchy distribution, cf. Example 1.2.8. There are two important special cases:

• Let $Z \approx N(0, w)$ ($w > 0$) be independent of X. Then, we see from (1.30) that $Y \stackrel{\text{def}}{=} Z^2 \approx \gamma(\frac{1}{2w}, \frac{1}{2})$ Thus, applying (1.69) with $(c, a) = (\frac{1}{2w}, \frac{1}{2})$, we have that

$$X/|Z| \approx (\sqrt{v/w})\text{-Cauchy distribution.} \tag{1.70}$$

• If $d = 1, X \approx N(0, 1)$, $Z \approx \chi_n^2 = \gamma(1/2, n/2)$ ($n \geq 1$) (cf. Example 1.2.5), and X and Z are independent. Then, $Z/n \approx \gamma(n/2, n/2)$. Thus, by (1.69) with $d = 1$, $v = 1$, $c = a = n/2$,

$$X/\sqrt{Z/n} \approx T_n \quad \text{cf. Example 1.2.8.} \tag{1.71}$$

In statistics, the r.v. on the left-hand saide of (1.71) is used to estimate the population mean, when n (the number of samples) is relatively small.

The proof of (1.69) goes as follows. Let $f : \mathbb{R}^d \to [0,\infty)$ be measurable. Then,

$$
\begin{aligned}
Ef(X/\sqrt{Y}) &= \int_0^\infty P(Y \in dy) \int_{\mathbb{R}^d} P(X \in dx) f(x/\sqrt{y}) \\
&= \frac{c^a}{\Gamma(a)(2\pi v)^{d/2}} \int_0^\infty y^{a-1} e^{-cy} dy \int_{\mathbb{R}^d} \exp\left(-\tfrac{|x|^2}{2v}\right) f(x/\sqrt{y}) dx \\
&= \frac{c^a}{\Gamma(a)(2\pi v)^{d/2}} \int_0^\infty y^{a+\frac{d}{2}-1} e^{-cy} dy \int_{\mathbb{R}^d} \exp\left(-\tfrac{y|z|^2}{2v}\right) f(z) dz \\
&= \frac{c^a}{\Gamma(a)(2\pi v)^{d/2}} \int_{\mathbb{R}^d} f(z) dz \int_0^\infty y^{a+\frac{d}{2}-1} \exp\left(-y\left(c + \tfrac{|z|^2}{2v}\right)\right) dy.
\end{aligned}
$$

We easily see from the definition of the Gamma-function that

$$
\int_0^\infty y^{a+\frac{d}{2}-1} \exp\left(-y\left(c + \tfrac{|z|^2}{2v}\right)\right) dy = \frac{\Gamma\left(a + \frac{d}{2}\right)}{\left(c + \frac{|z|^2}{2v}\right)^{a+\frac{d}{2}}}.
$$

Thus, we conclude that

$$
Ef(X/\sqrt{Y}) = \frac{(2cv)^a \Gamma\left(a + \frac{d}{2}\right)}{\pi^{d/2}\Gamma(a)} \int_{\mathbb{R}^d} \frac{f(z)dz}{(2cv + |z|^2)^{a+\frac{d}{2}}}.
$$

\(^□^)/

Exercise 1.7.1 Let Z be a r.v. defined on a probability space $(\Omega, \mathcal{F}, P)$ such that $Z \approx \mathrm{Bin}(n,p)$. Is it always true that there exist iid $X_j \approx \mathrm{Bin}(1,p)$ $(j = 1, ..., n)$ defined on $(\Omega, \mathcal{F}, P)$ such that $Z = X_1 + ... + X_n$?

Exercise 1.7.2 Let $X = (X_j)_{j=1}^n$ and $S_n = X_1 + \ldots + X_n$, where $X_1, ..., X_n$ are iid, $X_j \approx \mathrm{Bin}(1,p)$ $(j = 1, ..., n)$. Prove the following:
i) $P(X = x | S_n = m) = \binom{n}{m}^{-1}$, regardless of the value of p, for any $m = 0, 1, ..., n$ and $x = (x_j)_{j=1}^n \in \{0,1\}^n$ with $x_1 + ... + x_n = m$.
ii) $\frac{d}{dp} Ef(X) = \frac{1}{p(1-p)} \mathrm{cov}(f(X), S_n)$ for any $f : \{0,1\}^n \to \mathbb{R}$.

Exercise 1.7.3 Let X,Y and Z be r.v.'s with $(X,Y) \approx \gamma_{r,a} \otimes \gamma_{s,b}$. and $Z \approx \beta_{a,b}$. Prove then that

$$
\frac{X}{Y} \approx \frac{s}{r} \frac{Z}{1-Z} \approx \frac{(r/s)^a}{B(a,b)} \frac{x^{a-1} dx}{(1 + rx/s)^{a+b}}.
$$

When $r = a = m/2$ and $s = b = n/2$ $(m, n \in \mathbb{N})$, the above distribution is called the F_n^m *distribution* and is used in statistics.

Hint: Let $(X_1, Y_1) \approx \gamma_{1,a} \otimes \gamma_{1,b}$. Then, $(X,Y) \approx (X_1/r, Y_1/s)$ and $\frac{X_1}{Y_1} = \frac{\frac{X_1}{X_1+Y_1}}{1 - \frac{X_1}{X_1+Y_1}}$. Then use Example 1.7.5.

Exercise 1.7.4 Prove the following extension of Example 1.7.5. Let $X_j \approx \gamma_{c,a_j}$, $j = 1, .., n+1$ be independent r.v.'s and $S \stackrel{\text{def}}{=} X_1 + .. + X_{n+1}$. Then, S and $T \stackrel{\text{def}}{=} (\frac{X_j}{S})_{j=1}^n$ are independent r.v.'s such that $S \approx \gamma_{c, a_1 + .. + a_{n+1}}$ and

$$
T \approx \frac{\Gamma(a_1 + .. + a_{n+1})}{\Gamma(a_1) \cdots \Gamma(a_{n+1})} x_1^{a_1 - 1} \cdots x_n^{a_n - 1} \left(1 - \sum_{j=1}^n x_j\right)^{a_{n+1}-1} dx_1 \cdots dx_n.
$$

Exercise 1.7.5 Let e and U are independent r.v. such that $e \approx \gamma_{1,1}$ and U is uniformly distributed on $(0, 2\pi)$. Prove then that $\sqrt{2e}(\cos U, \sin U) \approx N(0,1) \otimes N(0,1)$.

Exercise 1.7.6 Let $X_i \approx \gamma_{c_i,1}$ ($i = 1, ..., n$ cf. (1.27)) be independent r.v.'s and $M_n = \min_{i=1,\ldots,n} X_i$. Prove then that for any $j = 1, \ldots, n$ and $x \geq 0$,

$$P(M_n = X_j \text{ and } X_j > x) = \frac{c_j}{\sum_{i=1}^n c_i} \exp\left(-x \sum_{i=1}^n c_i\right).$$

In particular, $M_n \approx \gamma_{c_1+\ldots+c_n,1}$

Exercise 1.7.7 (**Thinning of a Poisson r.v.**) Let N be a r.v. with $N \approx \pi_c$ and let $(X_n)_{n\geq 0}$ be i.i.d. with values in a finite set S. We suppose that N and $(X_n)_{n\geq 0}$ are independent. Prove then that $N_s = \sum_{j=0}^N \mathbf{1}\{X_j = s\}$ ($s \in S$) are independent and that $N_s \approx \pi_{p(s)c}$, where $p(s) = P(X_0 = s)$.

Exercise 1.7.8 (**Geometric distribution**) Let $G = \inf\{n \geq 1 \ ; \ X_n = 1\}$, where $(X_n)_{n\geq 1}$ are $\{0,1\}$-valued i.i.d. with $P(X_n = 1) = p$. Then, show that $P(G = n) = p(1-p)^{n-1}$, $EG = 1/p$, and var $G = (1-p)/p$. The distribution of G is called the *p-geometric distribution.*The geometric distribution can be thought of as a discrete analogue of the exponential distribution.

Exercise 1.7.9 (**n-th success in a Bernoulli trial**) Let $(X_k)_{k\geq 1}$ be as in Exercise 1.7.8, $S_k = X_1 + ... + X_k$, and

$$T_0 \equiv 0, \quad T_n = \inf\{k \geq 1 \ ; \ S_k = n\}, \quad n = 1, 2, ...$$

Then, prove the following:
i) $T_n - T_{n-1}$, $n = 1, 2, ...$ are iid with p-geometric distribution.
ii) $P(T_n = m) = \binom{m-1}{n-1} p^n (1-p)^{m-n}, \quad 1 \leq n \leq m.$
iii) $S_k = n$ for $T_n \leq k < T_{n+1}$, $n = 0, 1, ...$
In the Bernoulli trial $(X_k)_{k\geq 1}$, T_n is the time of n-th success, and $T_n - n$ is the number of failures before it. The distribution of the latter is called the (n,p)-*negative binomial distribution.* It follows from ii) above that

$$P(T_n - n = k) = \binom{n+k-1}{k} p^n (1-p)^k, \quad k \in \mathbb{N}.$$

On the other hand, the description of $(S_k)_{k\geq 1}$ in iii) above can be thought of as a discrete-time analogue of Poisson process (Example 1.7.6). This also shows that $(S_k)_{k\geq 1}$ (and hence $(X_k)_{k\geq 1}$) can be recoverd from $(T_n)_{n\geq 0}$.

Exercise 1.7.10 Let $G, \tau_1, \tau_2, ...$ be independent r.v.'s such that $P(G = n) = p(1-p)^{n-1}$ ($n = 1, 2, ...$) and $P(\tau_j \in \cdot) = \gamma_{c,1}$ (cf. (1.27)). Prove then that $P(\tau_1 + ... + \tau_G \in \cdot) = \gamma_{cp,1}$.

1.8 Applications to analysis

The following lemma is a weaker version of the law of large numbers (Theorem 1.10.2). This lemma will be applied to Example 1.8.2 and Example 1.8.3.

Lemma 1.8.1 *Let $I \subset \mathbb{R}$ be an interval, $X_n : \Omega \to I$ $(n \geq 1)$ be such that $X_n \in L^2(P)$, $\mathrm{cov}(X_\ell, X_n) = v\delta_{\ell,n}$, $EX_n = m$ for all $\ell, n \geq 1$. Then, for $S_n = X_1 + ... + X_n$, $f : I \to \mathbb{R}$, and $\delta > 0$,*

$$P\left(\left|\frac{S_n}{n} - m\right| \geq \delta\right) \leq \frac{v}{\delta^2 n}, \tag{1.72}$$

$$E\left|f\left(\frac{S_n}{n}\right) - f(m)\right| \leq \frac{2\|f\|v}{\delta^2 n} + \sup_{\substack{x \in I \\ |x-m|<\delta}} |f(x) - f(m)|, \tag{1.73}$$

where $\|f\| = \sup_{x \in I} |f(x)|$.

Proof: (1.72):

$$P\left(\left|\frac{S_n}{n} - m\right| \geq \delta\right) = P\left(|S_n - mn|^2 \geq \delta^2 n^2\right) \overset{\text{Chebyshev}}{\leq} \frac{\mathrm{var}\,(S_n)}{\delta^2 n^2} \overset{(1.55)}{=} \frac{v}{\delta^2 n}.$$

(1.73): We first observe that

1) $E\left[\left|f\left(\frac{S_n}{n}\right) - f(m)\right| : \left|\frac{S_n}{n} - m\right| \geq \delta\right] \leq 2\|f\|P\left(\left|\frac{S_n}{n} - m\right| \geq \delta\right) \overset{(1.72)}{\leq} \frac{2\|f\|v}{\delta^2 n}.$

On the other hand, it is clear that

2) $E\left[\left|f\left(\frac{S_n}{n}\right) - f(m)\right| : \left|\frac{S_n}{n} - m\right| < \delta\right] \leq \sup_{\substack{x \in I \\ |x-m|<\delta}} |f(x) - f(m)|.$

By 1) and 2), we get (1.73). \(^□^)/

Example 1.8.2 (**Weierstrass' approximation theorem**) Let $I = [0,1]$, $f \in C(I \to \mathbb{R})$ and

$$f_n(p) \overset{\text{def.}}{=} \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} p^k (1-p)^{n-k}.$$

Then,

1) $\qquad f_n \overset{n\to\infty}{\longrightarrow} f$ uniformly on I.

To prove this, we apply Lemma 1.8.1 for $X_n \approx \mathrm{Bin}(1,p)$. Then $S_n \approx \mathrm{Bin}(n,p)$ (Example 1.7.3) and hence

$$f_n(p) = Ef\left(\tfrac{S_n}{n}\right).$$

Since

$$EX_n = p, \text{ and } \mathrm{var}\, X_n = p(1-p) \leq 1/4,$$

we see from (1.73) with $\delta = n^{-1/3}$ that

$$|f_n(p) - f(p)| \le E\left|f\left(\frac{S_n}{n}\right) - f(p)\right| \le \frac{\|f\|}{2n^{1/3}} + \sup_{\substack{x\in I \\ |x-p|<n^{-1/3}}} |f(x) - f(p)|.$$

Since f is uniformly continuous on I, the right-hand side of the above inequality converges to zero uniformly in p, as $n \to \infty$, which proves 1). \(^□^)/

Example 1.8.3 (**Injectivity of the Laplace transform**) Let μ be a Borel signed measure on $[0, \infty)$. Then, the following are equivalent.

a) $\mu = 0$.

b) $\int_{[0,\infty)} e^{-\lambda x} d\mu(x) = 0$ for all $\lambda \ge 0$.

c) $\int_{[0,\infty)} x^k e^{-nx} d\mu(x) = 0$ for all $k \in \mathbb{N}$ and $n \in \mathbb{N}\backslash\{0\}$.

Proof: a) $\Rightarrow$ b): Obvious.
b) $\Rightarrow$ c): By differentiating the identity b) k times in λ, and then setting $\lambda = n \in \mathbb{N}$, we have c).
c) $\Rightarrow$ a): By Lemma 1.3.2, it is enough to prove that

1) $\int_{[0,\infty)} f d\mu = 0$ for $f \in C_{\mathrm{b}}([0, \infty))$,

Let $f \in C_{\mathrm{b}}([0, \infty))$ be arbitrary. We define $f_n : [0, \infty) \to \mathbb{R}$ ($n \in \mathbb{N}$) by

$$f_n(x) = e^{-nx} \sum_{k=0}^{\infty} \frac{(nx)^k}{k!} f\left(\frac{k}{n}\right), \quad x \ge 0.$$

We prove the following approximation:

2) $f_n(x) \stackrel{n\to\infty}{\longrightarrow} f(x)$ for any $x \in [0, \infty)$,

(As is explained in the remark after this proof, the convergence 2) is uniform in $x \in [0, M]$ for any $M > 0$. But, we do not need this fact to prove 1).) To prove 2), we fix $x \ge 0$ and apply Lemma 1.8.1 to $X_n \approx \pi_x$ (cf. (1.18)). Then $S_n \approx \pi_{nx}$ (Example 1.7.4) and hence

$$f_n(x) = Ef(\tfrac{S_n}{n}).$$

Since

$$EX_n = \mathrm{var}\, X_n = x,$$

we see from (1.73) with $\delta = n^{-1/3}$ that

$$|f_n(x) - f(x)| \le E\left|f\left(\frac{S_n}{n}\right) - f(x)\right| \le \frac{2\|f\|x}{n^{1/3}} + \sup_{\substack{y\ge 0 \\ |y-x|<n^{-1/3}}} |f(y) - f(x)|.$$

Since f is continuous, the right-hand side of the above inequality converges to zero as $n \to \infty$, which proves 2).
We now use 2) to prove 1). By multiplying both hands-sides of the identity c) by $\frac{n^k}{k!} f\left(\frac{k}{n}\right)$, and adding over $k \in \mathbb{N}$, we arrive at:

3) $\displaystyle\int_{[0,\infty)} f_n d\mu = 0.$

We obtain 1) from 2) and 3) via the bounded convergence theorem. \\(^□^)/

Remark: The convergence 2) in the proof of Example 1.8.3 is uniform in $x \in [0, M]$ for any $M > 0$. In fact, if $x \in [0, M]$, we see from the above proof that

$$|f_n(x) - f(x)| \leq E\left|f\left(\frac{S_n}{n}\right) - f(x)\right| \leq \frac{2\|f\|M}{n^{1/3}} + \sup_{\substack{y\in[0,M+1]\\|y-x|<n^{-1/3}}} |f(y) - f(x)|.$$

Since f is uniformly continuous on $[0, M + 1]$, the right-hand side of the above inequality converges to zero uniformly in $x \in [0, M]$ as $n \to \infty$. Note also that the function f_n can naturally be extended as a holomorhic function on $\mathbb{C}$. These prove that, for any $f \in C_{\rm b}([0, \infty))$, there exists a sequnece of holomorphic functions $f_n : \mathbb{C} \to \mathbb{C}$ $(n \in \mathbb{N})$ which converges uniformly to f on any bounded subset of $[0, \infty)$.

Exercise 1.8.1 (**Weierstrass' approximation theorem in higher dimensions**) Let $I = [0, 1]^d$ and $f \in C(I \to \mathbb{R})$. Prove that there exist polynomials $f_n : \mathbb{R}^d \to \mathbb{R}$ $(n \geq 1)$ such that $\lim\limits_{n\to\infty} \max\limits_{p\in I} |f_n(p) - f(p)| = 0$. Hint: Fix $p = (p_\nu)_{\nu=1}^d \in I$ and $n \in \mathbb{N}\backslash\{0\}$ for a moment. Let $S_n = (S_{n,\nu})_{\nu=1}^d$, where $S_n^1, ..., S_n^d$ are independent r.v.'s with $P(S_n^\nu = r) = \binom{n}{r} (p_\nu)^r(1 - p_\nu)^{n-r}$ $(0 \leq r \leq n,\ 1 \leq \nu \leq d)$. Then, $P(S_n = x) = \prod_{\nu=1}^d \binom{n}{x^\nu} (p_\nu)^{x^\nu}(1 - p_\nu)^{n-x^\nu}$.

Exercise 1.8.2 ($\star$) Show the following: (i) For any $n \in \mathbb{N}\backslash\{0\}$ and $z \in \mathbb{C}\backslash\{0\}$,

$$Q_n(z) \stackrel{\rm def.}{=} \frac{1}{n}\frac{2 - z^n - z^{-n}}{2 - z - z^{-1}} = 1 + \frac{1}{n}\sum_{\substack{1\leq \ell,m<n\\ \ell\neq m}} z^{\ell-\mu}. \tag{1.74}$$

where we define $Q_n(1) = n$. Hint: Let $s_n(z) = 1 + z + ... + z^{n-1}$. Then,

$$2 - z^n - z^{-n} = (1 - z^n)(1 - z^{-n}) = (1 - z)(1 - z^{-1})s_n(z)s_n(z^{-1}).$$

(ii) $F_n(\theta) \stackrel{\rm def.}{=} Q_n(e^{2\pi\mathbf{i}\theta}) \geq 0$ for all $\theta \in \mathbb{R}$, $\displaystyle\int_0^1 F_n(\theta)d\theta = 1.$

These show that F_n is a density of a probability measure on $[0, 1]$ with respect to the Lebesgue measure. F_n is called the *Fejér kernel*.

Exercise 1.8.3 ($\star$) (**Uniform approximation by trigonometric polynomials**) A function $Q : \mathbb{R} \to \mathbb{C}$ is called a *trigonometric polynomial*, if it is a finite linear combination of $\{\theta \mapsto e^{2\pi\mathbf{i}n\theta}\}_{n\in\mathbb{Z}}$. Let $f \in C(\mathbb{R} \to \mathbb{C})$ be of the period 1 and

$$f_n(\theta) = \int_0^1 f(\theta - \varphi)F_n(\varphi)d\varphi,$$

where F_n is the Fejér kernel (Exercise 1.8.2). Prove then that f_n is a trigonometric polynomial and that

$$\lim_{n\to\infty} \sup_{0\leq\theta\leq1} |f_n(\theta) - f(\theta)| = 0.$$

Hint: $f_n(\theta) = \int_0^1 f(\varphi)F_n(\theta - \varphi)d\varphi$ by the periodicity. Then, use (1.74) to see that f_n is a trigonometric polynomial.

1.9 Decimal Fractions

We begin by introducing the notation we use in this subsection. Let $q \geq 2$ be an integer. Recall that, for each $t \in (0,1]$, there exists a unique sequence $d_n(t) \in \{0, \ldots, q-1\}$ $(n \geq 1)$ such that

$$t = \sum_{n \geq 1} \frac{d_n(t)}{q^n} \text{ and } \sum_{n \geq 1} d_n(t) = \infty. \tag{1.75}$$

Thus, $d_n(t)$ stands for the n-th digit in the q-adic expansion of the number t, where the expansion is unique, thanks to the second condition of (1.75). As we describe below, the functions $d_1, \ldots, d_n$ are in correspondence to the partition $\{I_{s_1,\ldots,s_n}\}_{s_1,\ldots,s_n=0}^{q-1}$ of the interval $(0,1]$ into q^n smaller intervals of length q^{-n}. For each $s = 0, \ldots, q-1$,

$$I_s \stackrel{\text{def}}{=} \{t \in (0,1] \; ; \; d_1(t) = s\} = \frac{s}{q} + \left(0, \frac{1}{q}\right].$$

Similarly, for each $n \geq 1$ and $s_1, ..., s_n \in \{0, \ldots, q-1\}$,

$$I_{s_1,\ldots,s_n} \stackrel{\text{def}}{=} \{t \in (0,1] \; ; \; d_j(t) = s_j, \; 1 \leq \forall j \leq n\} = \sum_{j=1}^n \frac{s_j}{q^j} + \left(0, \frac{1}{q^n}\right]. \tag{1.76}$$

Example 1.9.1 (**Decimal fractions are i.i.d.**) Suppose that $(\Omega, \mathcal{F}, P)$ is a probability space and that $U : \Omega \to (0,1)$ is a r.v. with the uniform distribution on $(0,1)$. Then,

$$\{d_n(U)\}_{n \geq 1} \text{ are i.i.d. with } P(d_n(U) = s) = q^{-1}, \; s \in \{0, \ldots, q-1\}. \tag{1.77}$$

Proof: We see from the definition above that for all $s_1, ..., s_n \in \{0, \ldots, q-1\}$,

$$\bigcap_{j=1}^n \{d_j(U) = s_j\} \stackrel{(1.76)}{=} \{U \in I_{s_1 \cdots s_n}\}$$

and hence that

1) $P\left(\bigcap_{j=1}^n \{d_j(U) = s_j\}\right) = P(U \in I_{s_1 \cdots s_n}) = |I_{s_1 \cdots s_n}| = q^{-n}.$

In particular, for any $n \geq 1$,

2) $P\left(d_n(U) = s_n\right) = \displaystyle\sum_{s_1,\ldots,s_{n-1}=0}^{q-1} P\left(\bigcap_{j=1}^n \{d_j(U) = s_j\}\right) \stackrel{1)}{=} \sum_{s_1,\ldots,s_{n-1}=0}^{q-1} q^{-n} = q^{-1}.$

We conclude (1.77) from 1) and 2). \(^□^)/

Example 1.9.2 (**Cantor function**) We give an example of nondecreasing continuous function from $[0,1]$ onto $[0,1]$, whose associated Stieltjes measure is singular with respect to the Lebesgue measure. Let $q > q_0 \geq 2$ be integers, and S_0 be a subset of $\{0, \ldots, q-1\}$ with q_0 elements. We define

$$X = \sum_{n \geq 1} \frac{X_n}{q^n}, \quad F(t) = P(X \leq t) \;\; (0 \leq t \leq 1),$$

where $X_n : \Omega \to S_0$ $(n \geq 1)$ are i.i.d. with $P(X_n = s) = 1/q_0$ $(s \in S_0)$. Note then that the law $\mu \in \mathcal{P}([0,1])$ of the r.v. X is the Stieltjes measure associated to the function F. We prove the following

a) F is nondecreasing, continuous, $F(0) = 0$, $F(1) = 1$.

b) The measue μ is singular with respect to the Lebesgue measure λ, as can be described more precisely as follows. Let

$$C = \bigcap_{n \ge 1} \bigcup_{s_1,\ldots,s_n \in S_0} \overline{I_{s_1,\ldots,s_n}},$$

where $\{I_{s_1,\ldots,s_n}\}_{s_1,\ldots,s_n=0}^{q-1}$ $(n \ge 1)$ are the partition of $(0,1]$ defined by (1.76). Then, $\mu(C) = 1$ and $\lambda(C) = 0$.

Proof: a) We only need to prove the continuity, since the other properties can easily be seen from the definition. It is also not difficult to see that

$$\begin{aligned} F(t+) &= F(t) \ \text{ for } t \in [0,1), \\ F(t) - F(t-) &= P(X = t) \ \text{ for } t \in (0,1]. \end{aligned}$$

Thus, it is enough to verify that
1) $\quad P(X = t) = 0$ for all $t \in (0,1]$.
To do so, let us note the following.
2) $\quad P\left(\bigcap_{n \ge 1}\{X_n = d_n(X)\}\right) = 1$.
Indeed, $D \stackrel{\text{def}}{=} \{\sum_{n=1}^{\infty} X_n = \infty\} \subset \bigcap_{n \ge 1}\{X_n = d_n(X)\}$, thanks to the uniqueness of the digits in q-adic expansion (1.75). Moreover, $P(D) = 1$, since

$$P(D^{\mathsf{c}}) = P\left(\bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \{X_n = 0\}\right) = 0.$$

This proves 2). We conclude 1) from 2) as follows. By the uniqueness of the digits in q-adic expansion (1.75), $X = t$ if and only if $d_n(X) = d_n(t)$ for all $n \ge 1$. Therefore,

$$P(X = t) = P\left(\bigcap_{n \ge 1}\{d_n(X) = d_n(t)\}\right) \stackrel{2)}{=} P\left(\bigcap_{n \ge 1}\{X_n = d_n(t)\}\right) = 0.$$

b) Since $C_n \stackrel{\text{def}}{=} \bigcup_{s_1,\ldots,s_n \in S_0} \overline{I_{s_1,\ldots,s_n}} \searrow C$ as $n \to \infty$, it is enough to show that $\mu(C_n) = 1$ for all $n \ge 1$ and $\lambda(C_n) \stackrel{n\to\infty}{\longrightarrow} 0$.

$$\begin{aligned} 1 &\ge \mu(C_n) \ge \mu\left(\bigcup_{s_1,\ldots,s_n \in S_0} I_{s_1,\ldots,s_n}\right) \\ &\stackrel{(1.3)}{=} P\left(X \in \bigcup_{s_1,\ldots,s_n \in S_0} I_{s_1,\ldots,s_n}\right) \stackrel{(1.76)}{=} P\left(\bigcap_{j=1}^{n}\{d_j(X) \in S_0\}\right) \\ &\stackrel{2)}{=} P\left(\bigcap_{j=1}^{n}\{X_j \in S_0\}\right) \stackrel{(\text{definition of } X_n)}{=} 1. \end{aligned}$$

On the other hand,

$$\lambda(C_n) \le \sum_{s_1,\ldots,s_n \in S_0} \lambda(\overline{I_{s_1,\ldots,s_n}}) = \sum_{s_1,\ldots,s_n \in S_0} \lambda(I_{s_1,\ldots,s_n}) = q_0^n \cdot \frac{1}{q^n} \stackrel{n\to\infty}{\longrightarrow} 0.$$

\(^□^)/

Remark See Example 1.9.3 a) for an alternative expression of the function F. Also, under an additional assumption that $0 \in S_0$, the set C is identified with the support of the measure μ (cf. Remark after Example 1.9.3). For $q = 3$ and $S_0 = \{0, 2\}$, the set C is the *Cantor's middle thirds set*, and the function F is the *Cantor function*.

Example 1.9.3 ($\star$) We retain the setting of Example 1.9.2. We prove the following addtional properties.

a) For $t \in (0, 1]$, $F(t) = \sum_{n \geq 1} \frac{d_{0,n}(t)}{q_0^n}$, where $d_{0,n}(t) = |S_0 \cap [0, d_n(t))|$.

Suppose in addition that $S_0 \ni 0$. Then,

b) The set C has no isolated point. To put it more precisely, let

$$C_0 = \bigcap_{n \geq 1} \bigcup_{s_1,\ldots,s_n \in S_0} I_{s_1,\ldots,s_n}.$$

Then, $\forall t \in C$, $\exists \{t_N\}_{N \geq 1} \subset C_0 \backslash \{t\}$, $t_N \stackrel{N \to \infty}{\longrightarrow} t$.

c) For any $t \in C$, either t is a point of strict increase of F to the right ($\exists\, t_1 \in (t, 1]$, $\forall s \in (t, t_1]$, $F(t) < F(s)$), or t is a point of strict increase of F from the left ($\exists\, t_1 \in [0, t)$ $\forall\, s \in [t_1, t)$, $F(s) < F(t)$).

Proof: a) Note that

$$\{X < t\} = \bigcup_{n \geq 1} \{X_j = d_j(t) \text{ for } j < n \text{ and } X_n < d_n(t)\}.$$

Since $P(X = t) = 0$ as is shown in the proof of Example 1.9.2 a), we have

$$\begin{aligned} F(t) &= P(X < t) = \sum_{n \geq 1} P\left(X_j = d_j(t) \text{ for } j < n \text{ and } X_n < d_n(t)\right) \\ &= \sum_{n \geq 1} P\left(X_j = d_j(t) \text{ for } j < n\right) P\left(X_n < d_n(t)\right) \\ &= \sum_{n \geq 1} \frac{1}{q_0^{n-1}} \cdot \frac{d_{0,n}(t)}{q_0} = \sum_{n \geq 1} \frac{d_{0,n}(t)}{q_0^n}. \end{aligned}$$

b) Case 1, $t \in C_0$: By (1.76), $t = \sum_{n \geq 1} \frac{s_n}{q^n}$, where $s_n \in S_0$ for all $n \geq 1$ and $\sum_{n \geq 1} s_n = \infty$. For each $N \geq 1$, we choose $s'_N \in S_0 \backslash \{s_N\}$ ($\neq \emptyset$ since $q_0 \geq 2$) define $t_N = \sum_{j \geq 1} s_j^{(N)}/q^j$ ($N \geq 1$), where

$$s_j^{(N)} = \begin{cases} s_j & (j \neq N), \\ s'_N & (j = N). \end{cases}$$

Then, $\{s_j^{(N)}\}_{j \geq 1} \subset S_0$, $\sum_{j \geq 1} s_j^{(N)} = \infty$, and hence $\{t_N\}_{N \geq 1} \subset C_0 \backslash \{t\}$. Finally it is clear that $t_N \stackrel{N \to \infty}{\longrightarrow} t$.

Case 2, $t \in C \backslash C_0$: In this case, $t = \sum_{j=1}^{n} \frac{s_j}{q^j}$ for some $\{s_j\}_{j=1}^{n} \subset S_0$. We choose $s \in S_0 \backslash \{0\}$ ($\neq \emptyset$, since $q_0 \geq 2$), and define $t_N = \sum_{j \geq 1} s_j^{(N)}/q^j$ ($N \geq 1$), where

$$s_j^{(N)} = \begin{cases} s_j & (1 \leq j \leq n), \\ 0 & (n < j < n + N), \\ s & (j \geq n + N). \end{cases}$$

Then, $\{s_j^{(N)}\}_{j\ge 1} \subset S_0$, since $0 \in S_0$. Moreover, $\sum_{j\ge 1} s_j^{(N)} = \infty$, and hence $\{t_N\}_{N\ge 1} \subset C_0 \backslash \{t\}$. Finally it is clear that $t_N \stackrel{N\to\infty}{\longrightarrow} t$.
c) We start by observing that
3) $\displaystyle\sum_{n\ge 1} d_{0,n}(t) = \infty$ for all $t \in (0,1]$.
Indeed, $d_n(t) \ge 1$ implies that $S_0 \cap [0, d_n(t)) \ni 0$ (since $S_0 \ni 0$), and hence that $d_{0,n}(t) \ge 1$. Therefore, 1) follows from that $\sum_{n\ge 1} d_n(t) = \infty$.

We first prove that
4) F is strictly increasing on C_0.
Since F is nondecreasing, it is enough to prove that F is injective on the set C_0. To do so, suppose that $s, t \in C_0$ and $F(s) = F(t)$. Then, it follows from a) and 4) that $d_{0,n}(s) = d_{0,n}(t)$ for all $n \ge 1$. By the definition of $d_{0,n}$, this implies that for all $n \ge 1$,

$$S_0 \cap [d_n(s) \wedge d_n(t), d_n(s) \vee d_n(t)) = \emptyset.$$

However, since $d_n(s) \wedge d_n(t) \in S_0$, the above is possible only when $d_n(s) \wedge d_n(t) = d_n(s) \vee d_n(t)$, i. e., $d_n(s) = d_n(t)$. Therefore we have $s = t$.

Suppose that $t \in C$. Then, by b1) and b2), either there exists a decreasing sequence $t_1 > t_2 > \ldots$ in C_0 which converges to t, or there exists an increasing sequence $t_1 < t_2 < \ldots$ in C_0 which converges to t. In the former case, it follows from 4) that $F(t) < F(s)$ for all $s \in (t, t_1]$. Similarly, it follows in the latter case as well that $F(s) < F(t)$ for all $s \in [t_1, t)$. \(^□^)/

Remark If $0 \in S_0$, then, $\mathrm{supp}(\mu) = C$. Indeed, it follows from Example 1.9.2 b) that $\mathrm{supp}(\mu) \subset C$, whereas the opposite inclusion follows from Example 1.9.3 c).

Example 1.9.4 ($\star$) Construction of a sequence of independent random variables with discrete state spaces: Let $\mu_n \in \mathcal{P}(S_n, \mathcal{B}_n)$ $(n \ge 1)$ be a sequence of probability measures, where for each $n \ge 1$, S_n is a countable set and $\mathcal{B}_n$ is the collection of all subsets in S_n. We will construct a sequence $X_n : (\Omega, \mathcal{F}) \to (S_n, \mathcal{B}_n)$ of independent r.v.'s such that $X_n \approx \mu_n$ for all $n \ge 1$.

The construction is just a slight extension of Example 1.9.1. We first construct a sequence $I_{s_1 \cdots s_n}$ of sub-intervals of $[0,1)$ inductively as follows, where $n = 1, \cdots$ and $(s_1, \ldots, s_n) \in S_1 \times \cdots \times S_n$. We split $[0,1)$ into disjoint intervals $\{I_s\}_{s\in S_1}$ with length $|I_s| = \mu_1(s)$ for each $s \in S_1$. Suppose that we have disjoint intervals $I_{s_1\cdots s_{n-1}}$ such that $|I_{s_1\cdots s_{n-1}}| = \mu_1(s_1)\cdots\mu_{n-1}(s_{n-1})$ for $(s_1, \ldots, s_{n-1}) \in S_1 \times \cdots \times S_{n-1}$. We then split each $I_{s_1\cdots s_{n-1}}$ into disjoint intervals $\{I_{s_1\cdots s_{n-1}s_n}\}_{s_n\in S_n}$ so that $|I_{s_1\cdots s_{n-1}s_n}| = \mu_1(s_1)\cdots\mu_{n-1}(s_{n-1})\mu_n(s_n)$ for each $s_n \in S_n$. We now define

$$X_n(\omega) = s \quad \text{if } X(\omega) \in \textstyle\bigcup_{s_1,\ldots,s_{n-1}} I_{s_1\cdots s_{n-1}s}.$$

We see from the definition that

$$\bigcap_{j=1}^n \{\omega\ ;\ X_j(\omega) = s_j\} = \{\omega\ ;\ X(\omega) \in I_{s_1\cdots s_n}\}.$$

and hence that

1) $$P\left(\bigcap_{j=1}^n \{X_j = s_j\}\right) = |I_{s_1\cdots s_n}| = \mu_1(s_1)\cdots\mu_n(s_n).$$

We conclude from 1) that $(X_n)_{n\geq 1}$ are independent and that $X_n \approx \mu_n$ (cf. Exercise 1.5.3). \(^□^)/

Exercise 1.9.1 Referring to Example 1.9.2 with $S_0 = 0, q-1$, show the following.
(i) $F(t) = \sum_{n\geq 1} \frac{d_n(t)\wedge 1}{2^n}$ for $t \in (0,1]$. Hint: $\{X < t\} = \bigcup_{n\geq 1}\{X_j = d_j(t)$ for $j < n$ and $X_n < d_n(t)\}$.
(ii) F is strictly increasing on the set C.

1.10 The Law of Large Numbers

Let $\{X_n\}_{n\geq 1}$ be the outcome of independent coin tossings;

$$X_n = \begin{cases} 1 & \text{if the coin falls head by } n\text{-th toss,} \\ 0 & \text{if the coin falls tail by } n\text{-th toss.} \end{cases}$$

Then, $S_n = X_1 + \ldots + X_n$ is the number of tosses by which the coin falls head. For this reason, one would vaguely expect that

$$\frac{S_n}{n} \longrightarrow \frac{1}{2}(= EX_1), \quad \text{as } n \nearrow \infty. \tag{1.78}$$

The law of large numbers we will discuss in this section gives a mathematical justification for this intuition. However, here is one thing we should be careful about; there do exist exceptional events on which (1.78) fails, for example,

$$\bigcap_{n\geq 1}\{X_n = 0\} \ \text{ or } \ \bigcap_{n\geq 1}\{X_n = 1\}.$$

We first formulate a notion which is used to exclude such exceptions.

- Let $(\Omega, \mathcal{F}, P)$ be a probability space in what follows.

Definition 1.10.1 Let $A = \{\omega \in \Omega\ ;\ .....\} \subset \Omega$.

► We say "..... *almost surely*" ("..... *a.s.*" for short) if A^{c} is a null set.
Therefore, "almost surely" ("a.s.") just synonymizes "almost everywhere" ("a.e.") in measure theory.

Theorem 1.10.2 (The Law of Large Numbers) *Let $S_n = X_1 + ... + X_n$, where $\{X_n\}_{n\geq 1}$ are i.i.d. with $E|X_n| < \infty$. Then,*

$$\frac{S_n}{n} \stackrel{n\to\infty}{\longrightarrow} EX_1, \quad P\text{-a.s.} \tag{1.79}$$

Before proving Theorem 1.10.2, let us make a small (and useful) detour:

Lemma 1.10.3 (the first Borel-Cantelli lemma) *Let $X_n \geq 0$, $n \geq 1$ be r.v.'s*

$$\sum_{n\geq 1} EX_n < \infty \implies \sum_{n\geq 1} X_n < \infty, \ \text{a.s.} \implies X_n \stackrel{n\nearrow\infty}{\longrightarrow} 0, \ \text{a.s.} \tag{1.80}$$

Proof: a) Let $X = \sum_{n \geq 1} X_n$. Then,

$$EX \stackrel{\text{Fubini}}{=} \sum_{n \geq 1} EX_n < \infty.$$

Therefore $X < \infty$, a.s., which implies that $X_n \stackrel{n \nearrow \infty}{\longrightarrow} 0$, a.s. \(^□^)/

Here, we give a proof of Theorem 1.10.2 in a special case $X_i \in L^4(P)$, which is much simpler to prove and is enough in many applications. The proof for the general case is presented in Section 8.9. See also Exercise 1.10.6 below to see what happens if we do not assume $E|X_n| < \infty$.

Proof of Theorem 1.10.2 in a special case $X_i \in L^4(P)$: By considering $X_n - EX_n$ instead of X_n, we may assume that $EX_n \equiv 0$. Then, by (1.80), it is enough to prove that

1) $$\sum_{n \geq 1} E\left[S_n^4\right]/n^4 < \infty.$$

We have

2) $$E[S_n^4] = \sum_{i,j,k,\ell=1}^{n} E[X_i X_j X_k X_\ell] = \sum_{i=1}^{n} E[X_i^4] + 6 \sum_{1 \leq r < s \leq n} E[X_r^2] E[X_s^2].$$

Here is an explanation for the second equality of 2). The only terms in $\sum_{i,j,k,\ell=1}^{n}$ that do not vanish are those of the form either

- $E[X_i^4]$ $(i = 1, \ldots n)$, or
- $E[X_r^2 X_s^2] = E[X_r^2]E[X_s^2]$ $(1 \leq r < s \leq n)$. For given r and s, there are $\binom{4}{2} = 6$ possibility for (i, j, k, ℓ) such that two among them are r and the other are s.

Note also that there is a constant C such that

3) $$E[X_m^2]^2 \leq E[X_m^4] \leq C, \quad m = 1, 2, ...$$

Now, 1) follows from 2)–3), since

$$E[S_n^4] \stackrel{2\text{–}3)}{\leq} Cn + 3Cn(n-1) \leq 4Cn^2.$$ \(^□^)/

Example 1.10.4 (**Almost all numbers are normal.**) Let U be a r.v. with uniform distribution on $(0, 1)$ and $q \geq 2$ be integer. Let also $d_n(U) \in \{0, ..., q-1\}$ $(n \geq 1)$ be the digits of U in its q-adic expansion defined by (1.75). Then, *Borel's theorem* asserts that,

1) Almost surely, each number $s = 1, \ldots, q-1$ appears in $(d_n(U))_{n \geq 1}$ with equal frequency.

This will be formulated and proved as follows. We know from Example 1.9.1 that the digits $d_n(U)$ $(n \geq 1)$ are i.i.d. with $P(X_n = s) = 1/q$, $s = 0, \ldots, q-1$. We now fix any s and set $X_n = \mathbf{1}\{d_n(U) = s\}$. Then, X_n $(n \geq 1)$ are i.i.d. $\approx \mathrm{Bin}(1, 1/q)$ and hence $EX_n = 1/q$. Thus, by Theorem 1.10.2,

$$\frac{(\text{the number of } k = 1, \ldots, n \text{ with } d_k(U) = s)}{n} = \frac{X_1 + ... + X_n}{n} \stackrel{n \to \infty}{\longrightarrow} \frac{1}{q}, \quad P\text{-a.s.}$$

\(^□^)/

Example 1.10.5 (**Laws of distinct i.i.d's are mutually singular**.) Let $(S, \mathcal{B})$ be a measurable space, $\mu_1, \mu_2 \in \mathcal{P}(S, \mathcal{B})$, and $\mu_1 \neq \mu_2$. Then, for any infinite set Λ, the product measures $P_j = \otimes_{\lambda \in \Lambda} \mu_j$ $(j = 1, 2)$ are mutually singular.
Proof: Since $\mu_1 \neq \mu_2$, there exists $B \in \mathcal{B}$ such that $\mu_1(B) \neq \mu_2(B)$. Since Λ is an infinite set, we can choose an sequence $\Lambda_1 \subset \Lambda_2 \subset \ldots \subset \Lambda$ such that $|\Lambda_n| = n$ $(n \geq 1)$. We consider the following set.

$$C_j = \left\{ x = (x_\lambda)_{\lambda \in \Lambda} \in \prod_{\lambda \in \Lambda} S \; ; \; \frac{1}{n} \sum_{\lambda \in \Lambda_n} \mathbf{1}_B(x_\lambda) \stackrel{n \to \infty}{\longrightarrow} \mu_j(B) \right\}, \quad (j = 1, 2).$$

Under the measure P_j, $\{\mathbf{1}_B(x_\lambda)\}_{\lambda \in \Lambda}$ are i.i.d. with mean $\mu_j(B)$. Thus, it follows from Theorem 1.10.2 that $P_j(C_j) = 1$. Since $C_1 \cap C_2 = \emptyset$, P_1 and P_2 are mutually singular. \(^□^)/

Complement to section 1.10

Proposition 1.10.6 ($\star$)(**the second Borel-Cantelli lemma**) *Suppose that $X_n \geq 0$, $n \geq 1$ are independent r.v.'s and that there exists a constant M such that*

$$\sup_{n \geq 1} X_n \leq M, \quad a.s.$$

Then,

$$\sum_{n \geq 1} EX_n = \infty \implies \sum_{n \geq 1} X_n = \infty, \quad a.s. \tag{1.81}$$

Proof: We may assume that $M = 1/2$ (Consider $X_n/(2M)$, if necessary). We note that

1) $1 - x \leq e^{-x}$ for $x \geq 0$,

2) $e^{-2x} \leq 1 - x$ for $x \in [0, 1/2]$.

We have

3) $$E\left[\prod_{j=1}^{n}(1 - X_j)\right] \stackrel{(1.53)}{=} \prod_{j=1}^{n}(1 - EX_j) \stackrel{1)}{\leq} \exp\left(-\sum_{j=1}^{n} EX_j\right).$$

Letting $n \to \infty$ in 3), and applying the bounded convergence to the left-hand side,

$$E\left[\prod_{j=1}^{\infty}(1 - X_j)\right] \stackrel{3)}{\leq} \exp\left(-\sum_{j=1}^{\infty} EX_j\right) = 0,$$

hence

4) $$\prod_{j=1}^{\infty}(1 - X_j) = 0, \quad a.s.$$

On the other hand,

5) $$\exp\left(-2\sum_{j=1}^{\infty} X_j\right) = \prod_{j=1}^{\infty} \exp\left(-2X_j\right) \stackrel{2)}{\leq} \prod_{j=1}^{\infty}(1 - X_j).$$

We conclude from 4) and 5) that $\sum_{j=1}^{\infty} X_j = \infty$, a.s. \\(^□^)/

Exercise 1.10.1 Let $X, Y, X_n, Y_n \in L^1(P)$ $(n \in \mathbb{N})$ be such that $X_n \leq Y_n$ a.s. $(\forall n \in \mathbb{N})$ and that $X_n \to X$, $Y_n \to Y$ in probability. Prove then that $X \leq Y$ a.s.

Exercise 1.10.2 (**Shannon's theorem**) Let S be a finite set and $\mu \in \mathcal{P}(S)$ be such that $0 < \mu(s) < 1$ for all $s \in S$, we define the *entropy* $H(\mu)$ of μ by

$$H(\mu) = -\sum_{s \in S} \mu(s) \log \mu(s) > 0.$$

Let $\{X_n\}_{n \geq 1}$ be S-valued i.i.d. $\approx \mu$. Prove that

$$\left(\prod_{j=1}^{n} \mu(X_j) \right)^{1/n} \stackrel{n \to \infty}{\longrightarrow} e^{-H(\mu)}, \quad P\text{-a.s.}$$

Let us interpret S as the set of letters. Then, the above result says that the probability $\prod_{j=1}^{n} \mu(X_j)$ of almost all randomly generated sentence $X_1 X_2 .... X_n$ decays like $e^{-nH(\mu)}$ as $n \nearrow \infty$.

Exercise 1.10.3 (**LLN for renewal processes**) Let $N_t = \sup\{n \in \mathbb{N}\ ;\ T_n \leq t\}$, where $\{T_n - T_{n-1}\}_{n \geq 1}$ are positive r.v.'s with $T_0 \equiv 0$ and $ET_n < \infty$ for all n (cf. Example 1.7.6 for a special case). Prove then the following.
i) $N_\infty \stackrel{\text{def.}}{=} \lim_{t \nearrow \infty} N_t = \infty$, P-a.s.
Hint: $P(N_\infty < \infty) = P(\cup_{\ell \geq 1} \cap_{m \geq 1} \{N_m < \ell\})$ and $\{N_m < \ell\} \subset \{m < T_{\ell+1}\}$.
ii) If $\{T_n - T_{n-1}\}_{n \geq 1}$ are i.i.d., then $\lim_{t \nearrow \infty} N_t / t = 1/ET_1$, P-a.s.
Hint: $T_{N_t} \leq t < T_{N_t+1}$ and $\lim_{t \nearrow \infty} T_{N_t}/N_t = ET_1$ by Theorem 1.10.2.

Exercise 1.10.4 ($\star$) Let $q \geq 2$ be an integer and $\{p(s)\}_{s=0}^{q-1} \subset [0,1)$ be such that $\sum_{s=0}^{q-1} p(s) = 1$. For an i.i.d. $X_n \in \{0, \ldots, q-1\}$ $(n \geq 1)$, with $P(X_1 = s) = p(s)$ $(0 \leq s \leq q-1)$, we denote by μ the law of the r.v.$X = \sum_{n \geq 1} \frac{X_n}{q^n}$. Then, prove the following. **i)** If $p(s) \equiv 1/q$, then μ is the Lebesgue measure on $[0,1]$, **ii)** If $p(s) \not\equiv 1/q$, then μ is singular with respect to the Lebesgue measure. **Hint** Look at the set

$$C = \left\{ t \in (0,1]\ ;\ \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}\{d_k(t) = s\} \stackrel{n \to \infty}{\longrightarrow} p(s), \ \ 0 \leq \forall s \leq q-1 \right\},$$

where, for each $t \in (0,1]$, $d_n(t) \in \{0, 1, \ldots, q-1\}$ $(n \geq 1)$ is the unique sequence such that $t = \sum_{n \geq 1} \frac{d_n(t)}{q^n}$ and $\sum_{n \geq 1} d_n(t) = \infty$.

Remark Exercise 1.10.4 ii) shows that the function $F(t) = \mu([0,t])$ $(0 \leq t \leq 1)$ is singular with respect to the Lebesgue measure. If $q = 3$ and $p(0) = p(2) = 1/2$, then, F is the Cantor function (cf. Example 1.9.2). On the other hand, if $q = 2$ and $p(0) \neq p(1)$, then, F is called the *de Rham's singular function*.

Exercise 1.10.5 (⋆) (**functional equation which characterizes the generalized Cantor functions**) Referring to Exercise 1.10.4,consider the following functional equation for $f : [0,1] \to \mathbb{R}$.

$$f(1) = 1, \quad f\left(\frac{s+t}{q}\right) = \begin{cases} p(0)f(t), & (s=0), \\ p(0) + \cdots + p(s-1) + p(s)f(t), & (s = 1, \dots, q-1) \end{cases} \quad t \in [0,1].$$

Prove that F is the unique right-continuous solution to the above functional equation. Hint: To show that F is a solution, note that $\{X \le \frac{s+t}{q}\} = \{X_1 < s\} \cup \{X_1 = s, \ \sum_{n \ge 1} \frac{X_{n+1}}{q^n} \le t\}, \ \ s \in S$.

Exercise 1.10.6 (⋆) Let $S_n = X_1 + ... + X_n$, where $\{X_n\}_{n \ge 1}$ are i.i.d.
i) (**Infinite mean**) Suppose that $E[X_n^+] = \infty$ and $E[X_n^-] < \infty$. Prove then that $\frac{S_n}{n} \stackrel{n\to\infty}{\longrightarrow} \infty$ a.s. Hint: $X_n \wedge m \in L^1(P)$ for any fixed $m \in (0, \infty)$.
ii) (**Indefinite mean**) Suppose that $E[X_n^\pm] = \infty$. Prove then that $P(S_n/n \text{ converges}) = 0$. Hint: Use Proposition 1.10.6 to show that $\sum_{n \ge 1} \mathbf{1}\{X_n > n\} = \infty$, a.s. Then, note that $\frac{S_{n+1}}{n+1} - \frac{S_n}{n} = \frac{X_{n+1}}{n+1} - \frac{S_n}{n(n+1)}$.

1.11 (⋆) Ergodic theorems

The presentation of this subsection is based on [Dur95] and [Wal82].

Definition 1.11.1 Let $(\Omega, \mathcal{F}, P)$ be a probability space, and $T : \Omega \to \Omega$ be a measurable map.

► A r.v. $X : \Omega \to \mathbb{R}$ is said to be T**-invariant** if $X \circ T = X$, a.s. A event $A \in \mathcal{F}$ is said to be T-invariant if $\mathbf{1}_A$ is T-invariant. The totality of T-invariant events is denoted by $\mathcal{I}$.

► The map T is said to be P**-preserving** if $P \circ T^{-1} = P$, meaning that $P(T^{-1}A) = P(A)$ for all $A \in \mathcal{F}$.

► The map T is said to be P**-ergodic** if it is P-preserving and

$$X \in L^\infty(P), \ \ X \circ T = X, \text{ a.s.} \implies X = EX \text{ a.s.} \tag{1.82}$$

The main purpose of this subsection is to prove:

Theorem 1.11.2 (**Birkhoff Ergodic Theorem**) *Let $T : \Omega \to \Omega$ be P-preserving, $X \in L^1(P)$, and*

$$S_n = \sum_{j=0}^{n-1} X \circ T^j, \ \ n \ge 1.$$

Then, the following hold:

a) *There exists a T-invariant r.v. X^* such that*

$$\frac{S_n}{n} \stackrel{n\to\infty}{\longrightarrow} X^*, \ \ a.s. \tag{1.83}$$

b) *For $p \in [1, \infty]$, $\|S_n\|_p \le n\|X\|_p$ for all $n \ge 1$, and $\|X^*\|_p \le \|X\|_p$.*

c) *$E[X^* : A] = E[X : A]$ for all $A \in \mathcal{I}$. In particular, if T is ergodic, then $X^* = EX$, a.s.*

Remark: By part c) of Theorem 1.11.2, $X^* = E[X|\mathcal{I}]$ (cf. Proposition 4.1.3).

From Theorem 1.11.2, we easily deduce:

Corollary 1.11.3 (von Neumann Ergodic Theorem) *Let $T : \Omega \to \Omega$ be P-preserving, $X \in L^p(P)$ $(p \in [1, \infty))$, S_n, $n \geq 1$ and X^* be as in Theorem 1.11.2. Then,*

$$\frac{S_n}{n} \stackrel{n\to\infty}{\longrightarrow} X^*, \quad \textit{in } L^p(P). \tag{1.84}$$

Proof: Suppose first that $X \in L^\infty(P)$. Since $\|S_n/n\|_\infty \leq \|X\|_\infty$, (1.84) for $p \in [1, \infty)$ follows from the bounded convergence theorem. Note next that $L^\infty(P)$ is dense in $L^p(P)$. Combinning the observations made above, it is easy to prove that S_n/n is a Cauchy sequence in $L^p(P)$, via standard $\varepsilon/3$-argument. Since the convergence $S_n/n \longrightarrow X^*$ takes place a.s. by Theorem 1.11.2, this proves (1.84). \(^□^)/

Remark: Corollary 1.11.3 does not extend to the case of $p = \infty$. See the remark at the end of Example 1.11.4.

Example 1.11.4 (Shift of an i.i.d.) Let $(S_n, \mathcal{B}_n, \mu_n) = (S, \mathcal{B}, \mu)$ $(n \in \mathbb{N})$ be copies of a probability space and let $(\Omega, \mathcal{F}, P)$ be their product:

$$\Omega = \prod_{n\in\mathbb{N}} S_n, \quad \mathcal{F} = \bigotimes_{n\in\mathbb{N}} \mathcal{B}_n, \quad P = \bigotimes_{n\in\mathbb{N}} \mu_n.$$

We define $T : \Omega \to \Omega$ by

$$T\omega = (\omega_{j+1})_{j\in\mathbb{N}} \text{ for } \omega = (\omega_j)_{j\in\mathbb{N}}.$$

Then,

1) T is P-preserving,

since ω and $T\omega$ have the same law P. Moreover

2) T is P-ergodic.

To see this suppose that $X \in L^\infty(P)$ is T-invariant. Since $T^n\omega = (\omega_{n+j})_{j\in\mathbb{N}}$ and $X \circ T^n = X$, a.s., X is measurable by the σ-algebra $\sigma[\mathcal{T}_n, \mathcal{N}]$, where $\mathcal{N}$ is the totality of P-null sets and

$$\mathcal{T}_n \stackrel{\text{def}}{=} \sigma[\omega_{n+j} \ ; \ j \in \mathbb{N}].$$

Since n is arbitrary, X is measurable by the σ-algebra $\sigma[\mathcal{T}, \mathcal{N}]$, where $\mathcal{T}$ is the tail σ-algebra:

$$\mathcal{T} \stackrel{\text{def}}{=} \bigcap_{n\geq 1} \mathcal{T}_n.$$

The σ-algebra $\mathcal{T}$ is trivial by Kolmogorov 0-1 law, hence so is $\sigma[\mathcal{T}, \mathcal{N}]$. This implies that $X = EX$, a.s.

Finally, we apply Birkhoff ergodic theorem (Theorem 1.11.2) to give a proof of law of large numbers (Theorem 1.10.2). Let $(S, \mathcal{B}, \mu) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$, where the measure μ satisfies $\int |x| d\mu(x) < \infty$. We write $m = \int x d\mu(x)$. For $X(\omega) \stackrel{\text{def}}{=} \omega_0$, $X(T^n\omega) = \omega_n$, $(n \in \mathbb{N})$ are i.i.d. $\approx \mu$ and

$S_n = \sum_{j=0}^{n-1} \omega_j$. Moreover, $X^* = EX = m$ by 2). Thus, it follows from Birkhoff ergodic theorem that

$$\frac{1}{n}\sum_{j=0}^{n-1} \omega_j \stackrel{n\to\infty}{\longrightarrow} m, \quad \text{a.s.} \tag{1.85}$$

Remark: By von Neumann ergodic theorem (Corollary 1.11.3), the convergence (1.85) takes place in $L^p(P)$ if $p \in [1,\infty)$ and $|x| \in L^p(\mu)$. However, this is no longer true for $p = \infty$. Indeed, take $\mu = (\delta_{-1} + \delta_1)/2$. Then, $m = 0$ and $\|S_n/n\|_\infty = 1$ for all $n \geq 1$.

Example 1.11.5 (**Rotation of the circle**) Let $\Omega = \mathbb{R}/\mathbb{Z}$, which is identified with the interval $[0,1)$, $\mathcal{F} = \mathcal{B}([0,1))$ (the Borel σ-algebra), $P =$ the Lebesgue measure on $[0,1)$. For $\alpha \in (0,1)$, we define $T_\alpha : \Omega \to \Omega$ by:

$$T_\alpha \theta = \theta + \alpha - \lfloor \theta + \alpha \rfloor.$$

Then,

1) T_α is P-preserving.

To see this, we start by a simple observation. For a function $f : \Omega \to \mathbb{R}$, its periodic extension is defined as a unique function $F : \mathbb{R} \to \mathbb{R}$, such that $F|_{[0,1)} = f$ and $F(\theta) = F(\theta+1)$ ($\forall \theta \in \mathbb{R}$). Then, for $f : \Omega \to \mathbb{R}$, its periodic extention F, and $\theta \in [0,1)$,

2) $f(T_\alpha\theta) = F(T_\alpha\theta) = F(\theta + \alpha - \lfloor \theta + \alpha \rfloor) = F(\theta + \alpha)$,

Therefore, for $f \in L^1([0,1))$,

$$\int_0^1 f \circ T_\alpha = \int_0^1 F(\cdot + \alpha) \stackrel{F(\cdot+1)=F}{=} \int_0^1 F = \int_0^1 f.$$

This implies 1).
We next prove that

3) T_α is P-ergodic $\iff$ $\alpha \notin \mathbb{Q}$.

($\Rightarrow$) Suppose that $\alpha = p/q$ ($p, q \in \mathbb{N}$, $1 \leq p < q$). Take a bounded measurable function $f : [0,1) \to \mathbb{R}$ of period $1/q$, which is not a.s. constant. Then, f is $T_{1/q}$-invariant, and hence is T_α-invariant, since $T_\alpha = (T_{1/q})^p$.
($\Leftarrow$) Suppose that $f \in L^\infty(P)$ is T_α-invariant. Then, $F = F(\cdot + \alpha)$, a.s. by 2). We look at the Fourier coefficient $\widehat{F} \in \ell^\infty(\mathbb{Z})$:

$$\widehat{F}(n) = \int_0^1 F(\theta) \exp(-2\pi \mathbf{i} n\theta) d\theta.$$

On the other hand, let $F_\alpha \stackrel{\text{def}}{=} F(\cdot + \alpha)$. Then,

$$\begin{aligned}\widehat{F}(n) &= \widehat{F_\alpha}(n) = \int_0^1 F(\theta + \alpha) \exp(-2\pi \mathbf{i} n\theta) d\theta \\ &= \int_0^1 F(\theta) \exp(-2\pi \mathbf{i} n(\theta - \alpha)) d\theta = \exp(2\pi \mathbf{i} n\alpha)\widehat{F}(n).\end{aligned}$$

Since $\alpha \notin \mathbb{Q}$, $\exp(2\pi \mathbf{i} n\alpha) \neq 1$ for all $n \neq 0$, and hence $\widehat{F}(n) = 0$ for all $n \neq 0$. This implies that F is a.s. constant, and therefore f is a.s. constant.
As a consequence of Birkhoff ergodic theorem, we observe the equidistribution of the irrational rotation in the following form. Let $F : \mathbb{R} \to \mathbb{R}$ be measurable, of period 1, and $\int_0^1 |F| < \infty$. Then, for $\alpha \in \mathbb{R}\backslash\mathbb{Q}$, and for alomst all $\theta \in [0,1)$,

$$\frac{1}{n}\sum_{j=0}^{n-1} F(\theta + j\alpha) \stackrel{n\to\infty}{\longrightarrow} \int_0^1 F.$$

See also Exercise 2.4.12.

We now turn to the proof of Theorem 1.11.2, which is based on ths following:

Lemma 1.11.6 *For $\alpha \in \mathbb{R}$, let*

$$B_\alpha^+ = \bigcup_{n\geq 1}\{S_n > \alpha n\}, \quad B_\alpha^- = \bigcup_{n\geq 1}\{S_n < \alpha n\}.$$

Then, for any $A \in \mathcal{I}$,

$$E[X - \alpha : A \cap B_\alpha^+] \geq 0 \geq E[X - \alpha : A \cap B_\alpha^-].$$

The above inequalities remain true if $B_\alpha^\pm$ are replaced respectively by $B_{\alpha,n}^+ = \bigcup_{j=1}^n \{S_j > \alpha j\}$ and $B_{\alpha,n}^- = \bigcup_{j=1}^n \{S_j < \alpha j\}$ ($n \in \mathbb{N}\backslash\{0\}$).

Proof: Since $B_{\alpha,n}^\pm \nearrow B_\alpha^\pm$ as $n \nearrow \infty$, it is enough to consider the case of $B_{\alpha,n}^\pm$ instead of $B_\alpha^\pm$. Then, by replacing X by $X - \alpha$, we may assume that $\alpha = 0$. Finally, we may concentrate on the first inequality, since the second one follows from the first, by replacing X by $-X$. Therefore, it is enough to prove that

1) $\qquad E[X : A \cap B_{0,n}^+] \geq 0, \quad \text{for } n \in \mathbb{N}\backslash\{0\}.$

The inequality 1) is obvious for $n = 1$, since $S_1 = X$ and hence $B_{0,1}^+ = \{X > 0\}$. For $n \geq 2$, the inequality 1) is a consequnce of the following equality.

2) $\qquad (M_{n-1} \circ T)^+ = M_n - X$, where $M_n = \max_{1\leq j\leq n} S_j$.

Indeed, 2) implies 1) as follows. Note that $B_{0,n}^+ = \{M_n > 0\}$ and hence $M_n \mathbf{1}_{B_{0,n}^+} = M_n^+$. Therefore,

3) $\qquad (M_{n-1} \circ T)^+ \geq (M_{n-1} \circ T)^+ \mathbf{1}_{B_{0,n}^+} \stackrel{2)}{=} (M_n - X)\mathbf{1}_{B_n} = M_n^+ - X\mathbf{1}_{B_{0,n}^+}.$

On the othr hand, $E[(M_{n-1} \circ T)^+ : A] = E[M_{n-1}^+ : A]$, since $A \in \mathcal{I}$. Hence,

$$\begin{aligned} E[X : A \cap B_{0,n}^+] &\stackrel{3)}{\geq} E[M_n^+ : A] - E[(M_{n-1} \circ T)^+ : A] \\ &= E[M_n^+ : A] - E[M_{n-1}^+ : A] \geq 0. \end{aligned}$$

Let us turn to the proof of 2). Since $S_j \circ T = S_{j+1} - X$ ($\forall j \geq 1$), we have

4) $\qquad M_{n-1} \circ T = \max_{1\leq j\leq n-1} S_j \circ T = \max_{1\leq j\leq n-1} S_{j+1} - X.$

Takig a trivial equality $0 = S_1 - X$ into account, we obtain 2) as follows.

$$\begin{aligned}(M_{n-1} \circ T)^+ &= (M_{n-1} \circ T) \vee 0 = (M_{n-1} \circ T) \vee (S_1 - X) \\ &\overset{4)}{=} \max_{0 \le j \le n-1} S_{j+1} - X = M_n - X.\end{aligned}$$

\(^□^)/

Proof of Theorem 1.11.2: a) Let

$$\overline{X} = \overline{\lim_{n\to\infty}} \frac{S_n}{n}, \quad \underline{X} = \underline{\lim_{n\to\infty}} \frac{S_n}{n}.$$

Then,

1) $\overline{X} \circ T = \overline{X}$, and $\underline{X} \circ T = \underline{X}$.

Indeed, since $S_n \circ T = S_{n+1} - X$, we have,

$$\frac{S_n \circ T}{n} = \frac{n+1}{n} \frac{S_{n+1}}{n+1} - \frac{X}{n}.$$

By taking the upper and the lower limits, we obtain 1).
On the other hand, we have

$$\{\underline{X} < \overline{X}\} = \bigcup_{\substack{\alpha,\beta \in \mathbb{Q} \\ \alpha < \beta}} A_{\alpha,\beta}, \quad \text{with} \quad A_{\alpha,\beta} = \{\underline{X} < \alpha\} \cap \{\beta < \overline{X}\}.$$

Thus, to prove that the limit X^* exists a.s., it is enough to show that

2) $P(A_{\alpha,\beta}) = 0$ if $\alpha < \beta$.

By 1), we see that $A_{\alpha,\beta} \in \mathcal{I}$. Moreover, $A_{\alpha,\beta} \subset B_\alpha^- \cap B_\beta^+$ and hence $A_{\alpha,\beta} = A_{\alpha,\beta} \cap B_\alpha^- = A_{\alpha,\beta} \cap B_\beta^+$. Thus, by Lemma 1.11.6,

$$\beta P(A_{\alpha,\beta}) \le E[X : A_{\alpha,\beta}] \le \alpha P(A_{\alpha,\beta}),$$

which implies 2).
b) The first inequality follows from the triangle inequalty for L^p-norm. The second inequality follows from the first one via the Fatou's lemma (Note that Fatou's lemma is valid for L^∞-norm).
c) We next prove that $E[X^* : A] = E[X : A]$ for all $A \in \mathcal{I}$. Let

$$A_{n,k} = A \cap \left\{X^* \in \left(\frac{k}{n}, \frac{k+1}{n}\right]\right\} \in \mathcal{I}.$$

We observe that

2) $\frac{k}{n} P(A_{n,k}) \le E[X : A_{n,k}]$.

Indeed, $A_{n,k} \subset B^+_{k/n}$ and hence $A_{n,k} = A_{n,k} \cap B^+_{k/n}$. Thus, by Lemma 1.11.6, we obtain 2). It follows from 2) that

$$E[X^* : A_{n,k}] \leq \frac{k+1}{n} P(A_{n,k}) \stackrel{2)}{\leq} E[X : A_{n,k}] + \frac{P(A_{n,k})}{n}.$$

Thus, by summing over $k \in \mathbb{Z}$,

$$E[X^* : A] \leq E[X : A] + \frac{1}{n}.$$

Letting $n \to \infty$, we obtain $E[X^* : A] \leq E[X : A]$. Then, by replacing X by $-X$,

$$E[X : A] = -E[(-X) : A] \leq -E[(-X)^* : A] = E[X^* : A].$$

This finishes the proof. \(^□^)/

Exercise 1.11.1 Let Ω be a finite set with cardinality $q \geq 2$, $P = \frac{1}{q} \sum_{x \in \Omega} \delta_x$, and $T : \Omega \to \Omega$ be a bijection. **i)** Verify that T is P-preserving. **ii)** For each $x \in \Omega$, let $p(x)$ be the minimal $p \in \mathbb{N}$ such that $T^p x \in \{T^j x\}_{j=0}^{p-1}$. Then, verify that $T^{p(x)} x = x$. **iii)** Prove that the following conditions a)–c) are equivalent: a) $\forall x \in \Omega$, $p(x) = q$. b) $\exists x_0 \in \Omega$, $p(x_0) = q$. c) T is ergodic. **iv)** Let $f : \Omega \to \mathbb{R}$ and $x \in \Omega$ be arbitrary. Then, verify by direct computation the following special case of Birkhoff ergodic theorem.

$$\frac{1}{n} \sum_{j=0}^{n-1} f(T^j x) \stackrel{n \to \infty}{\longrightarrow} \frac{1}{p(x)} \sum_{j=0}^{p(x)-1} f(T^j x).$$

Exercise 1.11.2 Let $\Omega = (0,1)$, $P(A) = \frac{1}{\log 2} \int_A \frac{dx}{1+x}$ ($A \in \mathcal{B}(\Omega)$), and $Tx = \frac{1}{x} - \lfloor \frac{1}{x} \rfloor$ ($x \in \Omega$). **i)** Verify that T is P-preserving. **ii)** It is known that T is P-ergodic [Bil95, p.322]. Assuming this, use Theorem 1.11.2 to show that for any $k \geq 1$,

$$\frac{1}{n} \sum_{j=0}^{n-1} \mathbf{1}\{\lfloor 1/T^j x \rfloor = k\} \stackrel{n \to \infty}{\longrightarrow} \frac{1}{\log 2} \left(\log\left(1 + \frac{1}{k}\right) - \log\left(1 + \frac{1}{k+1}\right) \right), \quad P(dx)\text{-a.s.}$$

Remark For $x \in (0,1) \backslash \mathbb{Q}$, the numbers $a_n(x) = \lfloor 1/T^n x \rfloor$ ($n \geq 1$) give the digits in continued-fraction representation of x in the sense that $F(a_1(x), \ldots, a_n(x)) \stackrel{n \to \infty}{\longrightarrow} x$, where

$$F(a_1(x), \ldots, a_n(x)) = \cfrac{1}{a_1(x) + \cfrac{1}{a_2(x) + \cfrac{1}{\ddots \, a_{n-2}(x) + \cfrac{1}{a_{n-1}(x) + \frac{1}{a_n(x)}}}}}$$

cf. [Bil95, pp.319–320]. Therefore, the limit considered in ii) can be interpreted as the asymptotic frequency with which the number k appears in the continued-fraction representation of x.

2 Characteristic functions

2.1 Definitions and Elementary Properties

Definition 2.1.1 (Fourier transform)

► For a Borel signed measure μ on $\mathbb{R}^d$, the *Fourier transform* of μ is a function $\widehat{\mu} : \mathbb{R}^d \to \mathbb{C}$ defined by

$$\widehat{\mu}(\theta) = \int \exp(\mathbf{i}\theta \cdot x)\, d\mu(x). \tag{2.1}$$

Example 2.1.2 a) (Fourier transform of L^1-functions) Suppose that a signed measure μ is of the form:

$$d\mu(x) = f(x)dx, \quad f \in L^1(\mathbb{R}^d).$$

Then,

$$\widehat{\mu}(\theta) = \widehat{f}(\theta) \stackrel{\text{def}}{=} \int_{\mathbb{R}^d} \exp(\mathbf{i}\theta \cdot x)\, f(x)dx. \tag{2.2}$$

Thus, (2.1) is given by the classical Fourier transform $\widehat{f}$ of the L^1-function f.

b) (Fourier series of ℓ^1-series) Suppose that a set $S \subset \mathbb{R}^d$ is countable, $(c_x)_{x \in S} \in \ell^1(S)$, and that a signed measure μ is of the form:

$$\mu = \sum_{x \in S} c_x \delta_x.$$

Then,

$$\widehat{\mu}(\theta) \stackrel{\text{def}}{=} \sum_{x \in S} c_x \exp(\mathbf{i}\theta \cdot x). \tag{2.3}$$

If $S = \mathbb{Z}^d$, (2.1) is given by the classical Fourier series of a sequence in $\ell^1(\mathbb{Z}^d)$.

The following proposition states that a finite measure is uniquely characterized by its Fourier transform:

Proposition 2.1.3 (Injectivity of the Fourier transform) *For a Borel signed measure μ on $\mathbb{R}^d$,*

$$\mu = 0 \iff \widehat{\mu}(\theta) = 0 \text{ for all } \theta \in \mathbb{R}^d.$$

We will postpone the proof of this proposition until section 2.4.

Let $(\Omega, \mathcal{F}, P)$ be a probability space in what follows.

Proposition 2.1.4 (Characteristic function) *For $\mu \in \mathcal{P}(\mathbb{R}^d)$ and a r.v. $X : \Omega \to \mathbb{R}^d$, the following are equivalent:*

a) $E \exp(\mathbf{i}\theta \cdot X) = \widehat{\mu}(\theta)$ *for all* $\theta \in \mathbb{R}^d$*;*

b) $X \approx \mu$.

► *The expectation on the left-hand side of a) above is called the* **characteristic function** (**ch.f.** *for short*) *of* X.

Proof: Let $\nu = P(X \in \cdot)$. Then,

1) $$E\exp(\mathbf{i}\theta\cdot X) \stackrel{(1.9)}{=} \int \exp(\mathbf{i}\theta\cdot x)d\nu(x) \stackrel{(2.1)}{=} \widehat{\nu}(\theta).$$

Therefore

$$\text{a)} \stackrel{1)}{\Longleftrightarrow} \widehat{\mu}=\widehat{\nu} \stackrel{\text{Proposition 2.1.3}}{\Longleftrightarrow} \mu=\nu \Longleftrightarrow \text{b)}.$$

\(^▫^)/

Remark: By Proposition 2.1.4,

the ch.f. of a r.v. = the Fourier transform of its law.

Corollary 2.1.5 (Criterion of the independence) *Let $X_j:\Omega\to\mathbb{R}^{d_j}$ $(j=1,...,n)$ be r.v.'s. Then, the following are equivalent:*

a) $E\left[\prod_{j=1}^n \exp(\mathbf{i}\theta_j\cdot X_j)\right] = \prod_{j=1}^n E\exp(\mathbf{i}\theta_j\cdot X_j)$ *for all* $\theta_j\in\mathbb{R}^{d_j}$ $(j=1,...,n)$.

b) $\{X_j\}_{j=1}^n$ *are independent.*

Proof: Let $d=d_1+...+d_n$, $\theta_j\in\mathbb{R}^{d_j}$ and $\mu_j=P(X_j\in\cdot)\in\mathcal{P}(\mathbb{R}^{d_j})$ $(1\le j\le n)$. We write:

$$\theta=(\theta_j)_{j=1}^n\in\mathbb{R}^d,\quad X=(X_j)_{j=1}^n:\Omega\to\mathbb{R}^d,\quad \mu=\otimes_{j=1}^n\mu_j\in\mathcal{P}(\mathbb{R}^d).$$

Then,

1) $$\exp(\mathbf{i}\theta\cdot X)=\exp\left(\mathbf{i}\sum_{j=1}^n\theta_j\cdot X_j\right)=\prod_{j=1}^n\exp(\mathbf{i}\theta_j\cdot X_j).$$

Therefore,

2) $$E\exp(\mathbf{i}\theta\cdot X)\stackrel{1)}{=}E\left[\prod_{j=1}^n\exp(\mathbf{i}\theta_j\cdot X_j)\right],$$

and

3) $$\begin{cases} \widehat{\mu}(\theta) & \stackrel{(2.1)}{=} \displaystyle\int_{\mathbb{R}^d}\exp(\mathbf{i}\theta\cdot x)d\mu(x)=\int_{\mathbb{R}^d}\prod_{j=1}^n\exp(\mathbf{i}\theta_j\cdot x_j)d\mu_1(x_1)\cdots d\mu_n(x_n)\\ & \stackrel{\text{Fubini}}{=} \displaystyle\prod_{j=1}^n\int_{\mathbb{R}^{d_j}}\exp(\mathbf{i}\theta_j\cdot x_j)d\mu_j(x_j)\\ & \stackrel{(1.9)}{=} \displaystyle\prod_{j=1}^n E\exp(\mathbf{i}\theta_j\cdot X_j).\end{cases}$$

Therefore,

$$\begin{aligned}\text{a)} &\Longleftrightarrow \text{RHS 2)=RHS 3)} \stackrel{2),3)}{\Longleftrightarrow} \text{LHS 2)=LHS 3)} \stackrel{\text{Proposition 2.1.4}}{\Longleftrightarrow} X\approx\mu\\ &\stackrel{\text{Proposition 1.6.1}}{\Longleftrightarrow} \text{b)}.\end{aligned}$$

\(^▫^)/

Exercise 2.1.1 Let μ be a Borel signed measure on $\mathbb{R}^d$, and $|\mu|$ be its total variation. Prove that

$$|\widehat{\mu}(\theta)| \le |\mu|(\mathbb{R}^d), \quad |\widehat{\mu}(\theta) - \widehat{\mu}(\theta')| \le \int_{\mathbb{R}^d} |\exp(\mathbf{i}(\theta - \theta') \cdot x) - 1| d|\mu|$$

for $\theta, \theta' \in \mathbb{R}^d$. In particular, $\widehat{\mu}$ is bounded and uniformly continuous.

Exercise 2.1.2 Let $X = (X_\alpha)_{\alpha=1}^k$ be an $\mathbb{R}^k$ valued r.v. Prove that the following conditions are equivalent. (a) $UX \approx X$ for all $U \in \mathcal{O}_k$, where $\mathcal{O}_k$ denotes the totality of $k \times k$ real orthogonal matrices. (b) $E\exp(\mathbf{i}\theta \cdot X) = E\exp(\mathbf{i}|\theta|X_1)$ for all $\theta \in \mathbb{R}^k$.

Exercise 2.1.3 Let X be an $\mathbb{R}^k$ valued r.v. which satisfies the conditions stated in Exercise 2.1.2. Prove then that $AX \approx BX$ for $d \times k$ matirices A and B such that $AA^* = BB^*$. Hint: If $AA^* = BB^*$, then, $|A^*\theta| = |B^*\theta|$ for all $\theta \in \mathbb{R}^d$. Combine this observation with Exercise 2.1.2.

2.2 Basic Examples

Example 2.2.1 (ch.f. of binomial and Poisson r.v.'s) Let $\mu \in \mathcal{P}(\mathbb{N})$ and $X : \Omega \to \mathbb{N}$ be a r.v. with $X \approx \mu$. Recall that we have defined the generating function by

$$G(\mu; s) \stackrel{\text{def}}{=} Es^X = \sum_{n=0}^{\infty} \mu(n)s^n, \quad s \in \mathbb{C}, \ |s| \le 1,$$

where $\mu(n) = \mu(\{n\})$ (Definition 1.7.1). By plugging $s = \exp(\mathbf{i}\theta)$ in the above expression, we see that

$$\widehat{\mu}(\theta) = E\exp(\mathbf{i}\theta X) = G(\mu; \exp(\mathbf{i}\theta)). \tag{2.4}$$

Let $\mu_{n,p}$ be (n,p)-binomial distribution, and π_c be c-Poisson distribution. Then, we see from (1.61), (1.66) and (2.4) that

$$\widehat{\mu_{n,p}}(\theta) = G(\mu_{n,p}; \exp(\mathbf{i}\theta)) = (p\exp(\mathbf{i}\theta) + 1 - p)^n, \tag{2.5}$$

$$\widehat{\pi_c}(\theta) = G(\pi_c; \exp(\mathbf{i}\theta)) = \exp(c(\exp(\mathbf{i}\theta) - 1)). \tag{2.6}$$

Example 2.2.2 (ch.f. of a Uniform r.v.) Suppose that a r.v. U is uniformly distributed on an interval (a,b) (cf. (1.16)). Then,

$$E\exp(\mathbf{i}\theta \cdot U) = \frac{\exp(\mathbf{i}\theta b) - \exp(\mathbf{i}\theta a)}{\mathbf{i}(b-a)\theta}. \tag{2.7}$$

Proof: Since U has the density: $u(x) = (b-a)^{-1}1_{(a,b)}(x)$, we have that

$$E\exp(\mathbf{i}\theta \cdot U) \stackrel{(2.1)}{=} (b-a)^{-1}\int_a^b \exp(\mathbf{i}\theta x)dx = \text{ RHS (2.7)}.$$

\(^□^)/

Example 2.2.3 (ch.f. of $N(0, I_d)$) Let X be an $\mathbb{R}^d$-valued r.v. $\approx N(0, I_d)$. We will show that

$$E\exp(\mathbf{i}\theta \cdot X) = \exp\left(-\tfrac{1}{2}|\theta|^2\right). \tag{2.8}$$

Since X has the density : $h(x) \stackrel{\text{def}}{=} (2\pi)^{-d/2}\exp\left(-\frac{|x|^2}{2}\right)$, we have that

$$E\exp(\mathbf{i}\theta \cdot X) \stackrel{(2.1)}{=} \int_{\mathbb{R}^d} \exp(\mathbf{i}\theta \cdot x)h(x)dx.$$

Let us prove that

1) $$\int_{\mathbb{R}^d} \exp(z\theta \cdot x)h(x)dx = \exp\left(\tfrac{1}{2}z^2|\theta|^2\right), \ \forall \theta \in \mathbb{R}^d, \ \forall z \in \mathbb{C}$$

and hence (by setting $z = \mathbf{i}$) that (2.8) holds. Note first that both hand sides of 1) are holomorphic in z. Therefore, by the unicity theorem, it is enough to prove the equality for all $z = t \in \mathbb{R}$. Note that

$$t\theta \cdot x - \tfrac{1}{2}|x|^2 = \tfrac{1}{2}t^2|\theta|^2 - \tfrac{1}{2}|x - t\theta|^2,$$

and therefore,

2) $$\exp(t\theta \cdot x)h(x) = \exp(\tfrac{1}{2}t^2|\theta|^2)h(x - t\theta).$$

Thus,

$$\int_{\mathbb{R}^d} \exp(t\theta \cdot x)\, h(x)dx \stackrel{2)}{=} \exp\left(\tfrac{1}{2}t^2|\theta|^2\right) \underbrace{\int_{\mathbb{R}^d} h(x - t\theta)dx}_{=1} = \exp\left(\tfrac{1}{2}t^2|\theta|^2\right),$$

which implies 1). See Exercise 2.3.3 for an alternative proof. \(^□^)/

Example 2.2.4 (ch.f. of $N(m, V)$) For $d \in \mathbb{N}\backslash\{0\}$, we denote by $\mathcal{S}_d^+$ the totality of symmetric, non-negative definite $d \times d$ real matrices. Let $m \in \mathbb{R}^d$ and $V \in \mathcal{S}_d^+$ in what follows. In Example 1.2.4, we have defined multi-dimensional Gaussian distribution $N(m, V)$ when V is strictly positive definite. We now generalize the definition to the case where V is non-negative definite, but not necessarily strictly positive definite.

Let $k \in \mathbb{N}\backslash\{0\}$. We take a $d \times k$ matrix A such that $V = AA^*$. See Proposition 8.2.4 for a characterization of such A for a given V. Let Y be an $\mathbb{R}^k$-valued r.v. $\approx N(0, I_k)$. Then, we define $N(m, V)$ to be the law of the following r.v.

$$X \stackrel{\text{def}}{=} m + AY. \tag{2.9}$$

We will prove that:

$$E\exp(\mathbf{i}\theta \cdot X) = \exp\left(\mathbf{i}\theta \cdot m - \tfrac{1}{2}\theta \cdot V\theta\right), \quad \theta \in \mathbb{R}^d. \tag{2.10}$$

This, together with Proposition 2.1.4, shows that the law $N(m, V)$ is uniquely determined by m and V, without depending on the choice of A (See also Exercise 2.1.3). Note that:

1) $$\theta \cdot X \stackrel{(2.9)}{=} \theta \cdot m + \theta \cdot AY = \theta \cdot m + A^*\theta \cdot Y.$$

2) $$|A^*\theta|^2 = A^*\theta \cdot A^*\theta = \theta \cdot AA^*\theta = \theta \cdot V\theta.$$

We use these to see (2.10) as follows:

$$\begin{aligned} E\exp(\mathbf{i}\theta \cdot X) &\stackrel{1)}{=} \exp(\mathbf{i}\theta \cdot m)\, E\exp(\mathbf{i}A^*\theta \cdot Y) \\ &\stackrel{(2.8)}{=} \exp\left(\mathbf{i}\theta \cdot m - \tfrac{1}{2}|A^*\theta|^2\right) \stackrel{2)}{=} \exp\left(\mathbf{i}\theta \cdot m - \tfrac{1}{2}\theta \cdot V\theta\right). \end{aligned}$$

We will next use (2.10) to show the following. Let $X_j : \Omega \to \mathbb{R}^d$ ($j = 1, 2$) be independent r.v.'s such that $X_j \approx N(m_j, V_j)$, where $m_j \in \mathbb{R}^d$ and $V_j \in \mathcal{S}_d^+$. Then,

$$X \stackrel{\text{def}}{=} X_1 + X_2 \approx N(m, V), \text{ where } m = m_1 + m_2, \ V = V_1 + V_2. \tag{2.11}$$

We have for any $\theta \in \mathbb{R}^d$ that

$$
\begin{aligned}
E\exp(\mathbf{i}\theta\cdot X) &\stackrel{(1.53)}{=} \prod_{j=1}^{2} E\exp(\mathbf{i}\theta\cdot X_j) \\
&\stackrel{(2.10)}{=} \prod_{j=1}^{2} \exp\left(\mathbf{i}\theta\cdot m_j - \tfrac{1}{2}\theta\cdot V_j\theta\right) = \exp\left(\mathbf{i}\theta\cdot m - \tfrac{1}{2}\theta\cdot V\theta\right).
\end{aligned}
$$

This implies (2.11) via Proposition 2.1.4. \(^□^)/

Example 2.2.5 (**ch.f. of a Cauchy r.v.: dimension one**) Suppose that an $\mathbb{R}$-valued r.v. Y has (c)-Cauchy distribution: $Y \approx \frac{c}{\pi}\frac{dx}{c^2+x^2}$. Then,

$$
E\exp(\mathbf{i}\theta Y) = \exp(-c|\theta|), \quad \theta \in \mathbb{R}. \tag{2.12}
$$

Let $g_c(x) = \frac{e^{-c|x|}}{2c}$. Then,

$$
\begin{aligned}
\widehat{g_c}(\theta) &= \int_{-\infty}^{\infty} \frac{e^{-c|x|+\mathbf{i}\theta x}}{2c}dx = \int_0^{\infty} \frac{e^{-(c-\mathbf{i}\theta)x}}{2c}dx + \int_0^{\infty} \frac{e^{-(c+\mathbf{i}\theta)x}}{2c}dx \\
&= \frac{1}{2c}\left(\frac{1}{c-\mathbf{i}\theta} + \frac{1}{c+\mathbf{i}\theta}\right) = \frac{1}{c^2+\theta^2}.
\end{aligned} \tag{2.13}
$$

Thus,

$$
\begin{aligned}
\exp(-c|\theta|) &= 2cg_c(\theta) \stackrel{(2.37)}{=} 2c\cdot\frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(-\mathbf{i}\theta x)\widehat{g_c}(x)dx \\
&\stackrel{(2.13)}{=} \frac{c}{\pi}\int_{-\infty}^{\infty} \frac{\exp(-\mathbf{i}\theta x)}{c^2+x^2}dx = \frac{c}{\pi}\int_{-\infty}^{\infty} \frac{\exp(\mathbf{i}\theta x)}{c^2+x^2}dx = E\exp(\mathbf{i}\theta Y).
\end{aligned}
$$

\(^□^)/

Remark (*Relevance of (2.13) to functional analysis*) We see from (2.13) that $\frac{\widehat{f}(\theta)}{c^2+\theta^2} = \widehat{g_c}(\theta)\widehat{f}(\theta)$ for $f \in L^2(\mathbb{R})$. By the Fourier inversion, this implies that

$$
(c^2-\Delta)^{-1}f(x) = \int_{\mathbb{R}} g_c(x-y)f(y)dy,
$$

where $\Delta f = f''$ with the domain:

$$
\{f \in L^2(\mathbb{R})\ ;\ f \text{ and } f' \text{ are absolute continuous, } f'' \in L^2(\mathbb{R})\ \}.
$$

Exercise 2.2.1 Let U_1, U_2 be i.i.d. with uniform distribution on $(-1,1)$. **(i)** Show that $\frac{U_1+U_2}{2} \approx f(x)dx$, where $f(x) \stackrel{\text{def.}}{=} (1-|x|)^+$ and that $\widehat{f}(\theta) = \frac{\sin^2(\theta/2)}{(\theta/2)^2}$. **(ii)** Show that $\frac{1}{2\pi}\int_{-\infty}^{\infty}\widehat{f} = 1$. $\frac{1}{2\pi}\widehat{f}$ is the density of *Polya's distribution*. **Hint**: (2.37).

Exercise 2.2.2 Let $X_1, X_2, ..$ be iid such that $P(X_1 = \pm 1) = 1/2$. Prove the following. **(i)** $U \stackrel{\text{def}}{=} \sum_{n\geq 1}\frac{X_n}{2^n}$ is uniformly distributed on $[-1,1]$. **(ii)** $\dfrac{\sin\theta}{\theta} = \displaystyle\prod_{n=1}^{\infty}\cos\frac{\theta}{2^n}$ for $\theta \in \mathbb{R}$.

Exercise 2.2.3 Let V be a symmetric, non-negative definite $d\times d$ real matrix with eigenvalues $\{\lambda_\alpha\}_{\alpha=1}^d$ and let $X:\Omega\to\mathbb{R}^d$ be a r.v. $\approx N(0,V)$. Prove then that $|X|^2\approx\sum_{\alpha=1}^d\lambda_\alpha|Y_\alpha|^2$, where $Y=(Y_\alpha)_{\alpha=1}^d\approx N(0,I_d)$. **Hint** Let $D=(\sqrt{\lambda_\alpha}\delta_{\alpha,\beta})_{\alpha,\beta=1}^d$ and let U be an orthogonal matrix such that $V=UD^2U^*$. Then, $X\approx UDY$.

Exercise 2.2.4 (**Stability of Gaussian distribution**) Let X_1,X_2 be $\mathbb{R}^d$-valued independent r.v.'s such that $X_j\approx N(0,V_j)$, cf. (1.24) and A_1,A_2 be $d\times d$ matrices. Prove then that

$$X\stackrel{\text{def}}{=}A_1X_1+A_2X_2\approx N(0,V),\quad \text{where } V=A_1V_1A_1^*+A_2V_2A_2^*$$

Hint: Compute $E\exp(\mathbf{i}\theta\cdot X)$ and use Proposition 2.1.3.

Exercise 2.2.5 Let X be a mean-zero $\mathbb{R}^d$-valued r.v. Prove then that X is a Gaussian r.v. if and only if $X\cdot\theta$ is a Gaussian r.v. for any $\theta\in\mathbb{R}^d$. Hint: (2.10), Proposition 2.1.3.

Exercise 2.2.6 Suppose that $X=(X_\alpha)_{\alpha=1}^d$ is a mean-zero $\mathbb{R}^d$-valued Gaussian r.v. Prove then that coordinates $\{X_\alpha\}_{\alpha=1}^d$ are independent if and only if $E[X_\alpha X_\beta]=0$ for $\alpha\neq\beta$. This shows in particular that the independence for r.v.'s $\{X_\alpha\}_{\alpha=1}^d$ above follows from the pairwise independence. Hint: (2.10), Corollary 2.1.5.

Exercise 2.2.7 ($\star$) Suppose that X is a real r.v. $\approx\frac{2}{c\pi}\cosh(x/c)^{-1}dx$ $(c>0)$ (cf. Exercise 1.2.16). **(i)** Show that $E\exp(\mathbf{i}\theta X)=\cosh(c\pi\theta/2)^{-1}$ $(\forall\theta\in\mathbb{R})$. **Hint**: One can use residue theorem. **(ii)** Noting that $z\in\mathbb{C}\backslash\left(\frac{\pi}{2}\mathbf{i}+\pi\mathbf{i}\mathbb{Z}\right)\mapsto(\cosh z)^{-1}$ is holomorhic, we write its Taylor expansion around the origin as $(\cosh z)^{-1}=\sum_{k=0}^\infty(-1)^kE_kz^{2k}/(2k)!$ $(|z|<\pi/2)$, where the numbers E_k's are called *Euler numbers*. Then prove for $k\in\mathbb{N}$ that $E[X^{2k}]=(c\pi/2)^{2k}E_k$, and deduce therefrom the following celebrated formula.

$$\sum_{n=0}^\infty\frac{(-1)^n}{(2n+1)^{2k+1}}=\frac{\pi^{2k+1}}{E_k2^{2k+3}}.$$

Exercise 2.2.8 ($\star$) Suppose that $X=X_1+X_2$, where X_1 and X_2 are i.i.d. $\approx\frac{2}{c\pi}\cosh(x/c)^{-1}dx$ $(c>0)$. **(i)** Show that $E\exp(\mathbf{i}\theta X)=\cosh(c\pi\theta/2)^{-2}$ $(\forall\theta\in\mathbb{R})$. **Hint**: Exercise 2.2.7 (i). **(ii)** Show that $X\approx\frac{8}{c\pi^2}\frac{x}{\sinh(x/c)}dx$. **(iii)** Show that $E[|X|^{s-2}]=\frac{8c^{s-2}}{\pi^2}\Gamma(s)\sum_{n=0}^\infty(2n+1)^{-s}$ for $s\in(1,\infty)$. **(iii)** Noting that $z\in\mathbb{C}\backslash\left(\frac{\pi}{2}\mathbf{i}+\pi\mathbf{i}\mathbb{Z}\right)\mapsto\tanh z$ is holomorhic, we write its Taylor expansion around the origin as $\tanh z=\sum_{k=1}^\infty 2^{2k}(2^{2k}-1)(-1)^{k-1}B_kz^{2k-1}/(2k)!$ $(|z|<\pi/2)$, where the numbers B_k's are called *Bernoulli numbers*. Then prove for $k\in\mathbb{N}\backslash\{0\}$ that $E[X^{2k-2}]=2(2^{2k}-1)B_k(c\pi)^{2k-2}$ and deduce therefrom the following celebrated formula.

$$\sum_{n=0}^\infty\frac{1}{(2n+1)^{2k}}=\frac{(2^{2k}-1)B_k}{2}\frac{\pi^{2k}}{(2k)!}.$$

Exercise 2.2.9 Apply the residue theorem to a melomorphic function $\frac{\exp(i\theta z)}{c^2+z^2}$ to give an alternative proof of (2.12).

Exercise 2.2.10 (**Stability of Cauchy distribution**) **(i)** Suppose that Y_j $(j=1,2)$ has (c_j)-Cauchy distribution and that Y_1 and Y_2 are independent. Prove then that Y_1+Y_2 has (c_1+c_2)-Cauchy distribution. **(ii)** Let $S_n=Y_1+...+Y_n$, where $Y_1,Y_2,...$ are independent r.v.'s with (c)-Cauchy distribution. Prove then that $S_n/n\approx Y_1$ for all $n\geq 1$. This shows that S_n/n does not converge to a constant, even weakly (cf. Theorem 1.10.2).

2.3 ($\star$) Further Examples

Example 2.3.1 (ch.f. of a Gamma r.v.) Let X be a real r.v. such that $X \approx \gamma_{c,a}$. We will show that

$$\widehat{\gamma_{c,a}}(\theta) = \left(1 + \frac{\theta^2}{c^2}\right)^{-a/2} \exp\left(\mathbf{i}a\mathrm{Arctan}\,\frac{\theta}{c}\right). \tag{2.14}$$

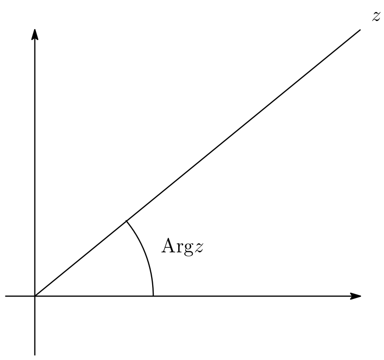
To prove this, we go thruough a little of complex analysis. For $z \in \mathbb{C}\backslash\{0\}$, we define Arg $z \in (-\pi, \pi]$ (argument of z) by

1) $\qquad z = |z| \exp(\mathbf{i}\mathrm{Arg}\ z),$

and Log$z \in \mathbb{C}$ by

$$\mathrm{Log}\ z = \log|z| + \mathbf{i}\mathrm{Arg}\ z.$$

By definition, Arg z is the angle, signed counter-clockwise, from the positive real axis to the vector representing z.



Finally we set:

$$z^s = \exp(s\mathrm{Log}\ z), \quad \text{for } z \in \mathbb{C}\backslash\{0\} \text{ and } s \in \mathbb{C}.$$

It is well-known that Log z is holomorphic in $z \in \mathbb{C}\backslash(-\infty, 0]$, and hence so is z^s. Note also that

2) $\qquad z^s = \exp(s\mathrm{Log}\ z) = \exp(s\log|z| + \mathbf{i}s\mathrm{Arg}\ z) = |z|^s \exp(\mathbf{i}s\mathrm{Arg}\ z).$

We first show that

3) $\qquad E\exp(-zX) = \left(1 + \frac{z}{c}\right)^{-a}$ for any $z \in \mathbb{C}$ with $\mathrm{Re}\, z > -c.$

To prove 3), note that both hand-sides are holomorphic in z for $\mathrm{Re}\, z > -c$. Therefore, by the unicity theorem, it is enough to prove the equality for all $z = t \in (-c, \infty)$. Then,

$$\begin{aligned} E\exp(-tX) &\stackrel{(1.27)}{=} \frac{c^a}{\Gamma(a)} \int_0^\infty x^{a-1} e^{-(t+c)x} dx \\ &\stackrel{x=y/(t+c)}{=} \frac{c^a}{\Gamma(a)} \left(\frac{1}{t+c}\right)^a \underbrace{\int_0^\infty y^{a-1} e^{-y} dy}_{=\Gamma(a)} = \left(1 + \frac{t}{c}\right)^{-a}. \end{aligned}$$

This proves 3).
Finally, we use 3) to derive (2.14). For $\theta \in \mathbb{R}$,

4) $$\left|1-\frac{\mathbf{i}\theta}{c}\right| = \left(1+\frac{\theta^2}{c^2}\right)^{1/2}, \quad \mathrm{Arg}\,\left(1-\frac{\mathbf{i}\theta}{c}\right) = -\mathrm{Arctan}\,\frac{\theta}{c}.$$

Therefore,

$$\begin{aligned}\widehat{\gamma_{c,a}}(\theta) &\stackrel{3)}{=} \left(1-\frac{\mathbf{i}\theta}{c}\right)^{-a} \stackrel{2)}{=} \left|1-\frac{\mathbf{i}\theta}{c}\right|^{-a} \exp\left(-\mathbf{i}a\mathrm{Arg}\,\left(1-\frac{\mathbf{i}\theta}{c}\right)\right) \\ &\stackrel{4)}{=} \left(1+\frac{\theta^2}{c^2}\right)^{-a/2} \exp\left(\mathbf{i}a\mathrm{Arctan}\,\frac{\theta}{c}\right).\end{aligned}$$

Example 2.3.2 (**Stieltjes' counterexample to the moment problem**) We consider the following question. Suppose that a function $f \in C([0,\infty))$ satisfies

$$\int_0^\infty t^n |f(t)| dt < \infty, \text{ and } \int_0^\infty t^n f(t) dt = 0 \text{ for all } n \in \mathbb{N}.$$

Then $f \equiv 0$? Stieltjes gave the following counterexample to this question (1894):

$$f(t) \stackrel{\text{def}}{=} \exp(-t^{1/4}) \sin t^{1/4}.$$

We can use (2.14) with $c = 1$, $a = 4n+4$ ($n \in \mathbb{N}$), $\theta = 1$ to verify that the above function is indeed a counterexample. Let $n \in \mathbb{N}$. Since Arctan $1 = \pi/4$, we have

1) $\exp(4(n+1)\mathbf{i}\mathrm{Arctan}\ 1) = \exp((n+1)\pi\mathbf{i}) = (-1)^{n+1}$.

Therefore, we see that

2) $$\frac{1}{\Gamma(4n+4)} \int_0^\infty x^{4n+3} \exp(-x+\mathbf{i}x) dx = \widehat{\gamma_{1,4n+4}}(1) \stackrel{(2.14),\ 1)}{=} (-1)^{n+1} 2^{-(2n+2)} \in \mathbb{R}.$$

Thus, taking the imaginary part, we have

$$0 \stackrel{2)}{=} \int_0^\infty x^{4n+3} \exp(-x) \sin x dx \stackrel{t=x^4}{=} \frac{1}{4} \int_0^\infty t^n \exp(-t^{1/4}) \sin t^{1/4} dt.$$

Example 2.3.3 (**Euler's complementary formula for the Gamma function**) We will use (2.14) to prove the following identity due to Euler:

$$\frac{1}{\Gamma(a)\Gamma(1-a)} = \frac{\sin(\pi a)}{\pi}, \quad a \in (0,1). \tag{2.15}$$

For $a = 1/2$, the above identity follows from $\Gamma(1/2) = \sqrt{\pi}$. Moreover, the identity is invariant under the replacement of a by $1-a$. Thus, to prove identity, we may and will assume that $a < 1/2$. Let $f_a(x) = \frac{1}{\Gamma(a)} x^{a-1} e^{-x} \mathbf{1}_{x>0}$ (the density of $\gamma(1,a)$). Note that $f_{1\pm a} \in L^2(\mathbb{R})$ for $a < 1/2$. Thus, we have by the Plancherel formula that:

1) $$\int_0^\infty f_{1+a}(x) f_{1-a}(x) dx = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{f_{1+a}}(\theta) \widehat{f_{1-a}}(-\theta) d\theta.$$

Since

$$f_{1+a}(x) f_{1-a}(x) = \frac{1}{\Gamma(1+a)\Gamma(1-a)} e^{-2x} \mathbf{1}_{x>0},$$

we see that

2) $$\int_0^\infty f_{1+a}(x)f_{1-a}(x)dx = \frac{1}{2\Gamma(1+a)\Gamma(1-a)}$$

On the other hand,

$$\begin{aligned}\widehat{f_{1+a}}(\theta)\widehat{f_{1-a}}(-\theta) &\stackrel{(2.14)}{=} \frac{1}{1+\theta^2}\exp(\mathbf{i}(1+a)\text{Arctan }\theta - \mathbf{i}(1-a)\text{Arctan }\theta) \\ &= (\text{Arctan }\theta)'\exp(2\mathbf{i}a\text{Arctan }\theta).\end{aligned}$$

Thus,

3) $$\begin{cases}\int_{\mathbb{R}} \widehat{f_{1+a}}(\theta)\widehat{f_{1-a}}(-\theta)d\theta \stackrel{t=\text{Arctan }\theta}{=} \int_{-\pi/2}^{\pi/2} \exp(2\mathbf{i}at)dt \\ \qquad = \frac{\exp(\mathbf{i}a\pi)-\exp(-\mathbf{i}\pi a)}{2\mathbf{i}a} = \frac{\sin(\pi a)}{a}\end{cases}$$

By 1)–3), we see that

$$\frac{1}{2\Gamma(1+a)\Gamma(1-a)} = \frac{\sin(\pi a)}{2\pi a},$$

which is equivalent to (2.15), since $\Gamma(1+a) = a\Gamma(a)$.

Before Example 2.3.5, we prepare the following Lemma.

Lemma 2.3.4 *For $a \in \mathbb{R}$, $c, \lambda > 0$,*

$$\begin{aligned}\int_0^\infty t^{a-1}\exp\left(-\tfrac{c^2t}{2} - \tfrac{\lambda^2}{2t}\right)dt &= \int_0^\infty t^{-a-1}\exp\left(-\tfrac{c^2}{2t} - \tfrac{\lambda^2 t}{2}\right)dt \\ &= 2(\lambda/c)^a K_a(c\lambda), \qquad (2.16)\end{aligned}$$

where K_a stands for the Macdonald's function, defined by (2.25). In particular, for $a = n+\frac{1}{2}$ $(n \in \mathbb{N}\cup\{-1\})$, the above integral takes the following more explicit form.

$$\sqrt{2\pi}c^{-(n+1)}\lambda^n p_n(1/c\lambda)e^{-c\lambda}, \qquad (2.17)$$

where $p_{-1}(x) = 1$ and $p_n(x) = \sum_{r=0}^n \frac{(n+r)!}{r!(n-r)!}\left(\frac{x}{2}\right)^r$ for $n \geq 0$.

Proof: The first equality is easily obtained by the change of variable $t \mapsto 1/t$. On the other hand, by the change of integral variable $t = (\lambda/c)e^x$, we have

$$\int_0^\infty t^{a-1}\exp\left(-\tfrac{c^2t}{2} - \tfrac{\lambda^2}{2t}\right)dt = (\lambda/c)^a\int_{-\infty}^\infty \exp(-c\lambda\cosh x)\exp(ax)dx.$$

Therefore, we obtain (2.16) and (2.17) from Lemma 2.3.8. which proves (3). \(^□^)/

Example 2.3.5 Le $a, c > 0$, and $X \approx \gamma(c, a)$. Then, the Laplace transform of $1/X$ is computed as:

$$E\exp(-\tfrac{\lambda}{X}) = \frac{2(c\lambda)^{\frac{a}{2}}}{\Gamma(a)}K_a(2\sqrt{c\lambda}), \quad \lambda > 0, \qquad (2.18)$$

where K_a stands for the Macdonald's function, defined by (2.25). In particular for $a = 1/2$, RHS of (2.18) equals $\exp(-2\sqrt{c\lambda})$. This result is used later, e.g., Example 2.3.6 and (6.50). We have:

$$E\exp(-\tfrac{\lambda}{X}) = \frac{c^a}{\Gamma(a)}\int_0^\infty t^{a-1}\exp\left(-ct - \tfrac{\lambda}{t}\right)dt$$

By Lemma 2.3.4, the above integral equals the RHS of (2.18). \(^▫^)/

Example 2.3.6 (**ch.f. of a Cauchy r.v.: higher dimensions**) With $c > 0$ and $a > 0$ fixed, we consider $\mu_{c,a} \in \mathcal{P}(\mathbb{R}^d)$ defined as follows.

$$\mu_{c,a}(B) = \frac{c^{2a}\Gamma(a+\frac{d}{2})}{\pi^{d/2}\Gamma(a)}\int_B \frac{dx}{(c^2+|x|^2)^{a+\frac{d}{2}}}, \quad B \in \mathcal{B}(\mathbb{R}^d).$$

Then, $\mu_{c,\frac{1}{2}}$ is the (c)-Cauchy distribution. We will show that

$$\int_{\mathbb{R}^d} \exp(\mathbf{i}x\cdot\theta)d\mu_{c,a}(x) = \frac{2(c|\theta|/2)^a}{\Gamma(a)}K_a(c|\theta|), \quad \theta \in \mathbb{R}^d, \tag{2.19}$$

where K_a stands for the Macdonald's function, defined by (2.25). In particular,

$$\int_{\mathbb{R}^d} \exp(\mathbf{i}x\cdot\theta)d\mu_{c,\frac{1}{2}}(x) = \exp(-c|\theta|), \quad \theta \in \mathbb{R}^d$$

Proof: We will use (1.69) to prove this. Let $X_1, X_2, ..., X_d, Y$ be independent r.v.'s with $X_j \approx N(0,1)$, $1 \le j \le d$ (cf. (1.24)) and $Y \approx \gamma_{c^2/2,a}$ (cf. (1.27)). Let us write $X = (X_j)_{j=1}^d$ for simplicity. Then,

$$\begin{aligned}\int_{\mathbb{R}^d} \exp(\mathbf{i}x\cdot\theta)d\mu_{c,a}(x) &\stackrel{(1.69)}{=} E\exp(\mathbf{i}\theta\cdot Y^{-1/2}X) = \int_0^\infty E\exp(\mathbf{i}\theta\cdot y^{-1/2}X)d\gamma_{\frac{c^2}{2},a}(y) \\ &\stackrel{(2.10)}{=} \int_0^\infty \exp\left(-\frac{|\theta|^2}{2y}\right)d\gamma_{\frac{c^2}{2},a}(y) \stackrel{(2.18)}{=} \frac{2(c|\theta|/2)^a}{\Gamma(a)}K_a(c|\theta|).\end{aligned}$$

Remark: An alternative proof of (2.19) for $d = 3$ can be given by applying the inversion formula (2.37) to $\widehat{F_{c,2}}(\theta) = \frac{c^4}{(c^2+|\theta|^2)^2}$ (Exercise 2.3.2) as in Example 2.2.5.

Before Example 2.3.7, we need some preparation. For $\nu \in (-1,\infty)$, we introduce the following power series.

$$F_\nu(z) = \sum_{n=0}^\infty c_n\left(\frac{z}{2}\right)^{2n}, \quad z \in \mathbb{C}, \text{ where } c_n = \frac{1}{\Gamma(\nu+n+1)n!}. \tag{2.20}$$

See (2.23)–(2.25) below for the relation of this power series to the Bessel functions. The series (2.20) converges for all $z \in \mathbb{C}$, since

$$\frac{c_{n+1}}{c_n} = \frac{1}{(\nu+n+1)(n+1)} \stackrel{n\to\infty}{\longrightarrow} 0.$$

Example 2.3.7 **(a)** For $z \in \mathbb{C}$, $\nu \in (-1,\infty)$, and $n \in \mathbb{N}$,

$$F_{\nu+n}(z) = \left(\frac{2}{z}\frac{d}{dz}\right)^n F_\nu(z), \quad F_{\nu+n}(\mathbf{i}z) = \left(-\frac{2}{z}\frac{d}{dz}\right)^n F_\nu(\mathbf{i}z). \tag{2.21}$$

In particular, settig $\nu = -1/2$ in (2.21),

$$F_{n-1/2}(z) = \frac{1}{\sqrt{\pi}}\left(\frac{2}{z}\frac{d}{dz}\right)^n \cosh z, \quad F_{n-1/2}(\mathbf{i}z) = \frac{1}{\sqrt{\pi}}\left(-\frac{2}{z}\frac{d}{dz}\right)^n \cos z. \tag{2.22}$$

(b) For $z \in \mathbb{C}$ and $\nu \in (-1/2,\infty)$,

$$\int_0^\pi \exp(z\cos\theta)\sin^{2\nu}\theta d\theta = \int_{-1}^1 \exp(zt)(1-t^2)^{\nu-\frac{1}{2}}dt = \sqrt{\pi}\Gamma(\nu+\tfrac{1}{2})F_\nu(z).$$

(c) For an integer $d \geq 2$, $z \in \mathbb{C}$, and $x \in \mathbb{R}^d$,

$$\int_{S^{d-1}} \exp(zx\cdot u)d\sigma_d(u) = 2\pi^{d/2}F_{\frac{d}{2}-1}(|x|z).$$

where σ_d stands for the surface measure on S^{d-1}.

Proof: (a) It is easy to see (2.21) for $n=1$, and hence they follow by induction. To see (2.22), it is enough to show that $F_{-1/2}(z) = \frac{1}{\sqrt{\pi}}\cosh z$ and $F_{-1/2}(\mathbf{i}z) = \frac{1}{\sqrt{\pi}}\cos z$. To do so, note first that

(1) $\Gamma\left(n+\frac{1}{2}\right) = \dfrac{(2n)!\sqrt{\pi}}{2^{2n}n!}$.

Then,

(2) $\dfrac{1}{\Gamma\left(n+\frac{1}{2}\right)n!} \stackrel{(1)}{=} \dfrac{2^{2n}n!}{\sqrt{\pi}(2n)!n!} = \dfrac{2^{2n}}{\sqrt{\pi}(2n)!}$,

Therefore,

$$F_{-1/2}(z) = \sum_{n=0}^\infty \frac{1}{\Gamma\left(n+\frac{1}{2}\right)n!}\left(\frac{z}{2}\right)^{2n} \stackrel{(2)}{=} \frac{1}{\sqrt{\pi}}\sum_{n=0}^\infty \frac{2^{2n}}{(2n)!}\left(\frac{z}{2}\right)^{2n} = \frac{1}{\sqrt{\pi}}\cosh z,$$

Hence $F_{-1/2}(\mathbf{i}z) = \frac{1}{\sqrt{\pi}}\cos z$.
(b) The first equality follows from the change of integral variable $t = \cos\theta$. To prove the second equality, we note that

(3) $\displaystyle\int_{-1}^1 t^{2n}(1-t^2)^{\nu-\frac{1}{2}}dt = \frac{(2n)!\sqrt{\pi}}{2^{2n}n!}\frac{\Gamma(\nu+\frac{1}{2})}{\Gamma(\nu+n+1)}$.

Indeed, by the change of integral variable $t = \sqrt{s}$, we have

$$\begin{aligned}\int_{-1}^1 t^{2n}(1-t^2)^{\nu-\frac{1}{2}}dt &= 2\int_0^1 t^{2n}(1-t^2)^{\nu-\frac{1}{2}}dt = \int_0^1 s^{n-\frac{1}{2}}(1-s)^{\nu-\frac{1}{2}}dt \\ &= B(n+\tfrac{1}{2},\nu+\tfrac{1}{2}) = \frac{\Gamma(n+\frac{1}{2})\Gamma(\nu+\frac{1}{2})}{\Gamma(\nu+n+1)} \\ &\stackrel{(1)}{=} \frac{(2n)!\sqrt{\pi}}{2^{2n}n!}\frac{\Gamma(\nu+\frac{1}{2})}{\Gamma(\nu+n+1)}.\end{aligned}$$

Since $(1-t^2)^{\nu-\frac{1}{2}}$ is an even function, we have

$$\begin{aligned}\sum_{n=0}^{\infty}\int_{-1}^{1}\exp(tz)(1-t^2)^{\nu-\frac{1}{2}}dt &= \sum_{n=0}^{\infty}\frac{z^{2n}}{(2n)!}\int_{-1}^{1}t^{2n}(1-t^2)^{\nu-\frac{1}{2}}dt \\ &\stackrel{(3)}{=} \sqrt{\pi}\Gamma(\nu+\tfrac{1}{2})\sum_{n=0}^{\infty}\frac{1}{\Gamma(\nu+n+1)n!}\left(\frac{z}{2}\right)^{2n}.\end{aligned}$$

This proves the second equality.
(c) Let $A_d=\sigma_d(S^{d-1})=2\pi^{\frac{d}{2}}/\Gamma\left(\frac{d}{2}\right)$. Since σ_d is invariant under rotation, we may assume that $x=|x|e_1$. Then,

$$\begin{aligned}\int_{S^{d-1}}\exp(zx\cdot u)d\sigma_d(u) &= \int_{S^{d-1}}\exp(|x|zu_1)d\sigma_d(u)=A_{d-1}\int_0^{\pi}\exp(|x|z\cos\theta)\sin^{d-2}\theta d\theta \\ &\stackrel{\text{(b)}}{=} A_{d-1}\sqrt{\pi}\Gamma\left(\tfrac{d}{2}-1\right)F_{\frac{d}{2}-1}(|x|z)=2\pi^{d/2}F_{\frac{d}{2}-1}(|x|z).\end{aligned}$$

\(^□^)/

Complement (Bessel functions): We have defined the power series (2.20) for $\nu\in(-1,\infty)$. We now extend its definition for $\nu\in\mathbb{R}$. To do so, recall that the Gamma function $\Gamma(z)=\int_0^\infty t^{z-1}\exp(-t)dt$ ($z\in\mathbb{C}$, $\operatorname{Re}z>0$) has a unique holomorphic extension on $\mathbb{C}\backslash(-\mathbb{N})$, which we denote by the same notation Γ. Recall also that the extension Γ satisfies the following properties.
(a) Γ has no zero's;
(b) $\Gamma(z)\to\infty$ as $z\to-n$ ($\forall n\in\mathbb{N}$);
(c) $\Gamma(z+1)=z\Gamma(z)$ ($\forall z\in\mathbb{C}\backslash(-\mathbb{N})$).
For $n\in\mathbb{N}$, we set $\Gamma(-n)=\infty$ and $1/\Gamma(-n)=0$, which is justified by the property (b) above. Using the extended Gamma function introduced now, and via the formula (2.20), we extend the definition of the power series $F_\nu(z)$ ($z\in\mathbb{C}$) for all $\nu\in\mathbb{C}$.

If $\nu\notin\{-m\ ;\ m\in\mathbb{N}\backslash\{0\}\}$, then $c_n\neq0$ ($\forall n\in\mathbb{N}$). On the other hand, if $\nu=-m$ for some $m\in\mathbb{N}\backslash\{0\}$, then $c_0=\ldots=c_{m-1}=0$ and $c_n\neq0$ ($\forall n\geq m$). In both cases, $c_n\neq0$ for all sufficiently large n's and

$$\frac{c_{n+1}}{c_n}=\frac{1}{(\nu+n+1)(n+1)}\stackrel{n\to\infty}{\longrightarrow}0.$$

Therefore, the series (2.20) converges for all $z\in\mathbb{C}$. Moreover, the recursion (2.21) extends to all $\nu\in\mathbb{R}$ and $z\in\mathbb{C}$.

For $\nu\in\mathbb{R}$ and $z\in(0,\infty)$, we introduce,

$$J_\nu(z) = \left(\frac{z}{2}\right)^{\nu}F_\nu(\mathbf{i}z)=\sum_{n=0}^{\infty}(-1)^n c_n\left(\frac{z}{2}\right)^{\nu+2n}, \tag{2.23}$$

$$I_\nu(z) = \left(\frac{z}{2}\right)^{\nu}F_\nu(z)=\sum_{n=0}^{\infty}c_n\left(\frac{z}{2}\right)^{\nu+2n}, \tag{2.24}$$

$$K_\nu(z) = \frac{\pi}{2}\frac{I_{-\nu}(z)-I_\nu(z)}{\sin\nu\pi}\ \text{ if } \nu\in\mathbb{C}\backslash\mathbb{Z}\ \text{ and }\ K_n(z)=\lim_{\substack{\nu\to n\\ \nu\in\mathbb{C}\backslash\mathbb{Z}}}K_\nu(z)\ \text{ if } n\in\mathbb{Z}. \tag{2.25}$$

The function J_ν is called the *Bessel function*. The function I_ν (resp. K_ν) is called respectively, the *modified Bessel function* of the first kind (resp. *Macdolald's function*). Note that we now have the recursions (2.21) for all $\nu \in \mathbb{C}$ and $z \in \mathbb{C}\backslash(-\infty, 0]$. They imply the raising operator relations.

$$J_{\nu+1}(z) = \left(-\frac{d}{dz} + \frac{\nu}{z}\right) J_\nu(z), \quad I_{\nu+1}(z) = \left(\frac{d}{dz} - \frac{\nu}{z}\right) I_\nu(z). \tag{2.26}$$

We also note the lowering operator relations.

$$J_{\nu-1}(z) = \left(\frac{d}{dz} + \frac{\nu}{z}\right) J_\nu(z), \quad I_{\nu-1}(z) = \left(\frac{d}{dz} + \frac{\nu}{z}\right) I_\nu(z). \tag{2.27}$$

Finally, by (2.25), (2.26) and (2.27),

$$K_{\nu+1}(z) = \left(-\frac{d}{dz} + \frac{\nu}{z}\right) K_\nu(z) \quad K_{\nu-1}(z) = -\left(\frac{d}{dz} + \frac{\nu}{z}\right) K_\nu(z). \tag{2.28}$$

It follows from (2.26)–(2.28) that both $I_\nu(z)$ and $K_\nu(z)$ solve the following differntial equation.

$$\left(\frac{d^2}{dz^2} + \frac{d}{dz} - 1 - \frac{\nu^2}{z^2}\right) u(z) = 0. \tag{2.29}$$

In fact, $I_\nu(z)$ and $K_\nu(z)$ are independent solution to (2.29), as can be seen from their assymptotic behavior as $z \to \infty$, cf. [Leb72, p.123, (5.11.8), (5.11.9)].

$$I_\nu(z) \sim \left(\frac{1}{2\pi z}\right)^{1/2} e^z, \quad K_\nu(z) \sim \left(\frac{\pi}{2z}\right)^{1/2} e^{-z}. \tag{2.30}$$

We also note the following formulas for $n \in \mathbb{N}$, which follow from (2.22).

$$J_{n-\frac{1}{2}}(z) = \left(\frac{2}{\pi}\right)^{1/2} z^{n-\frac{1}{2}} \left(-\frac{1}{z}\frac{d}{dz}\right)^n \cos z, \tag{2.31}$$

$$I_{n-\frac{1}{2}}(z) = \left(\frac{2}{\pi}\right)^{1/2} z^{n-\frac{1}{2}} \left(\frac{1}{z}\frac{d}{dz}\right)^n \cosh z. \tag{2.32}$$

We now prove the following representation formulas for $K_\nu(z)$.

Lemma 2.3.8

$$K_\nu(z) = \frac{1}{2}\int_{-\infty}^{\infty} \exp(-z\cosh x - \nu x) dx. \tag{2.33}$$

Moreover, for $n \in \mathbb{N}$,

$$K_{n+\frac{1}{2}}(z) = \left(\frac{\pi}{2z}\right)^{1/2} e^{-z} p_n\left(\frac{1}{z}\right), \tag{2.34}$$

where $p_{-1}(x) = 1$ *and* $p_n(x) = \sum_{r=0}^{n} \frac{(n+r)!}{r!(n-r)!}\left(\frac{x}{2}\right)^r$ *for* $n \geq 0$.

Proof: Let us donote the integral on the right-hand side of (2.33) by $u_\nu(z)$. To prove (2.33), it is enough to verify that

(1) u_ν solves (2.29). **(2)** $u_\nu(z) \sim \left(\frac{\pi}{2z}\right)^{1/2} e^{-z}$.

To verify (1), it is is enough to show the following two equations separetely.

(3) $u_{\nu+1}(z) = \left(-\frac{d}{dz} + \frac{\nu}{z}\right) u_\nu(z), \quad u_{\nu-1}(z) = -\left(\frac{d}{dz} + \frac{\nu}{z}\right) u_\nu(z).$

We have,

$$\begin{aligned}
-\frac{d}{dz}u_\nu(z) &= \frac{1}{2}\int_{-\infty}^{\infty} \exp(-z\cosh x)\cosh x \exp(-\nu x)dx = \frac{1}{2}u_{\nu-1}(z) + \frac{1}{2}u_{\nu+1}(z), \\
\frac{\nu}{z}u_\nu(z) &= -\frac{1}{2}\int_{-\infty}^{\infty} \frac{1}{z}\exp(-z\cosh x)\frac{d}{dx}(\exp(-\nu x))dx \\
&= -\frac{1}{2}\left[\frac{1}{z}\exp(-z\cosh x)\exp(-\nu x)\right]_{-\infty}^{\infty} - \frac{1}{2}\int_{-\infty}^{\infty} \exp(-z\cosh x)\sinh x \exp(-\nu x)dx \\
&= -\frac{1}{2}u_{\nu-1}(z) + \frac{1}{2}u_{\nu+1}(z),
\end{aligned}$$

from which (3) follow. Since $\cosh x \geq 1 + \frac{x^2}{2}$, we have

$$\begin{aligned}
u_\nu(z) &\leq \frac{e^{-z}}{2}\int_{-\infty}^{\infty} \exp\left(-\frac{zx^2}{2} - \nu x\right)dx = e^{-z}\left(\frac{\pi}{2z}\right)^{1/2}\exp\left(\frac{\nu^2}{2z}\right) \\
&= e^{-z}\left(\frac{\pi}{2z}\right)^{1/2}\left(1 + O(z^{-1})\right),
\end{aligned}$$

which gives the upper bound for (2). As for the lower bound, we note that $\cosh x \leq 1+\frac{x^2}{2}+C\varepsilon^4$, for $|x| \leq \varepsilon$. Therefore,

$$\begin{aligned}
u_\nu(z) &= \frac{1}{2}\int_{-\infty}^{\infty} \exp(-z\cosh x)\mathrm{ch}\,(\nu x)dx \geq \frac{1}{2}\int_{-\infty}^{\infty} \exp(-z\cosh x)dx \\
&\geq \frac{\exp(-z - Cz\varepsilon^4)}{2}\int_{-\varepsilon}^{\varepsilon} \exp\left(-\frac{zx^2}{2}\right)dx = \frac{\exp(-z - Cz\varepsilon^4)}{2\sqrt{z}}\int_{-\varepsilon\sqrt{z}}^{\varepsilon\sqrt{z}} \exp\left(-\frac{x^2}{2}\right)dx \\
&\geq \exp(-z - Cz\varepsilon^4)\left(\left(\frac{\pi}{2z}\right)^{1/2} - \frac{2}{\varepsilon\sqrt{z}}\exp\left(-\frac{\varepsilon^2 z}{2}\right)\right).
\end{aligned}$$

Choosing $\varepsilon = z^{-1/3}$, we get the desired lower bound.

For $n = 0$, (2.34) easily follows from (2.25) and (2.32). On the other hand, we see from tedious, but straightforward computations that

(4) $p_{n+1}\left(\frac{1}{z}\right) = \left(-\frac{d}{dz} + \frac{n+1}{z} + 1\right)p_n\left(\frac{1}{z}\right).$

Suppose that (2.34) is valid for some $n \in \mathbb{N}$. Then, by (2.28), the induction hypothesis (IH), and (4),

$$\begin{aligned}
K_{n+\frac{3}{2}}(z) &\overset{(2.28)}{=} \left(-\frac{d}{dz} + (n+\frac{1}{2})z^{-1}\right)K_{n+\frac{1}{2}}(z) \\
&\overset{(\mathrm{IH})}{=} \left(\frac{\pi}{2z}\right)^{1/2}e^{-z}\left(-\frac{d}{dz} + \frac{n+1}{z} + 1\right)p_n(1/z) \\
&\overset{(4)}{=} z^{-(n+2)}e^{-z}p_{n+1}(1/z).
\end{aligned}$$

\(^□^)/

Exercise 2.3.1 Let $a > 0$ and $f_{\frac{a+1}{2}}(x) = \frac{1}{\Gamma(\frac{a+1}{2})} x^{\frac{a-1}{2}} e^{-x} \mathbf{1}_{x>0}$ (the density of $\gamma(1, \frac{a+1}{2})$), $I_a = \int_0^\infty f_{\frac{a+1}{2}}(x)^2 dx$, $J_a = \int_{-\infty}^\infty (1+|x|^2)^{-\frac{a+1}{2}} dx$. Prove then the following. **(i)** $I_a = \frac{2^{-a}\Gamma(a)}{\Gamma(\frac{a+1}{2})^2}$. **(ii)** $J_a = \frac{\Gamma(\frac{a}{2})\sqrt{\pi}}{\Gamma(\frac{a+1}{2})}$. **Hint**: Exercise 1.2.13. **(iii)** $I_a = \frac{1}{2\pi} J_a$. **Hint**: Plancherel identity. **(iv)**
$\Gamma(a) = \frac{2^{a-1}}{\sqrt{\pi}} \Gamma\left(\frac{a}{2}\right) \Gamma\left(\frac{a+1}{2}\right) \quad a > 0.$

Exercise 2.3.2 **(i)** Let $f : [0, \infty) \to \mathbb{R}$ be a Borel function such that $\int_0^\infty r^2 |f(r)| dr < \infty$ and that $F(x) = f(|x|)$, $x \in \mathbb{R}^3$. Prove then that $F \in L^1(\mathbb{R}^3)$ and that for $\theta \in \mathbb{R}^3 \backslash \{0\}$,

$$\widehat{F}(\theta) = \frac{4\pi}{|\theta|} \int_0^\infty r f(r) \sin(r|\theta|) dr = \frac{4\pi}{|\theta|} \operatorname{Im}\left(\int_0^\infty r f(r) \exp(\mathbf{i} r|\theta|) dr\right).$$

(ii) Use (i) and Example 2.3.1 to show that

$$\widehat{F_{c,a}}(\theta) = \frac{c}{(|\theta|^2 + c^2)^{a/2}} \sin\left(a \operatorname{Arctan} \frac{|\theta|}{c}\right) \quad \text{for} \quad F_{c,a}(x) = \frac{c^{a+1}}{4\pi\Gamma(a+1)} |x|^{a-2} e^{-c|x|}, \quad a, c > 0.$$

In particular, $\widehat{F_{c,1}}(\theta) = \frac{c^2}{c^2 + |\theta|^2}$ and $\widehat{F_{c,2}}(\theta) = \frac{c^4}{(c^2+|\theta|^2)^2}$. $F_{c,1}$ is a constant $\times$ the Green function, while $\widehat{F_{c,2}}$ is a constant $\times$ the density of the Cauchy distribution, cf. the remark after Example 2.3.6.

Exercise 2.3.3 Give an alternative proof of (2.8) via polar coordiate transform and Example 2.3.7 (c).

2.4 Weak Convergence

The following fact has an important application to probability theory.

Proposition 2.4.1 (Weak convergence of measures) *Suppose that* $(\mu_n)_{n\geq 0}$ *are Borel finite measures on* $\mathbb{R}^d$. *Then the following are equivalent:*

a) $\widehat{\mu_n}(\theta) \stackrel{n\to\infty}{\longrightarrow} \widehat{\mu_0}(\theta)$ *for all* $\theta \in \mathbb{R}^d$ *(cf. (2.1)).*

b) *For all* $f \in C_{\mathrm{b}}(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} f d\mu_n \stackrel{n\to\infty}{\longrightarrow} \int_{\mathbb{R}^d} f d\mu_0. \tag{2.35}$$

► *The sequence* $(\mu_n)_{n\geq 1}$ *is said to* **converge weakly** *to* μ_0 *if one of (thus, both) a)–b) holds. We will henceforth denote this convergence by*

$$\mu_n \stackrel{\mathrm{w}}{\longrightarrow} \mu_0. \tag{2.36}$$

Here, the measure μ_0 *is called the* **weak limit** *of the sequence* $(\mu_n)_{n\geq 1}$.

Remark (i) For a sequence $(\mu_n)_{n\geq 1}$ of Borel finite measures on $\mathbb{R}^d$, its weak limit is unique. Indeed, if μ and ν are both weak limits, it follows from (2.35) that $\int_{\mathbb{R}^d} f d\mu = \int_{\mathbb{R}^d} f d\nu$, for $\forall f \in C_{\mathrm{b}}(\mathbb{R}^d)$, which implies that $\mu = \nu$ by Lemma 1.3.2. **(ii)** See Theorem 9.1.1 for some other equivalent conditions to a)–b) in Proposition 2.4.1.

The proof of Proposition 2.4.1 will be presented at the end of this section, followed by the proof of Proposition 2.1.3. We now look at a simple example to get familiar with the notion of weak convergence.

Example 2.4.2 (**Riemann sum**) Let $f \in C([0,1])$. Then, we know very well that

1) $\displaystyle \frac{1}{n}\sum_{k=0}^{n-1} f\left(\frac{k}{n}\right) \stackrel{n\to\infty}{\longrightarrow} \int_0^1 f d\mu_0,$

where μ_0 is the Lebesgue measure on $[0,1]$. However, the proof of 1) usually depends on the fact that

2) f is Riemann integrable.

Indeed, without resorting to 2), we would not even know the existence of the limit as $n \to \infty$ of the left-hand side of 1). On the other hand, as we see now, we can show 1) by Proposition 2.4.1, instead of 2).

Let $\mu_n = \frac{1}{n}\sum_{k=0}^{n-1} \delta_{k/n} \in \mathcal{P}(\mathbb{R})$ for $n \in \mathbb{N}\backslash\{0\}$, where δ_x is a point mass at $x \in \mathbb{R}$. We will show that

3) $\mu_n \stackrel{\mathrm{w}}{\longrightarrow} \mu_0,$

or equivalently,

$$\frac{1}{n}\sum_{k=0}^{n-1} f\left(\frac{k}{n}\right) = \int f d\mu_n \stackrel{n\to\infty}{\longrightarrow} \int_0^1 f d\mu_0,$$

which proves 1). By Proposition 2.4.1, 3) is equivalent to

4) $\widehat{\mu_n}(\theta) \stackrel{n\to\infty}{\longrightarrow} \widehat{\mu_0}(\theta)$, for all $\theta \in \mathbb{R}$.

This can be seen as follows. We have

$$\widehat{\mu_0}(\theta) = \begin{cases} \frac{\exp(\mathbf{i}\theta)-1}{\mathbf{i}\theta}, & \text{if } \theta \neq 0, \\ 1, & \text{if } \theta = 0, \end{cases}$$

$$\widehat{\mu_n}(\theta) = \frac{1}{n}\sum_{k=0}^{n-1} \exp\left(\frac{\mathbf{i}k\theta}{n}\right) = \begin{cases} \frac{1}{n}\frac{\exp(\mathbf{i}\theta)-1}{\exp(\mathbf{i}\theta/n)-1}, & \text{if } \theta \notin 2\pi n\mathbb{Z}, \\ 1, & \text{if } \theta \in 2\pi n\mathbb{Z}. \end{cases}$$

Let $\theta \in \mathbb{R}$ be arbitrary. Then, for $n > \frac{|\theta|}{2\pi}$,

$$\widehat{\mu_n}(\theta) = \frac{1}{n}\frac{\exp(\mathbf{i}\theta)-1}{\exp(\mathbf{i}\theta/n)-1} \stackrel{n\to\infty}{\longrightarrow} \frac{\exp(\mathbf{i}\theta)-1}{\mathbf{i}\theta} = \widehat{\mu_0}(\theta),$$

which proves 4).

Proposition 2.4.3 (Weak convergence of r.v.'s) *For $n = 0, 1, \ldots$, let X_n be $\mathbb{R}^d$-valued r.v.'s and that $X_n \approx \mu_n \in \mathcal{P}(\mathbb{R}^d)$. Then, the following are equivalent:*

a) $E\exp(\mathbf{i}\theta \cdot X_n) \longrightarrow E\exp(\mathbf{i}\theta \cdot X_0)$ *for all* $\theta \in \mathbb{R}^d$.

b) $\mu_n \stackrel{\text{w}}{\longrightarrow} \mu_0$.

▶ *The sequence $(X_n)_{n\ge 1}$ is said to* **converge weakly** *(or* **converge in law***) to X_0 if one (therefore all) of the above conditions is satisfied. We will henceforth denote this convergence by*

$$X_n \stackrel{\text{w}}{\longrightarrow} X_0 \quad or \quad X_n \stackrel{\text{w}}{\longrightarrow} \mu_0$$

Here, the r.v. X_0 is called the **weak limit** *(or* **limit in law***) of the sequence $(X_n)_{n\ge1}$.*

Proof:

$$E\exp(\mathbf{i}\theta \cdot X_n) = \widehat{\mu_n}(\theta), \quad n = 0, 1, \ldots$$

Thus,

$$\text{a)} \Longleftrightarrow \widehat{\mu_n}(\theta) \longrightarrow \widehat{\mu_0}(\theta),\ \forall \theta \in \mathbb{R}^d \stackrel{\text{Proposition 2.4.1}}{\Longleftrightarrow} \text{b)}.$$

\(^□^)/

Example 2.4.4 Let $(N_c)_{c>0}$ be r.v.'s such that $\pi_c(k) \stackrel{\text{def}}{=} P(N_c = k) = e^{-c}c^k/k!$ for all $k \in \mathbb{N}$ and $c > 0$. We will prove the following two facts, of which the first is probabilistic, the second purely analytic:

a) $\dfrac{N_c - c}{\sqrt{c}} \stackrel{\text{w}}{\longrightarrow} N(0,1)$, as $c \to \infty$.

b) $n! \stackrel{n\to\infty}{\sim} \sqrt{2\pi n}\,(n/e)^n$ (**Stirling's formula**).

Proof: Both a) and b) are based on the following observation.

1) $$\widehat{\pi_c}\left(\frac{\theta}{\sqrt{c}}\right)\exp\left(-\mathbf{i}\sqrt{c}\theta\right) \stackrel{c\to\infty}{\longrightarrow} \exp\left(-\frac{\theta^2}{2}\right).$$

To verify 1), note that

$$\exp(\mathbf{i}\theta) = 1 + \mathbf{i}\theta - \frac{\theta^2}{2} + O(|\theta|^3) \text{ as } \theta \to 0,$$

and hence that

2) $$\exp\left(\mathbf{i}\frac{\theta}{\sqrt{c}}\right) = 1 + \frac{\mathbf{i}\theta}{\sqrt{c}} - \frac{\theta^2}{2c} + O\left(\frac{|\theta|^3}{c^{3/2}}\right) \text{ as } c \to \infty \text{ for any } \theta \in \mathbb{R}.$$

Since $\widehat{\pi_c}(\theta) \stackrel{(2.6)}{=} \exp(c(\exp(\mathbf{i}\theta) - 1))$, we have

$$\begin{aligned}\widehat{\pi_c}\left(\frac{\theta}{\sqrt{c}}\right)\exp\left(-\mathbf{i}\sqrt{c}\theta\right) &= \exp\left(c\left(\exp\left(\mathbf{i}\frac{\theta}{\sqrt{c}}\right) - 1 - \mathbf{i}\frac{\theta}{\sqrt{c}}\right)\right)\\ &\stackrel{2)}{=} \exp\left(c\left(-\frac{\theta^2}{2c} + O\left(\frac{\theta^3}{c^{3/2}}\right)\right)\right) \stackrel{c\to\infty}{\longrightarrow} \exp\left(-\frac{\theta^2}{2}\right).\end{aligned}$$

This proves 1).
a) By 1), we have

3) $$E\exp\left(\mathbf{i}\theta\frac{N_c-c}{\sqrt{c}}\right)=\widehat{\pi_c}\left(\frac{\theta}{\sqrt{c}}\right)\exp\left(-\mathbf{i}\sqrt{c}\theta\right)\stackrel{c\to\infty}{\longrightarrow}\exp\left(-\frac{\theta^2}{2}\right).$$

Recall that $\exp\left(-\frac{\theta^2}{2}\right)$ is the Fourier transform of $N(0,1)$ (Example 2.2.4). We see the desired weak convergence from 3) and Proposition 2.4.3.
b) We will prove Stirling's formula in the following equivalent form.

4) $$\frac{\sqrt{n}}{n!}\left(\frac{n}{e}\right)^n\stackrel{n\to\infty}{\longrightarrow}\frac{1}{\sqrt{2\pi}}.$$

We have that
$$\widehat{\pi_c}(\theta)=\sum_{k\geq 0}\exp(\mathbf{i}k\theta)\pi_c(k),\quad \theta\in\mathbb{R}$$

Multiplying $\exp(-\mathbf{i}n\theta)/(2\pi)$ to both-hands sides of the above identity and integrating them over $\theta\in[-\pi,\pi]$, we obtain

5) $$\pi_c(n)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\widehat{\pi_c}(\theta)\exp(-\mathbf{i}n\theta)d\theta.$$

Moreover, since $1-\cos\theta\geq\frac{2\theta^2}{\pi^2}$, $|\theta|\leq\pi$, we have

6) $$\left|\widehat{\pi_c}\left(\tfrac{\theta}{\sqrt{c}}\right)\right|=\exp\left(-c\left(1-\cos\tfrac{\theta}{\sqrt{c}}\right)\right)\leq\exp\left(-\tfrac{2\theta^2}{\pi^2}\right),\quad |\theta|\leq\pi\sqrt{c}.$$

Finally, note that

7) $$\begin{cases}\dfrac{\sqrt{n}}{n!}\left(\dfrac{n}{e}\right)^n &= \sqrt{n}\pi_n(n)\stackrel{5)}{=}\dfrac{\sqrt{n}}{2\pi}\displaystyle\int_{-\pi}^{\pi}\widehat{\pi_n}(\theta)\exp(-\mathbf{i}n\theta)d\theta\\ &= \dfrac{1}{2\pi}\displaystyle\int_{-\pi\sqrt{n}}^{\pi\sqrt{n}}\widehat{\pi_n}\left(\dfrac{\theta}{\sqrt{n}}\right)\exp(-\mathbf{i}\sqrt{n}\theta)d\theta\end{cases}$$

By 1), 6) and the dominated convergence theorem, we conclude that, as $n\to\infty$, the right-hand side of 7) converges to
$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left(-\tfrac{\theta^2}{2}\right)d\theta=\frac{1}{\sqrt{2\pi}}.$$

This proves 4). \(^□^)/

Example 2.4.5 (**The Stirling's formula and a certain weak convergence**) Let X_a and Y_a $(a>0)$ be r.v's such that $X_a\approx\gamma(1,a)$ and $Y_a\approx\gamma(\sqrt{a},a)$. We will prove the following two facts, of which the first is probabilistic, the second purely analytic:

a) $\dfrac{X_a-a}{\sqrt{a}}\approx Y_a-\sqrt{a}\stackrel{\mathrm{w}}{\longrightarrow}N(0,1)$, as $a\to\infty$.

b) $\Gamma(a)\stackrel{a\to\infty}{\sim}\sqrt{(2\pi/a)}\,(a/e)^a$ (Stirling's formula).

Proof: a) It is easy to verify that $X_a/\sqrt{a}\approx Y_a$, and hence that $(X_a-a)/\sqrt{a}\approx Y_a-\sqrt{a}$. Let $f_{c,a}(x)=\frac{c^a x^{a-1}}{\Gamma(a)}e^{-cx}\mathbf{1}\{x>0\}$ (the density of $\gamma(c,a)$, $c,a>0$), and recall from Example 2.3.1 that

$$\textbf{1)} \qquad \widehat{f_{c,a}}(\theta) = \left(1 - \tfrac{\mathbf{i}\theta}{c}\right)^{-a} = \left(1 + \tfrac{\theta^2}{c^2}\right)^{-a/2} \exp\left(\mathbf{i}a\mathrm{Arctan}\,\tfrac{\theta}{c}\right).$$

On the other hand,

$$\textbf{2)} \qquad -\mathrm{Log}(1-z) = \sum_{n\geq 1} \frac{z^n}{n} = z + \frac{z^2}{2} + O(|z|^3), \quad \text{as } z \to 0.$$

Therefore,

$$\textbf{3)} \qquad \begin{cases} E\exp(\mathbf{i}\theta(Y_a - \sqrt{a})) & = & \widehat{f_{\sqrt{a},a}}(\theta)\exp(-\mathbf{i}\theta\sqrt{a}) \\ & \stackrel{1)}{=} & \left(1 - \frac{\mathbf{i}\theta}{\sqrt{a}}\right)^{-a} \exp(-\mathbf{i}\theta\sqrt{a}) \\ & = & \exp\left(-a\mathrm{Log}\left(1 - \frac{\mathbf{i}\theta}{\sqrt{a}}\right) - \mathbf{i}\theta\sqrt{a}\right) \\ & \stackrel{2)}{=} & \exp\left(a\left(\frac{\mathbf{i}\theta}{\sqrt{a}} - \frac{\theta^2}{2a} + O\left(\frac{|\theta|^3}{a^{3/2}}\right)\right) - \mathbf{i}\theta\sqrt{a}\right) \stackrel{a\to\infty}{\longrightarrow} \exp\left(-\frac{\theta^2}{2}\right). \end{cases}$$

Recall that $\exp\left(-\frac{\theta^2}{2}\right)$ is the Fourier transform of $N(0,1)$ (Example 2.2.4). We see from 3) and Proposition 2.4.3 that $Y_a - \sqrt{a} \xrightarrow{\mathrm{w}} N(0,1)$.
b) Suppose that $a \geq 2$. We then see from 1) that $f_{c,a} \in L^1(\mathbb{R}) \cap C(\mathbb{R})$ and $\widehat{f_{c,a}} \in L^1(\mathbb{R})$. Thus, we have by the inversion formula (Lemma 2.4.6 below) that

$$f_{c,a}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widehat{f_{c,a}}(\theta)\exp(-\mathbf{i}\theta x)d\theta, \quad \forall x \in \mathbb{R}.$$

In particular,

$$\textbf{4)} \qquad \frac{1}{\sqrt{a}\Gamma(a)}\left(\frac{a}{e}\right)^a = f_{\sqrt{a},a}(\sqrt{a}) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widehat{f_{\sqrt{a},a}}(\theta)\exp(-\mathbf{i}\theta\sqrt{a})d\theta$$

We know from 3) that

$$\textbf{5)} \qquad \widehat{f_{\sqrt{a},a}}(\theta)\exp(-\mathbf{i}\theta\sqrt{a}) \stackrel{a\to\infty}{\longrightarrow} \exp\left(-\tfrac{\theta^2}{2}\right), \quad \forall \theta \in \mathbb{R}^d.$$

Moreover, $\left(1 + \frac{\theta^2}{a}\right)^a$ is increasing in $a > 0$ and hence for $a \geq 2$,

$$\textbf{6)} \qquad |\widehat{f_{\sqrt{a},a}}(\theta)\exp(-\mathbf{i}\theta\sqrt{a})| \stackrel{1)}{=} \left(1 + \tfrac{\theta^2}{a}\right)^{-a/2} \leq \left(1 + \tfrac{\theta^2}{2}\right)^{-1} \in L^1(\mathbb{R}).$$

We now conclude from 4),5),6) and DCT that

$$\frac{1}{\sqrt{a}\Gamma(a)}\left(\frac{a}{e}\right)^a \stackrel{a\to\infty}{\longrightarrow} \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left(-\tfrac{\theta^2}{2}\right)d\theta = \frac{1}{\sqrt{2\pi}},$$

which is to be proved. \(^□^)/

To prove Proposition 2.4.1, we will use:

Lemma 2.4.6 *Suppose that $f, \widehat{f} \in L^1(\mathbb{R}^d)$.*

a) (Inversion formula) *For a.e. $x \in \mathbb{R}^d$,*

$$f(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} \exp(-\mathbf{i}\theta \cdot x)\widehat{f}(\theta)d\theta. \tag{2.37}$$

b) (Plancherel's formula) *Suppose in addition that f is continuous. Then, (2.37) holds for all $x \in \mathbb{R}^d$ and f is bounded. Moreover, for any Borel signed measure μ on $\mathbb{R}^d$,*

$$\int f d\mu = (2\pi)^{-d} \int_{\mathbb{R}^d} \widehat{f}(\theta)\widehat{\mu}(-\theta)d\theta. \tag{2.38}$$

Proof: a) We prepare

1) $\qquad h_t * f \longrightarrow f$ in $L^1(\mathbb{R}^d)$ as $t \to 0$, where $h_t(x) = (2\pi t)^{-d/2} \exp(-|x|^2/2t)$

We have that

$$|h_t * f - f|(x) \le \int_{\mathbb{R}^d} h_t(y)|f(x-y) - f(x)|dy = \int_{\mathbb{R}^d} h_1(y)|f(x - \sqrt{t}y) - f(x)|dy$$

and hence

2) $$\int_{\mathbb{R}^d} |h_t * f - f|(x)dx \le \int_{\mathbb{R}^d} h_1(y)g_t(y)dy \quad \text{where} \quad g_t(y) = \int_{\mathbb{R}^d} |f(x - \sqrt{t}y) - f(x)|dx.$$

We have for any $y \in \mathbb{R}^d$ that

$$\lim_{t \to 0} g_t(y) = 0 \quad \text{and} \quad 0 \le g_t(y) \le 2\int_{\mathbb{R}^d} |f(x)|dx.$$

Thus, by (2) and the dominated convergence theorem,

$$\lim_{t \to 0} \int_{\mathbb{R}^d} |h_t * f - f|(x)dx = 0.$$

We set $f^\vee(x) = (2\pi)^{-d}\widehat{f}(-x)$ $(x \in \mathbb{R}^d)$. We will next show that:

3) $\qquad f * h_t = (f^\wedge h_t^\wedge)^\vee$, where $h_t(x) = (2\pi t)^{-d/2} \exp(-|x|^2/2t)$ $(x \in \mathbb{R}^d,\ t > 0)$.

By (2.10),

4) $\qquad h_t^\wedge(\theta) = \exp(-t|\theta|^2/2)$.

Using (2.10) again, we see that $h_t = h_t^{\wedge\vee}$. Therefore,

$$\begin{aligned}
f * h_t(x) &= f * h_t^{\wedge\vee}(x) \\
&= (2\pi)^{-d} \int f(x-y)dy \int \underbrace{\exp(-\mathbf{i}\theta \cdot y)}_{=\exp(-\mathbf{i}\theta \cdot x)\exp(\mathbf{i}(\theta\cdot(x-y)))} h_t^\wedge(\theta)d\theta \\
&\overset{\text{Fubini}}{=} (2\pi)^{-d} \int \exp(-\mathbf{i}\theta \cdot x)h_t^\wedge(\theta)d\theta \underbrace{\int f(x-y)\exp(\mathbf{i}(\theta\cdot(x-y)))dy}_{=f^\wedge(\theta)} \\
&= (f^\wedge h_t^\wedge)^\vee(x).
\end{aligned}$$

We see from (4) and the dominated convergence theorem that

$$\lim_{t\to 0}(f^\wedge h_t^\wedge)^\vee(x) = f^{\wedge\vee}(x) \quad \text{for all } x \in \mathbb{R}^d.$$

Combining this, (1) and (3), we arrive at $f^{\wedge\vee} = f$, a.e., which is (2.37).
b) The right-hand side of (2.37) is bounded and continuous in x (Exercise 2.1.1). Thus, if f is continuous, it follows from a) that (2.37) is valid for all $x \in \mathbb{R}^d$, which also implies that f is bounded. Considering the positive and negative parts of the Jordan decomposition of μ, it is enough to prove (2.38), assuming that μ is a positive Borel measure. Then,

$$\begin{aligned}\int f d\mu &\stackrel{(2.37)}{=} \int_{\mathbb{R}^d} d\mu(x)(2\pi)^{-d}\int_{\mathbb{R}^d}\exp(-\mathbf{i}\theta\cdot x)\widehat{f}(\theta)d\theta \\ &\stackrel{\text{Fubini}}{=} (2\pi)^{-d}\int_{\mathbb{R}^d}\widehat{f}(\theta)d\theta \underbrace{\int \exp(-\mathbf{i}\theta\cdot x)d\mu(x)}_{=\widehat{\mu}(-\theta)}\end{aligned}$$

\(^□^)/

Now, we prove the following lemma which includes Proposition 2.4.1. The lemma can also be used in Exercise 2.4.11 and Exercise 2.4.12. To state the lemma, we introduce the following notation. For an open subset $G \subset \mathbb{R}^d$, let

$$\begin{aligned}C_{\rm c}(G) &= \{f \in C(\mathbb{R}^d)\ ;\ f \text{ has a compact support in } G\},\\ C_{\rm c}^\infty(G) &= C_{\rm c}(G)\cap C^\infty(G).\end{aligned}$$

Lemma 2.4.7 *Suppose that $(\mu_n)_{n\ge 0}$ are Borel finite measures on $\mathbb{R}^d$ such that $\mu_0(G^{\rm c}) = 0$ for an open subset $G \subset \mathbb{R}^d$ (To prove Proposition 2.4.1, it is enogh to take $G = \mathbb{R}^d$). Then, the following are equivalent:*

a) $\widehat{\mu_n}(\theta) \stackrel{n\to\infty}{\longrightarrow} \widehat{\mu_0}(\theta)$ *for all* $\theta \in \mathbb{R}^d$.

b) *(2.35) holds for all* $f \in C_{\rm b}(\mathbb{R}^d)$.

c) *(2.35) holds for all* $f \in C_{\rm c}^\infty(G)$ *and* $\overline{\lim\limits_{n\to\infty}}\, \mu_n(\mathbb{R}^d) \le \mu_0(\mathbb{R}^d)$.

Proof: a) ⇒ c): By setting $\theta = 0$ in the assumption a), we have $\mu_n(\mathbb{R}^d) \stackrel{n\to\infty}{\longrightarrow} \mu_0(\mathbb{R}^d)$, hence $\overline{\lim\limits_{n\to\infty}}\, \mu_n(\mathbb{R}^d) \le \mu_0(\mathbb{R}^d)$. Let us prove that (2.35) holds for all $f \in C_{\rm c}^\infty(\mathbb{R}^d)$ and therefore, for all $f \in C_{\rm c}^\infty(G)$. We have $\widehat{f} \in L^1(\mathbb{R}^d)$ for $f \in C_{\rm c}^\infty(\mathbb{R}^d)$, which is a well-known properties of the Fourier transform for the Schwartz space of rapidly decreasing functions (cf. [RS80, page 3, Theorem IX.1]), so that the Plancherel formula (2.38) is available [9]. On the other hand, we have

1) $$\sup_{n\ge 1}|\widehat{\mu_n}(-\theta)| \le \sup_{n\ge 1}\mu_n(\mathbb{R}^d) = \sup_{n\ge 1}\widehat{\mu_n}(0) \stackrel{\text{a)}}{<} \infty.$$

[9] The availability of the Plancherel formula is the very reason for which we go through the space $C_{\rm c}^\infty(\mathbb{R}^d)$, rather than working directly with the space $C_{\rm b}(\mathbb{R}^d)$. In fact, $\widehat{f}$ is not defined in general for $f \in C_{\rm b}(\mathbb{R}^d)$. Even for $f \in C_{\rm c}(\mathbb{R}^d)$, it is not true in general that $\widehat{f} \in L^1(\mathbb{R}^d)$ (A counterexample for $d = 1$ is provided by $f(x) = (1 - \log(1 - |x|))^{-1}\mathbf{1}_{\{|x|<1\}}$).

Therefore, by the dominated convergence theorem (DCT),

$$\int_{\mathbb{R}^d} f d\mu_n \stackrel{(2.38)}{=} (2\pi)^{-d} \int_{\mathbb{R}^d} \widehat{f}(\theta)\widehat{\mu_n}(-\theta)d\theta \stackrel{\text{a),1),DCT}}{\longrightarrow} (2\pi)^{-d} \int_{\mathbb{R}^d} \widehat{f}(\theta)\widehat{\mu}_0(-\theta)d\theta \stackrel{(2.38)}{=} \int_{\mathbb{R}^d} f d\mu_0.$$

c) $\Rightarrow$ b): We first verify that

2) any function $f \in C_{\rm c}(G)$ is uniformly approximated by an element of $C_{\rm c}^{\infty}(G)$.

Indeed, let $\varphi_\varepsilon(x) = \varepsilon^{-d}\varphi(x/\varepsilon)$, $\varepsilon > 0$ where $\varphi \in C^\infty(\mathbb{R}^d)$ is supported in the unit ball, $\int_{\mathbb{R}^d} \varphi = 1$. Set

$$(f * \varphi_\varepsilon)(x) \stackrel{\text{def}}{=} \int_{\mathbb{R}^d} \varphi_\varepsilon(x-y)f(y)dy.$$

Then, it is standard to verify that $f * \varphi_\varepsilon \in C_{\rm c}^{\infty}(G)$ for small enough ε and that

$$\sup_{x\in\mathbb{R}^d} |(f * \varphi_\varepsilon)(x) - f(x)| \stackrel{\varepsilon\to 0}{\longrightarrow} 0.$$

This proves 1).
By 2), we may assume that (2.35) holds for all $f \in C_{\rm c}(G)$. Let K_m, $m \geq 1$ be an increasing sequence of compact subsets in G such that $G = \bigcup_{m\geq 1} K_m$, and $h_m \in C_{\rm c}(G \to [0,1])$ be such that $h_m = 1$ on K_m. Then,

3) $\qquad h_m \stackrel{m\to\infty}{\longrightarrow} \mathbf{1}_G.$

Note also that for real sequences a_n and b_n,

4) $\qquad \underline{\lim}_{n\to\infty} (a_n + b_n) \leq \underline{\lim}_{n\to\infty} a_n + \overline{\lim}_{n\to\infty} b_n.$

Take $f \in C_{\rm b}(\mathbb{R}^d)$ with $M = \sup_x |f(x)|$. We then have by the dominated convergence theorem (DCT) that

$$\begin{aligned}
\int_{\mathbb{R}^d} f d\mu_0 + M\mu_0(\mathbb{R}^d) &= \int_{\mathbb{R}^d} (f+M)d\mu_0 = \int_G (f+M)d\mu_0 \\
&\stackrel{3),\ \text{DCT}}{=} \lim_{m\to\infty} \int_{\mathbb{R}^d} (f+M)h_m d\mu_0 \stackrel{\text{c)}}{=} \lim_{m\to\infty}\lim_{n\to\infty} \int_{\mathbb{R}^d} (f+M)h_m d\mu_n \\
&\leq \underline{\lim}_{n\to\infty} \int_{\mathbb{R}^d} (f+M)d\mu_n \stackrel{4)}{\leq} \underline{\lim}_{n\to\infty} \int_{\mathbb{R}^d} f d\mu_n + M\, \overline{\lim}_{n\to\infty} \mu_n(\mathbb{R}^d) \\
&\stackrel{\text{a)}}{\leq} \underline{\lim}_{n\to\infty} \int_{\mathbb{R}^d} f d\mu_n + M\mu_0(\mathbb{R}^d),
\end{aligned}$$

and hence that

5) $\qquad \int_{\mathbb{R}^d} f d\mu_0 \leq \underline{\lim}_{n\to\infty} \int_{\mathbb{R}^d} f d\mu_n.$

By replacing f by $-f$ in 5), we have

$$\int_{\mathbb{R}^d} f d\mu_0 \geq \overline{\lim}_{n\to\infty} \int_{\mathbb{R}^d} f d\mu_n,$$

which, together with 5), proves the desired convergence.
b) $\Rightarrow$ a): $x \mapsto \exp(\mathbf{i}\theta \cdot x)$ belongs to $C_{\mathrm{b}}(\mathbb{R}^d)$ for all $\theta \in \mathbb{R}^d$. \(^□^)/

Proof of Proposition 2.1.3: We only need to prove $\Leftarrow$. Thus, we have to prove that

$$\widehat{\mu^+}(\theta) = \widehat{\mu^-}(\theta) \text{ for all } \theta \in \mathbb{R}^d \implies \mu^+ = \mu^-,$$

where $\mu^{\pm}$ are positive and negative parts of the Jordan decomposition of μ. We consider a sequence $\nu_n = \mu^+$ ($\forall n \geq 1$), which is constant in n. Then we have by assumption that $\widehat{\nu_n}(\theta) = \widehat{\mu^+}(\theta) = \widehat{\mu^-}(\theta)$ for all $n \in \mathbb{N}$ and $\theta \in \mathbb{R}^d$, and hence that

$$\lim_{n\to\infty} \widehat{\nu_n}(\theta) \stackrel{(\nu_n = \mu^+)}{=} \widehat{\mu^+}(\theta) = \widehat{\mu^-}(\theta).$$

This implies by Proposition 2.4.1 that both $\mu^{\pm}$ are weak limits of the sequence μ_n, and hence $\mu^+ = \mu^-$ by the uniqueness of the weak limit (cf. Remark after Proposition 2.4.1). \(^□^)/

Exercise 2.4.1 Let $X, X_1, X_2, \ldots$ be $\mathbb{R}^d$-valued r.v.'s. Prove then that the following conditions are related as "a) or b) " $\Rightarrow$ c) $\Rightarrow$ d) $\Rightarrow$ e). a) $X_n \stackrel{n\to\infty}{\longrightarrow} X$, P-a.s. b) $X_n \stackrel{n\to\infty}{\longrightarrow} X$ in $L^p(P)$ for some $p \geq 1$. c) $X_n \stackrel{n\to\infty}{\longrightarrow} X$ in probability, i.e., $P(|X_n - X| > \varepsilon) \stackrel{n\to\infty}{\longrightarrow} 0$ for any $\varepsilon > 0$. d) $E|f(X_n) - f(X)| \stackrel{n\to\infty}{\longrightarrow} 0$ if $f : \mathbb{R}^d \to \mathbb{R}$ is bounded, uniformly continuous. e) $X_n \stackrel{n\to\infty}{\longrightarrow} X$ weakly.

Exercise 2.4.2 Show by an example that e) $\not\Rightarrow$ d) in Exercise 2.4.1. **Hint**: $X_n = (-1)^n X$, where $P(X = \pm 1) = 1/2$.

Exercise 2.4.3 Let $X, Y, X_1, X_2, \ldots$ be $\mathbb{R}^d$-valued r.v.'s such that $X_n \stackrel{\mathrm{w}}{\to} X$. Is it true in general that $X_n + Y \stackrel{\mathrm{w}}{\to} X + Y$?

Exercise 2.4.4 Let $X_1, X_2, \ldots$ be $\mathbb{R}^d$ valued r.v.'s and $c \in \mathbb{R}^d$. Prove then that $X_n \to c$ in probability if and only if $X_n \stackrel{\mathrm{w}}{\to} c$. **Hint**: $X_n \to c$ in probability if and only if $E\varphi(X_n) \to 0$, where $\varphi(x) = \frac{|x-c|}{1+|x-c|}$.

Exercise 2.4.5 Let (X_n, Y_n) be r.v.'s with values in $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. Suppose that X_n and Y_n are independent for each n and that $X_n \stackrel{\mathrm{w}}{\to} X$ and $Y_n \stackrel{\mathrm{w}}{\to} Y$. Prove then that $(X_n, Y_n) \stackrel{\mathrm{w}}{\to} (X, Y)$, and hence that $F(X_n, Y_n) \stackrel{\mathrm{w}}{\to} F(X, Y)$ for any $F \in C(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2})$.

Exercise 2.4.6 Let (X_n, Y_n) be r.v.'s with values in $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. Suppose that $X_n \stackrel{\mathrm{w}}{\to} X$ and $Y_n \stackrel{\mathrm{w}}{\to} c$ (Here, we do not assume that X_n and Y_n are independent for each n. Instead, we assume that c is a constant vector in $\mathbb{R}^{d_2}$). Prove then that $(X_n, Y_n) \stackrel{\mathrm{w}}{\to} (X, c)$, and hence that $F(X_n, Y_n) \stackrel{\mathrm{w}}{\to} F(X, c)$ for any $F \in C(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2})$. **Hint**: It is enough to show that

$$\lim_{n\to\infty} E\exp(\mathbf{i}\theta_1 \cdot X_n + \mathbf{i}\theta_2 \cdot Y_n) = E\exp(\mathbf{i}\theta_1 \cdot X + \mathbf{i}\theta_2 \cdot c) \quad \text{for } (\theta_1, \theta_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}.$$

In doing so, uniform continuity of the map $(x, y) \mapsto \exp(\mathbf{i}\theta_1 \cdot x + \mathbf{i}\theta_2 \cdot y)$ would help.

Exercise 2.4.7 Let $X, X_1, X_2, \ldots$ $\mathbb{R}^d$-valued r.v.'s. Suppose that X_n ($n = 1, 2, \ldots$) are mean-zero Gaussian r.v.'s and that they converge weakly to X. Prove then that X is a mean-zero Gaussian r.v. and that the covariance matrix $V = (v_{\alpha\beta})_{\alpha,\beta=1}^d$ is given by $v_{\alpha\beta} = \lim_{n\to\infty} E[X_{n\alpha} X_{n\beta}]$. **Hint**: Consider characteristic functions to see that limits $v_{\alpha\beta}$ ($1 \leq \alpha, \beta \leq n$) exist. Prove then that $E\exp(\mathbf{i}\theta \cdot X) = \exp(-\theta \cdot V\theta/2)$.

Exercise 2.4.8 Let $F_0(x) = \sum_{n=0}^{\infty} \frac{1}{(n!)^2} \left(\frac{x}{2}\right)^{2n}$, $x \geq 0$ (cf. (2.20)). Referring to Example 2.4.4, prove the following. **(i)** $F_0(x) = \frac{\exp(x)}{2\pi} \int_{-\pi}^{\pi} |\widehat{\pi_{\frac{x}{2}}}(\theta)|^2 d\theta$ for $x > 0$. **(ii)** $F_0(x) \stackrel{x\to\infty}{\sim} \frac{\exp(x)}{\sqrt{2\pi x}}$.

Exercise 2.4.9 Let $X, Y_1, Y_2, ...$ be r.v.'s with $X \approx \gamma_{c,a}$ and $Y_n \approx \beta_{a,n}$ ($n = 1, 2, ..$ cf. (1.27), (1.33)). Prove then that $nY_n \stackrel{\mathrm{w}}{\to} cX$. **Hint**: Let $S_n = X_1 + ... + X_n$ where $X_1, X_2, ...$ be i.i.d. such that $X_n \approx \gamma_{c,1}$. Then, $nY_n \approx \frac{nX}{X+S_n}$ by Example 1.7.5. Moreover, $\frac{nX}{X+S_n} \to cX$, P-a.s. by Theorem 1.10.2.

Exercise 2.4.10 Let $S_n = X_1 + ... + X_n$, where $(X_n)_{n\geq 1}$ are i.i.d. with Polya distributions (Exercise 2.2.1). Prove then that S_n/n converges weakly to (1)-Cauchy distribution as $n \to \infty$.

Exercise 2.4.11 Suppose that $(\mu_n)_{n\geq 0}$ are Borel finite measures on $\mathbb{R}$ such that $\mu_0(\mathbb{R}\backslash(0,1)) = 0$. Prove then that the following conditions (a) and (b) are equivalent. **(a)** $\widehat{\mu_n}(k) \stackrel{n\to\infty}{\longrightarrow} \widehat{\mu_n}(k)$ for all $k \in \mathbb{Z}$. **(b)** $\mu_n \stackrel{\mathrm{w}}{\longrightarrow} \mu_0$ as $n \to \infty$. **Hint**: It is enough to prove that a) implies b). Assume a). Then, by Lemma 2.4.7, it is enough to prove that $\int f d\mu_n \stackrel{n\to\infty}{\longrightarrow} \int f d\mu_0$ for $f \in C_{\mathrm{c}}((0,1))$, while $f \in C_{\mathrm{c}}((0,1))$ is uniformly approximated on $[0,1]$ by trigonometric polynomials (Exercise 1.8.3).

Exercise 2.4.12 (**Weyl's theorem**) Let $\alpha_n = n\alpha - \lfloor n\alpha \rfloor$, $n \in \mathbb{N}$, where $\alpha \in \mathbb{R}\backslash\mathbb{Q}$ and $\lfloor y \rfloor = \max\{n \in \mathbb{Z}\ ;\ n \leq y\}$ for $y \in \mathbb{R}$. Then, use Exercise 2.4.11 to prove that the measures $\mu_n = \frac{1}{n}\sum_{k=0}^{n-1} \delta_{\alpha_k}$ converges weakly to the uniform distribution on $(0,1)$.

Exercise 2.4.13 (**Benford law**) Let $q \geq 2$ be an integer. Then, each $x \in (0,\infty)$ is expressed as the q-adic expansion.

$$x = dq^n + \sum_{k=-\infty}^{n-1} d_k q^k,$$

where $n \in \mathbb{Z}$, $d \in \{1, \ldots q-1\}$, and $d_k \in \{0, \ldots q-1\}$ for $-\infty < k \leq n-1$. Moreover, n and d are uniquely determined. We call $d(x)$ the *initail digit* of x. Let $\pi(x) = x - \lfloor x \rfloor$ and suppose that $\{x_n\}_{n\geq 1} \subset (0,\infty)$ is a sequence for which the following measures converges to the uniform distribution on $(0,1)$.

$$\mu_n = \frac{1}{n}\sum_{j=1} \delta_{\pi(x_j)}, \quad n \geq 1.$$

Then, prove that

$$\frac{1}{n}\sum_{j=1} \mathbf{1}\{d(q^{x_j}) = d\} \stackrel{n\to\infty}{\longrightarrow} \log_q\left(\frac{d+1}{d}\right), \quad \text{for all } d = 1, \ldots, q-1.$$

Hint: Note that $d(q^x) = d \Leftrightarrow \pi(x) \in [\log_q d, \log_q(d+1))$. Then, the desired convergence follows immediately from the assumption.

Exercise 2.4.14 ($\star$) Let X be a real r.v.and $\varphi(\theta) = E\exp(\mathbf{i}\theta X)$. Then, $\varphi \in C^2 \iff X \in L^2(P)$. Prove this by assuming that X is symmetric (cf. Exercise 2.4.15 for the removal of this extra assumption). **Hint**: If $\varphi \in C^2$, then $\frac{1}{2}\varphi''(0) = \lim_{\theta\to 0} \frac{\varphi(\theta)+\varphi(-\theta)-2\varphi(0)}{\theta^2}$.

Exercise 2.4.15 ($\star$) Let X be a real r.v. **(i)** For $p \in [1,\infty)$, prove that $X - \widetilde{X} \in L^p(P) \iff X \in L^p(P)$, where $\widetilde{X}$ is an independent copy of X. **Hint**: $X \in L^p(P)$, if $X - c \in L^p(P)$ for some constant $c \in \mathbb{R}$. Combine this observation with Fubini's theorem. **(ii)** Use (i) to remove the assumption "symmetric" from Exercise 2.4.14

Exercise 2.4.16 ($\star$) Suppose that $X, X_1, X_2...$ are real r.v's and that $X_n \xrightarrow{w} X$. Prove then that ess.sup$\underline{X} \leq$ ess.sup$X \leq$ ess.sup$\overline{X}$, where $\underline{X} = \underline{\lim}_{n\to\infty} X_n$ and $\overline{X} = \overline{\lim}_{n\to\infty} X_n$ and, for a r.v. $Y \in [-\infty, \infty]$, ess.sup$Y$ is the supremum of $m \in \mathbb{R}$ such that $P(Y > m) > 0$.

Exercise 2.4.17 Referring to Proposition 2.4.1 and its proof, is it true that **c)** $\Rightarrow$ **b)**?
Hint $\mu_n = \delta_{x_n}$, where $|x_n| \to \infty$.

2.5 ($\star$) Convergence of Moments

Let $(Y_n)_{n\geq 0}$ be $\mathbb{R}^d$-valued r.v.'s such that $Y_n \xrightarrow{w} Y_0$, and let $f \in C(\mathbb{R}^d)$. If f is bounded, we have

$(*)$ $\lim_{n\to\infty} Ef(Y_n) = Ef(Y_0)$.

On the other hand, it is natural to ask under which condition we still have $(*)$ even when f is *unbounded*, e.g., $f(y) = |y|$. The following definition plays an important role in answering this question, where we have $X_n = f(Y_n)$ in mind.

Definition 2.5.1 (**uniform integrability**) Let Λ be a set. Real r.v.'s $(X_\lambda)_{\lambda\in\Lambda}$ are said to be *uniformly integrable* (*u.i.* in short) if

$$\sup_{\lambda\in\Lambda} E[|X_\lambda| : |X_\lambda| > m] \longrightarrow 0 \text{ as } m \to \infty.$$

The next lemma shows that the uniform integrability is close to, but slightly more than that

$$\sup_{\lambda\in\Lambda} E|X_\lambda| < \infty. \tag{2.39}$$

Lemma 2.5.2 *Let $(X_\lambda)_{\lambda\in\Lambda}$ be real r.v.'s.*

a) *If $(X_\lambda)_{\lambda\in\Lambda}$ are u.i., then (2.39) holds.*

b) *Suppose that there exists a non-decreasing $\varphi : [0, \infty) \to [0, \infty)$ such that*

$$\lim_{x\to\infty} \varphi(x) = \infty, \quad \sup_{\lambda\in\Lambda} E[|X_\lambda|\varphi(|X_\lambda|)] < \infty.$$

Then, $(X_\lambda)_{\lambda\in\Lambda}$ are u.i.

Proof: Let $\varepsilon_m = \sup_{\lambda\in\Lambda} E[|X_\lambda| : |X_\lambda| > m]$.
a):$\varepsilon_m \leq 1$ for large enough m, and for such m and for all $\lambda \in \Lambda$,

$$E|X_\lambda| \leq E[|X_\lambda| : |X_\lambda| \leq m] + E[|X_\lambda| : |X_\lambda| > m] \leq m + \varepsilon_m < m + 1.$$

b): By the monotonicity of φ and (a variant of) Chebychev's inequality (Proposition 1.1.9),

$$E[|X_\lambda| : |X_\lambda| > m] \leq E[|X_\lambda| : \varphi(|X_\lambda|) \geq \varphi(m)] \leq \varphi(m)^{-1}E[|X_\lambda|\varphi(|X_\lambda|)].$$

Thus, $\varepsilon_m \leq \varphi(m)^{-1}C \to 0$, as $m \to \infty$, where $C = \sup_{\lambda\in\Lambda} E[|X_\lambda|\varphi(X_\lambda)] < \infty$. \(^□^)/

Example 2.5.3 Let $S_n = X_1 + \ldots + X_n$, where $(X_n)_{n\geq 1}$ are real r.v.'s such that

$$\sup_{n\geq 1} \mathrm{var}\, X_n \leq M < \infty, \quad \mathrm{cov}(X_m, X_n) = 0 \text{ if } m \neq n.$$

Then, $Y_n = (S_n - ES_n)/\sqrt{n}$ are u.i. In fact,

$$E[|Y_n|^2] = \frac{1}{n}\mathrm{var}\, S_n = \frac{1}{n}\sum_{k=1}^{n} \mathrm{var}\, X_k \leq M.$$

Thus, Lemma 2.5.2b) applies.

Lemma 2.5.4 (Fatou's lemma for weak convergence) *Suppose that X, X_n $(n \in \mathbb{N})$ be real r.v.'s such that $X_n \to X$ weakly. Then,*

$$E|X| \leq \varliminf_{n\to\infty} E|X_n|. \tag{2.40}$$

Proof: Since $\mathbb{R} \ni x \mapsto |x| \wedge m$ is in $C_{\mathrm{b}}(\mathbb{R})$ for any $m > 0$, we have that

$$E|X| = \sup_{m\geq 0} E[|X| \wedge m] = \sup_{m\geq 0} \lim_{n\to\infty} E[|X_n| \wedge m] \leq \varliminf_{n\to\infty} E|X_n|.$$

\(^□^)/

Proposition 2.5.5 *Suppose that X, X_n $(n \in \mathbb{N})$ be real r.v.'s such that $X_n \to X$ weakly. Then, the following are equivalent.*

a) $(X_n)_{n\in\mathbb{N}}$ *are u.i.*

b) $X, X_n \in L^1(P)$ $(\forall n \in \mathbb{N})$, $EX_n \stackrel{n\to\infty}{\longrightarrow} EX$ *and* $E|X_n| \stackrel{n\to\infty}{\longrightarrow} E|X|$.

c) $X, X_n \in L^1(P)$ $(\forall n \in \mathbb{N})$ *and* $E|X_n| \stackrel{n\to\infty}{\longrightarrow} E|X|$.

Suppose in particular that $X_n \to X$ in probability. Then, the following is also equivalent to a)–c) above.

d) $X, X_n \in L^1(P)$ $(\forall n \in \mathbb{N})$ *and* $E|X_n - X| \stackrel{n\to\infty}{\longrightarrow} 0$.

Proof: a) ⇒ b): It follows from Lemma 2.5.2 and (2.40) that $X, X_n \in L^1(P)$ $(\forall n \in \mathbb{N})$. We prove that $EX_n \stackrel{n\to\infty}{\longrightarrow} EX$ and $E|X_n| \stackrel{n\to\infty}{\longrightarrow} E|X|$, by showing that $E[X_n^{\pm}] \stackrel{n\to\infty}{\longrightarrow} E[X^{\pm}]$. We note that

$$\begin{aligned} X_n \longrightarrow X \text{ weakly} \quad &\Longrightarrow \quad X_n^{\pm} \longrightarrow X^{\pm} \text{ weakly}, \\ (X_n)_{n\in\mathbb{N}} \text{ are u.i.} \quad &\Longrightarrow \quad \text{so are } (X_n^{\pm})_{n\in\mathbb{N}}. \end{aligned}$$

Thus, it is enough to prove that $EX_n \stackrel{n\to\infty}{\longrightarrow} EX$ assuming that $X, X_n \geq 0$ $(n \in \mathbb{N})$. By (2.40), it is enough to show that

1) $\quad \varlimsup_{n\to\infty} EX_n \leq EX.$

Note that

2) $$\varlimsup_{n\to\infty} E[X_n : X_n \le m] \le EX \quad \text{for any } m > 0.$$

In fact,

$$\varlimsup_{n\to\infty} E[X_n : X_n \le m] \le \varlimsup_{n\to\infty} E[X_n \wedge m] = E[X \wedge m] \le EX.$$

Then, with $\varepsilon_m \stackrel{\text{def}}{=} \sup_{n\ge 1} E[X_n : X_n > m]$,

$$\varlimsup_{n\to\infty} EX_n = \varlimsup_{n\to\infty} (E[X_n : X_n \le m] + E[X_n : X_n > m]) \stackrel{2)}{\le} EX + \varepsilon_m.$$

Since m is arbitrary, we get 1).
b) $\Rightarrow$ c): Obvious.
c) $\Rightarrow$ a): Let $\varepsilon > 0$ be arbitrary. Since $E|X_n| \to E|X|$, there exists an $n_1 = n_1(\varepsilon) \in \mathbb{N}$ such that

3) $E|X_n| < E|X| + \varepsilon/4$ for $n \ge n_1$.

For $m > 0$, let $f_m \in C_{\mathrm{b}}(\mathbb{R})$ be defined by

$$f_m(x) = \begin{cases} x & \text{if } x \in [0, m/2] \\ m - x & \text{if } x \in [m/2, m] \\ 0 & \text{if } x \notin [0, m] \end{cases}$$

Then, by MCT, there exists an $\ell = \ell(\varepsilon) > 0$ such that

4) $E|X| < Ef_\ell(|X|) + \varepsilon/4$.

Since $X_n \to X$ weakly, there exists an $n_2 = n_2(\varepsilon)$ such that

5) $E[|X_n| : |X_n| \le \ell] \ge Ef_\ell(|X_n|) \ge Ef_\ell(|X|)| - \varepsilon/4$ for $n \ge n_2$.

By 3)–5), we have for $n \ge n_3 \stackrel{\text{def}}{=} n_1 \vee n_2$ and $m \ge \ell$ that

$$\begin{aligned} E[|X_n| : |X_n| > m] &\le E[|X_n| : |X_n| > \ell] = E|X_n| - E[|X_n| : |X_n| \le \ell] \\ &\stackrel{3),5)}{\le} E|X| - Ef_\ell(|X|) + \varepsilon/2 \stackrel{4)}{\le} 3\varepsilon/4. \end{aligned}$$

Note that n_3 depends only on ε. Thus, there exists an $m_0 = m_0(\varepsilon)$ such that

$$\max_{n \le n_3} E[|X_n| : |X_n| > m] < \varepsilon/4 \text{ for } m \ge m_0.$$

Putting these together, we conclude that

$$\sup_{n\in\mathbb{N}} E[|X_n| : |X_n| > m] < \varepsilon \text{ for } m \ge \ell \vee m_0.$$

We suppose from here on that $X_n \to X$ in probability.
a) $\Rightarrow$ d): Let $\varepsilon > 0$ be arbitrary. By a) and the integrability of X, there exists an $m = m(\varepsilon)$ such that

6) $$\sup_{n\in\mathbb{N}} E[|X_n| : |X_n| > m] + E[|X| : |X| > m] < \varepsilon/2.$$

Let $g_m \in C_{\mathrm{b}}(\mathbb{R})$ be defined by

$$g_m(x) = \begin{cases} -m & \text{if } x \in (-\infty, -m] \\ x & \text{if } x \in [-m, m] \\ m & \text{if } x \in [m, \infty) \end{cases}$$

Since $X_n \to X$ in probability, we have that

$$E|g_m(X_n) - g_m(X)| \stackrel{n\to\infty}{\longrightarrow} 0,$$

(Exercise 2.4.1) and hence, there exists an $n_0 = n_0(\varepsilon)$ such that

7) $\qquad \sup_{n \geq n_0} E|g_m(X_n) - g_m(X)| < \varepsilon/2.$

Note that $|g_m(x) - x| = (|x| - m)\mathbf{1}_{|x|>m} \leq |x|\mathbf{1}_{|x|>m}$. Thus, for $n \geq n_0$,

$$\begin{aligned} E|X_n - X| &\leq E|X_n - g_m(X_n)| + E|g_m(X_n) - g_m(X)| + E|g_m(X) - X| \\ &\leq E[|X_n| : |X_n| > m] + E|g_m(X_n) - g_m(X)| + E[|X| : |X| > m] \\ &\stackrel{6),7)}{<} \varepsilon/2 + \varepsilon/2 = \varepsilon. \end{aligned}$$

d) ⇒ c): Obvious. \(^□^)/

Remarks: Let everything be as in Proposition 2.5.5.
1) The following condition does not imply a)–c).

c') $X, X_n \in L^1(P)$ $(\forall n \in \mathbb{N})$, $EX_n \stackrel{n\to\infty}{\longrightarrow} EX$.

For example, Let U be a r.v. uniformly distributed on $(-1, 1)$, and let $X \equiv 0$, and $X_n = n^2 U \mathbf{1}\{|U| \leq 1/n\}$. Then, $X, X_n \in L^1(P)$ $(\forall n \in \mathbb{N})$, $X_n \to X$ a.s. Moreover, $EX = EX_n = 0$, hence $EX_n \to EX$. However, $E|X| = 0$, $E|X_n| = 1/2$, hence $E|X_n| \not\to E|X|$.
2) a)–c) do not imply d) without assuming that $X_n \to X$ in probability. For example, let $P(X = \pm 1) = 1/2$ and $X_n = (-1)^n X$. Since $X_n \approx X$, $X_n \to X$ weakly and $(X_n)_{n\in\mathbb{N}}$ are u.i. But for odd n's, $|X_n - X| = 2$ and hence $E|X_n - X| = 2$.

Exercise 2.5.1 Disprove the converse to Lemma 2.5.2a) with the following example: let P be the Lebesgue measure on $(\Omega, \mathcal{F}) \stackrel{\text{def}}{=} ([0, 1], \mathcal{B}([0, 1]))$ and $X_n(\omega) = n\mathbf{1}\{\omega \leq 1/n\}$, $n \geq 1$.

Exercise 2.5.2 Prove that real r.v.'s $(X_n)_{n\geq 1}$ are u.i. if $E[\sup_{n\geq 1} |X_n|] < \infty$.

Exercise 2.5.3 Suppose that $X_n > 0$, $n \geq 1$ are i.i.d. and that $E[X_1^{-\varepsilon}] < \infty$ for some $\varepsilon > 0$. Prove then that the r.v.'s $n/(X_1 + \ldots + X_n)$ converge as $n \to \infty$ to $1/EX_1$ a.s. and in $L^1(P)$ (with convention $1/\infty = 0$). **Hint**: Show the convergence in $L^1(P)$ via the uniform integrability.

2.6 The Central Limit Theorem

Recall that we have introduced in Example 2.2.4 the Gaussian distribution $N(m, V)$, where $m \in \mathbb{R}^d$, and V is a $d \times d$ symmetric, non-negative definite matrix. Recall also that we have introduced in Proposition 2.4.3 the notion of weak convergence of r.v.'s. In this section, we will discuss the following

Theorem 2.6.1 (The Central Limit Theorem) *Let $(\Omega, \mathcal{F}, P)$ be a probability space and $X_n : \Omega \to \mathbb{R}^d$ $(n \geq 1)$ be i.i.d. with $E[|X_1|^2] < \infty$. Define*

$$\begin{aligned} S_n &= X_1 + ... + X_n, \\ m &= (E[X_{1,\alpha}])_{\alpha=1}^d \in \mathbb{R}^d \quad \text{and} \quad V = (\text{cov}(X_{1,\alpha}, X_{1,\beta}))_{\alpha,\beta=1}^d. \end{aligned}$$

Then,

$$\frac{S_n - nm}{\sqrt{n}} \xrightarrow{\text{w}} N(0, V) \quad \text{as } n \to \infty, \tag{2.41}$$

Remarks : **1)** Theorem 2.6.1 tells us the following information on the distribution of S_n for large n. Let Y be r.v. such that $Y \approx N(0, V)$. Roughly speaking, Theorem 2.6.1 says that for large n,

$$\frac{S_n - nm}{\sqrt{n}} \overset{\text{approximately}}{\approx} Y$$

or

$$S_n \overset{\text{approximately}}{\approx} nm + \sqrt{n}\, Y.$$

2) The "central limit theorem" is often abbreviated as CLT.

Although it requires some work to prove CLT in the generality of Theorem 2.6.1, the proof is remarkably easy in some examples:

Example 2.6.2 (**CLT for Poisson r.v.'s**) Let π_c denote the (c)-Poisson distribution and suppose that $X_n \approx \pi_1$ in Theorem 2.6.1. Recall that

$$EX_n = \text{var}\, X_n = 1, \quad \text{(Exercise 1.2.2)}.$$

Recall also that:

$$S_n = X_1 + ... + X_n \approx \pi_n \ \ \text{(cf. (1.65))}.$$

Therefore, by Example 2.4.4,

$$\frac{S_n - n}{\sqrt{n}} \xrightarrow{\text{w}} N(0, 1) \ \ (n \to \infty).$$

Thus we have verified Theorem 2.6.1 in this special case. \(^□^)/

Example 2.6.3 ($\star$) (**Stirling's formula**) Let us prove as an application of CLT for Poisson r.v.'s (Example 2.6.2) that

1) $\qquad n! \sim \sqrt{2\pi n}(n/e)^n \quad \text{as } n \to \infty.$

Proof: Let N be a r.v. with $P(N = n) = \frac{r^n e^{-r}}{n!}$ ((r)-Poisson r.v.), Then,

2) $$E[(N-r)^-] = r\,\frac{r^{\lfloor r\rfloor}e^{-r}}{\lfloor r\rfloor!}.$$

In fact,

$$\begin{aligned} E[(N-r)^-] &= \sum_{n=0}^{\lfloor r\rfloor}(r-n)\frac{r^n e^{-r}}{n!} = r\sum_{n=0}^{\lfloor r\rfloor}\frac{r^n e^{-r}}{n!} - \sum_{n=1}^{\lfloor r\rfloor}\frac{r^n e^{-r}}{(n-1)!} \\ &= r\sum_{n=0}^{\lfloor r\rfloor}\frac{r^n e^{-r}}{n!} - r\sum_{n=0}^{\lfloor r\rfloor-1}\frac{r^n e^{-r}}{n!} = r\,\frac{r^{\lfloor r\rfloor}e^{-r}}{\lfloor r\rfloor!}. \end{aligned}$$

Now, let S_n be an (n)-Poisson r.v. Then,

3) $$E\left[\left(\frac{S_n-n}{\sqrt{n}}\right)^-\right] = n^{-1/2}E\left[(S_n-n)^-\right] \stackrel{(2)}{=} n^{-1/2}\cdot n\cdot\frac{n^n e^{-n}}{n!} = \frac{n^{n+\frac{1}{2}}e^{-n}}{n!}.$$

Since $(S_n-n)/\sqrt{n}$ $(n\geq 1)$ are uniformly integrable by Example 2.5.3, so are their negative parts. Thus, we conclude 1) from 3), CLT (Example 2.6.2) and Proposition 2.5.5 as follows:

$$\begin{aligned} \lim_{n\to\infty}\frac{n^{n+\frac{1}{2}}e^{-n}}{n!} &\stackrel{3)}{=} \lim_{n\to\infty}E\left[\left(\frac{S_n-n}{\sqrt{n}}\right)^-\right] = \sqrt{\frac{1}{2\pi}}\int_{-\infty}^0 x^- e^{-x^2/2}dx \\ &= \sqrt{\frac{1}{2\pi}}\int_0^\infty xe^{-x^2/2}dx = \sqrt{\frac{1}{2\pi}} \end{aligned}$$

\(^□^)/

Exercise 2.6.1 Suppose that $X_n \approx N(m,V)$ in Theorem 2.6.1. Prove then that $\frac{S_n-mn}{\sqrt{n}} \approx N(0,V)$ for any $n\geq 1$. Thus the theorem in this special case is trivial.

Exercise 2.6.2 (**A generalization of CLT**) Let $(S_n)_{n\geq 0}$ be as in Theorem 2.6.1 and $Y \approx N(0,V)$ Suppose that $f:\mathbb{R}^d\to\mathbb{R}^m$ be measurable, differentiable at m, and that

$$|f(m+x)-f(m)-f'(m)x| \leq C|x|^2 \quad \text{for all } x\in\mathbb{R}^d$$

where C is a constant. Use (2.41) to show that

$$\sqrt{n}\,(f(S_n/n)-f(m)) \stackrel{\text{w}}{\longrightarrow} f'(m)Y \quad \text{as } n\to\infty,$$

This result includes (2.41) as a special case that $f(x)=x$.
Hint: Set $Y_n=(S_n-mn)/\sqrt{n}$ and $g(x)=f(m+x)-f(m)-f'(m)x$ to write

$$\sqrt{n}\,(f(S_n/n)-f(m)) = f'(m)Y_n + \sqrt{n}g(Y_n/\sqrt{n}).$$

Then, apply Exercise 2.4.6 to $F(x,y)=x+y$.

Exercise 2.6.3 ($\star$) Let $X_1,X_2,...$ be mean-zero, real iid with $E[X_1^2]\in(0,\infty)$ and let $S_n = X_1+\ldots+X_n$. Prove then that $P(\overline{\lim}_{n\to\infty}\frac{S_n}{\sqrt{n}}=\infty) = P(\underline{\lim}_{n\to\infty}\frac{S_n}{\sqrt{n}}=-\infty)=1$. [Hint: Use the CLT and Fatou's lemma to show that $P(\overline{\lim}_{n\to\infty}\frac{S_n}{\sqrt{n}}\geq x)>0$ and that $P(\underline{\lim}_{n\to\infty}\frac{S_n}{\sqrt{n}}\leq x)>0$ for any $x\in\mathbb{R}$. Then, combine these with Kolmogorov's zero-one law (Lemma 1.6.4) to deduce the conclusion.]

Exercise 2.6.4 ($\star$) (*Wallis' formula*) Prove that $4^{-n}\binom{2n}{n} \stackrel{n\to\infty}{\sim} \frac{1}{\sqrt{\pi n}}$ in two different way as follows. (i) Prove Wallis' formula by applying Stirling's formula (cf. (2.50)). (ii) Let S_n be r.v. with $P(S_n = r) = 2^{-n}\binom{n}{r}$. Prove first that $E[(S_{2n} - n)^-] = \frac{n}{2}4^{-n}\binom{2n}{n}$ and then use CLT to conclude Wallis' formula as in Example 2.6.3.

Exercise 2.6.5 ($\star$) (*chi-square test*) Referring to Theorem 2.6.1, suppose in addition that $E[X_{1,\alpha}X_{1,\beta}] = q_\alpha \delta_{\alpha,\beta}$ with $q_\alpha > 0$ ($\alpha, \beta = 1, \dots, d$). Then,

$$(*) \qquad \sum_{\alpha=1}^{d} \frac{(S_{n,\alpha} - m_\alpha n)^2}{q_\alpha n} \stackrel{\mathrm{w}}{\longrightarrow} \sum_{\alpha=1}^{d-1} |Y_\alpha|^2 + (1 - |\ell|^2)|Y_d|^2 \quad \text{as } n \to \infty,$$

where $Y_1, \dots, Y_d$ are i.i.d. $\approx N(0,1)$ and $\ell = (m_\alpha/\sqrt{q_\alpha})_{\alpha=1}^d$. Prove this by successively verifying the following. **i)** $V = D(I_d - \ell \otimes \ell)D$, where $D = (q_\alpha^{1/2}\delta_{\alpha,\beta})_{\alpha,\beta=1}^d$ and $\ell \otimes \ell = (\ell_\alpha \ell_\beta)_{\alpha,\beta=1}^d$. **ii)** $\ell \in \mathrm{Ker}(|\ell|^2 - \ell \otimes \ell)$ and $(\mathbb{R}\ell)^\perp \subset \mathrm{Ker}(\ell \otimes \ell)$. **iii)** $|Z|^2 \approx \sum_{\alpha=1}^{d-1} |Y_\alpha|^2 + (1 - |\ell|^2)|Y_d|^2$ for a r.v. $Z \approx N(0, I_d - \ell \otimes \ell)$. **iv)** $D^{-1}\left(\frac{S_n - nm}{\sqrt{n}}\right) \stackrel{\mathrm{w}}{\longrightarrow} N(0, I_d - \ell \otimes \ell)$. **v)** $(*)$ holds.

Remark Here is a typical setting to which the result of Exercise 2.6.5 can be applied. Let ξ_n be i.i.d. with values in a measurable space $(S, \mathcal{B})$, and $B_1, \dots, B_d \in \mathcal{B}$ be disjoint sets with $q_\alpha \stackrel{\text{def}}{=} P(\xi_1 \in B_\alpha) > 0$ ($\alpha = 1, \dots, d$). Then, the assumption of Exercise 2.6.5 is satisfied by $X_n \stackrel{\text{def}}{=} (\mathbf{1}\{\xi_n \in B_\alpha\})_{\alpha=1}^d$. Moreover, if $q_1 + \cdots + q_d = 1$, then $\lambda_0 = 1$, and therefore, the limit law for $(*)$ is χ^2_{d-1}.

2.7 Proof of the Central Limit Theorem

We start by explainning the outline of the proof. We will prove that

$$E\exp\left(\mathbf{i}\theta \cdot \frac{S_n - nm}{\sqrt{n}}\right) \stackrel{n\to\infty}{\longrightarrow} \exp\left(-\tfrac{1}{2}\theta \cdot V\theta\right) \quad \text{for all } \theta \in \mathbb{R}^d. \tag{2.42}$$

By Proposition 2.4.1, (2.42) finishes the proof of Theorem 2.6.1. We set $Y = X_1 - m$ and $\varphi(\theta) = E\exp(\mathbf{i}\theta \cdot Y)$. Then,

$$\begin{aligned} E\exp\left(\mathbf{i}\theta \cdot \frac{S_n - mn}{\sqrt{n}}\right) &= E\prod_{j=1}^{n} \exp\left(\mathbf{i}\theta \cdot \frac{X_j - m}{\sqrt{n}}\right) \stackrel{(1.53)}{=} \left(E\exp\left(\mathbf{i}\theta \cdot \frac{Y}{\sqrt{n}}\right)\right)^n \\ &= \varphi\left(\frac{\theta}{\sqrt{n}}\right)^n \end{aligned} \tag{2.43}$$

We will show in Lemma 2.7.2 below that:

$$\varphi(\theta) = 1 - \tfrac{1}{2}\theta \cdot V\theta + o(|\theta|^2), \quad \theta \longrightarrow 0. \tag{2.44}$$

We will see by Lemma 2.7.3 below that

$$\varphi\left(\frac{\theta}{\sqrt{n}}\right)^n = \left(1 - \frac{\theta \cdot V\theta}{2n} + o\left(\frac{|\theta|^2}{n}\right)\right)^n \stackrel{n\to\infty}{\longrightarrow} \exp\left(-\tfrac{1}{2}\theta \cdot V\theta\right),$$

which proves (2.42).

We first prepare an elementary estimate.

Lemma 2.7.1 *For $t \in \mathbb{R}$,*

$$\left| \exp(\mathbf{i}t) - 1 - \mathbf{i}t + \frac{t^2}{2} \right| \leq |t|^3 \wedge |t|^2. \tag{2.45}$$

Proof: We will prove for $z \in \mathbb{C}$ and $n \in \mathbb{N}\backslash\{0\}$ that

1) $$\left| \exp z - \sum_{m=0}^{n} \frac{z^m}{m!} \right| \leq \frac{|z|^{n+1} \exp((\operatorname{Re} z)^+)}{(n+1)!} \wedge \frac{|z|^n (\exp((\operatorname{Re} z)^+) + 1)}{n!}$$

By setting $z = \mathbf{i}t$ and $n = 2$ in 1), we obtain (2.45). We fix $z \in \mathbb{C}$ and introduce $f(t) = e^{tz}$, $t \in \mathbb{R}$. By Taylor's theorem,

$$\begin{aligned} g_n(z) &\stackrel{\text{def}}{=} \exp z - \sum_{m=0}^{n} \frac{z^m}{m!} = f(1) - \sum_{m=0}^{n} \frac{f^{(m)}(0)}{m!} \\ &= \frac{1}{n!} \int_0^1 (1-t)^n f^{(n+1)}(t) dt = \frac{z^{n+1}}{n!} \int_0^1 (1-t)^n \exp(tz) dt. \end{aligned}$$

Since $|\exp(tz)| = \exp(t \operatorname{Re} z) \leq \exp((\operatorname{Re} z)^+)$, we obtain

2) $$|g_n(z)| \leq \frac{|z|^{n+1} \exp((\operatorname{Re} z)^+)}{(n+1)!}$$

On the other hand,

$$|g_n(z)| = \left| g_{n-1}(z) + \frac{z^n}{n!} \right| \stackrel{2)}{\leq} \frac{|z|^n (\exp((\operatorname{Re} z)^+) + 1)}{n!}.$$

\(^□^)/

We now present a lemma which implies (2.44). This lemma will also play an important role in the proof of Theorem 3.2.2.

Lemma 2.7.2 *Let $Y = X_1 - m$. Then,*

$$\left| E \exp(\mathbf{i}\theta \cdot Y) - (1 - \tfrac{1}{2}\theta \cdot V\theta) \right| = o(|\theta|^2) \quad \text{as } |\theta| \searrow 0. \tag{2.46}$$

Proof: We have that

$$E[\theta \cdot Y] = \sum_{\alpha=1}^{d} \theta_\alpha E[Y_\alpha] = 0, \quad E[(\theta \cdot Y)^2] = \sum_{\alpha,\beta=1}^{d} \theta_\alpha \theta_\beta E[Y_\alpha Y_\beta] = \theta \cdot V\theta,$$

and hence that

1) $$\begin{cases} E \exp(\mathbf{i}\theta \cdot Y) - (1 - \frac{1}{2}\theta \cdot V\theta) &= E\left[\exp(\mathbf{i}\theta \cdot Y) - 1 - \mathbf{i}\theta \cdot Y + \frac{1}{2}(\theta \cdot Y)^2\right] \\ &= E[f(\theta \cdot Y)], \end{cases}$$

where $f(t) = \exp(\mathbf{i}t) - 1 - \mathbf{i}t + \frac{t^2}{2}$ $(t \in \mathbb{R})$. Therefore,

$$
\textbf{2)} \left\{ \begin{array}{lcl} \text{left-hand side of (2.46)} & \overset{1)}{=} & |E\left[f(Y\cdot\theta)\right]| \leq E\left[|f(Y\cdot\theta)|\right] \\ & \overset{(2.45)}{\leq} & E\left[|Y\cdot\theta|^3 \wedge |Y\cdot\theta|^2\right] \leq |\theta|^2 E\left[|Y|^2(|Y||\theta|\wedge 1)\right]. \end{array} \right.
$$

We see by the dominated convergence theorem that

$$
\lim_{|\theta|\searrow 0} E\left[|Y|^2(|Y||\theta|\wedge 1)\right] = 0
$$

which, together with 2), proves (2.46). \(^□^)/

We now conclude (2.42) by (2.43), (2.44) and the following lemma with $\alpha = 2$, $h(\theta) = -\frac{1}{2}\theta\cdot V\theta$.

Lemma 2.7.3 *Let $h, \varphi : \mathbb{R}^d \to \mathbb{C}$, $\alpha > 0$ be such that*

a) *$h(r\theta) = r^\alpha h(\theta)$ for all $\theta \in \mathbb{R}^d$ and $r \in (0,1]$,*

b) *$\varphi(\theta) = 1 + h(\theta) + o(|\theta|^\alpha)$ as $\theta \to 0$.*

Then,

$$
\varphi\left(\frac{\theta}{n^{1/\alpha}}\right)^n \overset{n\to\infty}{\longrightarrow} \exp(h(\theta)) \ \text{for all } \theta \in \mathbb{R}^d.
$$

If additionally h is locally bounded, then the convergence above is locally uniform in θ.

Remark: Suppose that h and φ are as in Lemma 2.7.3 and that $\varphi = \widehat{\mu}$ for some $\mu \in \mathcal{P}(\mathbb{R}^d)$. Then $\varphi(n^{-1/\alpha}\theta)^n$ is the ch.f. of

$$
Y_n = \frac{X_1 + ... + X_n}{n^{1/\alpha}},
$$

where $X_1, X_2, ...$ are i.i.d. $\approx \mu$. Thus, Lemma 2.7.3, together with Lévy's convergence theorem (Theorem 9.2.1) shows that e^h is a ch.f. of a random variable Y and $Y_n \overset{\text{w}}{\to} Y$.

Proof of Lemma 2.7.3: Recall that

1) $$|z^n - w^n| \leq n(|z|\vee|w|)^{n-1}|z-w|, \quad z, w \in \mathbb{C}, \ n \geq 1.$$

We will apply this inequality to $z \overset{\text{def}}{=} \varphi\left(\frac{\theta}{n^{1/\alpha}}\right)$ and $w \overset{\text{def}}{=} \exp\left(\frac{h(\theta)}{n}\right)$, so that

2) $$z^n - w^n = \varphi\left(\frac{\theta}{n^{1/\alpha}}\right)^n - \exp\left(h(\theta)\right).$$

We have that

3) $$z = \varphi\left(\frac{\theta}{n^{1/\alpha}}\right) \overset{b)}{=} 1 + h\left(\frac{\theta}{n^{1/\alpha}}\right) + o\left(\frac{|\theta|^\alpha}{n}\right) \overset{a)}{=} 1 + \frac{h(\theta)}{n} + o\left(\frac{|\theta|^\alpha}{n}\right).$$

Since $e^z = 1 + z + O(|z|^2)$ as $|z| \to 0$,

4) $$w = \exp\left(\frac{h(\theta)}{n}\right) = 1 + \frac{h(\theta)}{n} + O\left(\frac{|h(\theta)|^2}{n^2}\right).$$

Therefore,

5) $$z - w \overset{3),4)}{=} o\left(\frac{|\theta|^\alpha}{n}\right) + O\left(\frac{|h(\theta)|^2}{n^2}\right).$$

Moreover, for large enough n's,

6) $$|z| = \left|\varphi\left(\frac{\theta}{n^{1/\alpha}}\right)\right| \overset{3)}{\leq} 1 + \frac{|h(\theta)| + |\theta|^\alpha}{n}.$$

Hence,

7) $$\begin{cases} |z|^{n-1} = \left|\varphi\left(\frac{\theta}{n^{1/\alpha}}\right)\right|^{n-1} \overset{6)}{\leq} \left(1 + \frac{|h(\theta)| + |\theta|^\alpha}{n}\right)^{n-1} \leq \exp(|h(\theta)| + |\theta|^\alpha), \\ |w|^{n-1} = \left|\exp\left(\frac{h(\theta)}{n}\right)\right|^{n-1} = \exp\left(\frac{n-1}{n}\operatorname{Re} h(\theta)\right) \leq \exp(|h(\theta)|). \end{cases}$$

Therefore,

$$\begin{aligned} &\left|\varphi\left(\frac{\theta}{n^{1/\alpha}}\right)^n - \exp(h(\theta))\right| \\ &\overset{2)}{=} \quad |z^n - w^n| \overset{1)}{\leq} n(|z| \vee |w|)^{n-1}|z - w| \\ &\overset{5),7)}{\leq} \quad n\exp(|h(\theta)| + |\theta|^\alpha)\left(o\left(\frac{|\theta|^\alpha}{n}\right) + O\left(\frac{|h(\theta)|^2}{n^2}\right)\right) \overset{n\to\infty}{\longrightarrow} 0. \end{aligned}$$

Moreover, the final estimate shows that the convergence above is uniform in θ, if h is locally bounded. \(^□^)/

Exercise 2.7.1 (**CLT for continuous-time RW**) Let $S_n = X_1 + ... + X_n$ be as in Theorem 2.6.1 and $(N_t)_{t\geq 0}$ be Poisson process with parameter $c > 0$ (Example 1.7.6). We suppose that $(X_n)_{n\geq 1}$ and $(N_t)_{t\geq 0}$ are independent and define $\widetilde{S}_t = S_{N_t}$. Then, show the following:
(i) $E\exp(\mathbf{i}\theta \cdot \widetilde{S}_t) = \exp\left((E\exp(\mathbf{i}\theta \cdot X_1) - 1)\, ct\right)$.
(ii) $\dfrac{\widetilde{S}_t - mct}{\sqrt{t}} \overset{\text{w}}{\longrightarrow} N(0, c\widetilde{V})$ as $t \to \infty$, where the matrix $\widetilde{V}$ is given by $\widetilde{V}_{\alpha\beta} = E[X_{1,\alpha}X_{1,\beta}]$ $(\alpha, \beta = 1, ..., d)$.

Exercise 2.7.2 (**More than** L^2) Use the argument in the proof of Lemma 2.7.2 to prove the following:
(i) If $X_1 \in L^{2+q}(P)$ for some $q \in [0, 1]$, then,

$$\left|E\exp(\mathbf{i}Y \cdot \theta) - 1 + \tfrac{1}{2}\theta \cdot V\theta\right| \leq |\theta|^{2+q}P[|Y|^{2+q}] = O(|\theta|^{2+q}) \quad \text{as } |\theta| \searrow 0.$$

Hint: $\min\{|Y||\theta|,\ 1\} \leq |Y|^q|\theta|^q$.
(ii) If Y is symmetric and $X_1 \in L^{3+q}(P)$ for some $q \in [0, 1]$, then,

$$\left|E\exp(\mathbf{i}Y \cdot \theta) - 1 + \tfrac{1}{2}\theta \cdot V\theta\right| \leq |\theta|^{3+q}P(|Y|^{3+q}) = O(|\theta|^{3+q}) \quad \text{as } |\theta| \searrow 0.$$

Exercise 2.7.3 ($\star$)(**Less than** L^2) Let X real r.v. with the density $\frac{\alpha}{2}|x|^{-(\alpha+1)}\mathbf{1}\{|x| \geq 1\}$, where $0 < \alpha \leq 2$. Show that

$$\begin{aligned} \varphi(\theta) \overset{\text{def}}{=} E\exp(\mathbf{i}\theta X) &= \cos\theta - |\theta|^\alpha \int_{|\theta|}^\infty \frac{\sin y}{y^\alpha}dy \\ &= \begin{cases} 1 - \theta^2\ln(1/|\theta|) + O(\theta^2) & \text{if } \alpha = 2 \\ 1 - c(\alpha)|\theta|^\alpha + o(|\theta|^\alpha) & \text{if } 0 < \alpha < 2 \end{cases} \quad \text{as } \theta \to 0, \end{aligned}$$

where $c(\alpha) = \int_0^\infty \frac{\sin y}{y^\alpha}dy = \frac{\pi}{2\Gamma(\alpha)\sin\frac{\alpha\pi}{2}}$.

Exercise 2.7.4 ($\star$)(**Logarithmic correction to Lemma 2.7.3**) Replace the condition (c) in Lemma 2.7.3 by:

$$\varphi(\theta) = 1 - \, h(\theta)\ln(1/|\theta|) + O(|\theta|^\alpha), \quad \theta \to 0,$$

while keeping all the other assumptions. Prove then that

$$\varphi\left(\frac{\theta}{(n\ln n)^{1/\alpha}}\right)^n \stackrel{n\to\infty}{\longrightarrow} \exp(-h(\theta)) \text{ for all } \theta \in \mathbb{R}^d.$$

Exercise 2.7.5 ($\star$) (**α-stable law**) Let $S_n = X_1 + ... + X_n$, where $(X_n)_{n\geq 1}$ are real i.i.d. with the density $\frac{\alpha}{2}|x|^{-(\alpha+1)}\mathbf{1}\{|x| \geq 1\}$, where $0 < \alpha \leq 2$.
(i) For $\alpha = 2$, use Exercise 2.7.3 and Exercise 2.7.4 to prove that $\frac{S_n}{\sqrt{n\ln n}} \stackrel{\mathrm{w}}{\longrightarrow} N(0,1)$.
(ii) ($\star$) For $0 < \alpha < 2$ and $c > 0$, use Exercise 2.7.3 and Lemma 2.7.3 to show that

$$\varphi(n^{-1/\alpha}\theta)^n \stackrel{n\to\infty}{\longrightarrow} \exp(-c(\alpha)|\theta|^\alpha) \text{ uniformly in } |\theta| < R \text{ for any } R > 0,$$

or equivalently, for any $c > 0$,

$$\varphi(n^{-1/\alpha}r\theta)^n \stackrel{n\to\infty}{\longrightarrow} \exp(-c|\theta|^\alpha) \text{ uniformly in } |\theta| < R \text{ for any } R > 0,$$

where $r = (c/c(\alpha))^{1/\alpha}$. This shows that there exists $\mu_{c,\alpha} \in \mathcal{P}(\mathbb{R})$ such that

$$\widehat{\mu_{c,\alpha}}(\theta) = \exp(-c|\theta|^\alpha), \quad \theta \in \mathbb{R}, \quad \text{and that } \frac{rS_n}{n^{1/\alpha}} \stackrel{\mathrm{w}}{\longrightarrow} \mu_{c,\alpha},$$

(cf. the remark after Lemma 2.7.3). $\mu_{c,\alpha}$ is called the *symmetric α-stable law* (For $\alpha = 2$, it is $N(0, 2c)$, and for $\alpha = 1$, it is the (c)-Cauchy distribution).

2.8 ($\star$) Local Central Limit Theorem

Example 2.8.1 (**Local CLT for Poisson distribution**) Let $\pi_c(n) = \frac{e^{-c}c^n}{n!}$, $n \in \mathbb{N}$, $c > 0$. If c is large enough, then the histogram of the function $n \mapsto \pi_c(n)$ looks like the density of Gaussian distribution (In Example 1.2.2, we see a picture for $c = 14$). Here is a mathematical explication.

$$\pi_c(n) = \frac{1}{\sqrt{2\pi c}}\exp\left(-\frac{(n-c)^2}{2c}\right) + O\left(\frac{1}{c}\right), \quad \text{as } c \to \infty, \text{ uniformly in } n \in \mathbb{N}. \tag{2.47}$$

This shows that $n \mapsto \pi_c(n)$ is well approximated by the density of $N(c, c)$ as $c \to \infty$. As we will see now, (2.10) and (2.6) can be used to prove (2.47).

Proof: We see from (2.6) that

$$\sum_{n\geq 0}\pi_c(n)\exp(\mathbf{i}\theta n) = \exp((e^{\mathbf{i}\theta} - 1)c),$$

which is the Fourier series of the sequence $\pi_c(n)$. Therefore, by inverting the Fourier series, we have that

1) $$\pi_c(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\exp\left(-\mathbf{i}\theta n + (e^{\mathbf{i}\theta} - 1)c\right)d\theta.$$

Let $q(\theta) = 1 + \mathbf{i}\theta - e^{\mathbf{i}\theta}$ and $\widetilde{n}_c = (n-c)/\sqrt{c}$. We then have that

2) $$\begin{cases} \pi_c(n) \stackrel{1)}{=} \dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} \exp\left(-\mathbf{i}\theta(n-c) - cq(\theta)\right) d\theta \\ \qquad = \dfrac{1}{2\pi\sqrt{c}}\displaystyle\int_{-\pi\sqrt{c}}^{\pi\sqrt{c}} \exp\left(-\mathbf{i}\theta\widetilde{n}_c - cq\left(\tfrac{\theta}{\sqrt{c}}\right)\right) d\theta. \end{cases}$$

On the other hand,

$$\exp\left(-\frac{c\theta^2}{2}\right) \stackrel{(2.10)}{=} \frac{1}{\sqrt{2\pi c}}\int_{-\infty}^{\infty} \exp\left(\mathbf{i}\theta x - \tfrac{x^2}{2c}\right) dx = \frac{1}{\sqrt{2\pi c}}\int_{-\infty}^{\infty} \exp\left(-\mathbf{i}\theta x - \tfrac{x^2}{2c}\right) dx.$$

Replacing c by $1/c$, and interchanging the letters θ and x, we have that

3) $$\exp\left(-\tfrac{x^2}{2c}\right) = \sqrt{\frac{c}{2\pi}}\int_{-\infty}^{\infty} \exp\left(-\mathbf{i}\theta x - \tfrac{c\theta^2}{2}\right) d\theta.$$

Let $h_c(x) = \frac{1}{\sqrt{2\pi c}}\exp\left(-\frac{x^2}{2c}\right)$ $(x \in \mathbb{R})$. Then,

4) $$\begin{cases} h_c(n-c) \stackrel{3)}{=} \dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty} \exp\left(-\mathbf{i}\theta(n-c) - \tfrac{c\theta^2}{2}\right) d\theta \\ \qquad = \dfrac{1}{2\pi\sqrt{c}}\displaystyle\int_{-\infty}^{\infty} \exp\left(-\mathbf{i}\theta\widetilde{n}_c - \tfrac{\theta^2}{2}\right) d\theta. \end{cases}$$

By dividing the integral $\int_{-\infty}^{\infty}$ in 4) into $\int_{-\pi\sqrt{c}}^{\pi\sqrt{c}}$ and $\int_{|\theta|\geq\pi\sqrt{c}}$, we see from 3) and 4) that

5) $$\sup_{n\in\mathbb{N}} |\pi_c(n) - h_c(n-c)| \leq \frac{1}{2\pi\sqrt{c}}(I_1 + I_2),$$

where

$$I_1 = \int_{-\pi\sqrt{c}}^{\pi\sqrt{c}} \left|\exp\left(-cq\left(\tfrac{\theta}{\sqrt{c}}\right)\right) - \exp\left(-\tfrac{\theta^2}{2}\right)\right| d\theta, \quad I_2 = \int_{|\theta|\geq\pi\sqrt{c}} \exp\left(-\tfrac{\theta^2}{2}\right) d\theta.$$

The integral I_2 can easily be bounded.

6) $$I_2 \stackrel{(1.37)}{\leq} \frac{2}{\pi\sqrt{c}}\exp\left(-\tfrac{c\pi^2}{2}\right).$$

To bound the integral I_1, we recall that

7) $$|\exp z - \exp w| \leq |z-w| \exp(\mathrm{Re}\, z \vee \mathrm{Re}\, w), \quad z, w \in \mathbb{C}.$$

We will apply this inequality to $z = -cq\left(\frac{\theta}{\sqrt{c}}\right)$ and $w = -\theta^2/2$. By expanding the exponential,

8) $$\exp(\mathbf{i}\theta) = 1 + \mathbf{i}\theta - \tfrac{\theta^2}{2} + r(\theta), \ \text{ with } \ |r(\theta)| \leq \tfrac{|\theta|^3}{6} \leq |\theta|^3.$$

Hence,

9) $$\left|cq\left(\tfrac{\theta}{\sqrt{c}}\right) - \tfrac{\theta^2}{2}\right| = \left|c\left(1 + \tfrac{\mathbf{i}\theta}{\sqrt{c}} - \exp\left(\tfrac{\mathbf{i}\theta}{\sqrt{c}}\right) - \tfrac{\theta^2}{2c}\right)\right| \stackrel{8)}{=} c\left|r\left(\tfrac{\theta}{\sqrt{c}}\right)\right| \leq \tfrac{|\theta|^3}{\sqrt{c}}.$$

Moreover, we note that

$$1-\cos\theta \geq \tfrac{2\theta^2}{\pi^2}, \quad |\theta| \leq \pi,$$

and hence

10) $\quad \mathrm{Re}\, cq\left(\tfrac{\theta}{\sqrt{c}}\right) = c\left(1-\cos\tfrac{\theta}{\sqrt{c}}\right) \geq \tfrac{2\theta^2}{\pi^2}, \quad |\theta| \leq \pi\sqrt{c},$

Therefore, putting together 7), 9), 10) and noting $\frac{2}{\pi^2} < \frac{1}{2}$, we have for $|\theta| \leq \pi\sqrt{c}$ that

$$\left|\exp\left(-cq\left(\tfrac{\theta}{\sqrt{c}}\right)\right) - \exp\left(-\tfrac{\theta^2}{2}\right)\right| \leq \tfrac{|\theta|^3}{\sqrt{c}}\exp\left(-\tfrac{2\theta^2}{\pi^2}\right).$$

Hence,

11) $\quad I_1 \leq \dfrac{1}{\sqrt{c}}\displaystyle\int_{-\infty}^{\infty} |\theta|^3 \exp\left(-\tfrac{2\theta^2}{\pi^2}\right) d\theta = O(1/\sqrt{c}).$

Finally, we conclude from 5), 6), 11) that

$$\sup_{n\in\mathbb{N}} |\pi_c(n) - h_c(n-c)| = O(1/c).$$

\(^□^)/

Example 2.8.2 (**Local CLT for trinomial distribution**) Let $p, q, r \in [0,1)$ be such that $p+q+r=1$. We assume either $r \in (0,1)$ or $p=q=1/2$. Let also X_n, $n \in \mathbb{N}\backslash\{0\}$ be i.i.d. such that $X_n = 1, -1, 0$ with probabilities, p, q, r, respectively. Then, $m \stackrel{\text{def}}{=} EX_1 = p-q$ and $v \stackrel{\text{def}}{=} \mathrm{var}\, X_1 = 4pq + r(1-r) > 0$. For $n \in \mathbb{N}\backslash\{0\}$, we define $S_n = X_1 + ... + X_n$ and

$$\mu_n(k) = P(S_n = k), \ \ |k| \leq n.$$

If $r > 0$, then

$$\max_{|k|\leq n}\left|\mu_n(k) - \frac{1}{\sqrt{2\pi vn}}\exp\left(-\frac{(k-mn)^2}{2vn}\right)\right| = O\left(n^{-\alpha}\right), \ \text{ as } n \to \infty, \tag{2.48}$$

where $\alpha = 3/2$ if $E[(X_1-m)^3] = (2m^2+3r-2)m = 0$, and $\alpha = 1$ if otherwise. On the other hand, if $p = q = 1/2$, then,

$$\max_{\substack{|k|\leq n \\ n+k \text{ is even}}}\left|\mu_n(k) - \sqrt{\frac{2}{\pi n}}\exp\left(-\frac{k^2}{2n}\right)\right| = O\left(n^{-3/2}\right), \ \text{ as } n \to \infty. \tag{2.49}$$

Proof: We start by preparing some equalities/inequalities which will be needed later. Let

$$\begin{aligned}\varphi(\theta) &= E\exp(\mathbf{i}\theta(X_1-m)) = pe^{\mathbf{i}(1-m)\theta} + qe^{-\mathbf{i}(1+m)\theta} + re^{-m\mathbf{i}\theta},\\ \psi(\theta) &= \varphi(\theta) - \exp\left(-\tfrac{v\theta^2}{2}\right).\end{aligned}$$

Let also $\widetilde{k}_n = (k-mn)/\sqrt{n}$. We first show that

1) $\quad \mu_n(k) = \dfrac{1}{2\pi\sqrt{n}}\displaystyle\int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} \exp\left(-\mathbf{i}\theta\widetilde{k}_n\right)\varphi\left(\tfrac{\theta}{\sqrt{n}}\right)^n d\theta,$

and that

1') $$\mu_n(k) = \frac{1+(-1)^{k+n}}{2\pi\sqrt{n}} \int_{-\pi\sqrt{n}/2}^{\pi\sqrt{n}/2} \exp\left(-\mathbf{i}\theta\widetilde{k}_n\right)\varphi\left(\tfrac{\theta}{\sqrt{n}}\right)^n d\theta \quad \text{if } p=q=1/2.$$

We have

$$\sum_{k=-n}^{n} \mu_n(k)\exp(\mathbf{i}\theta k) = E\exp(\mathbf{i}\theta S_n) = (pe^{\mathbf{i}\theta} + qe^{-\mathbf{i}\theta} + r)^n = \exp(\mathbf{i}nm)\varphi(\theta)^n.$$

Therefore, by inverting the Fourier series, we have

$$\begin{aligned}\mu_n(k) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \exp\left(-\mathbf{i}\theta(k-mn)\right)\varphi(\theta)^n d\theta \\ &= \frac{1}{2\pi\sqrt{n}}\int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} \exp\left(-\mathbf{i}\theta\widetilde{k}_n\right)\varphi\left(\tfrac{\theta}{\sqrt{n}}\right)^n d\theta.\end{aligned}$$

If $p=q=1/2$, then $\varphi(\theta)=\cos\theta$ and hence $\varphi(\pi-\theta) = -\varphi(\theta)$. Thus,

$$\begin{aligned}\mu_n(k) &\overset{1)}{=} \frac{1}{2\pi}\int_{-\pi/2}^{\pi/2} \exp\left(-\mathbf{i}\theta k\right)\varphi(\theta)^n d\theta + \frac{1}{2\pi}\int_{-\pi/2}^{\pi/2} \exp\left(-\mathbf{i}(\pi-\theta)k\right)\varphi(\pi-\theta)^n d\theta \\ &= \frac{1+(-1)^{k+n}}{2\pi}\int_{-\pi/2}^{\pi/2} \exp\left(-\mathbf{i}\theta k\right)\varphi(\theta)^n d\theta \\ &= \frac{1+(-1)^{k+n}}{2\pi\sqrt{n}}\int_{-\pi\sqrt{n}/2}^{\pi\sqrt{n}/2} \exp\left(-\mathbf{i}\theta\widetilde{k}_n\right)\varphi\left(\tfrac{\theta}{\sqrt{n}}\right)^n d\theta.\end{aligned}$$

Next, we show that there exists a constant $c_1 \in (0,\infty)$ which depends only on p,q,r such that

2) $$|\psi(\theta)| \le c_1|\theta|^\beta, \quad |\theta| \le \pi,$$

where $\beta = 4$ if $E[(X_1-m)^3] = 0$, and $\beta = 3$ if otherwise. By expanding the exponential,

$$\exp(\mathbf{i}\theta) = 1 + \mathbf{i}\theta - \tfrac{\theta^2}{2} - \mathbf{i}\tfrac{\theta^3}{3!} + O(\theta^4).$$

Hence

$$\begin{aligned}\psi(\theta) &= \left(1 - \tfrac{v}{2}\theta^2 - \tfrac{\mathbf{i}}{3!}E[(X_1-m)^3]\theta^3 + O(\theta^4)\right) - \left(1 - \tfrac{v}{2}\theta^2 + O(\theta^4)\right) \\ &= -\tfrac{\mathbf{i}}{3!}E[(X_1-m)^3]\theta^3 + O(\theta^4).\end{aligned}$$

This implies 2).
We next show that there exists a constant $c_2 \in (0,\infty)$ which depends only on p,q,r such that

3) $$|\varphi(\theta)| \le \exp(-c_2\theta^2), \quad \begin{cases} \text{for } |\theta| \le \pi \text{ if } 0 < r < 1, \\ \text{for } |\theta| \le \pi/2 \text{ if } pq > 0. \end{cases}$$

We note that

$$|\sin\theta| \ge \tfrac{2|\theta|}{\pi}, \text{if } |\theta| \le \pi/2,$$

On the other hand, we see from a direct computation that

$$|\varphi(\theta)| = \sqrt{1 - 4pq\sin^2\theta - 4r(1-r)\sin^2\tfrac{\theta}{2}}.$$

If $r \in (0,1)$, then, for $|\theta| \le \pi$,

$$|\varphi(\theta)| \le \sqrt{1-4r(1-r)\sin^2\tfrac{\theta}{2}} \le \sqrt{1-4r(1-r)\theta^2/\pi^2} \le \exp(-2r(1-r)\theta^2/\pi^2).$$

If $pq > 0$, then, for $|\theta| \le \pi/2$,

$$|\varphi(\theta)| \le \sqrt{1-4pq\sin^2\theta} \le \sqrt{1-16pq\theta^2/\pi^2} \le \exp(-8pq\theta^2/\pi^2).$$

These imply 3).
Let $h_n(x) = \frac{1}{\sqrt{2\pi vn}}\exp\left(-\frac{x^2}{2vn}\right)$ $(x \in \mathbb{R})$. We will next show that

4) $$h_n(k-mn) = \frac{1}{2\pi\sqrt{n}}\int_{-\infty}^{\infty}\exp\left(-\mathbf{i}\theta\widetilde{k}_n - \tfrac{v\theta^2}{2}\right)d\theta.$$

We know from the proof of Example 2.8.1 that

$$\exp\left(-\tfrac{x^2}{2c}\right) = \sqrt{\frac{c}{2\pi}}\int_{-\infty}^{\infty}\exp\left(-\mathbf{i}\theta x - \tfrac{c\theta^2}{2}\right)d\theta, \quad x \in \mathbb{R},\ c > 0.$$

Setting $c = vn$, we have that

5) $$\exp\left(-\tfrac{x^2}{2vn}\right) = \sqrt{\frac{vn}{2\pi}}\int_{-\infty}^{\infty}\exp\left(-\mathbf{i}\theta x - \tfrac{vn\theta^2}{2}\right)d\theta.$$

Thus

$$\begin{aligned} h_n(k-mn) &\stackrel{5)}{=} \frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left(-\mathbf{i}\theta(k-mn) - \tfrac{vn\theta^2}{2}\right)d\theta \\ &= \frac{1}{2\pi\sqrt{n}}\int_{-\infty}^{\infty}\exp\left(-\mathbf{i}\theta\widetilde{k}_n - \tfrac{v\theta^2}{2}\right)d\theta. \end{aligned}$$

We combine 1)–4) above to prove (2.48) and (2.49). Let us first consider (2.48). We have that We see from 1) and 4) that

6) $$\max_{|k|\le n}|\mu_n(k) - h_n(k-mn)| \le \frac{1}{2\pi\sqrt{n}}(I_1 + I_2),$$

where

$$I_1 = \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}}\left|\varphi\left(\tfrac{\theta}{\sqrt{n}}\right)^n - \exp\left(-\tfrac{v\theta^2}{2}\right)\right|d\theta, \quad I_2 = \int_{|\theta|\ge\pi\sqrt{n}}\exp\left(-\tfrac{v\theta^2}{2}\right)d\theta.$$

The integral I_2 can easily be bounded.

7) $$I_2 = \frac{1}{\sqrt{v}}\int_{|\theta|\ge\pi\sqrt{vn}}\exp\left(-\tfrac{\theta^2}{2}\right)d\theta \stackrel{(1.37)}{\le} \frac{2}{\pi v\sqrt{n}}\exp\left(-\tfrac{\pi^2 vn}{2}\right).$$

We now estimate the integral I_1. Recall that

$$|z^n - w^n| \le n|z-w|(|z|\vee|w|)^{n-1}, \quad z, w \in \mathbb{C},\ n = 1, 2, ...$$

We will apply this inequality to $z = \varphi\left(\frac{\theta}{\sqrt{n}}\right)$ and $w = \exp\left(-\frac{v\theta^2}{2n}\right)$. Then, if $|\theta| \le \pi\sqrt{n}$,

$$\begin{aligned}
&\left|\varphi\left(\tfrac{\theta}{\sqrt{n}}\right)^n - \exp\left(-\tfrac{v\theta^2}{2}\right)\right| \\
&= \quad |z^n - w^n| \le n|z-w|(|z| \vee |w|)^{n-1} \\
&\le \quad n\left|\psi\left(\tfrac{\theta}{\sqrt{n}}\right)\right|\left(\left|\varphi\left(\tfrac{\theta}{\sqrt{n}}\right)\right| \vee \exp\left(-\tfrac{v\theta^2}{2n}\right)\right)^{n-1} \\
&\stackrel{7,\,8)}{\le} \quad c_1 n^{1-\frac{\beta}{2}}|\theta|^\beta \exp\left(-c_3\theta^2\right),
\end{aligned}$$

for some $c_3 > 0$. Therefore, we obtain that

8) $$I_1 \le c_1 n^{1-\frac{\beta}{2}} \int_{-\infty}^{\infty} |\theta|^\beta \exp\left(-c_3\theta^2\right) d\theta = O(n^{1-\frac{\beta}{2}}).$$

Finally, we conclude from 6), 7), 8) that

$$\max_{|k|\le n} |\mu_n(k) - h_n(k - mn)| = O(n^{-\frac{\beta-1}{2}}),$$

which proves (2.48). Using 1') instead of 1), (2.49) can be obtained similarly as above. \\(^□^)/

Exercise 2.8.1 We refer to Example 2.8.1 and suppose that $n, c \to \infty$ and that $n = c + O(\sqrt{c})$. Prove the *de Moivre-Laplace theorem* for Poisson distribution:

$$\pi_c(n) \sim \frac{1}{\sqrt{2\pi c}} \exp\left(-\frac{(n-c)^2}{2c}\right).$$

Also, by setting $c = n$, deduce *Stirling's formula*:

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n. \tag{2.50}$$

Exercise 2.8.2 We refer to Example 2.8.2 and suppose that $n, k \to \infty$ and that $k = mn + O(\sqrt{n})$. Prove the *de Moivre-Laplace theorem* for trinomial distribution:

$$\mu_n(k) \sim \begin{cases} \frac{1}{\sqrt{2\pi v n}} \exp\left(-\frac{(k-mn)^2}{2vn}\right), & \text{if } r \in (0,1) \\ \sqrt{\frac{2}{\pi n}} \exp\left(-\frac{k^2}{2n}\right) & \text{if } p = q = 1/2. \end{cases}$$

3 Random Walks

3.1 Definition

Definition 3.1.1 Suppose that $(X_n)_{n\geq 1}$ are $\mathbb{R}^d$-valued i.i.d. defined on a probability space $(\Omega, \mathcal{F}, P)$. A *random walk* is a sequence $(S_n)_{n\geq 0}$ of $\mathbb{R}^d$-valued r.v.'s defined by $S_0 = 0$, and

$$S_n = X_1 + \ldots + X_n \text{ for } n \geq 1.$$

Remarks
1) Note that the iid $(X_n)_{n\geq 1}$ referred to above certainly exists by Proposition 8.3.1 and so does the random walk $(S_n)_{n\geq 0}$.
2) Our definition of "random walk" is the same as in [Dur95]. This definiton however is rather wider than traditional ones (e.g., [Spi76]) which will be called, in our language, the $\mathbb{Z}^d$-valued random walk.

Theorem 1.10.2 implies;

Theorem 3.1.2 *Let $(S_n)_{n\geq 0}$ be a random walk such that $E[|X_1|] < \infty$. We define its mean vector by*

$$m = (m_\alpha)_{\alpha=1}^d = (E[X_{1,\alpha}])_{\alpha=1}^d\,, \tag{3.1}$$

where $X_{1,\alpha}$ is the α-th coordinate of $X_1 \in \mathbb{R}^d$. Then,

$$S_n/n \stackrel{n\to\infty}{\longrightarrow} m, \quad P\text{-a.s.} \tag{3.2}$$

Remark: If we write S_n in a silly expression:

$$S_n = nm + (S_n - nm),$$

then (3.2) says that $\{S_n\}_{n\geq 1}$ almost surely follows a deterministic constant velocity motion $\{nm\}_{n\geq 1}$ by the correction term $S_n - nm$ which is of order $o(n)$. In this sense, one can conclude that the random walk travels in the direction of the vector m.

Exercise 3.1.1 Suppose that the random walk satisfies $P(X_1 \in \{0, \pm e_1, \ldots, \pm e_d\}) = 1$. Prove the following.
i) $m_\alpha = p(e_\alpha) - p(-e_\alpha)$ and $v_{\alpha\beta} = \delta_{\alpha\beta}(p(e_\alpha) + p(-e_\alpha)) - m_\alpha m_\beta$, where $p(x) = P(X_1 = x)$.
ii) Two different coordinates $X_{n,\alpha}$, $X_{n,\beta}$ ($\alpha \neq \beta$) of X_n are not independent of each other, even though they are uncorrelated if $m = 0$.

Exercise 3.1.2 Consider a $\mathbb{Z}$-valued random walk such $P(X_1 = \pm 1) = p_\pm > 0$, $P(X_1 = 0) = p_0 = 1 - p_+ - p_-$. Show the following. **(i)** For $y_0, y_1, \ldots, y_n \in \mathbb{Z}$, let $N(0) = \sum_{j=1}^n 1\{y_j - y_{j-1} = 0\}$ $(x \in \mathbb{Z})$. Then,

$$\begin{aligned} P(S_1 = y_1, \ldots S_n = y_n) &= p_+^{\frac{n-N(0)+y_n}{2}} p_-^{\frac{n-N(0)-y_n}{2}} p_0^{N(0)} \\ &= (p_+/p_-)^{y_n} P(S_1 = -y_1, \ldots S_n = -y_n). \end{aligned}$$

(ii) $P(S_n = y) = \displaystyle\sum_{\substack{|y| \leq m \leq n \\ m \pm y \text{ are even}}} \binom{m}{\frac{m+y}{2}} p_+^{\frac{m+y}{2}} p_-^{\frac{m-y}{2}} p_0^{n-m} = (p_+/p_-)^y P(S_n = -y)$

Exercise 3.1.3 An $\mathbb{R}^d$-valued r.v. X is said to be *symmetric* if $-X \approx X$. A random walk is said to be *symmetric* if X_1 is symmetric. Check that a symmetric random walks with $E|X_1| < \infty$ has the mean vector $m = 0$.

Exercise 3.1.4 Let $(S_n)_{n\geq 0}$ be a symmetric random walk (cf. Exercise 3.1.3). For $m \geq 0$, define $(S_n^{(m)})_{n\geq 0}$ by $S_n^{(m)} = S_n$ for $n \leq m$ and $S_n^{(m)} = 2S_m - S_n$ for $n \geq m$. Prove then that $(S_n^{(m)})_{n\geq 0}$ has the same distribution as $(S_n)_{n\geq 0}$ for each m.

Exercise 3.1.5 Consider a random walk such that $E|X_1| < \infty$. Use Theorem 3.1.2 to prove that, if $m_\alpha > 0$ (resp. $m_\alpha < 0$), for some $\alpha = 1, \ldots, d$, then

$$P(S_{n,\alpha} \stackrel{n\to\infty}{\longrightarrow} +\infty) = 1, \quad (\text{resp. } P(S_{n,\alpha} \stackrel{n\to\infty}{\longrightarrow} -\infty) = 1.)$$

Exercise 3.1.6 (**LLN for continuous-time RW**) Let $S_n = X_1 + ... + X_n$ be as in Theorem 3.1.2 and $(N_t)_{t\geq 0}$ be Poisson process with parameter $c > 0$ (Example 1.7.6). We suppose that $(X_n)_{n\geq 1}$ and $(N_t)_{t\geq 0}$ are independent and define $\widetilde{S}_t = S_{N_t}$. Then, show that $\widetilde{S}_t/t \stackrel{t\to\infty}{\longrightarrow} cm$, a.s.

3.2 Transience and Recurrence

In this section, we will take up a question whether a random walk $(S_n)_{n\geq 0}$ comes back to its starting point with probability one.

Definition 3.2.1 Let $(S_n)_{n\geq 0}$ be a random walk in $\mathbb{R}^d$, and $X_n = S_n - S_{n-1}$ ($n \geq 1$).

• If $P(X_1 \in \mathbb{Z}^d) = 1$, or equivalently, $P(S_n \in \mathbb{Z}^d) = 1$ for all $n \geq 0$, we say that the random walk is $\mathbb{Z}^d$*-valued*.

• A $\mathbb{Z}^d$-valued random walk is said to be *simple* if

$$P(X_1 = \pm e_\alpha) = (2d)^{-1} \quad \text{for all } \alpha = 1, \ldots, d. \tag{3.3}$$

- Throughout this section, we will restrict ourselves to $\mathbb{Z}^d$-valued random walks.

This is to avoid being bothered by inessential complication. We will prove the following

Theorem 3.2.2 *Consider a $\mathbb{Z}^d$-valued random walk with:*

$$E[|X_1|^2] < \infty, \ E[X_{1,\alpha}] = 0 \ (\forall \alpha = 1, \ldots, d). \tag{3.4}$$

$$\det V > 0, \ \textit{where } V = (\mathrm{cov}(X_{1,\alpha}, X_{1,\beta}))_{\alpha,\beta=1}^d. \tag{3.5}$$

Then,

$$h(0) \stackrel{\text{def}}{=} P\left(S_n = 0 \textit{ for some } n \geq 1\right) \begin{cases} = 1 & \textit{if } d \leq 2, \\ < 1 & \textit{if } d \geq 3. \end{cases}$$

Example 3.2.3 Supppose that $P(X_1 \in \{0, \pm e_1, ..., \pm e_d\}) = 1$ and set $p(x) = P(X_1 = x)$ ($x \in \mathbb{Z}^d$). Then, (3.6) & (3.7) $\iff$ (3.5), where

$$p(e_\alpha) \vee p(-e_\alpha) > 0 \text{ for all } \alpha = 1, ..., d, \tag{3.6}$$

$$p(0) + \sum_{\alpha=1}^d p(e_\alpha) \wedge p(-e_\alpha) > 0. \tag{3.7}$$

(See also Example 10.1.3.)

Proof: (3.5) is equivalent to that

1) $\theta \cdot V\theta > 0$ for $\theta \in \mathbb{R}^d \backslash \{0\}$.

To simplify the notation, We write

$$\begin{aligned} v_\alpha &= p(e_\alpha) \vee p(-e_\alpha), \quad w_\alpha = p(e_\alpha) \wedge p(-e_\alpha), \\ q_\alpha &= p(e_\alpha) + p(-e_\alpha) = v_\alpha + w_\alpha. \end{aligned}$$

Then, $\mathrm{cov}(X_{1,\alpha}, X_{1,\beta}) = q_\alpha \delta_{\alpha,\beta} - m_\alpha m_\beta$, cf. (0.18). Thus,

2) $$\theta \cdot V\theta = \sum_{\alpha,\beta=1}^{d} (q_\alpha \delta_{\alpha,\beta} - m_\alpha m_\beta)\theta_\alpha \theta_\beta = \sum_{\alpha=1}^{d} q_\alpha \theta_\alpha^2 - \left(\sum_{\alpha=1}^{d} m_\alpha \theta_\alpha\right)^2.$$

If we suppose (3.6), then $q_\alpha \geq v_\alpha \stackrel{(3.6)}{>} 0$ for all $\alpha = 1, ..., d$, so that we can define:

$$\delta = \sum_{\alpha=1}^{d} \frac{m_\alpha^2}{q_\alpha} = \sum_{\alpha=1}^{d} \frac{(v_\alpha - w_\alpha)^2}{v_\alpha + w_\alpha}.$$

(3.6) & (3.7) $\Rightarrow$ (3.5): Since $\frac{v_\alpha - w_\alpha}{v_\alpha + w_\alpha} \leq 1$, it follows that

3) $$\delta \leq \sum_{\alpha=1}^{d} (v_\alpha - w_\alpha) \stackrel{(3.7)}{<} p(0) + \sum_{\alpha=1}^{d} v_\alpha \leq p(0) + \sum_{\alpha=1}^{d} (v_\alpha + w_\alpha) = 1,$$

4) $$\left(\sum_{\alpha=1}^{d} m_\alpha \theta_\alpha\right)^2 = \left(\sum_{\alpha=1}^{d} \frac{m_\alpha}{\sqrt{q_\alpha}} \sqrt{q_\alpha}\theta_\alpha\right)^2 \stackrel{\text{Schwarz}}{\leq} \delta \sum_{\alpha=1}^{d} q_\alpha \theta_\alpha^2.$$

Suppose that $\theta \neq 0$. Then, $\sum_{\alpha=1}^{d} q_\alpha \theta_\alpha^2 > 0$. We thus obtain 1):

$$\theta \cdot V\theta \stackrel{2),4)}{\geq} (1 - \delta) \sum_{\alpha=1}^{d} q_\alpha \theta_\alpha^2 \stackrel{3)}{>} 0.$$

(3.6) & (3.7) $\Leftarrow$ (3.5): Suppose that (3.6) fails, i.e., that $v_\alpha = 0$ for some $\alpha = 1, ..., d$. Then, $q_\alpha = m_\alpha = 0$, and hence the α-th row and the α-th column of the matrix V vanish. Thus (3.5) fails. Suppose on the other hand that (3.6) holds but (3.7) fails. In this case, we have $w_\alpha = 0$, $q_\alpha = v_\alpha = |m_\alpha| > 0$ for all $\alpha = 1, ..., d$, so that

5) $$\delta = \sum_{\alpha=1}^{d} \frac{m_\alpha^2}{q_\alpha} = \sum_{\alpha=1}^{d} v_\alpha \stackrel{(3.7)\text{ fails}}{=} p(0) + \sum_{\alpha=1}^{d} (v_\alpha + w_\alpha) = 1.$$

Now, choosing $\theta \in \mathbb{R}^d$ with $\theta_\alpha = m_\alpha / q_\alpha \neq 0$, $\alpha = 1, ..., d$,

$$\theta \cdot V\theta \stackrel{2)}{=} \sum_{\alpha=1}^{d} m_\alpha^2 / q_\alpha - \left(\sum_{\alpha=1}^{d} m_\alpha^2 / q_\alpha\right)^2 = \delta - \delta^2 \stackrel{5)}{=} 0.$$

Thus, 1) fails. \\(^□^)/

It is convenient to introduce the following notations. For $x \in \mathbb{Z}^d$, we set

$$V(x) = \sum_{n \geq 1} \mathbf{1}\{S_n = x\} = \text{"the number of visits to } x\text{"}. \tag{3.8}$$

$$h^{(m)}(x) = P(V(x) \geq m), \quad m = 1, \ldots, \infty \tag{3.9}$$
$$= \text{"probability that } x \text{ is visited at least } m \text{ times"}.$$

$$h(x) = h^{(1)}(x) \tag{3.10}$$
$$= \text{"probability that } x \text{ is visited at least once"}.$$

$$g(x) = \sum_{n \geq 0} P(S_n = x) \in [0, \infty], \quad 0 \leq s \leq 1, \tag{3.11}$$

The function $g(x)$ above is called the *Green function* of the random walk.

Proposition 3.2.4 (Transience/Recurrence) *Let $(S_n)_{n \geq 0}$ be a $\mathbb{Z}^d$-valued random walk. Then, the following conditions T1)–T5) are equivalent:*

T1) $h(0) < 1$.

T2) $g(0) < \infty$.

T3) $g(x) < \infty$ *for all* $x \in \mathbb{Z}^d$.

T4) $h^{(\infty)}(0) = 0$.

T5) $h^{(\infty)}(x) = 0$ *for all* $x \in \mathbb{Z}^d$.

$(S_n)_{n \geq 0}$ *is said to be* **transient** *if one of (therefore all of) conditions T1)–T5) are satisfied. On the other hand, the following conditions R1)–R5) are equivalent:*

R1) $h(0) = 1$.

R2) $g(0) = \infty$.

R3) $g(x) = \infty$ *if* $h(x) > 0$.

R4) $h^{(\infty)}(0) = 1$.

R5) $h^{(\infty)}(x) = 1$ *if* $h(x) > 0$.

$(S_n)_{n \geq 0}$ *is said to be* **recurrent** *if one of (therefore all of) conditions R1)–R5) are satisfied.*

Example 3.2.5 Suppose that you and one of your friends perform simple random walks independently from $0 \in \mathbb{Z}^d$. Then, you will meet each other infinitely many times if $d \leq 2$ and you will eventually be separated forever if $d \geq 3$.

Proof: Let $(S'_n)_{n \geq 0}$ and $(S''_n)_{n \geq 0}$ be independent random walks. Then, $S_n = S'_n - S''_n$, $n \geq 0$ is again a random walk and

1) $\quad P(S_n = 0) = P(S'_n - S''_n = 0) = P(S'_{2n} = 0)$

Let g and g' be the Green functions of $S.$ and $S'_.$ respectively. Then,

$$g(0) = \sum_{n \geq 0} P(S_n = 0) \stackrel{(1)}{=} \sum_{n \geq 0} P(S'_{2n} = 0) = g'_1(0),$$

where the reason for the last identity is that $P(S'_{2n+1} = 0) = 0$. Thus, we see the claim from Theorem 3.2.2 and Proposition 3.2.4. \(^□^)/

Exercise 3.2.1 Prove that

$$P(\lim_{n \to \infty} |S_n| = +\infty) = \begin{cases} 0 & \text{for a recurrent RW,} \\ 1 & \text{for a transient RW.} \end{cases}$$

Exercise 3.2.2 Prove that $P(H \subset \{S_n\}_{n \geq 1}) = 1$ for any recurrent RW, where $H = \{x \in \mathbb{Z}^d \ ; \ h(x) > 0\}$. It would be interesting to compare this with Exercise 3.4.1 below.

Exercise 3.2.3 Prove that for all $z \in \mathbb{Z}^d$,

$$g(z) = \delta_{0,z} + Eg(z - X_1), \tag{3.12}$$
$$h(z) = (1 - h(0))P\{X_1 = z\} + Eh(z - X_1), \tag{3.13}$$
$$h^{(\infty)}(z) = Eh^{(\infty)}(z - X_1). \tag{3.14}$$

Exercise 3.2.4 ($\star$)(*Green function for continuous-time RW*) Let $S_n = X_1 + ... + X_n$ be a $\mathbb{Z}^d$-valued random walk and $(N_t)_{t \geq 0}$ be Poisson process with parameter $c > 0$ (Example 1.7.6). We suppose that $(X_n)_{n \geq 1}$ and $(N_t)_{t \geq 0}$ are independent. Then, show that $\int_0^\infty P(S_{N_t} = x)dt = \frac{1}{c}g(x), \quad x \in \mathbb{Z}^d$, where g is the Green function for $(S_n)_{n \in \mathbb{N}}$.

3.3 Proof of Proposition 3.2.4 for T1)–T3), R1)–R3)

In this section, we will prove the equivalence of (T1)–(T3), and that of (R1)–(R3). These are simpler than the other part of the equivalence, and still enough to proceed to the proof of Theorem 3.2.2. (T4), (T5), (R4) and (R5) will be discussed in section 3.5.

We begin by proving the following

Lemma 3.3.1 *For $x \in \mathbb{Z}^d$,*

$$h^{(m)}(x) = h(x)h(0)^{m-1}, \tag{3.15}$$
$$g(x) = \begin{cases} \frac{1}{1-h(0)} & \text{if } x = 0, \\ \frac{h(x)}{1-h(0)} & \text{if } x \neq 0. \end{cases} \tag{3.16}$$
$$g(x) = h(x)g(0) \ \text{ if } x \neq 0. \tag{3.17}$$

Remark Intuition behind (3.15) can be explained as follows; A trajectory of a random walk which visits a point x m times can be decomposed into m segments; a segment starting from the origin until its first visit to x and $m-1$ "loops" (or "excursions") starting from x until their next return to x. One can vaguely imagine that these m segments should be independent for the following reason; each time the random walk visits x, it starts afresh from there *independently* from the past.

Proof: Define the m^{th}-hitting time to $x \in \mathbb{Z}^d$ by

$$T_x^{(m)} = \inf\left\{ n \geq 1 \ ; \ \sum_{k=1}^{n} \mathbf{1}\{S_k = x\} = m \right\}. \tag{3.18}$$

Then,

1) $$h^{(m)}(x) = P(T_x^{(m)} < \infty) = \sum_{\ell \geq 1} P\left(T_x^{(m-1)} = \ell, \ \exists n \geq 1, \ S_{n+\ell} - S_\ell = 0\right)$$

We observe that

$$\begin{aligned} E_\ell &\stackrel{\text{def}}{=} \{T_x^{(m-1)} = \ell\} \in \sigma[X_j \ ; \ j \leq \ell], \\ F_\ell &\stackrel{\text{def}}{=} \{\exists n \geq 1, \ S_{n+\ell} - S_\ell = 0\} \in \sigma[X_j \ ; \ j > \ell], \end{aligned}$$

and therefore that

2) E_ℓ and F_ℓ are independent.

We also see that

3) $$\sum_{\ell \geq 1} P(E_\ell) = P(T_x^{(m-1)} < \infty) \stackrel{(3.9)}{=} h^{(m-1)}(x).$$

Note on the other hand that

$$(S_{n+\ell} - S_\ell)_{n=1}^{\infty} \approx (S_n)_{n=1}^{\infty}.$$

This implies that

4) $$P(F_\ell) = P(F_0) \stackrel{(3.10)}{=} h(0).$$

Combinning 1)–4), we have that

$$\begin{aligned} h^{(m)}(x) &\stackrel{1)}{=} \sum_{\ell \geq 1} P(E_\ell \cap F_\ell) \stackrel{2)}{=} \sum_{\ell \geq 1} P(E_\ell) P(F_\ell) \\ &\stackrel{4)}{=} \sum_{\ell \geq 1} P(E_\ell) h(0) \stackrel{3)}{=} h^{(m-1)}(x) h(0). \end{aligned}$$

We then get (3.15) by induction. Equality (3.16) can be seen as follows;

$$\begin{aligned} g(x) &\stackrel{(3.26)}{=} \delta_{0,x} + \sum_{n \geq 1} P(S_n = x) \stackrel{\text{Fubini}}{=} \delta_{0,x} + EV(x) \\ &\stackrel{(1.12)}{=} \delta_{0,x} + \sum_{m \geq 1} \underbrace{P(V(x) \geq m)}_{= h^{(m)}(x)} \stackrel{(3.15)}{=} \delta_{0,x} + \frac{h(x)}{1 - h(0)}. \end{aligned}$$

(3.17) follows immediately from (3.16). \\(^□^)/

Proof of T1) $\Longleftrightarrow$ *T2)* $\Longleftrightarrow$ *T3)*:

T1) $\Leftrightarrow$ **T2)**: This follows from the identity $g(0) \stackrel{(3.16)}{=} 1/(1 - h(0))$.

T2) $\Rightarrow$ **T3)**: This follows from the identity $g(x) \stackrel{(3.17)}{=} h(x)g(0)$ for $x \neq 0$.
T3) $\Rightarrow$ **T2)**: Obvious. \(^□^)/

Proof of R1) $\Longleftrightarrow$ *R2)* $\Longleftrightarrow$ *R3)*:
R1) $\Leftrightarrow$ **R2)**: This follows from the identity $g(0) \stackrel{(3.16)}{=} 1/(1-h(0))$.
R2) $\Rightarrow$ **R3)**: This follows from the identity $g(x) \stackrel{(3.17)}{=} h(x)g(0)$ for $x \neq 0$.
R3) $\Rightarrow$ **R2)**: It is clear that $\exists x \in \mathbb{Z}^d$, $h(x) > 0$. Then, it follows from R3) that $g(x) = \infty$. If $x = 0$, we are done. If $x \neq 0$, $g(0) \stackrel{(3.17)}{=} g(x)/h(x) = \infty$. \(^□^)/

Exercise 3.3.1 Conclude from (3.15) that $V(0)$ for a transient RW is a r.v. with geometric distribution with the parameter $1 - h(0)$ (cf. Exercise 1.7.8).

Exercise 3.3.2 **(i)** Show that $h(x+y) \geq h(x)h(y)$ for all $x, y \in \mathbb{Z}^d$. This implies that the set $H = \{x \in \mathbb{Z}^d \; ; \; h(x) > 0\}$ has the property that $x, y \in H \;\; \Rightarrow \;\; x + y \in H$. Hint: Apply the argument in the proof of (3.15) above. **(ii)** Use (i) and (3.16) to show that $g(x+y)g(0) \geq g(x)g(y)$ for all $x, y \in \mathbb{Z}^d$.

Exercise 3.3.3 Prove the following for $\mathbb{Z}$-valued random walk. **(i)** If $P(X_1 \geq 2) = 0$, then, $P(\sup_{n \geq 0} S_n \geq x) = h(x) = h(1)^x$ for all $x \geq 1$. **Hint** Apply the argument in the proof of (3.15) to verify that $h(x+1) = h(x)h(1)$ for all $x \geq 1$. **(ii)** If $P(X_1 \leq -2) = 0$, then, $P(\inf_{n \geq 0} S_n \leq -x) = h(-x) = h(-1)^x$ for all $x \geq 1$. **(iii)**[10] If $P(|X_1| \geq 2) = 0$ and $p_\pm \stackrel{\text{def}}{=} P(X_1 = \pm 1) > 0$, then $h(x) = \left(\frac{p_+}{p_-} \wedge 1\right)^x$ and $h(-x) = \left(\frac{p_-}{p_+} \wedge 1\right)^x$ for all $x \geq 1$.

3.4 Proof of Theorem 3.2.2

Let $S_n = X_1 + ... + X_n$ be a random walk in $\mathbb{Z}^d$ such $X_1 \approx \mu \in \mathcal{P}(\mathbb{Z}^d)$. As before, we write:

$$\widehat{\mu}(\theta) = E \exp(\mathbf{i}\theta \cdot X_1) = \sum_{x \in \mathbb{Z}^d} \exp(\mathbf{i}\theta \cdot x)\mu(x), \quad \theta \in \mathbb{R}^d. \tag{3.19}$$

The following proposition relates the transience/recurrence of the random walk to the behaviour of $\widehat{\mu}(\theta)$ as $\theta \to 0$:

Proposition 3.4.1 *Let* $\alpha, \delta > 0$.

a) *Suppose that there exists a constant* $c_1 \in (0, \infty)$ *such that*

$$c_1 |\theta|^\alpha \leq |1 - \widehat{\mu}(\theta)| \;\; \text{for } |\theta| \leq \delta. \tag{3.20}$$

Then, $h(0) < 1$ *if* $d > \alpha$.

b) *Suppose that there exist constants* $c_2, c_3 \in (0, \infty)$ *such that*

$$c_2 |\theta|^\alpha \leq 1 - \mathrm{Re}\, \widehat{\mu}(\theta) \text{ and } |1 - \widehat{\mu}(\theta)| \leq c_3 |\theta|^\alpha \text{ for } |\theta| \leq \delta. \tag{3.21}$$

Then, $h(0) = 1$ *if* $d \leq \alpha$.

[10]See also Exercise 3.7.1 and (4.69).

Proof of Theorem 3.2.2 assuming Proposition 3.4.1: It follows from Lemma 2.7.2 that:

1) $\quad 1-\widehat{\mu}(\theta)=\frac{1}{2}\theta\cdot V\theta+o(|\theta|^2)\quad\text{as } |\theta|\to 0.$

Since $\det V\neq 0$, 1) implies (3.20), (3.21) with $\alpha=2$ and small enough $\delta>0$. Thus, the conclusion follows from Proposition 3.4.1. \(^□^)/

For $\delta>0$, we write

$$\delta B=\{x\in\mathbb{R}^d\ ;\ |x|\le\delta\}.$$

Lemma 3.4.2 *For any $\delta>0$, there exists a constant $C_\delta\in(0,\infty)$ and $w_\delta\in C(\mathbb{R}^d)$ such that*

$$0 \le w_\delta\le C_\delta\mathbf{1}_{\delta B}, \tag{3.22}$$

$$f(x) \le \int_{\delta B}\exp(-\mathbf{i}\theta\cdot x)w_\delta(\theta)\widehat{f}(\theta)d\theta, \tag{3.23}$$

for all nonnegative $f\in\ell^1(\mathbb{Z}^d)$, where $\widehat{f}(\theta)=\sum_{x\in\mathbb{Z}^d}\exp(\mathbf{i}\theta\cdot x)f(x)$.

Proof: We first take an even, continuous function $v:\mathbb{R}^d\to[0,\infty)$ and $C\in(0,\infty)$ such that :

$$0\le v\le C\mathbf{1}_{\frac{1}{2}B},\quad \int_{\mathbb{R}^d}v=1.$$

We then define:

$$v_\delta(x)=\delta^{-d}v(x/\delta),\quad w_\delta(x)=\int_{\mathbb{R}^d}v_\delta(x-y)v_\delta(y)dy.$$

Then, w_δ is even and continuous. Moreover, we have

$$0\le v_\delta\le C\delta^{-d}\mathbf{1}_{\frac{\delta}{2}B},\quad \widehat{v_\delta}(\theta)\in\mathbb{R},\quad \int_{\mathbb{R}^d}v_\delta=1.$$

Thus,

1) $\quad 0\le w_\delta(x)\le\|v_\delta\|_\infty\int_{\mathbb{R}^d}v_\delta\le C\delta^{-d},$

2) $\quad \operatorname{supp} w_\delta\subset\{x+y\ ;x,y\in\operatorname{supp} v_\delta\}\subset\delta B,$

3) $\quad \widehat{w_\delta}(\theta)=\widehat{v_\delta}(\theta)^2\ge 0,$

4) $\quad \widehat{w_\delta}(0)=\int_{\mathbb{R}^d}w_\delta=\left(\int_{\mathbb{R}^d}v_\delta\right)^2=1.$

We see (3.22) from 1) and 2), whereras (3.23) is obtained as follows.

$$\begin{aligned}
f(x) &\overset{4)}{=} f(x)\widehat{w_\delta}(0)\overset{3)}{\le}\sum_{y\in\mathbb{Z}^d}f(y)\widehat{w_\delta}(y-x)\\
&= \sum_{y\in\mathbb{Z}^d}f(y)\int_{\mathbb{R}^d}\exp(\mathbf{i}(y-x)\cdot\theta)w_\delta(\theta)d\theta\\
&\overset{\text{Fubini}}{=} \int_{\mathbb{R}^d}\exp(-\mathbf{i}x\cdot\theta)w_\delta(\theta)d\theta\sum_{y\in\mathbb{Z}^d}f(y)\exp(\mathbf{i}y\cdot\theta)\\
&= \int_{\mathbb{R}^d}\exp(-\mathbf{i}x\cdot\theta)w_\delta(\theta)\widehat{f}(\theta)d\theta=\int_{\delta B}\exp(-\mathbf{i}x\cdot\theta)w_\delta(\theta)\widehat{f}(\theta)d\theta.
\end{aligned}$$

Lemma 3.4.3

$$g(0) \geq \frac{1}{(2\pi)^d} \int_{(\pi I)\setminus\Gamma(\mu)} \frac{1 - \mathrm{Re}\,\widehat{\mu}(\theta)}{|1-\widehat{\mu}(\theta)|^2} d\theta, \tag{3.24}$$

where $\pi I = [-r, r]^d$ *and* $\Gamma(\mu) = \{\theta \in \mathbb{R}^d \; ; \; \widehat{\mu}(\theta) = 1\}$. *On the other hand, for* $\delta > 0$,

$$g(0) \leq C_\delta \int_{\delta B} \frac{d\theta}{|1-\widehat{\mu}(\theta)|}. \tag{3.25}$$

where C_δ *is from Lemma 3.4.2.*

Proof: For $s \in (0, 1]$, we introduce

$$g_s(x) = \sum_{n\geq 0} s^n P(S_n = x), \quad x \in \mathbb{Z}^d. \tag{3.26}$$

Then, for $s \in (0, 1)$, $g_s(x)$ converges absolutely and $g_s(x) \nearrow g(x)$ as $s \nearrow 1$. We first prove that

$$g_s(x) = \frac{1}{(2\pi)^d} \int_{\pi I} \frac{\exp(-\mathbf{i}\theta \cdot x)}{1 - s\widehat{\mu}(\theta)} d\theta \;\text{ for } x \in \mathbb{Z}^d \text{ and } 0 \leq s < 1. \tag{3.27}$$

Note that

$$\sum_{x\in\mathbb{Z}^d} \exp(\mathbf{i}\theta \cdot x) P(S_n = x) = E \exp(\mathbf{i}\theta \cdot S_n) \stackrel{\text{Corollary 2.1.5}}{=} \widehat{\mu}(\theta)^n. \tag{3.28}$$

Thus, by inverting the Fourier series, we get[11]:

$$P(S_n = x) = \frac{1}{(2\pi)^d} \int_{\pi I} \exp(-\mathbf{i}\theta \cdot x) \widehat{\mu}(\theta)^n d\theta, \quad \text{for } x \in \mathbb{Z}^d \text{ and } n \in \mathbb{N}. \tag{3.29}$$

For $x \in \mathbb{Z}^d$ and $0 \leq s < 1$,

$$\begin{aligned} g_s(x) &= \sum_{n\geq 0} s^n P(S_n = x) \stackrel{(3.29)}{=} \frac{1}{(2\pi)^d} \sum_{n\geq 0} s^n \int_{\pi I} \exp(-\mathbf{i}\theta \cdot x) \widehat{\mu}(\theta)^n d\theta \\ &\stackrel{\text{Fubini}}{=} \frac{1}{(2\pi)^d} \int_{\pi I} \exp(-\mathbf{i}\theta \cdot x) \sum_{n\geq 0} s^n \widehat{\mu}(\theta)^n d\theta = \frac{1}{(2\pi)^d} \int_{\pi I} \frac{\exp(-\mathbf{i}\theta \cdot x)}{1 - s\widehat{\mu}(\theta)} d\theta. \end{aligned}$$

(3.24): Since the left-hand side of (3.27) is a real number, we may replace the integrand in the right-hand side by its real part. We therefore see that

1) $$g_s(0) = \frac{1}{(2\pi)^d} \int_{\pi I} \mathrm{Re}\, \frac{1}{1 - s\widehat{\mu}(\theta)} d\theta = \frac{1}{(2\pi)^d} \int_{\pi I} \frac{1 - s\,\mathrm{Re}\,\widehat{\mu}(\theta)}{|1 - s\widehat{\mu}(\theta)|^2} d\theta.$$

[11] The equality (3.29) will also be used in the proof of Proposition 3.6.1 below.

We use 1) to obtain (3.24) as follows.

$$\begin{aligned}
g(0) &\overset{\text{MCT}}{=} \lim_{s\nearrow 1} g_s(0) \overset{1)}{=} \frac{1}{(2\pi)^d} \lim_{s\nearrow 1} \int_{\pi I} \frac{1 - s\,\mathrm{Re}\,\widehat{\mu}(\theta)}{|1 - s\widehat{\mu}(\theta)|^2} d\theta \\
&\geq \frac{1}{(2\pi)^d} \lim_{s\nearrow 1} \int_{(\pi I)\setminus\Gamma(\mu)} \frac{1 - s\,\mathrm{Re}\,\widehat{\mu}(\theta)}{|1 - s\widehat{\mu}(\theta)|^2} d\theta \\
&\overset{\text{Fatou}}{\geq} \frac{1}{(2\pi)^d} \int_{(\pi I)\setminus\Gamma(\mu)} \lim_{s\nearrow 1} \frac{1 - s\,\mathrm{Re}\,\widehat{\mu}(\theta)}{|1 - s\widehat{\mu}(\theta)|^2} d\theta \\
&= \frac{1}{(2\pi)^d} \int_{(\pi I)\setminus\Gamma(\mu)} \frac{1 - \mathrm{Re}\,\widehat{\mu}(\theta)}{|1 - \widehat{\mu}(\theta)|^2} d\theta.
\end{aligned}$$

(3.25): Let w_δ be from Lemma 3.4.2. We apply (3.23) to $f(x) = P(S_n = x)$. Since $\widehat{f} = \widehat{\mu}^n$ by (3.28), we have for $x \in \mathbb{Z}^d$ and $n \in \mathbb{N}$ that[12]:

$$P(S_n = x) \leq \int_{\delta B} \exp(-\mathbf{i}x \cdot \theta) w_\delta(\theta) \widehat{\mu}(\theta)^n d\theta. \tag{3.30}$$

Let $0 \leq s < 1$. Note that for $z \in \mathbb{C}$ with $\mathrm{Re}\, z \leq 1$

$$|1 - sz| \geq s|1 - z|. \tag{3.31}$$

In fact, with $x = \mathrm{Re}\, z$ and $y = \mathrm{Im}\, z$, we have

$$1 - sx \geq s(1 - x) \geq 0.$$

Hence

$$|1 - sz|^2 = (1 - sx)^2 + (sy)^2 \geq (s(1 - x))^2 + (sy)^2 = (s|1 - z|)^2.$$

Thus,

$$\begin{aligned}
g_s(0) &= \sum_{n\geq 0} s^n P(S_n = 0) \overset{(3.30)}{\leq} \sum_{n\geq 0} s^n \int_{\delta B} w_\delta(\theta) \widehat{\mu}(\theta)^n d\theta \\
&\overset{\text{Fubini}}{=} \int_{\delta B} w_\delta(\theta) \sum_{n\geq 0} s^n \widehat{\mu}(\theta)^n d\theta = \int_{\delta B} \frac{w_\delta(\theta) d\theta}{1 - s\widehat{\mu}(\theta)} \\
&\leq \int_{\delta B} \frac{w_\delta(\theta) d\theta}{|1 - s\widehat{\mu}(\theta)|} \overset{(3.31)}{\leq} \frac{1}{s} \int_{\delta B} \frac{w_\delta(\theta) d\theta}{|1 - \widehat{\mu}(\theta)|} \\
&\overset{(3.22)}{\leq} \frac{C_\delta}{s} \int_{\delta B} \frac{d\theta}{|1 - \widehat{\mu}(\theta)|}.
\end{aligned}$$

Hence,

$$g(0) \overset{\text{MCT}}{=} \lim_{s\nearrow 1} g_s(0) \leq C_\delta \int_{\delta B} \frac{d\theta}{|1 - \widehat{\mu}(\theta)|}.$$

\(^▫^)/

Remark Concerning (3.25), the following inequality is easier to prove.

$$g(x) \leq \frac{1}{(2\pi)^d} \int_{\pi I} \frac{d\theta}{|1 - \widehat{\mu}(\theta)|}, \quad x \in \mathbb{Z}^d. \tag{3.32}$$

[12]The bound (3.30) will also be used in the proof of Proposition 3.6.1 below.

In fact,

$$g(x) \overset{\text{MCT}}{=} \lim_{s\nearrow 1} g_s(x) \overset{(3.27)}{\leq} \lim_{s\nearrow 1} \frac{1}{(2\pi)^d}\int_{\pi I} \frac{d\theta}{|1-s\widehat{\mu}(\theta)|} \overset{(3.31)}{\leq} \frac{1}{(2\pi)^d}\int_{\pi I} \frac{d\theta}{|1-\widehat{\mu}(\theta)|}.$$

An advantage of (3.25) over (3.32) is that the integral on the right-hand side is only over a small neighborhood of $\theta = 0$, cf. the proof of Proposition 3.4.1.

Proof of Proposition 3.4.1: We begin with a simple observation. Let $A_d = 2\pi^{\frac{d}{2}}/\Gamma(\frac{d}{2})$ (the area of the unit sphere in $\mathbb{R}^d$). Using the polar coordinate transform, we see that

1) $$\int_{\delta B} \frac{d\theta}{|\theta|^\alpha} = A_d \int_0^\delta r^{d-\alpha-1} dr \begin{cases} <\infty & \text{if } d>\alpha, \\ =\infty & \text{if } d \leq \alpha. \end{cases}$$

a) Let $d > \alpha$. Then,

$$g(0) \overset{(3.25)}{\leq} C_\delta \int_{\delta B} \frac{d\theta}{|1-\widehat{\mu}(\theta)|} \overset{(3.20)}{\leq} \frac{C_\delta}{c_1}\int_{\delta B} \frac{d\theta}{|\theta|^\alpha} \overset{1)}{<} \infty.$$

Thus, $h(0) < 1$ by Proposition 3.2.4.
b) Let $d \leq \alpha$. Let also c_2, c_3 and δ be from (3.21). We may suppose that $\delta \leq \pi$. By the first estimate of (3.21), we see that $(\delta B)\backslash\{0\} \subset (\pi I)\backslash\Gamma(\mu)$. Therefore,

$$\begin{aligned} g(0) &\overset{(3.24)}{\geq} \frac{1}{(2\pi)^d}\int_{(\delta B)\backslash\{0\}} \frac{1-\mathrm{Re}\,\widehat{\mu}(\theta)}{|1-\widehat{\mu}(\theta)|^2} d\theta \\ &\overset{(3.21)}{\geq} \frac{c_2}{(2\pi)^d c_3^2}\int_{(\delta B)\backslash\{0\}} \frac{d\theta}{|\theta|^\alpha} \overset{1)}{=} \infty. \end{aligned}$$

Thus, $h(0) = 1$ by Proposition 3.2.4. \(^□^)/

Example 3.4.4 Let $\alpha \in (0,2)$. We will present an example of $\mu \in \mathcal{P}(\mathbb{Z}^d)$ for which (3.20) and (3.21) hold true. Let $\mu_1 \in \mathcal{P}(\mathbb{Z})$ such that

$$\mu_1(0) = 0 \ \text{ and } \ \mu_1(x) = \frac{|x|^{-1-\alpha}}{2c_1} \ \text{for } x \neq 0,$$

where $c_1 = 2\sum_{n\geq 1} n^{-1-\alpha}$. We define $\mu \in \mathcal{P}(\mathbb{Z}^d)$ by

$$\mu(x) = \begin{cases} \frac{1}{d}\mu_1(x_\beta), & \text{if } x = (\delta_{\beta,\gamma}x_\beta)_{\gamma=1}^d \text{ for some } \beta = 1, ..., d, \\ 0, & \text{if otherwise.} \end{cases}$$

Then, it is easy to see that

$$\widehat{\mu}(\theta) = \frac{1}{d}\sum_{\beta=1}^d \widehat{\mu_1}(\theta_\beta).$$

Thus, (3.20) and (3.21) follow from those for μ_1. In fact, we will prove that:

1) $$\frac{1-\widehat{\mu_1}(\theta)}{|\theta|^\alpha} \overset{\theta\to 0}{\longrightarrow} \frac{c_2}{c_1} \in (0,\infty), \text{ where } c_2 = 2\int_0^\infty \frac{1-\cos x}{x^{1+\alpha}} dx = \frac{\pi}{\Gamma(\alpha+1)\sin\frac{\alpha\pi}{2}}.$$

We may assume that $\theta \neq 0$. By symmetry, we may also assume that $\theta > 0$. It is convenient to introduce

$$f(x) = \frac{1-\cos x}{x^{1+\alpha}}, \quad x > 0$$

and its approximation $f_\theta(x) \stackrel{\theta \to 0}{\longrightarrow} f(x)$ $(x > 0)$ defined by:

$$f_\theta(x) = f(n\theta) \ \text{ if } x \in ((n-1)\theta, n\theta],\ n = 1, 2, ...$$

We compute:

$$\widehat{\mu_1}(\theta) = \frac{1}{2c_1} \sum_{\substack{x \in \mathbb{Z} \\ x \neq 0}} |x|^{-1-\alpha} \exp(\mathbf{i}x\theta) = \frac{1}{c_1} \sum_{n \geq 1} n^{-1-\alpha} \cos(n\theta)$$

Thus,

2) $$\frac{1-\widehat{\mu_1}(\theta)}{\theta^\alpha} = \frac{\theta}{c_1} \sum_{n \geq 1} (n\theta)^{-\alpha-1}(1-\cos(n\theta)) = \frac{1}{c_1} \int_0^\infty f_\theta(x)dx.$$

We will check that

3) there exists a $g \in L^1((0,\infty))$ such that $f_\theta(x) \leq g(x)$ for $x > 0$ and $\theta \in (0,1]$.

Then, 1) follows from 2) and the dominated convergence theorem. Note that

4) $$0 \leq 1 - \cos\theta \leq 2 \wedge \frac{\theta^2}{2} \quad \text{for } \theta \in \mathbb{R}.$$

Suppose that $x \in (0,1)$ and that $x \in ((n-1)\theta, n\theta]$. Then,

$$f_\theta(x) = (n\theta)^{-\alpha-1}(1-\cos(n\theta)) \stackrel{4)}{\leq} (n\theta)^{1-\alpha} \leq \left\{ \begin{array}{ll} x^{1-\alpha} & \text{if } \alpha \in [1,2) \\ (1+x)^{1-\alpha} & \text{if } \alpha \in (0,1) \end{array} \right\} \in L^1((0,1)).$$

Suppose on the other hand that $x \in [1,\infty)$ and that $x \in ((n-1)\theta, n\theta]$. Then,

$$f_\theta(x) = (n\theta)^{-\alpha-1}(1-\cos(n\theta)) \stackrel{4)}{\leq} 2(n\theta)^{-1-\alpha} \leq 2x^{-1-\alpha} \in L^1([1,\infty)).$$

These prove 3). \(^□^)/

($\star$) Completion Referring to (3.9) and (3.18), we now define

$$h_s^{(m)}(x) = \begin{cases} E[s^{T_x^{(m)}}], & \text{if } 0 \leq s < 1, \\ h^{(m)}(x), & \text{if } s = 1. \end{cases} \tag{3.33}$$

and

$$h_s(x) = h_s^{(1)}(x), \quad 0 \leq s \leq 1. \tag{3.34}$$

Note that, by the monotone convergence theorem,

$$h_1^{(m)}(x) = \lim_{s \nearrow 1} h_s^{(m)}(x), \tag{3.35}$$

$$h^{(\infty)}(x) = \lim_{m \nearrow \infty} h^{(m)}(x). \tag{3.36}$$

We now prove (3.15) in the following generalized form.

Lemma 3.4.5 *Consider a random walk $(S_n)_{n\geq 0}$ on $\mathbb{Z}^d$. For all $s \in [0,1]$, $x \in \mathbb{Z}^d$ and $m \geq 1$,*

$$h_s^{(m)}(x) = h_s(x)h_s(0)^{m-1}, \tag{3.37}$$

$$g_s(x) = \delta_{0,x} + \frac{h_s(x)}{1-h_s(0)}. \tag{3.38}$$

Proof: It is enough to prove (3.37) and (3.38) for $s < 1$. The results for $s = 1$ can be obtained by passing to the limit $s \nearrow 1$. We begin by proving (3.37) for $s < 1$. To do so, we may assume that $P\{T_x < \infty\} > 0$. In fact, (3.37) is just "0=0" if otherwise. For $1 \leq k < \infty$, define

$$T_0^{(m-1,k)} = \inf\left\{ n \geq 1 \; ; \; \sum_{j=1}^n 1\{S_{k+j} - S_k = 0\} = m-1 \right\}.$$

Then,
1) $T_0^{(m-1,k)} \approx T_0^{(m-1)}$.
2) $T_0^{(m-1,k)}$ is independent of $\{X_j\}_{j=1}^k$ and thus, independent of $\{T_x = k\}$.
3) $\{T_x = k\} \subset \{T_x^{(m)} = k + T_0^{(m-1,k)}\}$.
Note also that

$$s^{T_x^{(m)}} = s^{T_x^{(m)}} \mathbf{1}\{T_x < \infty\}.$$

We therefore have that

$$\textbf{4)} \begin{cases} E\left[s^{T_x^{(m)}}\right] = E\left[s^{T_x^{(m)}} : T_x < \infty\right] \stackrel{3)}{=} \sum_{k=1}^\infty s^k E\left[s^{T_0^{(m-1,k)}} : T_x = k\right] \\ \stackrel{1),2)}{=} \sum_{k=1}^\infty s^k E\left[s^{T_0^{(m-1)}}\right] P(T_x = k) = E\left[s^{T_x}\right] E\left[s^{T_0^{(m-1)}}\right]. \end{cases}$$

By applying 4) to $x = 0$ inductively, we see that

$$E\left[s^{T_0^{(m-1)}}\right] = E\left[s^{T_0}\right]^{m-1},$$

which, in conjunction with 4), proves (3.37). We next prove (3.38) for $s < 1$ as follows:

$$\begin{aligned} g_s(x) &= \delta_{0,x} + \sum_{n=1}^\infty s^n P\{S_n = x\}, \\ \sum_{n=1}^\infty s^n P\{S_n = x\} &= \sum_{n=1}^\infty s^n \sum_{m=1}^\infty P\{T_x^{(m)} = n\} = \sum_{m=1}^\infty E\left[\sum_{n=1}^\infty s^{T_x^{(m)}} \mathbf{1}\{T_x^{(m)} = n\}\right] \\ &= \sum_{m=1}^\infty E\left[s^{T_x^{(m)}}\right] \stackrel{(3.37)}{=} \sum_{m=1}^\infty h_s(x)h_s(0)^{m-1} = \frac{h_s(x)}{1-h_s(0)}. \end{aligned}$$

\(^□^)/

Exercise 3.4.1 Suppose that $\int_{\pi I} \frac{d\theta}{1-\mathrm{Re}\widehat{\mu}(\theta)} < \infty$, which is true for the simple random walk with $d \geq 3$. Prove then the following.
i) $\frac{1}{1-\widehat{\mu}} \in L^1(\pi I)$, $\quad g(x) = (2\pi)^{-d} \int_{\pi I} d\theta \frac{\exp(-\mathbf{i}\theta\cdot x)}{1-\widehat{\mu}(\theta)}$, $\quad x \in \mathbb{Z}^d$.
ii) $g(x) \to 0$ and $h(x) \to 0$ as $|x| \to \infty$. Hint: The Riemann-Lebesgue lemma.
iii) $P(H \not\subset \{S_n\}_{n\geq 1}) = 1$, where $H = \{x \in \mathbb{Z}^d \; ; \; h(x) > 0\}$. This is in contrast with Exercise 3.2.2. Hint: $P(H \subset \{S_n\}_{n\geq 1}) \leq h(x)$ for any $x \in H$.

Exercise 3.4.2 Prove that

$$E[T_x^{(m)} : T_x^{(m)} < \infty] = \lim_{s \nearrow 1} \tfrac{\partial}{\partial s} h_s^{(m)}(x). \tag{3.39}$$

Exercise 3.4.3 Consider a $\mathbb{Z}$-valued random walk such that

$$P(X_1 = \pm 1) = p_\pm > 0 \text{ and } P(X_1 = 0) = p_0 = 1 - p_+ - p_-.$$

i) Use residue theorem to compute the integral (3.27) and conclude that

$$g_s(x) = \begin{cases} \delta(s)^{-1/2} f_-(s)^x & \text{if } x \geq 0, \\ \delta(s)^{-1/2} f_+(s)^{|x|} & \text{if } x \leq 0, \end{cases} \tag{3.40}$$

where $\delta(s) = (1 - p_0 s)^2 - 4p_+ p_- s^2$ and $f_\pm(s) = \frac{1 - p_0 s - \delta(s)^{1/2}}{2p_\pm s}$. **ii)**[13] Use (3.38) and (3.40) to prove that

$$h_s(x) = \begin{cases} f_-(s)^x & \text{if } x > 0, \\ 1 - \delta(s)^{1/2} & \text{if } x = 0, \\ f_+(s)^{|x|} & \text{if } x < 0. \end{cases} \tag{3.41}$$

iii)[14] Use (3.35), (3.39) and (3.41) to prove that

$$h(x) = \begin{cases} 1 \wedge (p_+/p_-)^x & \text{if } x \neq 0, \\ 1 - |p_+ - p_-| & \text{if } x = 0. \end{cases}$$

$$E[T_x] = \begin{cases} |x|/|p_+ - p_-| & \text{if } x(p_+ - p_-) > 0, \\ \infty & \text{if otherwise.} \end{cases}$$

$$E[T_x | T_x < \infty] = \begin{cases} |x|/|p_+ - p_-| & \text{if } x(p_+ - p_-) < 0, \\ (1 - |p_+ - p_-|)(p_+ + p_- + 4p_+ p_-)/|p_+ - p_-| & \text{if } p_+ \neq p_- \text{ and } x = 0, \\ \infty & \text{if } p_+ = p_-. \end{cases}$$

Exercise 3.4.4 (**Green function in a subset**) Suppose that $(S_n)_{n \in \mathbb{N}}$ is a $\mathbb{Z}^d$-valued random walk and that $0, x \in A \subset \mathbb{Z}^d$. Define

$$\begin{aligned} T(A^{\mathsf{c}}) &= \inf\{n \geq 1 \, ; \, S_n \notin A\}, \quad T_x = \inf\{n \geq 1 \, ; \, S_n = x\}, \\ g_s^A(x) &= \sum_{n=0}^{\infty} s^n P(S_n = x, n < T(A^{\mathsf{c}})), \end{aligned}$$

$$h_s^A(x) = \begin{cases} E\left[s^{T_x} : T_x < T(A^{\mathsf{c}})\right], & \text{if } 0 \leq s < 1, \\ P(T_x < T(A^{\mathsf{c}})) & \text{if } s = 1. \end{cases}$$

$$H_s^A(x) = \begin{cases} E\left[s^{T(A^{\mathsf{c}})} : S_{T(A^{\mathsf{c}})} = x\right], & \text{if } 0 \leq s < 1, \\ P(T(A^{\mathsf{c}}) < \infty, \ S_{T(A^{\mathsf{c}})} = x) & \text{if } s = 1. \end{cases}$$

Then, prove that[15]

$$g_s^A(x) = \delta_{x,0} + \frac{h_s^A(x)}{1 - h_s^A(0)}, \quad 0 < s \leq 1, \tag{3.42}$$

$$g_s(x) = g_s^A(x) + \sum_{y \in \mathbb{Z}^d \setminus A} H_s^A(y) g_s(x - y), \quad 0 < s \leq 1. \tag{3.43}$$

[13] See also (4.68) below.

[14] See also Exercise 3.3.3, Exercise 3.7.1, and Proposition 4.5.3.

[15] Special case of these identities can be found in [Law91]; See Exercise 1.5.7 and Proposition 1.5.8. of that book.

Exercise 3.4.5 [16] Prove the following for the random walk considered in Exercise 3.4.3. For $a, b \in \mathbb{N}\backslash\{0\}$ and $s \in (0, 1]$,

$$E[s^{T_{-a}} : T_{-a} < T_b] = \frac{f_-(s)^{-b} - f_+(s)^b}{f_+(s)^{-a} f_-(s)^{-b} - f_+(s)^b f_-(s)^a}, \tag{3.44}$$

$$E[s^{T_b} : T_b < T_{-a}] = \frac{f_+(s)^{-a} - f_-(s)^a}{f_+(s)^{-a} f_-(s)^{-b} - f_+(s)^b f_-(s)^a}. \tag{3.45}$$

In particular, if $p_+ < p_-$, then as special cases of (3.44) and (3.45) with $s = 1$,

$$P(T_{-a} < T_b) = \frac{(p_-/p_+)^b - 1}{(p_-/p_+)^b - (p_-/p_+)^{-a}}, \quad P(T_b < T_{-a}) = \frac{1 - (p_-/p_+)^{-a}}{(p_-/p_+)^b - (p_-/p_+)^{-a}}. \tag{3.46}$$

Hint: Referring to Exercise 3.4.4, for $A = \mathbb{Z} \cap (-\infty, b)$, $h_s^A(-a) = E[s^{T_{-a}} : T_{-a} < T_b]$. Similarly, or $A = \mathbb{Z} \cap (-a, \infty)$, $h_s^A(b) = E[s^{T_b} : T_b < T_{-a}]$.

3.5 (⋆) Completion of the Proof of Proposition 3.2.4

We will finish the proof of Proposition 3.2.4 by taking care of T4),T5),R4) and R5). To do so, we prepare a couple of lemmas.

Lemma 3.5.1 *For $y, z \in \mathbb{Z}^d$,*

$$1 - h_\infty(y) \geq h(z)(1 - h_\infty(y - z)). \tag{3.47}$$

Proof: Define the first hitting time to $x \in \mathbb{Z}^d$ by

$$\eta(x) = \inf\{n \geq 1 \mid S_n = x\}.$$

Then,

$$\begin{aligned} 1 - h_\infty(y) &= P\left(\bigcup_{m \geq 1} \bigcap_{n \geq m} \{S_n \neq y\}\right) \\ &\geq P\left(\eta(z) < \infty, \bigcup_{m \geq 1} \bigcap_{n \geq m} \{S_{n+\eta(z)} \neq y\}\right) \\ &= \sum_{\ell \geq 1} P\left(\underbrace{\eta(z) = \ell}_{=:E_\ell}, \underbrace{\bigcup_{m \geq 1} \bigcap_{n \geq m} \{S_{n+\ell} - S_\ell \neq y - z\}}_{=:F_\ell}\right). \end{aligned}$$

We observe that

$$E_\ell \in \sigma[X_j \,;\, j \leq \ell], \quad F_\ell \in \sigma[X_j \,;\, j > \ell],$$

and therefore that

1) E_ℓ and F_ℓ are independent.

[16] See also Proposition 4.5.5.

We also see that

2) $$\sum_{\ell\geq 1} P(E_\ell) = P(\eta(z)<\infty) \stackrel{(3.9)}{=} h(z).$$

Note on the other hand that

$$(S_{n+\ell}-S_\ell)_{n=1}^\infty \approx (S_n)_{n=1}^\infty.$$

This implies that

3) $$P(F_\ell) = P(F_0) = 1-h_\infty(y-z).$$

Combinning 1)–3), we have that

$$\begin{aligned}\sum_{\ell\geq 1} P(E_\ell\cap F_\ell) &\stackrel{(1)}{=} \sum_{\ell\geq 1} P(E_\ell)P(F_\ell) \stackrel{(3)}{=} \sum_{\ell\geq 1} P(E_\ell)(1-h_\infty(y-z)) \\ &\stackrel{(2)}{=} h(z)(1-h_\infty(y-z)).\end{aligned}$$

Putting things together, we obtain (3.47). \(^□^)/

The equivalence of T1),T4),T5) and that of R1),R4),R5) are immediate from the following

Lemma 3.5.2 *For* $x\in\mathbb{Z}^d$,

$$h_\infty(x) = \begin{cases} 0 & \Longleftrightarrow \quad h(0)<1 \text{ or } h(x)=0, \\ 1 & \Longleftrightarrow \quad h(0)=1 \text{ and } h(x)>0. \end{cases} \tag{3.48}$$

$$h_\infty(0) = \begin{cases} 0 & \Longleftrightarrow \quad h(0)<1, \\ 1 & \Longleftrightarrow \quad h(0)=1. \end{cases} \tag{3.49}$$

Proof: By the monotone convergence theorem (MCT) and (3.15), we have that:

1) $$h_\infty(x) \stackrel{\text{MCT}}{=} \lim_{m\nearrow\infty} h_m(x) \stackrel{(3.15)}{=} \begin{cases} 0 & \text{if } h(0)<1, \\ h(x) & \text{if } h(0)=1. \end{cases}$$

By setting $x=0$ in 1), we see that

$$h_\infty(0) = \begin{cases} 0 & \text{if } h(0)<1, \\ 1 & \text{if } h(0)=1. \end{cases}$$

This implies (3.49). Observe that (3.48) follows from 1) and the following

2) $$h(0)=1,\ h(x)>0 \implies h_\infty(x)=1.$$

To see this, suppose that $h(0)=1, h(x)>0$. Then, $h_\infty(0)=1, h(x)>0$ by (3.49). Then, by taking $(y,z)=(0,x)$ in Lemma 3.5.1, we have

$$0 = 1-h_\infty(0) \geq h(x)(1-h_\infty(-x)), \text{ hence } h_\infty(-x)=1.$$

This in particular implies that $h(-x)>0$. Then, by taking $(y,z)=(0,-x)$ in Lemma 3.5.1, we have

$$0 = 1-h_\infty(0) \geq h(-x)(1-h_\infty(x)), \text{ hence } h_\infty(x)=1.$$

This proves 2). \(^□^)/

Exercise 3.5.1 Conclude from Lemma 3.5.1 that the set $\{x\in\mathbb{Z}^d\ ;\ h_\infty(x)=1\}$ is either empty or a subgroup of $\mathbb{Z}^d$.

3.6 ($\star$) Bounds on the Transition Probabilities

In section 3.4, we have used the characteristic function to estimate the Green function. In this section, we will estimate the transition probabilities by similar argument. We will prove:

Proposition 3.6.1 *Let* $\alpha > 0$.

a) *Suppose that there exists a constant* $c_1, \delta \in (0, \infty)$ *such that*

$$1 - |\widehat{\mu}(\theta)| \geq c_1 |\theta|^\alpha \ \text{for} \ |\theta| \leq \delta. \tag{3.50}$$

Then, there exists a constant $b_1 \in (0, \infty)$ *such that*

$$\sup_{x \in \mathbb{Z}^d} P(S_n = x) \leq \frac{b_2}{n^{d/\alpha}} \ \text{for all} \ n \geq 1. \tag{3.51}$$

b) *Suppose that* $X_1 \approx -X_1$ *and that there exists a constant* $c_2, \delta \in (0, \infty)$ *such that (3.21) holds. Then, there exists a constant* $b_2 \in (0, \infty)$ *such that*

$$P(S_{2n} = 0) \geq \frac{b_1}{n^{d/\alpha}} \ \text{for all} \ n \geq 1. \tag{3.52}$$

Remark The condition (3.50) is slightly stronger than (3.20) in general, but they are equivalent if $X_1 \approx -X_1$, since $\widehat{\mu}(\theta) > 0$ for θ close to the origin. The random walk considered in Theorem 3.2.2 satisfies the conditions for Proposition 3.6.1 with $\alpha = 2$. Example 3.4.4 provides an example for which the conditions for Proposition 3.6.1 hold for $\alpha \in (0, 2)$.

We prepare a technical estimate:

Lemma 3.6.2 *Suppose that* $\alpha, c, \delta > 0$ *and* $c\delta^\alpha \leq 1$. *Then, there exist* $b_1, b_2 \in (0, \infty)$ *such that:*

$$\frac{b_1}{t^{d/\alpha}} \leq \int_{x \in \mathbb{R}^d,\ |x| \leq \delta} (1 - c|x|^\alpha)^t dx \leq \frac{b_2}{t^{d/\alpha}} \ \text{for all} \ t \geq 1 \tag{3.53}$$

Proof: We write the integral in (3.53) by I_t. Then,

1) $I_t \stackrel{x = yt^{-1/\alpha}}{=} t^{-d/\alpha} J_t$ with $J_t = \int_{|y| \leq \delta t^{1/\alpha}} \left(1 - \frac{c|y|^\alpha}{t}\right)^t dy.$

Since the integrand of J_t is increasing in $t \geq 1$ and converges to $\exp(-c|y|^\alpha)$ as $t \to \infty$, we have

2) $0 < J_1 \leq J_t \leq \int_{\mathbb{R}^d} \exp(-c|y|^\alpha) dy < \infty.$ for all $t \geq 1$.

(3.53) follows from 1)–2). \(^□^)/

Proof of Proposition 3.6.1:

a) We have that

$$P(S_n = x) \stackrel{(3.30),\,(3.22)}{\leq} C_\delta \int_{|\theta| \leq \delta} |\widehat{\mu}(\theta)|^n d\theta \stackrel{(3.50)}{\leq} C_\delta \int_{|\theta| \leq \delta} (1 - c_2|\theta|^\alpha)^n d\theta \stackrel{(3.53)}{\leq} \frac{b_2}{n^{d/\alpha}}.$$

This proves (3.51).
b) Note that $\widehat{\mu}(\theta) \in \mathbb{R}$ and hence that $\widehat{\mu}(\theta)^{2n} \geq 0$. Thus,

$$\begin{aligned} P(S_{2n}=0) &\stackrel{(3.29)}{=} \frac{1}{(2\pi)^d}\int_{\pi I}\widehat{\mu}(\theta)^{2n}d\theta \geq \frac{1}{(2\pi)^d}\int_{|\theta|\leq\delta}\widehat{\mu}(\theta)^{2n}d\theta \\ &\stackrel{(3.21)}{\geq} \frac{1}{(2\pi)^d}\int_{|\theta|\leq\delta}(1-c_1|\theta|^\alpha)^{2n}d\theta \stackrel{(3.53)}{\geq} \frac{b_1}{n^{d/\alpha}}. \end{aligned}$$

This proves (3.52). \(^□^)/

3.7 Reflection Principle and its Applications

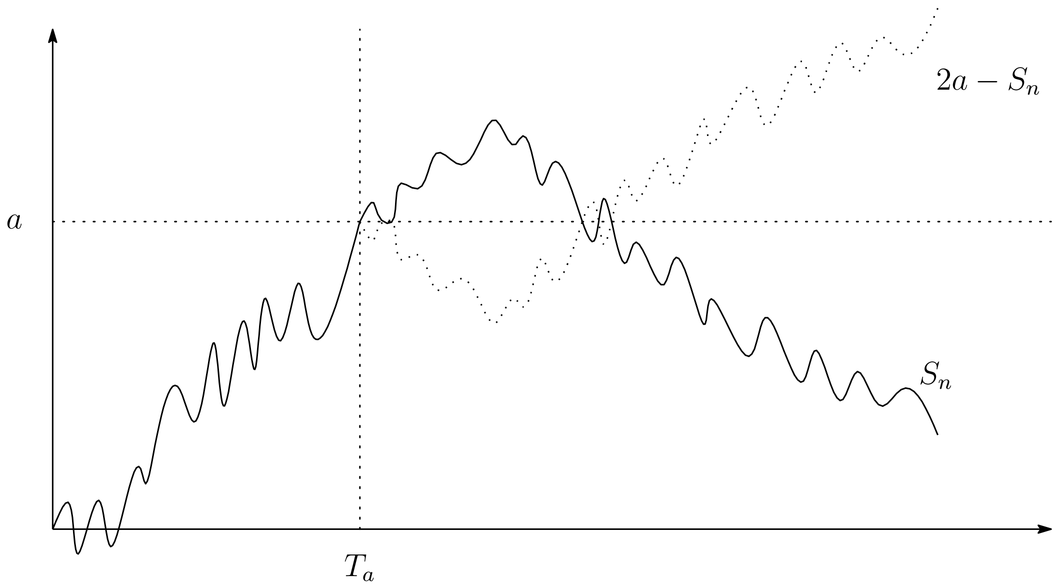
Reflection principle (Proposition 3.7.1) is an important tool to study nearest-neighbor random walks in $\mathbb{Z}$. In this subsection, we will focus on the reflection principle and its applications. Throughout this subsection, we consider a $\mathbb{Z}$-valued random walk $S_0 = 0$, $S_n = X_1 + ... + X_n$, $n \geq 1$ such that

$$P(X_1 = \pm 1) = p_\pm > 0, \text{ and } P(X_1 = 0) = p_0 = 1 - p_+ - p_-.$$

For $a \in \mathbb{Z}$, define

$$T_a = \inf\{n \geq 0 \; ; \; S_n = a\}.$$

Then, we have



Proposition 3.7.1 (**Reflection principle**). *For $x \in \mathbb{Z}^k$ and $y \in \mathbb{Z}^n$,*

$$\begin{aligned} &P\left(T_a = k, (S_j)_{j=1}^k = x, (S_{k+j})_{j=1}^n = y\right) \\ &= P\left(T_a = k, (S_j)_{j=1}^k = x, (2a - S_{k+j})_{j=1}^n = y\right)(p_+/p_-)^{y_n - a}. \end{aligned} \quad (3.54)$$

In particular, letting $a = 0$,

$$P\left((S_j)_{j=1}^n = y\right) = P\left((S_j)_{j=1}^n = -y\right)(p_+/p_-)^{y_n}. \quad (3.55)$$

Proof: We define the events A, $B_\pm$ by

$$A = \{T_a = k, (S_j)_{j=1}^k = x\}, \quad B_\pm = \{(a \pm (S_{k+j} - S_k))_{j=1}^n = y\}.$$

Note that $A \in \sigma(X_1, \ldots, X_k)$ and that $B_\pm \in \sigma(X_{k+1}, \ldots, X_{k+n})$. Therefore, A is independent of $B_\pm$. Moreover, $A \subset \{S_k = a\}$. Therefore,

1) $$\left\{ \begin{array}{rcl} \text{the LHS of (3.54)} & = & P(A \cap \{(S_{k+j})_{j=1}^n = y\}) \\ & = & P(A \cap \{(a + S_{k+j} - S_k)_{j=1}^n = y\}) \\ & = & P(A)P(B_+). \end{array} \right.$$

Similarly,

2) $$\left\{ \begin{array}{l} (p_+/p_-)^{-(y_n - a)} \times \text{the RHS of (3.54)} \\ \quad = \; P(A \cap \{(2a - S_{k+j})_{j=1}^n = y\}) \\ \quad = \; P(A \cap \{(a - (S_{k+j} - S_k))_{j=1}^n = y\}) \\ \quad = \; P(A)P(B_-). \end{array} \right.$$

Note that $P(X_j = \pm 1) = (p_+/p_-)^{\pm 1} P(X_j = \mp 1)$. Thus, with the convention $y_k = a$, we have

3) $$\left\{ \begin{array}{rcl} P(B_+) & = & \prod_{j=1}^n P(X_{k+j} = y_j - y_{j-1}) \\ & = & (p_+/p_-)^{y_n - a} \prod_{j=1}^n P(-X_{k+j} = y_j - y_{j-1}) \\ & = & (p_+/p_-)^{y_n - a} P(B_-). \end{array} \right.$$

Therefore,

$$\begin{array}{rcl} \text{the LHS of (3.54)} & \stackrel{1)}{=} & P(A)P(B_+) \\ & \stackrel{3)}{=} & (p_+/p_-)^{y_n - a} P(A)P(B_-) \stackrel{2)}{=} \text{the RHS of (3.54)} \end{array}.$$

\(^□^)/

Corollary 3.7.2 *For $a \in \mathbb{Z} \backslash \{0\}$, $n \geq 1$, and $x \in \mathbb{Z}$ with $a(a - x) > 0$,*

$$P(T_a > n, S_n = x) = P(S_n = x) - (p_+/p_-)^a P(S_n = x - 2a). \tag{3.56}$$

Moreover,

$$P(T_a > n) = \left\{ \begin{array}{ll} P(S_n < a) - (p_+/p_-)^a P(S_n < -a), & \text{if } a > 0, \\ P(-|a| < S_n) - (p_+/p_-)^a P(|a| < S_n), & \text{if } a < 0. \end{array} \right. \tag{3.57}$$

Proof: (3.56). If $a > 0$ and $x < a$, then, $2a - x > a$. If $a < 0$ and $a < x$, then $2a - x < a$. Thus we have the following inclusion in both cases.

1) $\{S_n = 2a - x\} \subset \{T_a \leq n\}$.

On the other hand, it follows from (3.55) that

$$P(S_n = x) = (p_+/p_-)^x P(S_n = -x). \tag{3.58}$$

Therefore,

2) $$\left\{ \begin{array}{rcl} P(S_n = x - 2a) & \stackrel{(3.58)}{=} & (p_+/p_-)^{x-2a} P(S_n = 2a - x) \\ & \stackrel{1)}{=} & (p_+/p_-)^{x-2a} P(T_a \leq n, S_n = 2a - x) \\ & \stackrel{(3.54)}{=} & (p_+/p_-)^{-a} P(T_a \leq n, S_n = x) \\ & = & (p_+/p_-)^{-a} \{P(S_n = x) - P(T_a > n, S_n = x)\}, \end{array} \right.$$

which proves (3.56).

(3.57): If $a > 0$, then, taking the summation of both-hands side of (3.56) over $x < a$, we have

$$P(T_a > n) = P(S_n < a) - (p_+/p_-)^a P(S_n < -a).$$

which proves (3.57) for $a > 0$. If $a < 0$, then, taking the summation of both-hands side of 3) over $x > a$, we have

$$P(T_a > n) = P(-|a| < S_n) - (p_+/p_-)^a P(|a| < S_n).$$

which proves (3.57) for $a < 0$. \\(^□^)/

Remark: Suppose that $p_+ = p_-$. Then, we see from Example 2.8.2 the following. If $p_0 > 0$,

$$P(S_n = x) = \frac{1}{\sqrt{2\pi v n}} + O\left(n^{-1}\right), \quad \text{as } n \to \infty,$$

where $v = 2p_+ + 4p_0(1 - p_0)$.

If $p_0 = 0$, then for $x \in \mathbb{Z}$ and $n \in \mathbb{N}$ such that $n + x$ is even,

$$P(S_n = x) = \sqrt{\frac{2}{\pi n}} + O\left(n^{-3/2}\right), \quad \text{as } n \to \infty.$$

These, together with (3.57), imply that

$$P(T_a > n) = |a|\sqrt{\frac{2}{\pi v n}} + O\left(n^{-3/2}\right), \quad \text{as } n \to \infty.$$

Exercise 3.7.1 Use (3.57) to prove the following.
(i)[17] For $a > 0$, $P(T_a < \infty) = (p_+/p_-)^a \wedge 1$, $P(T_{-a} < \infty) = (p_-/p_+)^a \wedge 1$.

(ii) $$\begin{cases} P(T_1 > n) &= \left(1 - \frac{p_0}{2p_+}\right) P(S_n = 0) + \frac{1}{2p_+} P(S_{n+1} = 0) + \left(1 - \frac{p_-}{p_+}\right) P(S_n > 1), \\ P(T_{-1} > n) &= \left(1 - \frac{p_0}{2p_-}\right) P(S_n = 0) + \frac{1}{2p_-} P(S_{n+1} = 0) + \left(1 - \frac{p_+}{p_-}\right) P(S_n < -1). \end{cases}$$

(iii) $$\begin{cases} P(T_0 > n) &= p_+ P(T_{-1} > n-1) + p_- P(T_1 > n-1) \\ &= \frac{p_-}{p_+} P(S_n = 0) + \left(1 - \frac{p_0(1-p_0)^2}{2p_+p_-}\right) P(S_{n-1} = 0) \\ &\quad + (p_+ - p_-)(P(S_{n-1} > 1) - P(S_{n-1} < -1)). \end{cases}$$

Exercise 3.7.2 Prove the following. **(i)** For $x \in \mathbb{Z}$ and an even function $F : \mathbb{Z}^n \to \mathbb{R}$,

$$E[F(S_1, \ldots, S_n) : S_n = x] = E[F(S_1, \ldots, S_n) : S_n = -x](p_+/p_-)^x.$$

(ii) Let $A_n = \bigcap_{j=1}^n \{|S_j| = r_j\}$ for $r_1, \ldots, r_n \in \mathbb{N}$ with $|r_j - r_{j-1}| \leq 1$ ($r_0 \stackrel{\text{def}}{=} 0$). Then, $P(S_n = r_n | A_n) = p_+^{r_n}/(p_+^{r_n} + p_-^{r_n})$. **(iii)** $P(A_n) = \prod_{j=1}^n p(r_{j-1}, r_j)$, where

$$p(r, s) = \begin{cases} (p_+^{r+1} + p_-^{r+1})/(p_+^r + p_-^r) & \text{if } s = r+1, \\ (p_- p_+^r + p_+ p_-^r)/(p_+^r + p_-^r) & \text{if } r \geq 1 \text{ and } s = r-1, \\ p_0 & \text{if } s = r. \end{cases}$$

[17] See Exercise 3.3.3 and (4.69) for alternative proofs.

4 Martingales

4.1 Conditional Expectation

Let $(\Omega, \mathcal{G})$ be a measurable space, μ be a measure on $(\Omega, \mathcal{G})$, and ν be either a measure or a signed measure on $(\Omega, \mathcal{G})$. ν is said to be *absolutely continuous* with respect to μ, and denoted by $\nu \ll \mu$ if

$$A \in \mathcal{G},\ \mu(A) = 0 \implies \nu(A) = 0. \tag{4.1}$$

We start by recalling

Theorem 4.1.1 (The Radon-Nikodym theorem) *Let $(\Omega, \mathcal{G})$ be a measurable space, μ be a σ-finite measure on $(\Omega, \mathcal{G})$. Suppose that a signed measure ν on $(\Omega, \mathcal{G})$ is absolutely continuous with respect to μ. Then, there exists a unique $\rho \in L^1(\mu)$ such that*

$$\nu(A) = \int_A \rho d\mu \ \text{ for all } A \in \mathcal{G}. \tag{4.2}$$

The function ρ is called the **Radon-Nikodym derivative** *and is denoted by $\frac{d\nu}{d\mu}$.*

Lemma 4.1.2 *Let $(\Omega, \mathcal{G})$ and μ be as in Theorem 4.1.1, Suppose that signed measures ν, ν_1, ν_2 on $(\Omega, \mathcal{G})$ are absolutely continuous with respect to μ and that $\rho = \frac{d\nu}{d\mu}$, $\rho_j = \frac{d\nu_j}{d\mu}$ $(j = 1, 2)$. Then,*

$$\nu = \alpha\nu_1 + \beta\nu_2 \implies \rho = \alpha\rho_1 + \beta\rho_2, \ \ \mu\text{-a.e. for } \alpha, \beta \in \mathbb{R}, \tag{4.3}$$

$$\nu_1 \le \nu_2 \implies \rho_1 \le \rho_2, \ \ \mu\text{-a.e.}, \tag{4.4}$$

$$|\rho| \le \frac{d|\nu|}{d\mu}, \ \ \mu\text{-a.e., where } |\nu| \text{ denotes the total variation of } \nu. \tag{4.5}$$

Proof: (4.3): Let $A \in \mathcal{G}$ be arbitrary. Then,

$$\nu(A) = \alpha\nu_1(A) + \beta\nu_2(A) \stackrel{(4.2)}{=} \int_A (\alpha\rho_1 + \beta\rho_2) d\mu.$$

Thus, $\rho = \alpha\rho_1 + \beta\rho_2$, μ-a.e. by the uniqueness of the Radon-Nikodym derivative.
(4.4): Let $A \in \mathcal{G}$ be arbitrary. Then,

$$\int_A \rho_1 d\mu \stackrel{(4.2)}{=} \nu_1(A) \le \nu_2(A) \stackrel{(4.2)}{=} \int_A \rho_2 d\mu.$$

Thus, $\rho_1 \le \rho_2$, μ-a.e.
(4.5): Since $\pm\nu \le |\nu|$, it follows from (4.4) that $\pm\rho \le \frac{d|\nu|}{d\mu}$, μ-a.e. \\(^□^)/

For the rest of this subsection, we suppose that $(\Omega, \mathcal{F}, P)$ is a probability space, and that $\mathcal{G}$ is a sub σ-algebra of $\mathcal{F}$.

Proposition 4.1.3 (**Conditional expectation**) *Let* $X \in L^1(P)$.

a) *There exists a unique* $Y \in L^1(\Omega, \mathcal{G}, P|_{\mathcal{G}})$ *such that*

$$E[X : A] = E[Y : A] \text{ for all } A \in \mathcal{G}. \tag{4.6}$$

The r.v. Y *is called the* **conditional expectation** *of* X *given* $\mathcal{G}$*, and is denoted by* $E[X|\mathcal{G}]$.

b) *For* $X, X_n \in L^1(P)$ $(n \in \mathbb{N})$,

$$E[\alpha X_1 + \beta X_2|\mathcal{G}] = \alpha E[X_1|\mathcal{G}] + \beta E[X_2|\mathcal{G}], \quad \text{a.s. for } \alpha, \beta \in \mathbb{R}, \tag{4.7}$$
$$X_1 \le X_2, \quad \text{a.s.} \implies E[X_1|\mathcal{G}] \le E[X_2|\mathcal{G}], \quad \text{a.s.,} \tag{4.8}$$
$$|E[X|\mathcal{G}]| \le E[|X||\mathcal{G}], \quad \text{a.s.,} \tag{4.9}$$
$$X \text{ is } \mathcal{G}\text{-measurable} \iff E[X|\mathcal{G}] = X, \quad \text{a.s.} \tag{4.10}$$
$$X \text{ is independent of } \mathcal{G} \implies E[X : A] = EXP(A), \forall A \in \mathcal{G} \tag{4.11}$$
$$\iff E[X|\mathcal{G}] = EX, \quad \text{a.s.} \tag{4.12}$$
$$X_n \stackrel{n\to\infty}{\longrightarrow} X \text{ in } L^1(P) \iff E[|X_n - X||\mathcal{G}] \stackrel{n\to\infty}{\longrightarrow} 0 \text{ in } L^1(P). \tag{4.13}$$

Proof: a) Let Q be a signed measure on $(\Omega, \mathcal{F})$ defined by $Q(A) \stackrel{\text{def}}{=} E[X : A]$ $(A \in \mathcal{F})$. Then, $Q|_{\mathcal{G}} \ll P|_{\mathcal{G}}$ and $|Q|(A) = E[|X| : A]$ $(A \in \mathcal{F})$. Thus, by Theorem 4.1.1, there exists a unique $Y \in L^1(\Omega, \mathcal{G}, P|_{\mathcal{G}})$ such that

$$Q(A) = \int_A Y dP, \text{ for all } A \in \mathcal{G}.$$

which however is nothing but (4.6). In particular,

$$E[X|\mathcal{G}] = \frac{dQ|_{\mathcal{G}}}{dP|_{\mathcal{G}}}. \tag{4.14}$$

b) (4.7), (4.8), (4.9) follow respectively from (4.3), (4.4), (4.5).
(4.10) $\Rightarrow$: Suppose that X is $\mathcal{G}$-measurable. Since the relation (4.6) is trivially true for $Y = X$, it follows from the uniqueness of the conditional expectation that $X = E[X|\mathcal{G}]$ a.s.
(4.10) $\Leftarrow$: This is obvious, since $E[X|\mathcal{G}]$ is $\mathcal{G}$-measurable by definition.
(4.11): Obvious.
(4.12) $\Rightarrow$: Let $Y = EX$ and $A \in \mathcal{G}$. Then,

$$E[X : A] = EXP(A) = E[Y : A].$$

This implies, via the uniqueness of the conditional expectation, that $Y = E[X|\mathcal{G}]$ a.s.
(4.11) $\Leftarrow$: Suppose that $E[X|\mathcal{G}] = EX$, a.s. Then, by taking the expectation of the both-side hands over the event $A \in \mathcal{G}$, we have that $E[X : A] = EXP(A)$.
(4.12): This follows from (4.11).
(4.13): Let $Y_n = E[|X_n - X||\mathcal{G}]$. Then,

$$E|Y_n| \stackrel{(4.6)}{=} E|X_n - X|.$$

Thus $E|X_n - X| \stackrel{n\to\infty}{\longrightarrow} 0 \iff E|Y_n| \stackrel{n\to\infty}{\longrightarrow} 0$. \(^□^)/

Remark Referring to Proposition 4.1.3, if $\mathcal{G} = \sigma(Y_1, Y_2, ...)$ for r.v's $Y_1, Y_2, ...$, we write $E[X|\mathcal{G}] = E[X|Y_1, Y_2, ...]$.

Example 4.1.4 For $X \in L^1(P)$, the *conditional expectation* of X, given an event $A \in \mathcal{F}$ with $P(A) > 0$ is defined by

$$E[X|A] = E[X : A]/P(A). \tag{4.15}$$

Let J be an at most countable set and $\{G_j\}_{j\in J} \subset \mathcal{F}$ be such that $P(G_j) > 0$ for all $j \in J$, $\Omega = \bigcup_{j\in J} G_j$ and $G_j \cap G_k = \emptyset$ if $j \neq k$. Finally, let $\mathcal{G} = \sigma[\{G_j\}_{j\in J}]$. Then, for $X \in L^1(P)$,

$$E[X|\mathcal{G}] = \sum_{j\in J} E[X|G_j]\mathbf{1}_{G_j}, \text{ a.s.}, \tag{4.16}$$

To verify this identity, we take an arbitrary $A \in \mathcal{G}$ and let

1) $Y = \sum_{j\in J} E[X|G_j]\mathbf{1}_{G_j}$.

Since there exists $K \subset J$ such that

2) $A = \bigcup_{j\in K} G_j$,

we have for any $j \in J$ that

3) $G_j \cap A = \begin{cases} G_j, & \text{if } j \in K, \\ \emptyset, & \text{if } j \notin K. \end{cases}$

By putting these together, we see that Y satisfies (4.6) as follows.

$$\begin{aligned} E[Y : A] &\stackrel{1)}{=} \sum_{j\in J} E[X|G_j]P(G_j \cap A) \stackrel{3)}{=} \sum_{j\in K} E[X|G_j]P(G_j) \\ &\stackrel{(4.15)}{=} \sum_{j\in K} E[X : G_j] \stackrel{2)}{=} E[X : A]. \end{aligned}$$

This implies (4.16).

Example 4.1.5 For $j = 1, 2$, let $(S_j, \mathcal{B}_j)$ be a measurable space, $f : S_1 \times S_2 \to \mathbb{R}$ be measurable, $X_j : \Omega \to S_j$ be a r.v. Suppose that X_1 and X_2 are independent and that $f(X_1, X_2) \in L^1(P)$. Then,

$$E[f(X_1, X_2)|X_2] = \int_{S_1} f(x_1, X_2)P(X_1 \in dx_1) \text{ a.s.}$$

Let

$$F_2(x_2) = \int_{S_1} f(x_1, x_2)P(X_1 \in dx_1), \quad x_2 \in X_2.$$

Then, we should prove that

1) $\forall A \in \sigma[X_2]$, $E[f(X_1, X_2) : A] = E[F_2(X_2) : A]$.

Note that $\forall A \in \sigma[X_2]$, $\exists B \in \mathcal{B}(\mathbb{R})$, $A_2 = \{X_2 \in B\}$. Thus,

$$\begin{aligned} E[F_2(X_2) : A] &= E[F_2(X_2) : X_2 \in B] = \int_B P(x_2 \in dx_2) \int_{S_1} f(x_1, x_2) P(X_1 \in dx_1) \\ &\stackrel{\text{Fubini}}{=} = E[f(X_1, X_2) : X_2 \in B] = E[f(X_1, X_2) : A]. \end{aligned}$$

Proposition 4.1.6 (**The projection property**) *Let $X \in L^1(P)$. Then, for σ-algebras $\mathcal{G}_1$, $\mathcal{G}_2$ such that $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{F}$,*

$$E[E[X|\mathcal{G}_1]|\mathcal{G}_2] = E[X|\mathcal{G}_1], \quad a.s. \tag{4.17}$$

$$E[E[X|\mathcal{G}_2]|\mathcal{G}_1] = E[X|\mathcal{G}_1], \quad a.s. \tag{4.18}$$

Proof: Let $Y_j = E[X|\mathcal{G}_j]$ $(j = 1, 2)$.
(4.17): Since Y_1 is $\mathcal{G}_1$-measurable, it is also $\mathcal{G}_2$-measurable. Thus, we see from (4.10) that $E[Y_1|\mathcal{G}_2] = Y_1$,a.s.
(4.18): Let $A \in \mathcal{G}_1$ be arbitrary. Then, since $A \in \mathcal{G}_2$,

$$E[Y_2 : A] \stackrel{(4.6)}{=} E[X : A] \stackrel{(4.6)}{=} E[Y_1 : A],$$

Thus, we see from (4.6) that $E[Y_2|\mathcal{G}_1] = Y_1$, a.s. \(^□^)/

Remark: $E[E[X|\mathcal{G}_1]|\mathcal{G}_2]$ and $E[E[X|\mathcal{G}_1]|\mathcal{G}_2]$ are not always the same if we do not assume either $\mathcal{G}_1 \subset \mathcal{G}_2$ or $\mathcal{G}_2 \subset \mathcal{G}_1$ (cf. Exercise 4.1.5).

Proposition 4.1.7 *Let X, Z be r.v.'s such that Z is $\mathcal{G}$-measurable, $X, ZX \in L^1(P)$. Then,*

$$E[ZX|\mathcal{G}] = ZE[X|\mathcal{G}], \quad a.s. \tag{4.19}$$

Proof: **a)** We first consider the case where $Z = \mathbf{1}_B$ with $B \in \mathcal{G}$. Let $A \in \mathcal{G}$ be arbitrary. Since $A \cap B \in \mathcal{G}$, we have

$$E[ZX : A] = E[X : A \cap B] \stackrel{(4.6)}{=} E[E[X|\mathcal{G}] : A \cap B] = E[ZE[X|\mathcal{G}] : A].$$

Thus, (4.19) holds.
b) We now consider the general case. There exists a sequence Z_n of $\mathcal{G}$-measurable simple r.v.'s such that $Z_n \stackrel{n\to\infty}{\longrightarrow} Z$ and that $|Z_n| \le |Z|$. By a) and (4.7), we have for each $n \in \mathbb{N}$ that

1) $\qquad E[Z_n X|\mathcal{G}] = Z_n E[X|\mathcal{G}]$, a.s.

Since $Z_n X \stackrel{n\to\infty}{\longrightarrow} ZX$ in $L^1(P)$ by DCT, we see from (4.13) that $E[Z_n X|\mathcal{G}] \stackrel{n\to\infty}{\longrightarrow} E[ZX|\mathcal{G}]$ in $L^1(P)$. Therefore, there exists a subsequence $\{Z_{n(k)}\}_{k\in\mathbb{N}}$ such that

2) $\qquad E[Z_{n(k)} X|\mathcal{G}] \stackrel{k\to\infty}{\longrightarrow} E[ZX|\mathcal{G}]$, a.s.

On the other hand, since $Z_n \stackrel{n\to\infty}{\longrightarrow} Z$, a.s., we have $Z_{n(k)} \stackrel{k\to\infty}{\longrightarrow} Z$, a.s., and hence,

3) $\qquad Z_{n(k)} E[X|\mathcal{G}] \stackrel{k\to\infty}{\longrightarrow} ZE[X|\mathcal{G}]$, a.s.

Thus, we obtain (4.19) from 1),2),3). \(^□^)/

Proposition 4.1.8 (**Hölder's inequality**) *Let* $p, q \in (1, \infty)$, $\frac{1}{p} + \frac{1}{q} = 1$, $X \in L^p(P)$ *and* $Y \in L^q(P)$. *Then,*

$$E[|XY||\mathcal{G}] \le E[|X|^p|\mathcal{G}]^{1/p} E[|Y|^q|\mathcal{G}]^{1/q} \quad a.s. \tag{4.20}$$

In particular,

$$E[|X||\mathcal{G}]^p \le E[|X|^p|\mathcal{G}] \quad a.s. \tag{4.21}$$

Proof: Thanks to (4.7), (4.8), and (4.19), the proof of (4.20) goes in the same way as that of usual Hölder's inequality (cf. Proposition 8.1.1). \(^□^)/

Proposition 4.1.9 (**The orthogonal projection property**) *Let* $M = L^2(\Omega, \mathcal{G}, P|_{\mathcal{G}})$ *and* $M^{\perp}$ *be its orthogonal complement in* $L^2(P)$. *Then, for* $X \in L^2(P)$,

$$E[X|\mathcal{G}] \in M, \quad X - E[X|\mathcal{G}] \in M^{\perp}, \tag{4.22}$$

that is, the map $X \mapsto E[X|\mathcal{G}]$ $(L^2(P) \to M)$ *is the orthogonal projection from* $L^2(P)$ *to* M.

Proof: $Y \stackrel{\text{def}}{=} E[X|\mathcal{G}]$ is $\mathcal{G}$-measurable by Proposition 4.1.3 and it is square integrable by (4.21). Hence, $Y \in M$. On the other hand, let $Z \in M$ be arbitrary. Then,

$$Z(X - Y) \stackrel{(4.19)}{=} ZX - E[ZX|\mathcal{G}], \text{ and hence } E[Z(X - Y)] = 0.$$

Therefore, $X - Y \in M^{\perp}$. \(^□^)/

Proposition 4.1.10 (**Jensen's inequality**) *Let* $I \subset \mathbb{R}$ *be an open interval and* $\varphi : I \to \mathbb{R}$ *be convex. Suppose that* $X : \Omega \to I$ *satisfies* $X, \varphi(X) \in L^1(P)$. *Then,*

$$\varphi\left(E[X|\mathcal{G}]\right) \le E[\varphi(X)|\mathcal{G}], \quad a.s. \tag{4.23}$$

Proof: We set $Y = E[X|\mathcal{G}]$ to simplify the notation.
a) We first consider the case where $Y \in J$ a.s., where $J \subset I$ is a compact interval. As is well known, for $y \in I$, the following limit (the right derivative of φ at y) exists and is non decreasing in y.

$$\varphi'_+(y) \stackrel{\text{def}}{=} \lim_{\substack{h \to 0 \\ h > 0}} \frac{\varphi(y + h) - \varphi(y)}{h}$$

Moreover,

$$\varphi(x) \ge \varphi(y) + \varphi'_+(y)(x - y), \quad \text{for all } x, y \in I.$$

Thus,

$$\varphi(X) \ge \varphi(Y) + \varphi'_+(Y)(X - Y), \quad \text{a.s.}$$

Since φ is continuous, and φ'_+ is monotone on I, both φ, φ'_+ are bounded on J. As a consequence, the right-hand side of the last inequality is integrable. Therefore, by taking the conditional expectation, and by using Proposition 4.1.7, we have that a.s.,

$$E[\varphi(X)|\mathcal{G}] \ge \varphi(Y) + \varphi'_+(Y)(E[X|\mathcal{G}] - Y) = \varphi(Y).$$

b) We now consider the general case. By translation, if necessary, we may assume that $0 \in I$. Let J_n $(n \geq 1)$ be an increasing sequence of compact intervals such that $J_1 \ni 0$ and $\bigcup_{n\geq 1} J_n = I$. Let also $Z_n = \mathbf{1}\{Y \in J_n\}$. Then, by Proposition 4.1.7,

$$E[Z_n X|\mathcal{G}] = Z_n Y \in J_n, \text{ a.s.}$$

Hence, we may apply the result of a) to $Z_n X$, in place of X, to obtain that

1) $\varphi(Z_n Y) = \varphi(E[Z_n X|\mathcal{G}]) \leq E[\varphi(Z_n X)|\mathcal{G}]$, a.s.

As for the left-hand side of 1), note that $Z_n Y \stackrel{n\to\infty}{\longrightarrow} Y$, a.s. Thus, by the continuity of φ,

2) $\varphi(Z_n Y) \stackrel{n\to\infty}{\longrightarrow} \varphi(Y)$, a.s.

As for the right-hand side of 1), note that

3) $\varphi(Z_n X) = Z_n \varphi(X) + (1 - Z_n)\varphi(0)$,

and hence, a.s.,

4) $E[\varphi(Z_n X)|\mathcal{G}] \stackrel{3)}{=} Z_n E[\varphi(X)|\mathcal{G}] + (1 - Z_n)\varphi(0) \stackrel{n\to\infty}{\longrightarrow} E[\varphi(X)|\mathcal{G}]$.

Thus, (4.23) follows from 1),2) and 4). \(^□^)/

Lemma 4.1.11 ($\star$) (**MCT**) *Let $X_n \in L^1(P)$ be such that $X_n \leq X_{n+1}$ $(\forall n \in \mathbb{N})$ and that $X = \sup_{n\in\mathbb{N}} X_n \in L^1(P)$. Then,*

$$E[X_n|\mathcal{G}] \stackrel{n\to\infty}{\longrightarrow} E[X|\mathcal{G}], \quad \textit{a.s. and in } L^1(P). \tag{4.24}$$

Proof: (4.24) is equivalently stated as $Y_n \stackrel{\text{def}}{=} E[X - X_n|\mathcal{G}] \stackrel{n\to\infty}{\longrightarrow} 0$, a.s. and in $L^1(P)$. As for the L^1-convergence, we note that $0 \leq X - X_n \leq X - X_1 \in L^1(P)$, which, via DCT implies that $X - X_n \stackrel{n\to\infty}{\longrightarrow} 0$ in $L^1(P)$. Thus, we see from (4.13) that

1) $Y_n \stackrel{n\to\infty}{\longrightarrow} 0$ in $L^1(P)$.

We next show that $Y_n \stackrel{n\to\infty}{\longrightarrow} 0$, a.s. We see from (4.8) that $Y_n \geq Y_{n+1} \geq 0$, a.s. for $\forall n \in \mathbb{N}$. Thus, there exists a $\mathcal{G}$-measurable r.v. $Y_\infty \geq 0$ such that $Y_n \stackrel{n\to\infty}{\longrightarrow} Y_\infty$, a.s. We combine this with 1) to coclude that $Y_\infty = 0$, a.s. \(^□^)/

Proposition 4.1.12 ($\star$) (**Fatou's lemma and DCT**) *Consider the following conditions for$X_n \in L^1(\Omega, \mathcal{F}, P)$.*

a) $\sup_{n\in\mathbb{N}} |X_n| \in L^1(P)$,

b) $X_n \stackrel{n\to\infty}{\longrightarrow} X$*,a.s. for some r.v.* X.

Then, under the assumption a),

$$E[\underline{\lim}_{n\to\infty} X_n|\mathcal{G}] \leq \underline{\lim}_{n\to\infty} E[X_n|\mathcal{G}] \leq \overline{\lim}_{n\to\infty} E[X_n|\mathcal{G}] \leq E[\overline{\lim}_{n\to\infty} X_n|\mathcal{G}], \quad \textit{a.s.} \tag{4.25}$$

Moreover, under the assumptions a) and b),

$$X \in L^1(\Omega, \mathcal{F}, P) \textit{ and } E[|X - X_n||\mathcal{G}] \stackrel{n\to\infty}{\longrightarrow} 0, \quad \textit{a.s. and in } L^1(P). \tag{4.26}$$

Proof: If we assume a), then, the inequality (4.25) follows from Lemma 4.1.11, exactly in the same way as Fatou's lemma follows from MCT in the theory of Lebesgue integration. To see (4.26), let $Y_n \stackrel{\rm def}{=} E[|X-X_n||\mathcal{G}]$. Note that a) and b) imply that $X_n \stackrel{n\to\infty}{\longrightarrow} X$ in $L^1(P)$ via DCT. Thus,

$$EY_n \stackrel{(4.6)}{=} E|X-X_n| \stackrel{n\to\infty}{\longrightarrow} 0.$$

Hence $Y_n \stackrel{n\to\infty}{\longrightarrow} 0$ in $L^1(P)$. On the other hand, by using (4.25) with X_n replaced by $|X-X_n|$, and by applying condition b), we have

$$\overline{\lim_{n\to\infty}} Y_n \stackrel{(4.25)}{\leq} E[\overline{\lim_{n\to\infty}} |X-X_n||\mathcal{G}] = 0.$$

Hence $Y_n \stackrel{n\to\infty}{\longrightarrow} 0$, P-a.s. \(^□^)/

Lemma 4.1.13 *($\star$) Let $X \in L^1(P)$. Then, the family of r.v.'s defined as follows is u.i.*

$$\{E[X|\mathcal{G}] \ ; \ \mathcal{G} \text{ is a sub } \sigma\text{-algebra of } \mathcal{F}\}.$$

Proof: Let $\varepsilon > 0$ be arbitrary. Recall from Exercise 1.1.5 that there exists $\delta > 0$ such that $E[|X| : A] < \varepsilon$ for all $A \in \mathcal{F}$ with $P(A) < \delta$. Let $m > E|X|/\delta$. Then, for any sub σ-algebra $\mathcal{G}$ of $\mathcal{F}$,

$$P(E[|X||\mathcal{G}] > m) \stackrel{\text{Chebyshev}}{\leq} E[E[|X||\mathcal{G}]]/m = E|X|/m < \delta.$$

Thus,

$$\begin{aligned} E[|E[X|\mathcal{G}]| : |E[X|\mathcal{G}]| > m] &\leq E[E[|X||\mathcal{G}]| : E[|X||\mathcal{G}] > m] \\ &= E[|X| : E[|X||\mathcal{G}] > m] < \varepsilon. \end{aligned}$$

\(^□^)/

Exercise 4.1.1 Let λ be a σ-finite measure, μ be a finite measure, and ν be a signed measure, such that $\nu \ll \mu \ll \lambda$. Prove then that λ-a.e. $\frac{d\nu}{d\lambda}$, $\frac{d\nu}{d\mu}$, $\frac{d\mu}{d\lambda}$ are well-defined and $\frac{d\nu}{d\lambda} = \frac{d\nu}{d\mu}\frac{d\mu}{d\lambda}$.

Exercise 4.1.2 Suppose that $X_1, X_2 \in L^1(P)$, $B \in \mathcal{G}$, and that $X_1 \leq X_2$ a.s. on B. Then, prove that $E[X_1|\mathcal{G}] \leq E[X_2|\mathcal{G}]$ a.s. on B.

Exercise 4.1.3 Is the converse to (4.12) true? **Hint** Let $\Omega = \{-1, 0, 1\}$, $\mathcal{F} = 2^\Omega$, $P(\{j\}) = 1/3$ $(j = 0, \pm 1)$, $\mathcal{G} = \sigma[\{0\}, \{-1, 1\}]$ and $X(j) = j$.

Exercise 4.1.4 Let $X \in L^1(P)$, $X \geq 0$, a.s., and $Y = E[X|\mathcal{G}]$. Then, prove for any $\alpha, \beta > 0$ that $P(X \geq \alpha|\mathcal{G}) \leq \beta\mathbf{1}\{Y > 0\} + \mathbf{1}\{Y \geq \alpha\beta\}$, a.s. and hence in particular that $P(X \geq \alpha) \leq \beta P(Y > 0) + P(Y \geq \alpha\beta)$.

Exercise 4.1.5 Let $\Omega = \{1, 2, 3\}$, $\mathcal{F} = 2^\Omega$, $P(\{i\}) = 1/3$, $\mathcal{G}_i = \sigma[\{i\}]$, $\chi_i(\omega) = \mathbf{1}\{\omega = i\}$ for $i = 1, 2, 3$. Then, for $X : \Omega \to \mathbb{R}$ and for $(i, j) = (1, 2), (2, 1)$, verify that

$$\begin{aligned} E[X|\mathcal{G}_i] &= X(i)\chi_i + \frac{X(j)+X(3)}{2}(1-\chi_i), \\ E[E[X|\mathcal{G}_i]|\mathcal{G}_j] &= \frac{X(j)+X(3)}{2}\chi_j + \left(\frac{X(i)}{2} + \frac{X(j)+X(3)}{4}\right)(1-\chi_j). \end{aligned}$$

Conclude from this that $E[E[X|\mathcal{G}_1]|\mathcal{G}_2] \neq E[E[X|\mathcal{G}_2]|\mathcal{G}_1]$, unless X is a constant.

Exercise 4.1.6 Suppose that $X, Y \in L^1(P)$, and that $\mathcal{G}$, $\mathcal{H}$ are sub σ-algebras of $\mathcal{F}$. Then, show the following. (i) $\sigma(X)\vee\mathcal{G}$ and $\sigma(Y)\vee\mathcal{H}$ are independent $\Rightarrow E[XY|\mathcal{G}\vee\mathcal{H}] = E[X|\mathcal{G}]E[Y|\mathcal{H}]$, a.s. (ii) $\sigma(X)\vee\mathcal{G}$ and $\mathcal{H}$ are independent $\Rightarrow E[X|\mathcal{G}\vee\mathcal{H}] = E[X|\mathcal{G}]$, a.s. (iii) $\sigma(X)$ and $\mathcal{H}$ are independent, and $\mathcal{G}$ and $\mathcal{H}$ are independent $\not\Rightarrow E[X|\mathcal{G}\vee\mathcal{H}] = E[X|\mathcal{G}]$, a.s. [Hint. $\Omega = \{0, 1, 2, 3\}$, $\mathcal{F} = 2^\Omega$, $P(\{\omega\}) = 1/4$ $(\forall\omega \in \Omega)$, $X = \mathbf{1}_{\{1,2\}}$, $\mathcal{G} = \sigma[\{2, 3\}]$, $\mathcal{H} = \sigma[\{1, 3\}]$.]

4.2 Filtrations and Stopping Times I

Throughout this subsection, we assume that

- $(\Omega, \mathcal{F}, P)$ is a probability space and $\mathbb{T} \subset \mathbb{R}$.

The set $\mathbb{T}$ is considered as the set of time parameters, typical examples of which are $\mathbb{N}$ and $[0, \infty)$. In section 5.5, we consider the case of $\mathbb{T} = -\mathbb{N}$.

Definition 4.2.1 (**Filtration, Stopping times**)
▶ A sequence $(\mathcal{F}_t)_{t\in\mathbb{T}}$ of sub σ-algebras of $\mathcal{F}$ is called a *filtration* if

$$\mathcal{F}_s \subset \mathcal{F}_t \text{ for all } s, t \in \mathbb{T} \text{ with } s < t. \tag{4.27}$$

▶ Given a filtration $(\mathcal{F}_t)_{t\in\mathbb{T}}$, a r.v. $T : \Omega \to \mathbb{T} \cup \{\infty\}$ is called a *stopping time* if

$$\{T \le t\} \in \mathcal{F}_t \text{ for all } t \in \mathbb{T}. \tag{4.28}$$

▶ Given a filtration $(\mathcal{F}_t)_{t\in\mathbb{T}}$, and a stopping time T, we define a sub σ-algebra $\mathcal{F}_T$ of $\mathcal{F}$ by

$$A \in \mathcal{F}_T \iff A \cap \{T \le t\} \in \mathcal{F}_t \text{ for all } t \in \mathbb{T}. \tag{4.29}$$

Remark It is easy to verify that $\mathcal{F}_T$ defined by (4.29) is indeed a sub σ-algebra of $\mathcal{F}$ and that, if $T \equiv t$ (a constant), then $\mathcal{F}_T = \mathcal{F}_t$.

Example 4.2.2 (**First entry/hitting time**) Let $(S, \mathcal{B})$ be a measurable space and $X_t : \Omega \to S$, $t \in \mathbb{T}$ be a sequence of r.v.'s. We set

$$\mathcal{F}_t^0 = \sigma(X_s : s \in \mathbb{T},\ s \le t). \tag{4.30}$$

Then, $(\mathcal{F}_t^0)_{t\in\mathbb{T}}$ is a filtration, which we refer to in this example. Now, suppose that $\mathbb{T} \subset [0, \infty)$. For $A \in \mathcal{B}$, we define

$$T_A = \inf\{t \in \mathbb{T}\ ;\ X_t \in A\}, \tag{4.31}$$
$$T_A^+ = \inf\{t \in \mathbb{T} \cap (0, \infty)\ ;\ X_t \in A\}. \tag{4.32}$$

T_A and T_A^+ are called, the *first entry time* and the *first hitting time*. Let us now assume for simplicity that

$$\text{every bounded subset of } \mathbb{T} \text{ is a finite set.} \tag{4.33}$$

Then, T_A and T_A^+ are stopping times w.r.t.the filtraiton (4.30). To see this, we observe that (4.33) implies the following properties.

1) $\mathbb{T}$ is at most countable.

2) Any subset of $\mathbb{T}$ is closed in $\mathbb{R}$.

We will then, verify that the following are equivalent for any $t \in \mathbb{T}$.

a) $\exists s \in \mathbb{T} \cap [0,t], \ X_s \in A$.

b) $T_A \le t$.

Indeed, it is obvious that a) implies b). To show the converse, let $U_A = \{s \in \mathbb{T}, \ X_s \in A\}$ so that $T_A = \inf U_A$ by definition. This does not directly mean[18] that b) implies a). We will verify that $T_A = \min U_A$, which does mean that b) implies a). U_A is bounded from below by definition, and is closed by 2). Moreover, $T_A < \infty \iff U_A \ne \emptyset$. Thus, if $T_A < \infty$, then $T_A = \inf U_A = \min U_A$.

Thanks to the equivalence of a) and b), togeter with the property 1), we have

$$\{T_A \le t\} = \bigcup_{s \in \mathbb{T} \cap [0,t]} \{X_s \in A\} \stackrel{1)}{\in} \mathcal{F}_t^0.$$

Similarly, $\{T_A^+ \le t\} \in \mathcal{F}_t^0$. Therefore, T_A and T_A^+ are stopping times by (4.28).

We summarize some basic properties of stopping times in the following

Lemma 4.2.3 *Let S, T and T_n $(n = 1, 2, ...)$ be stopping times. Then,*

$$T \text{ is } \mathcal{F}_T\text{-measurable}, \tag{4.34}$$
$$S \le T \implies \mathcal{F}_S \subset \mathcal{F}_T, \tag{4.35}$$
$$\sup_{n \ge 1} T_n(\omega) \in \mathbb{T} \cup \{\infty\}, \ \forall \omega \in \Omega \implies \sup_{n \ge 1} T_n \text{ is a stopping time}, \tag{4.36}$$
$$\min_{1 \le j \le n} T_j \ (n = 1, 2, ...) \text{ are stopping times}, \tag{4.37}$$
$$\mathcal{F}_{S \wedge T} = \mathcal{F}_S \cap \mathcal{F}_T, \tag{4.38}$$
$$\{S \le t < T\}, \ \{S \le T \le t\} \in \mathcal{F}_t, \ \ \forall t \in \mathbb{T}, \tag{4.39}$$
$$\{S \le T\} \in \mathcal{F}_{S \wedge T} \tag{4.40}$$

Moreover, for a r.v. X,

$$X \text{ is } \mathcal{F}_S\text{-measurable} \implies X\mathbf{1}_{\{S \le T\}} \text{ is } \mathcal{F}_{S \wedge T}\text{-measurable}. \tag{4.41}$$

Proof: (4.34):It is enough to show that $A \stackrel{\text{def}}{=} \{T \le s\} \in \mathcal{F}_T$ for $\forall s \in \mathbb{T}$. We take an arbitrary $t \in \mathbb{T}$ to verify the condition (4.29). Then,

$$A \cap \{T \le t\} = \{T \le s \wedge t\} \stackrel{(4.28)}{\in} \mathcal{F}_{s \wedge t} \stackrel{(4.27)}{\subset} \mathcal{F}_t.$$

Hence $A \in \mathcal{F}_T$.
(4.35): We take an arbitrary $A \in \mathcal{F}_S$ and show that $A \in \mathcal{F}_T$. Let us take $t \in \mathbb{T}$ to verify the condition (4.29). Note that

1) $A \cap \{S \le t\} \stackrel{(4.29)}{\in} \mathcal{F}_t$ and $\{T \le t\} \stackrel{(4.28)}{\in} \mathcal{F}_t$.

[18]For example, let $U = \{t + n^{-1}\}_{n \ge 1}$. Then, $t = \inf U$, but $U \cap [0,t] = \emptyset$.

Since $\{T \le t\} = \{S \le t\} \cap \{T \le t\}$, we see that

$$A \cap \{T \le t\} = (A \cap \{S \le t\}) \cap \{T \le t\} \stackrel{1)}{\in} \mathcal{F}_t.$$

Hence $A \in \mathcal{F}_T$ by (4.29).
(4.36): By assumption, $\sup_{n \in \mathbb{N}} T_n$ defines a measurable function from Ω to $\mathbb{T} \cup \{\infty\}$. Moreover, for all $t \in \mathbb{T}$,

$$\left\{\sup_{n \in \mathbb{N}} T_n \le t\right\} = \bigcap_{n \in \mathbb{N}} \{T_n \le t\} \in \mathcal{F}_t.$$

Hence $\sup_{n \in \mathbb{N}} T_n$ is a stopping time by (4.28).
(4.37): For $t \in \mathbb{T}$,

$$\left\{\min_{1 \le j \le n} T_j \le t\right\} = \bigcup_{1 \le j \le n} \{T_j \le t\} \in \mathcal{F}_t.$$

Hence $\min_{1 \le j \le n} T_j$ is a stopping time by (4.28).
(4.38): The inclusion $\subset$ follows from (4.35). To prove the opposite inclusion, we take an arbitrary $A \in \mathcal{F}_S \cap \mathcal{F}_T$ and $t \in \mathbb{T}$ and verify that $A \cap \{S \wedge T \le t\} \in \mathcal{F}_t$. Since $\{S \wedge T \le t\} = \{S \le t\} \cup \{T \le t\}$,

$$A \cap \{S \wedge T \le t\} = (A \cap \{S \le t\}) \cup (A \cap \{T \le t\}) \stackrel{(4.28)}{\in} \mathcal{F}_t.$$

Thus $A \in \mathcal{F}_{S \wedge T}$ by (4.29).
(4.39): As for the first set,

$$\{S \le t < T\} = \{S \le t\} \backslash \{T \le t\} \stackrel{(4.28)}{\in} \mathcal{F}_t.$$

As for the second, note that $S \wedge t$ is $\mathcal{F}_{S \wedge t}$-measurable by (4.34) and hence $\mathcal{F}_t$-measurable by (4.35). Similary $T \wedge t$ is $\mathcal{F}_t$-measurable. These imply that $\{S \wedge t \le T \wedge t\} \in \mathcal{F}_t$. Hence,

$$\{S \le T \le t\} = \{S \wedge t \le T \wedge t\} \cap \{S \le t\} \cap \{T \le t\} \in \mathcal{F}_t.$$

(4.40): We verify that the set $A \stackrel{\text{def}}{=} \{S \le T\}$ satisfies $A \cap \{S \wedge T \le t\} \in \mathcal{F}_t$ for all $t \in \mathbb{T}$ as follows.

$$A \cap \{S \wedge T \le t\} = \{S \le T, S \le t\} = \{S \le t < T\} \cup \{S \le T \le t\} \stackrel{(4.39)}{\in} \mathcal{F}_t.$$

This proves (4.40).
(4.41): Since it is enough to consider the case where $X = \mathbf{1}_A$ for some $A \in \mathcal{F}_S$, we have only to prove that $A \cap \{S \le T\} \in \mathcal{F}_{S \wedge T}$ for $A \in \mathcal{F}_S$. Note first that

$$\{S \le T\} \stackrel{(4.40)}{\in} \mathcal{F}_{S \wedge T} \stackrel{(4.35)}{\subset} \mathcal{F}_S,$$

and hence $A \cap \{S \le T\} \in \mathcal{F}_S$. On the other hand, $\mathcal{F}_{S \wedge T} = \mathcal{F}_S \cap \mathcal{F}_T$, by (4.38). Therefore, it only remains to prove that

2) $A \cap \{S \le T\} \in \mathcal{F}_T$.

To do so, we take an arbitray $t \in \mathbb{T}$. Then, $A \cap \{S \le t\} \overset{(4.29)}{\in} \mathcal{F}_t$ and $\{S \le T \le t\} \overset{(4.39)}{\in} \mathcal{F}_t$. Therefore,

$$(A \cap \{S \le T\}) \cap \{T \le t\} = (A \cap \{S \le t\}) \cap \{S \le T \le t\} \in \mathcal{F}_t,$$

which proves 2) by (4.29). \(^□^)/

Remark Referring to (4.37), it is not true in general that

$$\inf_{n \ge 1} T_n(\omega) \in \mathbb{T},\ \forall \omega \in \Omega \implies \inf_{n \ge 1} T_n \text{ is a stopping time.} \tag{4.42}$$

See Example 6.9.5 for a counterexample. On the other hand, (4.42) holds true under either of the following assumptions.

- The set $\mathbb{T}$ consists only of isolated points (To see that this implies (4.42), apply Exercise 4.2.6 to the sequence $S_n \overset{\text{def}}{=} \min_{1 \le j \le n} T_j$ of stopping times).
- $\mathbb{T} = [0, \infty)$ and the filtration is right-continuous (Exercise 6.9.2).

Lemma 4.2.4 *Let $\mathbb{T} = \mathbb{N}$, or $[0, \infty)$. If S and T are stopping times, then, so is $S + T$.*

Proof: If $\mathbb{T} = \mathbb{N}$ and $n \in \mathbb{N}$, then

1) $\{S + T \le n\} = \bigcup_{j=0}^{n} \{S \le j,\ T \le n - j\} \in \mathcal{F}_n.$

Hence, $S + T$ is a stopping time by (4.28).
Suppose that $\mathbb{T} = [0, \infty)$. By (4.28), it is enough to prove that $\{t < S + T\} \in \mathcal{F}_t$ for all $t \ge 0$. By dividing the event $\{t < S + T\}$ into the three possibilities $S = 0$, $0 < S \le t$, $t < S$, we have

$$\{t < S + T\} = \{S = 0,\ t < T\} \cup \{0 < S \le t,\ t < S + T\} \cup \{t < S\}.$$

It is easy to see that, the first, and third events on the right-hand side are in $\mathcal{F}_t$. As for the second event, we note that for $r \in (0, t)$

$$\{r < S \le t,\ t < r + T\} = \{r < S \le t,\ t - r < T\} \in \mathcal{F}_t.$$

Thus,

$$\{0 < S \le t, t - S < T\} = \bigcup_{\substack{r \in \mathbb{Q} \\ 0 < r < t}} \{r < S \le t,\ t < r + T\} \in \mathcal{F}_t.$$

Hence, $\{t < S + T\} \in \mathcal{F}_t$. See also Exercise 6.9.3 for an alternative proof assuming the right-continuity of the filtration. \(^□^)/

Exercise 4.2.1 Let S and T be stopping times and let $A \in \mathcal{F}_{S \wedge T}$. Prove then that $S\mathbf{1}_A + T\mathbf{1}_{A^c}$ is a stopping time.

Exercise 4.2.2 Referring to Example 4.2.2, let U be a stopping time and define

$$T_{A,U} = \inf\{t \in \mathbb{T}\ ;\ U \le t,\ X_t \in A\}.$$

Assuming (4.33), prove that $T_{A,U}$ is a stopping time.

Exercise 4.2.3 Let T_A be defined by (4.31), $A_n \subset S$, $n \in \mathbb{N}$ and $A = \bigcup_{n \in \mathbb{N}} A_n$. Prove then that $T_A = \inf_{n \in \mathbb{N}} T_{A_n}$.

Exercise 4.2.4 Referring to Example 4.2.2, suppose that S is a metric space, $\mathbb{T} = [0, \infty)$, and that $t \mapsto X_t(\omega)$ is left-continuous for all $\omega \in \Omega$. Suppose also that $(A_n)_{n \in \mathbb{N}}$ is a decreasing sequence of closed subsets of S and that $A = \bigcap_{n \in \mathbb{N}} A_n$. Prove then that $T_A = \sup_{n \in \mathbb{N}} T_{A_n}$.

Exercise 4.2.5 Let T_n ($n \in \mathbb{N}$) be stopping times and suppose that, for each $\omega \in \Omega$, there exists $m = m(\omega) \in \mathbb{N}$ such that $T_n = T_m$ for all $n \geq m$. Then, prove that $T \stackrel{\text{def}}{=} \lim\limits_{n \to \infty} T_n$ is a stopping time. Hint: Note that $\Omega = \bigcup\limits_{m \in \mathbb{N}} A_m$, where $A_m = \bigcap\limits_{n \geq m} \{T_n = T_m\}$ and that $T = T_m$ on A_m. Therefore, it is enough to show that $A_m \cap \{T_m \leq t\} \in \mathcal{F}_t$ for all $t \in \mathbb{T}$.

Exercise 4.2.6 Let T_n ($n \in \mathbb{N}$) be stopping times. Suppose that the set $\mathbb{T}$ consists only of isolated points and that, for all $\omega \in \Omega$, $T(\omega) = \lim_{n \to \infty} T_n(\omega)$ exists and belongs to $\mathbb{T}$. Then, prove that T is a stopping time. Hint: Check that the assumption for Exercise 4.2.5 is satisfied.

4.3 Martingales, Definition and Examples

Throughout this section, we assume that

- $(\Omega, \mathcal{F}, P)$ is a probability space and $\mathbb{T} \subset \mathbb{R}$;
- $(\mathcal{F}_t)_{t \in \mathbb{T}}$ is a filtration, cf. Definition 4.2.1;
- $X = (X_t)_{t \in \mathbb{T}}$ is a sequence of real r.v.'s defined on $(\Omega, \mathcal{F}, P)$.

Definition 4.3.1 Referring to the notation introduced at the beginning of this section, $X = (X_t, \mathcal{F}_t)_{t \in \mathbb{T}}$ is called a *martingale* if the following hold true.

- (**adapted**) X_t is $\mathcal{F}_t$-measurable for all $t \in \mathbb{T}$;
- (**integrable**) $X_t \in L^1(P)$ for all $t \in \mathbb{T}$;
- (**martingale property**)

$$E[X_t | \mathcal{F}_s] = X_s \text{ a.s. if } s, t \in \mathbb{T} \text{ and } s < t. \tag{4.43}$$

If the equality in (4.43) is replaced by $\geq$ (resp. $\leq$), X is called a *submartingale* (resp. *supermartingale*).

Remark When we simply say that $(X_t)_{t \in \mathbb{T}}$ is a martingale, it means that $(X_t, \mathcal{F}_t)_{t \in \mathbb{T}}$ is a martingale for some filtration $(\mathcal{F}_t)_{t \in \mathbb{T}}$. This applies similarly to submartingales and supermartingales.

Example 4.3.2 a) If $t \mapsto X_t$ is a non random function of t, then it is a submartingale (resp. supermartingale) iff it is nondecreasing (resp. nonincreasing).

b) Let $Y \in L^1(P)$. Then, the process defined by $X_t = E[Y | \mathcal{F}_t]$, $t \in \mathbb{T}$ is a martingale.

c) Let $\mathcal{F}_\infty = \sigma(\mathcal{F}_t, t \in \mathbb{T})$, Q be a signed measure on $(\Omega, \mathcal{F}_\infty)$, and $P_t = P|_{\mathcal{F}_t}$, $Q_t = Q|_{\mathcal{F}_t}$. Suppose that $Q_t \ll P_t$ for all $t \in \mathbb{T}$. Then, $X_t \stackrel{\text{def}}{=} \frac{dQ_t}{dP_t}$, $t \in \mathbb{T}$ is a martingale.

Proof: a) Obvious.
b) It follows from the definition of the conditional expectation that $(X_t)_{t\in\mathbb{T}}$ is adapted and integrable. Moreover, let $s, t \in \mathbb{T}$, $s < t$ and $A \in \mathcal{F}_s$. Then,

$$E[X_t|\mathcal{F}_s] = E[E[Y|\mathcal{F}_t]|\mathcal{F}_s] \stackrel{(4.18)}{=} E[Y|\mathcal{F}_s] = X_s.$$

Hence, $(X_t)_{t\in\mathbb{T}}$ is a martingale.
c) X_t is $\mathcal{F}_t$-measurable and $X_t \in L^1(P)$. Let $s, t \in \mathbb{T}$, $s < t$ and $A \in \mathcal{F}_s$. Then, since $A \in \mathcal{F}_t$,

$$E[X_t : A] = Q_t(A) = Q(A) = Q_s(A) = E[X_s : A].$$

Thus, $E[X_t|\mathcal{F}_s] = X_s$, a.s. \(^□^)/

Remark: Example 4.3.2 b) is a special case of c), where $Q(A) = E[Y : A]$. One might then ask:

> For all martingale $(X_t)_{t\in\mathbb{T}}$, does there exist a signed measure Q such that $Q_t \ll P_t$ and $X_t = \frac{dQ_t}{dP_t}$ for all $t \in \mathbb{T}$?

See Proposition 4.7.1 below for the answer.

Lemma 4.3.3 *Let $X = (X_t, \mathcal{F}_t)_{t\in\mathbb{T}}$ be a submartingale and $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex function such that $\varphi(X_t) \in L^1(P)$ for all $t \in \mathbb{T}$. Then, $(\varphi(X_t), \mathcal{F}_t)_{t\in\mathbb{T}}$ is a submartingale if either φ is increasing or X is a martingale.*

Proof: Let $s, t \in \mathbb{T}$, $s < t$. We will prove that

1) $E[\varphi(X_t)|\mathcal{F}_s] \geq \varphi(X_s)$ a.s.

By Proposition 4.1.10,

2) $E[\varphi(X_t)|\mathcal{F}_s] \geq \varphi(E[X_t|\mathcal{F}_s])$ a.s.

If φ is increasing, then 2) implies 1), since $E[X_t|\mathcal{F}_s] \geq X_s$, a.s. If X is a martingale, then 2) implies 1) again, since $E[X_t|\mathcal{F}_s] = X_s$, a.s. Therefore, $(\varphi(X_t), \mathcal{F}_t)_{t\in\mathbb{T}}$ is a submartingale in both cases. \(^□^)/

Remark: For a submartingale $X = (X_t, \mathcal{F}_t)_{t\in\mathbb{T}}$ and a convex function $\varphi : \mathbb{R} \to \mathbb{R}$, $(\varphi(X_t), \mathcal{F}_t)_{t\in\mathbb{T}}$ is not necessarily a submartingale. In fact, let $t \mapsto X_t$ be a non random, strictly increasing positive function and $\varphi(x) = 1/x$. Then, X is a submartingale (Example 4.3.2 a)) and φ is convex. However, $\varphi(X_t) = 1/X_t$ is not a submartingale, since it is a non random, strictly decreasing function.

In what follows, we consider the case of $\mathbb{T} = \mathbb{N}$. For r.v.'s X_n, $n \in \mathbb{N}$, we set

$$\Delta X_n = X_n - X_{n-1}, \quad n \geq 1. \tag{4.44}$$

Lemma 4.3.4 *Suppose that $X = (X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ is adapted, integrable. Then, the following are equivalent.*

$$X \text{ is a martingale;} \tag{4.45}$$
$$E[X_{n+1}|\mathcal{F}_n] = X_n, \quad \text{a.s. for all } n \in \mathbb{N}; \tag{4.46}$$
$$E[\Delta X_{n+1}|\mathcal{F}_n] = 0, \quad \text{a.s. for all } n \in \mathbb{N}. \tag{4.47}$$

Moreover, submartingale (resp. supermartingale) are characterized by similar conditions as (4.46) and (4.47) with equalities replaced by $\geq$ (resp. $\leq$).

Proof: (4.45) $\Longrightarrow$ (4.46) $\Longleftrightarrow$(4.47): Obvious.
(4.47) $\Rightarrow$ (4.45): Let $m, n \in \mathbb{N}$, $m < n$. Since $X_n - X_m = \sum_{j=m}^{n-1} \Delta X_{j+1}$, we have

$$\begin{aligned} E[X_n|\mathcal{F}_m] - X_m &= E[X_n - X_m|\mathcal{F}_m] = \sum_{j=m}^{n-1} E[\Delta X_{j+1}|\mathcal{F}_m] \\ &\stackrel{(4.18)}{=} \sum_{j=m}^{n-1} E[E[\Delta X_{j+1}|\mathcal{F}_j]|\mathcal{F}_m] \stackrel{(4.47)}{=} 0 \text{ a.s.} \end{aligned}$$

The case of submartingale (resp. supermartingale) can be treated similarly. \(^□^)/

As a direct consequence of the preceeding lemma, we have

Example 4.3.5 (**summation of conditionally mean-zero r.v's**) Let $(\xi_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ be adapted, integrable. We define $X = (X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ by

$$X_n = \sum_{j=0}^{n} \xi_j.$$

Then,

$$E[\xi_{n+1}|\mathcal{F}_n] = 0 \text{ a.s. for } n \in \mathbb{N} \iff X \text{ is a martingale.} \tag{4.48}$$

Moreover,

$$E[\xi_{n+1}|\mathcal{F}_n] \geq 0 \text{ (resp. } \leq 0) \text{ a.s.} \iff X \text{ is a submartingale (resp. supermartingale).} \tag{4.49}$$

Example 4.3.6 (**product of conditionally mean-one r.v's**) Let $(\xi_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ be adapted, integrable. We define $X = (X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ by

$$X_n = \prod_{j=0}^{n} \xi_j.$$

We assume that $X_n \in L^1(P)$ for all $n \in \mathbb{N}$. Then,

$$E[\xi_{n+1}|\mathcal{F}_n] = 1 \text{ a.s. for } n \in \mathbb{N} \Longrightarrow X \text{ is a martingale.} \tag{4.50}$$

The converse is true if $X_n \neq 0$ a,s. for $n \in \mathbb{N}$.

Suppose in addition that $\xi_n \geq 0$ a.s. for all $n \in \mathbb{N}$. Then,

$$E[\xi_{n+1}|\mathcal{F}_n] \geq 1 \text{ (resp. } \leq 1) \text{ a.s.} \Longrightarrow X \text{ is a submartingale (resp. supermartingale).} \tag{4.51}$$

The converse is true if $X_n \neq 0$ a,s. for $n \in \mathbb{N}$.

Proof: Before go into (4.50), let us observe the consequence of the preamble. X is adapted by the definition and is integrable by the assumption. Let $n \in \mathbb{N}$. Since $X_{n+1} = X_n \xi_{n+1}$, we have

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

(4.50) ($\Rightarrow$) The right-hand side of 1) is $= X_n$ a.s., if $E[\xi_{n+1}|\mathcal{F}_n] = 1$, a. s.
(4.50) ($\Leftarrow$) If X is a martingale, then, it follows from 1) that $X_n = X_n E[\xi_{n+1}|\mathcal{F}_n]$ a.s. Thus, $1 = E[\xi_{n+1}|\mathcal{F}_n]$ a.s. if $X_n \neq 0$ a,s.
Proofs of (4.51) and its converse are similar. \(^□^)/

Remark: Referring to Example 4.3.6 a), suppose that $\xi_0, \xi_1, \ldots$ are independent, $\mathcal{F}_n = \sigma(\xi_0, \ldots, \xi_n)$, $n \in \mathbb{N}$ and $E\xi_n = 1$, $n \geq 1$. Then, $E[\xi_{n+1}|\mathcal{F}_n] \stackrel{(4.12)}{=} E\xi_{n+1} = 1$ a.s. for $n \in \mathbb{N}$. Hence X is a martingale.

($\star$) Complement to section 4.3: **Analogy between martingales and harmonic fnctions** Let us briefly review some basic properties of harmonic function on the open unit disc $D \subset \mathbb{C}$.

Soppose that a function $u : D \to \mathbb{R}$ is Borel measurable and locally bounded. u is called *harmonic* if

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} u(a + re^{\mathbf{i}\theta})d\theta = u(a), \tag{4.52}$$

whenever $a + r\overline{D} \subset D$ ($a \in D$, $r \in (0,1)$). Similarly u is called *subharmonic* (resp. *superharmonic*) if the equality in the definition (4.52) is replaced by the inequality $\geq$ (resp. $\leq$). Suppose in particular that $u \in C^2(D)$. Then u is harmonic (resp. subharmonic, superharmonic) if and only if

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) u = 0 \ \text{ resp. } (\geq 0, \leq 0) \text{ on } D.$$

cf. [MP10, p.65, Theorem 3.2].

In what follows, we identify the unit circle $\mathbb{S}^1$ with the interval $(-\pi, \pi]$, equipped with the Borel σ-algebra and the normalized Lebesgue measure. For $0 < r \leq 1$ and a Borel measurable and integrable function $f : \mathbb{S}^1 \to \mathbb{R}$, we define the *Poisson integral* $H_r f : r\overline{D} \to \mathbb{R}$ by

$$(H_r f)(z) = \begin{cases} \frac{1}{2\pi}\int_{-\pi}^{\pi} h(z, re^{\mathbf{i}\varphi}) f(e^{\mathbf{i}\varphi}) d\varphi, & \text{if } z \in rD, \\ f(\sigma), & \text{if } z = r\sigma,\ \sigma \in \mathbb{S}^1, \end{cases} \tag{4.53}$$

where $h(z, w)$ denotes the *Poisson kernel*:

$$h(z, w) = \mathrm{Re}\,\frac{w - z}{w + z} = \frac{|w|^2 - |z|^2}{|w - z|^2}. \tag{4.54}$$

It is easy to see that for $z \in rD$ and $w \in r\mathbb{S}^1$,

$$0 \leq h(z, w) \leq r + |z|, \quad \frac{1}{2\pi}\int_{-\pi}^{\pi} h(z, re^{\mathbf{i}\theta})d\theta = 1. \tag{4.55}$$

Therefore, the function $z \mapsto H_r f(z)$ on the the disc $r\overline{D}$ is well-defined and is obtained by averaging $f : \mathbb{S}^1 \to \mathbb{R}$ by the probability measure $\frac{1}{2\pi}h(z, re^{\mathbf{i}\theta})d\theta$. It is known that

$$H_r f \text{ is continuous on } r\overline{D}, \text{ harmonic on } rD. \tag{4.56}$$

cf. [Rud87, p.112, 5.25]. For $0 < r \le 1$, and a Borel measurable function $u : r\overline{D} \to \mathbb{R}$, let $u_r : \mathbb{S}^1 \to \mathbb{R}$ be defined by $u_r(\sigma) = u|_{r\mathbb{S}^1}(r\sigma)$. Then,

$$\text{If } u \in C(r\overline{D}) \text{ is harmonic on } rD, \text{ then } H_r u_r = u \text{ on } rD. \tag{4.57}$$

$$\text{If } u \in C(r\overline{D}) \text{ is subharmonic (resp. superharmonic) on } rD, \text{ then } H_r u_r \ge u \text{ (resp. } H_r u_r \le u) \text{ on } rD. \tag{4.58}$$

cf. [Rud87, p.112, 5.25, p.234, 11.8, p.338, 17.9].

If $u \in C(D)$ is harmonic on D, then, it follows from (4.57) that

$$(H_t u_t)_s = u_s \quad 0 \le s < t < 1,$$

This can be thought of as an analogy of the martingale property $E[X_t|\mathcal{F}_s] = X_s$ $(0 \le s < t)$. Similarly, if $u \in C(D)$ is subharmonic (resp. superharmonic) on D, then, it follows from (4.58) that

$$(H_t u_t)_s \ge u_s \text{ (resp. } (H_t u_t)_s \le u_s) \text{ if } 0 \le s < t < 1.$$

This can be thought of as an analogy of the submartingale (resp. supermartingale) property $E[X_t|\mathcal{F}_s] \ge X_s$ (resp. $E[X_t|\mathcal{F}_s] \le X_s$).

Exercise 4.3.1 Let $(X_t, \mathcal{F}_t)_{t\in\mathbb{T}}$ be a martingale, $s, t \in \mathbb{T}$, $s < t$. Suppose that a r.v. Y is $\mathcal{F}_s$-measurable and that $X_t Y \in L^1(P)$. Prove then that $X_s Y \in L^1(P)$ and that $E[X_t Y] = E[X_s Y]$.

Exercise 4.3.2 Let $s, t \in \mathbb{T}$, $s < t$. Prove the following. i) If $(X_t)_{t\in\mathbb{T}}$ is a nonnegative supermartingale, then, $X_t = 0$ a.s. on $\{X_s = 0\}$. ii) If $(X_t)_{t\in\mathbb{T}}$ is a nonnegative submartingale and $X_t = 0$ a.s. then, $X_s = 0$ a.s. [Here, it is not true in general that $X_s = 0$ a.s. on $\{X_t = 0\}$. For example, consider a nonnegative submartingale $X_n = |S_n|$, where S_n is a simple random walk with $S_0 \equiv 0$. Then, $\{X_{2n} = 0\} \subset \{X_{2n-1} = 1\}$ for $n \ge 1$.]

Exercise 4.3.3 Let $(Y_t, \mathcal{F}_t)_{t\in\mathbb{T}}$ be a martingale, $a \in \mathbb{T}$, and Z_t, $t \in \mathbb{T} \cap (-\infty, a]$ be $\mathcal{F}_a$-measurable r.v.'s. Suppose that $Y_t = 0$ for $t \le a$, and that $Y_t Z_a \in L^1(P)$ for $t \ge a$. Prove then that $X_t = Y_t Z_{t\wedge a}$ is a martingale. [Hint: Prove (4.43) separately for $s \le a$ and for $s \ge a$.]

Exercise 4.3.4 Let $\xi_1, \xi_2, ... \in L^1(P)$ be mean-zero, independent, $\mathcal{F}_0 = \{\emptyset, \Omega\}$, and $\mathcal{F}_n = \sigma(\xi_1, ..., \xi_n)$, $n \ge 1$. For $k \in \mathbb{N}\backslash\{0\}$, prove that $(X_n^{(k)}, \mathcal{F}_n)_{n\in\mathbb{N}}$ defined as follows is a martingale.

$$X_0^{(k)} = 0, \quad X_n^{(k)} = \sum_{1\le j_1 < ... < j_k \le n} \xi_{j_1} \cdots \xi_{j_k}, \quad n \ge 1.$$

Exercise 4.3.5 Let $X_0, \xi_n, \eta_n \in L^1(P)$, $n \in \mathbb{N}\backslash\{0\}$ be such that $E\xi_n = 0$, $E\eta_n = 1$ for all $n \in \mathbb{N}\backslash\{0\}$ and that $X_0, \zeta_1, \zeta_2, ...$ are independent, where $\zeta_n = (\xi_n, \eta_n)$. We define X_n, $n \in \mathbb{N}\backslash\{0\}$ by $X_n = \xi_n + \eta_n X_{n-1}$ for $n \ge 1$. Then, prove that $(X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ is a martingale, where $\mathcal{F}_n = \sigma(X_0, \zeta_1, ..., \zeta_n)$.

Exercise 4.3.6 Suppoese that X_t $(t \ge 0)$ is a nonnegative submartigale and $b \in (0, \infty)$. Then, prove that $(X_t)_{t\le b}$ is uniformly integrable.

4.4 Discrete Stochastic Integral

Definition 4.4.1 A sequence of r.v.'s $H = (H_n)_{n\in\mathbb{N}}$ is said to be *predictable* if H_n is $\mathcal{F}_{n-1}$-measurable for all $n \geq 1$.

Proposition 4.4.2 *For sequences $X = (X_n)_{n\in\mathbb{N}}$, $H = (H_n)_{n\geq 1}$ of r.v.'s, we define $H\cdot X = ((H \cdot X)_n)_{n\in\mathbb{N}}$ by*

$$(H \cdot X)_0 = 0 \ \text{ and } \ (H \cdot X)_n = \sum_{j=1}^{n} H_j \Delta X_j \ \text{ for } n \geq 1. \tag{4.59}$$

cf. (4.44). Suppose that H is predictable and that $H_n\Delta X_n \in L^1(P)$ for $n \geq 1$.

a) *If X is a martingale w.r.t. $(\mathcal{F}_n)_{n\in\mathbb{N}}$, then, so is $H \cdot X$.*

b) *Suppose that $H_n \geq 0$ a.s. for all $n \geq 1$. If X is a submartingale (resp. supermartingale) w.r.t. $(\mathcal{F}_n)_{n\in\mathbb{N}}$, then, so is $H \cdot X$.*

The process $H \cdot X$ is called the the **discrete stochastic integral** *of H by X.*

Proof: $H \cdot X$ is adapted by the definition and is integrable by the assumption. Let $n \in \mathbb{N}$. Since $\Delta(H \cdot X)_{n+1} = H_{n+1}\Delta X_{n+1}$ and H_{n+1} is $\mathcal{F}_n$-measurable, we have

1) $E[\Delta(H \cdot X)_{n+1}|\mathcal{F}_n] = H_{n+1}E[\Delta X_{n+1}|\mathcal{F}_n]$.

The right-hand side of 1) is $= 0$ a.s., if X is a martingale. Suppose that $H_n \geq 0$ a.s. for all $n \geq 1$. Then, the right-hand side of 1) is ≥ 0 (resp. ≤ 0) a.s., if X is a submartingale (resp. supermartingale). Thus, we obtain a) and b) by Lemma 4.3.4. \(^□^)/

The following corollary to Proposition 4.4.2 will be applied to proof of the upcrossing inequality (Lemma 5.1.6), which is a key lemma for the martingale convergence theorem (Theorem 5.1.1).

Corollary 4.4.3 *Suppose that $X = (X_n)_{n\in\mathbb{N}}$ is a submartingale and that $H = (H_n)_{n\geq 1}$, $K = (K_n)_{n\geq 1}$ are predictable, $H_n\Delta X_n, K_n\Delta X_n \in L^1(P)$, $H_n \leq K_n$ a.s., $\forall n \geq 1$. Then,*

$$E(H \cdot X)_n \leq E(K \cdot X)_n, \quad \forall n \in \mathbb{N}. \tag{4.60}$$

If X is replaced by a supermartingale, then the inequality $\leq$ in (4.60) is replaced by $\geq$.

Proof: $(K - H) \cdot X$ is a submartingale by Proposition 4.4.2. Thus,

$$E(K \cdot X)_n - E(H \cdot X)_n = E((K - H) \cdot X)_n \geq E((K - H) \cdot X)_0 = 0.$$

\(^□^)/

Corollary 4.4.4 (**stopped processes**) *Let S and T be stopping times w.r.t. $(\mathcal{F}_n)_{n\in\mathbb{N}}$ such that $S(\omega) \le T(\omega)$ for all $\omega \in \Omega$. If $X = (X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ is a submartingale (resp. supermartingale), then, so is*

$$(X_{T\wedge n} - X_{S\wedge n}, \mathcal{F}_n)_{n\in\mathbb{N}}.$$

In particular, taking $S \equiv 0$, $(X_{T\wedge n}, \mathcal{F}_n)_{n\in\mathbb{N}}$ is a submartingale (resp. supermartingale).

Proof: Let $H_n = \mathbf{1}\{S < n \le T\}$. Then, H_n, $n \ge 1$ is predictable, since,

$$\{S < n \le T\} = \{S \le n-1\}\backslash\{T \le n-1\} \in \mathcal{F}_{n-1}.$$

Thus, $H \cdot X$ is a submartingale by Proposition 4.4.2. Moreover, for $n \ge 1$,

$$\begin{aligned}(H\cdot X)_n &= \sum_{j=1}^n \mathbf{1}_{\{S<j\le T\}}\Delta X_j = \sum_{j=1}^n \mathbf{1}_{\{j\le T\}}\Delta X_j - \sum_{j=1}^n \mathbf{1}_{\{j\le S\}}\Delta X_j \\ &= \sum_{j=1}^{T\wedge n}\Delta X_j - \sum_{j=1}^{S\wedge n}\Delta X_j = X_{T\wedge n} - X_{S\wedge n}.\end{aligned}$$

\(^▫^)/

4.5 Hitting Times for One-dimensional Random Walks

Let ξ_n, $n \in \mathbb{N}\backslash\{0\}$ be i.i.d. such that $\xi_n = 0, \pm 1$ with probabilities, p_0, $p_\pm$, respectively, where $p_0 \ge 0$, $p_\pm > 0$, $p_0 + p_+ + p_- = 1$. We define $(S_n)_{n\in\mathbb{N}}$ by

$$S_0 = 0, \quad S_{n+1} = S_n + \xi_{n+1}, \quad n \in \mathbb{N}.$$

We consider a filtration defined by $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_n = \sigma(\xi_1, ..., \xi_n)$ for $n \ge 1$. In this subsection, we investigate the following stopping time.

$$T_a = \inf\{n \ge 0\ ;\ S_n = a\} \quad a \in \mathbb{Z}.$$

For this purpose, we introdce the following function.

$$g(s,t) \stackrel{\text{def}}{=} stEt^{\xi_1} - t = p_+st^2 - (1-p_0s)t + p_-s, \quad \text{for } s > 0 \text{ and } t \in \mathbb{R}. \tag{4.61}$$

As for the discriminant of the quadratic function $t \mapsto g(s,t)$, we have

$$\delta(s) \stackrel{\text{def}}{=} (1-p_0s)^2 - 4p_+p_-s^2 \ge 0 \quad \text{for } s \in (0, s_*],$$

where

$$s_* = \frac{1}{2\sqrt{p_+p_-} + p_0}\begin{cases} > 1, & \text{if } p_+ \ne p_-, \\ = 1, & \text{if } p_+ = p_-. \end{cases} \tag{4.62}$$

For $s \in (0, s_*]$, we define

$$f_\pm(s) = \frac{1 - p_0s - \sqrt{\delta(s)}}{2p_\pm s}. \tag{4.63}$$

Then, for any fixed $s \in (0, s_*]$, the equation $g(s,t) = 0$ has real solutions

$$t = f_+(s) \ \text{ and } \ t = \frac{1 - rs + \sqrt{\delta(s)}}{2p_+s} = f_-(s)^{-1}.$$

Let us quickly collect some information on $f_\pm(s)$, which we will need. To do so, it is enough to look at $f_+(s)$ only, since $f_\pm(s)$ are essentially the same, with only the roles of $p_\pm$ interchanged. The function f_+ is differentiable on $(0, s_*)$ and

$$f'_+(s) = \frac{f_+(s)}{s\sqrt{\delta(s)}}, \quad s \in (0, s_*). \tag{4.64}$$

This can be computed for example as follows. Since $g(s, f_+(s)) \equiv 0$, we have

$$\begin{aligned} 0 &= \frac{d}{ds} g(s, f_+(s)) = \frac{\partial g}{\partial s}(s, f_+(s)) + \frac{\partial g}{\partial t}(s, f_+(s)) f'_+(s) \\ &= p_+ f_+(s)^2 + p_0 f_+(s) + p_- + (2p_+ s f_+(s) - (1 - p_0 s)) f'_+(s) \\ &= f_+(s)/s - \sqrt{\delta(s)} f'_+(s). \end{aligned}$$

By (4.64), the functions f_+ behave as we summarize in the following table.

s	0	$\nearrow$	1	$\nearrow$	s_*
$f_+(s)$	0	$\nearrow$	$(p_-/p_+) \wedge 1$	$\nearrow$	$(p_-/p_+)^{1/2}$

In particular, we note that

$$f_+(s) < f_+(1) = (p_-/p_+) \wedge 1, \quad \text{for all } s \in (0, 1). \tag{4.65}$$

Lemma 4.5.1 *Let $0 < s \le s_*$, $t > 0$, and $X_n = t^{S_n} s^n$, $n \in \mathbb{N}$. Then,*

a) $(X_n)_{n \in \mathbb{N}}$ *is a* $\begin{cases} \textit{supermartingale if } t \in [f_+(s), f_-(s)^{-1}], \\ \textit{submartingale if } t \notin (f_+(s), f_-(s)^{-1}). \end{cases}$ *In particular, the following processes are martingales.*

$$X_\pm(n) \stackrel{\text{def}}{=} f_\pm(s)^{\pm S_n} s^n, \quad n \in \mathbb{N}. \tag{4.66}$$

b) *Suppose that T is a stopping time such that $(X_\pm(n \wedge T))_{n \in \mathbb{N}}$ is bounded, and $X_\pm(n \wedge T) \stackrel{n \to \infty}{\longrightarrow} Y_\pm$, a.s. for some r.v $Y_\pm$. Then,*

$$EY_\pm = 1. \tag{4.67}$$

Proof: a) We compute

$$X_{n+1} - X_n = t^{S_n + \xi_{n+1}} s^{n+1} - t^{S_n} s^n = t^{S_n - 1} s^n (t^{\xi_{n+1}+1} s - t).$$

Since ξ_{n+1} is independent of $\mathcal{F}_n$, we have

$$\begin{aligned} E[X_{n+1} | \mathcal{F}_n] - X_n &\stackrel{(4.12)}{=} t^{S_n - 1} s^n (st E[t^{\xi_{n+1}}] - t) \\ &= t^{S_n - 1} s^n (p_+ s t^2 + p_- s + p_0 s t - t) = t^{S_n - 1} s^n g(s, t). \end{aligned}$$

Note that

$$g(s, t) \begin{cases} \le 0 & \text{if } t \in [f_-(s)^{-1}, f_+(s)], \\ \ge 0 & \text{if } t \notin (f_-(s)^{-1}, f_+(s)). \end{cases}$$

Therefore, we arrive at the conclusion via Lemma 4.3.4.
b) By a) and Corollary 4.4.4, $X_\pm(n \wedge T)$, $n \in \mathbb{N}$ are martingales, so that

$$EX_\pm(n \wedge T) = EX_\pm(0) = 1.$$

Then, (4.67) follows from BCT. \\(^□^)/

Corollary 4.5.2 *Let $\varphi : \mathbb{Z} \to \mathbb{R}$ be defined by $\varphi(x) = x$ if $p_+ = p_-$ and $\varphi(x) = (p_-/p_+)^x$ if $p_+ \neq p_-$. Then, $(\varphi(S_n), \mathcal{F}_n)_{n\in\mathbb{N}}$ is a martingale.*

Proof: If $p_+ = p_-$, then $\varphi(S_n) = S_n$ is a martingale, since it is the summation of mean-zero i.i.d ξ_n. If $p_+ \neq p_-$, then

$$(p_-/p_+) \stackrel{(4.65)}{=} \begin{cases} f_+(1) & \text{if } p_+ > p_+, \\ f_-(1)^{-1} & \text{if } p_+ < p_-. \end{cases}$$

Therefore,

$$\varphi(S_n) = (p_-/p_+)^{S_n} = \begin{cases} f_+(1)^{S_n} & \text{if } p_+ > p_+, \\ f_-(1)^{-S_n} & \text{if } p_+ < p_-, \end{cases}$$

which is a martingale by Lemma 4.5.1. \\(^□^)/

Proposition 4.5.3 [a] *For $a \in \mathbb{N}\backslash\{0\}$ and $0 < s < 1$,*

$$Es^{T_a} = f_-(s)^a, \quad Es^{T_{-a}} = f_+(s)^a, \tag{4.68}$$

$$P(T_a < \infty) = ((p_+/p_-) \wedge 1)^a, \quad P(T_{-a} < \infty) = ((p_-/p_+) \wedge 1)^a. \tag{4.69}$$

with the convention that $s^\infty = 0$. Moreover, if $p_+ < p_-$, then

$$ET_{-a} = E[T_a | T_a < \infty] = \frac{a}{p_- - p_+}. \tag{4.70}$$

On the other hand, if $p_+ = p_-$, then

$$ET_{-a} = ET_a = \infty. \tag{4.71}$$

[a] See also Exercise 3.7.1, Exercise 3.3.3, and Exercise 3.4.3

Proof: (4.68): To prove the first equality, note that $S(n \wedge T_a) \leq a$, and that

$$1 \leq (p_+/p_-) \vee 1 \stackrel{(4.65)}{<} f_-(s)^{-1}.$$

Thus,

$$0 \leq X_-(n \wedge T_a) \leq f_-(s)^{-S(n\wedge T_a)} \leq f_-(s)^{-a}.$$

If $T_a < \infty$, then, $S(n \wedge T_a) \stackrel{n\to\infty}{\longrightarrow} S(T_a) = a$, and hence,

$$X_-(n \wedge T_a) = f_-(s)^{-S(n\wedge T_a)} s^{n\wedge T_a} \stackrel{n\to\infty}{\longrightarrow} f_-(s)^{-a} s^{T_a}.$$

On the other hand, if $T_a = \infty$, then, $0 \leq f_-(s)^{-S_n} \leq f_-(s)^{-a}$, $\forall n \in \mathbb{N}$, and hence

$$X_-(n \wedge T_a) = f_-(s)^{-S_n} s^n \stackrel{n\to\infty}{\longrightarrow} 0 = f_-(s)^{-a} s^{T_a}.$$

We now apply (4.67) to X_- and $T = T_a$:

$$1 = f_-(s)^{-a} E s^{T_a}.$$

This proves the first equality. The second equality is obtained in the same way.
(4.69): We have for any r.v. $T : \Omega \to [0, \infty]$ that

$$\lim_{\substack{s \to 1 \\ s<1}} E s^T = P(T < \infty).$$

Thus, we see (4.69) from (4.65) and (4.68).
(4.70), (4.71): We compute the limits $f'_-(1-) \stackrel{\text{def}}{=} \lim_{\substack{s \to 1 \\ s<1}} f'_-(s)$. We see from (4.64) and (4.65) that

$$\textbf{1)} \qquad f'_-(1-) = \begin{cases} \frac{1}{p_+ - p_-} & \text{if } p_+ > p_-, \\ \frac{1}{p_- - p_+} \cdot \frac{p_+}{p_-} & \text{if } p_+ < p_-, \\ \infty & \text{if } p_+ = p_-. \end{cases}$$

It follows from (4.68) and Exercise 1.1.6 that

$$\begin{aligned} E[T_a : T_a < \infty] &= \lim_{\substack{s \to 1 \\ s<1}} \frac{d}{ds} E s^{T_a} \stackrel{(4.68)}{=} \lim_{\substack{s \to 1 \\ s<1}} \frac{d}{ds} f_+(s)^{-a} \\ &= a f_-(1)^{-a-1} f'_-(1-) \stackrel{1),\,(4.65)}{=} \frac{a}{p_- - p_+} \left(\frac{p_+}{p_-} \right)^a. \end{aligned}$$

Since $P(T_a < \infty) = (p_+/p_-)^a$ by (4.69), we obtain the second equality of (4.70). The other equalities can be obtained in the same way. \(^□^)/

Remark (i) See Exercise 3.3.3 and Exercise 3.7.1 for alternative proofs for (4.69). **(ii)** If $p_+ < p_-$, the validity of the first identity of (4.68) extends to all $s \in (0, s_*]$ (cf. (4.62)). In particular, T_{-a} is exponentially integrable. To see this, we note that $X_n = f_-(s_*)^{-S_n} s_*^n$ is a martingale by Lemma 4.5.1. Thus $E s_*^{T_{-a}} \le f_-(s_*)^{-a}$ by Exercise **??**. This implies that $E s^{T_{-a}}$ for $s \in \mathbb{C}$, $|s| < s_*$ can be expressed as an absolutely converging power seris. Therefore, by the unicity theorem, the first identity of (4.68) extends to all $s \in (0, s_*)$. Finally, the case of $s = s_*$ is obtained by the monotone convergence theorem.

Corollary 4.5.4 *Suppose that $p_+ < p_-$. Then, the following r.v. is geometrically distributed with parameter p_+/p_-.*

$$M \stackrel{\text{def}}{=} \max_{n \in \mathbb{N}} S_n.$$

Proof: $P(M \ge a) = P(T_a < \infty) \stackrel{(4.69)}{=} (p_+/p_-)^a$. \(^□^)/

Proposition 4.5.5 [a] *For $a, b \in \mathbb{N}\backslash\{0\}$ and $s \in (0,1]$,*

$$E[s^{T_{-a}} : T_{-a} < T_b] = \frac{f_-(s)^{-b} - f_+(s)^b}{f_+(s)^{-a} f_-(s)^{-b} - f_+(s)^b f_-(s)^a}, \tag{4.72}$$

$$E[s^{T_b} : T_b < T_{-a}] = \frac{f_+(s)^{-a} - f_-(s)^a}{f_+(s)^{-a} f_-(s)^{-b} - f_+(s)^b f_-(s)^a}. \tag{4.73}$$

In particular, if $p_+ < p_-$, then as special cases of (4.72) and (4.73) with $s = 1$,

$$P(T_{-a} < T_b) = \frac{(p_-/p_+)^b - 1}{(p_-/p_+)^b - (p_-/p_+)^{-a}}, \quad P(T_b < T_{-a}) = \frac{1 - (p_-/p_+)^{-a}}{(p_-/p_+)^b - (p_-/p_+)^{-a}}. \tag{4.74}$$

On the other hand, if $p_+ = p_-$, then

$$P(T_{-a} < T_b) = \frac{b}{a+b}, \quad P(T_b < T_{-a}) = \frac{a}{a+b}. \tag{4.75}$$

[a] See also Exercise 3.4.5.

Proof: (4.72) and (4.73): As in the proof of Proposition 4.5.3, we consider the martingales (4.66). This time, we take $T = T_{-a} \wedge T_b$. Then,

1) $0 \le X_-(n \wedge T) \le f_-(s)^{-S(n\wedge T)} \le f_-(s)^{-b}$, $0 \le X_+(n \wedge T) \le f_+(s)^{S(n\wedge T)} \le f_+(s)^a$.

We now note that

2) $T_{-a} \neq T_b$ a.s.

This can be seen as follows. If $T_{-a} = T_b < \infty$, then, $-a = S(T_{-a}) = S(T_b) = b$, which is impossible. Hence, $\{T_{-a} = T_b < \infty\} = \emptyset$. On the other hand, we see from (4.69) that

$$p_+ \le p_- \implies P(T_{-a} < \infty) = 1, \quad p_+ \ge p_- \implies P(T_b < \infty) = 1.$$

Thus, $P(T_{-a} = T_b = \infty) = 0$.
It follows from 2) that almost surely,

3) $$\begin{cases} X_+(n \wedge T) & = X_+(n \wedge T_{-a})\mathbf{1}\{T_{-a} < T_b\} + X_+(n \wedge T_b)\mathbf{1}\{T_b < T_{-a}\} \\ & \stackrel{n\to\infty}{\longrightarrow} f_+(s)^{-a} s^{T_{-a}} \mathbf{1}\{T_{-a} < T_b\} + f_+(s)^b s^{T_b} \mathbf{1}\{T_b < T_{-a}\}. \end{cases}$$

4) $$\begin{cases} X_-(n \wedge T) & = X_-(n \wedge T_{-a})\mathbf{1}\{T_{-a} < T_b\} + X_-(n \wedge T_b)\mathbf{1}\{T_b < T_{-a}\} \\ & \stackrel{n\to\infty}{\longrightarrow} f_-(s)^{a} s^{T_{-a}} \mathbf{1}\{T_{-a} < T_b\} + f_-(s)^{-b} s^{T_b} \mathbf{1}\{T_b < T_{-a}\}. \end{cases}$$

Now, by applying (4.67), we have

$$\begin{aligned} 1 &= f_+(s)^{-a} E[s^{T_{-a}} : T_{-a} < T_b] + f_+(s)^b E[s^{T_b} : T_b < T_{-a}], \\ 1 &= f_-(s)^{a} E[s^{T_{-a}} : T_{-a} < T_b] + f_-(s)^{-b} E[s^{T_b} : T_b < T_{-a}] \end{aligned}$$

from which we obtain (4.72) and (4.73).
(4.75): This follows easily from the above argument applied to the (much simpler) martingale S_n, instead of $X_\pm(n)$. \(^□^)/

Corollary 4.5.6 *Suppose that* $p_- \geq p_+$*. Then, the law of the r.v.*

$$Z \stackrel{\text{def}}{=} \max_{n \leq T_{-a}} S_n$$

(Note that $T_{-a} < \infty$ *a.s. by (4.69)) is given by*

$$P(Z \geq b) = \begin{cases} \frac{1-(p_-/p_+)^{-a}}{(p_-/p_+)^{b}-(p_-/p_+)^{-a}} & \text{if } p_+ < p_-, \\ a/(a+b) & \text{if } p_+ = p_-. \end{cases} \tag{4.76}$$

In particular,

$$EZ = \sum_{b=1}^{\infty} P(Z \geq b) \begin{cases} < \infty & \text{if } p_+ < p_-, \\ = \infty & \text{if } p_+ = p_-. \end{cases}$$

Proof: $P(Z \geq b) = P(T_b < T_{-a}) \stackrel{(4.74),(4.75)}{=}$ the right-hand side of (4.76). \(^□^)/

Exercise 4.5.1 Let $\chi_n = \mathbf{1}_{\{S_n=0\}}$. Then, prove the following.
(i) $\Delta|S_n| = \chi_{n-1}|\xi_n| + (1-\chi_{n-1})(\xi_n S_{n-1}/|S_{n-1}|)$ for $n \geq 1$, cf. (4.44). (ii) If $p_+ = p_-$, then $|S_n| - (1-p_0)\sum_{j=0}^{n-1}\chi_j$, $n \in \mathbb{N}$ is a martingale. Hint: Proposition 4.6.2.

Exercise 4.5.2 Prove that

$$E[T_{-a} \wedge T_b] = \begin{cases} \frac{b(p_+/p_-)^{-a}+a(p_+/p_-)^{b}-(a+b)}{(p_--p_+)\left((p_+/p_-)^{-a}-(p_+/p_-)^{b}\right)} & \text{if } p_+ < p_-, \\ = \frac{ab}{1-p_0} & \text{if } p_+ = p_-. \end{cases}$$

[Hint: For $p_+ < p_-$, use the martingale $S_n - (p_+ - p_-)n$, and for $p_+ = p_-$, use the martingale $S_n^2 - (1-p_0)n$.]

Remark By Proposition 4.5.7 below, $T_{-a} \wedge T_b$ is exponentially integrable, whenever $p_0 < 1$.

Exercise 4.5.3 (Position of the first decrease by length ℓ) Let $s \in (0,1]$, $M_n = \max_{0\leq j\leq n} S_j$, and

$$X_n = (p_- + (1-s)p_+(M_n - S_n))s^{M_n}.$$

Prove the following.
i) $E[X_{n+1}|\mathcal{F}_n] = X_n + (1-s)p_+(p_- - p_+)s^{M_n}\mathbf{1}\{M_n > S_n\}, \quad n \in \mathbb{N}$. As a consequence, $(X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ is a submartingale (resp. supermartingale) if $p_+ \leq p_-$ (resp. $p_+ \geq p_-$).
ii) If $p_+ \leq p_-$, and $a \in \mathbb{N}\backslash\{0\}$, then $T \stackrel{\text{def}}{=} \inf\{n \geq 0 \,;\, M_n - S_n = \ell\} < \infty$ a.s. and

$$Es^{M_T} \geq \frac{p_-}{p_- + (1-s)p_+\ell}.$$

In particular, if $p_+ = p_-$, then the above inequality becomes an equality, which implies that the r.v. $M_T + 1 (= S_T + \ell + 1)$ is geometrically distributed with parameter $1/(a+1)$. See Example 7.6.4 for an analogy in the case of the Brownian motion.

(⋆) Complement to section 4.5

Let $(\xi_n)_{n\geq 1}$ be i.i.d. with values in $\mathbb{Z}^d$ such that $P(\xi_1 = 0) \neq 1$. We define $(S_n)_{n\in\mathbb{N}}$ by

$$S_0 = 0, \quad S_{n+1} = S_n + \xi_{n+1}, \quad n \in \mathbb{N}.$$

For $x \in \mathbb{Z}^d$ and $A \subset \mathbb{Z}^d$, we set

$$T(x, A) = \inf\{n \geq 1 \; ; \; x + S_n \in A\}.$$

Proposition 4.5.7 *(***Exit time from a finite set***) For a finite set $A \subset \mathbb{Z}^d$, there is an $\varepsilon > 0$ such that*

$$E \exp\left(\varepsilon T(x, A^{\mathsf{c}})\right) < \infty \quad \text{for all } x \in A. \tag{4.77}$$

Proof: We first pick $z \neq 0$ such that $\alpha \stackrel{\text{def.}}{=} P(\xi_1 = z) > 0$. Since $A - A = \{x - x' \; ; \; x, x' \in A\}$ is a finite set, there exists $m \in \mathbb{N}\backslash\{0\}$ such that $mz \notin A - A$. We then set $\beta = 1 - \alpha^m < 1$. We will prove by induction that

1) $\sup_{x\in A} P(T(x, A^{\mathsf{c}}) > km) \leq \beta^k, \quad k = 1, 2, \ldots.$

We begin with $k = 1$.

$$\begin{aligned} P(T(x, A^{\mathsf{c}}) \leq m) &\geq P(x + S_m \notin A) \\ &\geq P(S_m \notin A - A) \\ &\geq P(S_m = mz) \\ &\geq P(X_1 = \ldots = X_m = z) = \alpha^m. \end{aligned}$$

This proves 1) for $k = 1$. We now suppose 1) for some k. Then,

$$P(T(x, A^{\mathsf{c}}) > (k+1)m) \leq \sum_{y\in A} P(T(x, A^{\mathsf{c}}) > km, \; x + S_{km} = y, \; \widetilde{T}(y, A^{\mathsf{c}}) > m),$$

where

$$\widetilde{T}(y, A^{\mathsf{c}}) = \inf\{n \geq 1 \; ; \; y + S_{n+km} - S_{km} \in A^{\mathsf{c}}\}.$$

Let $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_n = \sigma(\xi_1, ..., \xi_n)$ for $n \geq 1$. Then,

2) $\{T(x, A^{\mathsf{c}}) > km, \; x + S_{km} = y\} \in \mathcal{F}_{km}$,

3) $\widetilde{T}(y, A^{\mathsf{c}})$ is independent of $\mathcal{F}_{km}$,

4) $\widetilde{T}(y, A^{\mathsf{c}})$ has the same distribution as $T(y, A^{\mathsf{c}})$.

Thus, we have

$$\begin{aligned} &\sum_{y\in A} P(T(x, A^{\mathsf{c}}) > km, \; x + S_{km} = y, \; \widetilde{T}(y, A^{\mathsf{c}}) > m) \\ &= \sum_{y\in A} P(T(x, A^{\mathsf{c}}) > km, \; x + S_{km} = y) P(T(y, A^{\mathsf{c}}) > m) \quad \text{by 2),3),4)} \\ &\leq \beta \sum_{y\in A} P(T(x, A^{\mathsf{c}}) > km, \; x + S_{km} = y) \quad \text{by 1) for } k = 1, \\ &= \beta P(T(x, A^{\mathsf{c}}) > km) \\ &\leq \beta^{k+1} \quad \text{by the induction hypothesis.} \end{aligned}$$

This completes the induction and proves 1).

Now, 1) can be used to prove that there are $C > 0$ and $\varepsilon > 0$ such that

$$P(T(x, A^c) > n) \le C \exp(-\varepsilon n), \quad \text{for all } n \ge 1,$$

which proves (4.77) (cf. Exercise 1.1.3). \(^□^)/

4.6 Quadratic variation and discrete stochastic integrals

Lemma 4.6.1 *Let $X = (X_n, \mathcal{F}_n)_{n \in \mathbb{N}}$ be a predictable martingale. Then, $X_n \equiv X_0$, a.s., $\forall n \in \mathbb{N}$.*

Proof: Since X_{n+1} is $\mathcal{F}_n$-measurable, we have

$$X_n \stackrel{(4.43)}{=} E[X_{n+1} | \mathcal{F}_n] \stackrel{(4.10)}{=} X_{n+1}.$$

Thus, we arrive at the conclusion by induction. \(^□^)/

Proposition 4.6.2 *Let $X = (X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$, $Y = (Y_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ be adapted, integrable.*

a) *There exists a unique predictable, integrable process $A = (A_n)_{n\in\mathbb{N}}$ with $A_0 \equiv 0$ such that*

$$M \stackrel{\text{def}}{=} (X_n - A_n, \mathcal{F}_n)_{n\in\mathbb{N}}$$

is a martingale. Moreover, A_n for $n \geq 1$ is given by

$$A_n = \sum_{j=1}^{n} E[\Delta X_j | \mathcal{F}_{j-1}]. \tag{4.78}$$

The processes M and A are called respectively the **martingale part** *and the* **predictable part** *of X.*

b) *Suppose that $X_m Y_n \in L^1(P)$ for all $m, n \in \mathbb{N}$. Then, there exists a unique predictable, integrable process $\langle X, Y \rangle = (\langle X, Y \rangle_n)_{n\in\mathbb{N}}$ with $\langle X, Y \rangle_0 \equiv 0$ such that*

$$\widetilde{M} \stackrel{\text{def}}{=} (X_n Y_n - \langle X, Y \rangle_n, \mathcal{F}_n)_{n\in\mathbb{N}}$$

is a martingale. Moreover, $\langle X, Y \rangle_n$ for $n \geq 1$ is given by

$$\langle X, Y \rangle_n = \sum_{j=1}^{n} E[\Delta(X_j Y_j) | \mathcal{F}_{j-1}]. \tag{4.79}$$

Suppose in particular that X and Y are martingales. Then,

$$\langle X, Y \rangle_n = \sum_{j=1}^{n} E[\Delta X_j \Delta Y_j | \mathcal{F}_{j-1}], \quad n \geq 1. \tag{4.80}$$

The process $\langle X, Y \rangle$ is called the **bracket** *of X and Y. In particular, when $X = Y$, the process $\langle X \rangle \stackrel{\text{def}}{=} \langle X, X \rangle$ is called the* **quadratic variation** *of X.*

Proof: a) We first verify the uniqueness of A. If both A and A' are such processes, then, $M_n \stackrel{\text{def}}{=} X_n - A_n$ and $M'_n \stackrel{\text{def}}{=} X_n - A'_n$ are martingales and $M_n - M'_n = A'_n - A_n$. Thus, by Lemma 4.6.1, $A_n - A'_n \equiv A_0 - A'_0 = 0$ for all $n \in \mathbb{N}$.
Next, let $M_n = X_n - A_n$, where $A_0 \equiv 0$ and A_n for $n \geq 1$ is given by (4.78). Since

$$\Delta M_{n+1} = \Delta X_{n+1} - E[\Delta X_{n+1} | \mathcal{F}_n], \quad n \in \mathbb{N},$$

we have $E[\Delta M_{n+1} | \mathcal{F}_n] = 0$, a.s. Thus, M is a martingale by Lemma 4.3.4.
b) This is a special case of a) in which X_n is replaced by $X_n Y_n$. If X, Y are martingales, then,

$$\begin{aligned}
E[\Delta X_j \Delta Y_j | \mathcal{F}_{j-1}] &\stackrel{(4.44)}{=} E[X_j Y_j - X_{j-1} Y_j - X_j Y_{j-1} + X_{j-1} Y_{j-1} | \mathcal{F}_{j-1}] \\
&= E[X_j Y_j | \mathcal{F}_{j-1}] - X_{j-1} E[Y_j | \mathcal{F}_{j-1}] - Y_{j-1} E[X_j | \mathcal{F}_{j-1}] + X_{j-1} Y_{j-1} \\
&\stackrel{(4.43)}{=} E[X_j Y_j | \mathcal{F}_{j-1}] - X_{j-1} Y_{j-1} \stackrel{(4.44)}{=} E[\Delta(X_j Y_j) | \mathcal{F}_{j-1}].
\end{aligned}$$

This implies (4.80). \(^□^)/

Remark: By Proposition 4.6.2 a), any adapted, integrable process X is decomposed into a martingale M and a predictable process A. This decomposition is called the **Doob's decomposition**.

Corollary 4.6.3 *Suppose that* $X = (X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ *and* $Y = (Y_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ *are martingales such that* $X_mY_n \in L^1(P)$ *for all* $m, n \in \mathbb{N}$. *Then,*

$$E[X_mY_n] = E[X_0Y_0] + E\langle\, X, Y\,\rangle_{m\wedge n}, \quad m, n \in \mathbb{N}. \tag{4.81}$$

Proof: Suppose for example that $m \le n$. Then,

1) $E[X_mY_n] = E[E[X_mY_n|\mathcal{F}_m]] \stackrel{(4.19)}{=} E[X_mE[Y_n|\mathcal{F}_m]] \stackrel{(4.43)}{=} E[X_mY_m]$.

On the other hand, since $\widetilde{M}_n = X_nY_n - \langle\, X, Y\,\rangle_n$ is a martingale, we have

2) $E[X_mY_m] - E\langle\, X, Y\,\rangle_m = E\widetilde{M}_m = E\widetilde{M}_0 = E[X_0Y_0]$

By 1) and 2),

$$E[X_mY_n] = E[X_0Y_0] + E\langle\, X, Y\,\rangle_m.$$

\(^□^)/

The following special case of Proposition 4.6.2 is well worth being stated as

Corollary 4.6.4 *Referring to Proposition 4.6.2, suppose in particular that*

$$X_0, \Delta X_1, \Delta X_2, \ldots \text{ are independent and } \mathcal{F}_n = \sigma(X_0, \ldots, X_n),\ n \in \mathbb{N}. \tag{4.82}$$

Then, the following hold true.

a)

$$A_n = \sum_{j=1}^{n} m_j, \quad n \in \mathbb{N}, \quad \text{where } m_j = E[\Delta X_j]. \tag{4.83}$$

As a consequence, $\left(X_n - \sum_{j=1}^n m_j,\ \mathcal{F}_n\right)_{n\in\mathbb{N}}$ *is a martingale.*

b) *Suppose that* $X_n \in L^2(P)$ *for all* $n \in \mathbb{N}$, *and that* $m_n = 0$, $n \ge 1$. *Then,*

$$\langle\, X\,\rangle_n = \sum_{j=1}^{n} v_j, \quad n \in \mathbb{N}, \quad \text{where } v_j = E[(\Delta X_j)^2]. \tag{4.84}$$

As a consequence, $\left(X_n^2 - \sum_{j=1}^n v_j,\ \mathcal{F}_n\right)_{n\in\mathbb{N}}$ *is a martingale.*

Proof: a) ΔX_j is independent of $\mathcal{F}_{j-1}$ for all $j \ge 1$. Therefore,

$$E[\Delta X_j|\mathcal{F}_{j-1}] \stackrel{(4.12)}{=} E[\Delta X_j] = m_j.$$

This implies (4.83).
b) X is a martingale by a). Moreover,

$$E[(\Delta X_j)^2|\mathcal{F}_{j-1}] \stackrel{(4.12)}{=} E[(\Delta X_j)^2] = v_j.$$

Thus, we see (4.84) from (4.80). \\(^□^)/

Proposition 4.6.5 *Suppose that $X = (X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ is a martingale such that $X_n \in L^2(P)$ for all $n \in \mathbb{N}$. Then, the following are equivalent.*

a) $E\langle X \rangle_\infty < \infty$.

b) *X_n converges to a r.v X_∞ in L^2 as $n \to \infty$.*

Moreover, these imply that

$$E[X_\infty^2] = E[X_0^2] + E\langle X \rangle_\infty. \tag{4.85}$$

Proof: a) $\Rightarrow$ b): It is enough to prove that X_n is a Cauchy sequence in L^2. Let $m \le n$. Then,

1) $E[X_m X_n] \stackrel{\text{Exercise 4.3.1}}{=} E[X_m^2] \stackrel{(4.81)}{=} E[X_0^2] + E\langle X \rangle_m$.

Thus,

$$\begin{aligned} E[|X_n - X_m|^2] &= E[X_n^2] + E[X_m^2] - 2E[X_m X_n] \\ &\stackrel{1)}{=} E\langle X \rangle_n - E\langle X \rangle_m \stackrel{m,n\to\infty}{\longrightarrow} 0. \end{aligned}$$

a) $\Leftarrow$ b): Since $\langle X \rangle_n$ is nondecreasing in n, we have by monotone convergence theorem that

$$E\langle X \rangle_\infty = \lim_{n\to\infty} E\langle X \rangle_n = \lim_{n\to\infty} E[X_n^2] - E[X_0^2] = E[X_\infty^2] - E[X_0^2]$$

\\(^□^)/

Proposition 4.6.6 *Let $X = (X_n)_{n\in\mathbb{N}}$, $Y = (Y_n)_{n\in\mathbb{N}}$ be martingales, $H = (H_n)_{n\ge1}$, $K = (K_n)_{n\ge1}$ be predictable. Suppose that $H_n \in L^\infty(P)$, $K_n \in L^\infty(P)$, and $X_m Y_n \in L^1(P)$ for all $m, n \in \mathbb{N}\backslash\{0\}$. Then, referring to (4.59),*

$$\langle H \cdot X, K \cdot Y \rangle_n = \sum_{j=1}^n H_j K_j \Delta\langle X, Y \rangle_j \quad n \in \mathbb{N}\backslash\{0\}. \tag{4.86}$$

In particular,

$$E[(H \cdot X)_m (K \cdot Y)_n] = \sum_{j=1}^{m\wedge n} E[H_j K_j \Delta X_j \Delta Y_j], \quad m, n \in \mathbb{N}\backslash\{0\}. \tag{4.87}$$

Proof: For $j \ge 1$,

1) $$\begin{cases} \Delta\langle H \cdot X, K \cdot Y \rangle_j & \stackrel{(4.80)}{=} E[\Delta(H\cdot X)_j \Delta(K \cdot Y)_j|\mathcal{F}_{j-1}] \\ & \stackrel{(4.59)}{=} H_j K_j E[\Delta X_j \Delta Y_j|\mathcal{F}_{j-1}] \stackrel{(4.80)}{=} H_j K_j \Delta\langle X, Y \rangle_j \end{cases}$$

By taking summation over $j = 1, ..., n$, this implies (4.86). The equality (4.87) is obtained as follows

$$E[(H\cdot X)_m(K\cdot Y)_n] \stackrel{(4.81)}{=} E\langle H\cdot X, K\cdot Y\rangle_{m\wedge n} \stackrel{(4.86),\,1)}{=} \sum_{j=1}^{m\wedge n} E[H_jK_j\Delta X_j\Delta Y_j].$$

\(^□^)/

Lemma 4.6.7 *Let $X = (X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ be adapted, integrable, and T be a stopping time w.r.t $(\mathcal{F}_n)_{n\in\mathbb{N}}$, such that $ET < \infty$ and that*

$$\sup_{n\ge 1} E[|\Delta X_n||\mathcal{F}_{n-1}] \le C_1 \text{ for a constant } C_1 \in [0,\infty).$$

Then, $X_T \in L^1(P)$ and

$$E|X_T - X_{n\wedge T}| \stackrel{n\to\infty}{\longrightarrow} 0.$$

Proof: Note that

1) $\{T \ge n\} = \{T \le n-1\}^{\mathsf{c}} \in \mathcal{F}_{n-1}$

Thus,

$$\begin{aligned} E|X_T - X_{n\wedge T}| &= E[|X_T - X_n| : T > n] \le E[\sum_{j=n+1}^{T} |\Delta X_j| : T > n] \\ &= E[\sum_{j=n+1}^{\infty} |\Delta X_j| \mathbf{1}_{\{T\ge j\}}] = \sum_{j=n+1}^{\infty} E[|\Delta X_j| : T \ge j] \\ &\stackrel{1)}{=} \sum_{j=n+1}^{\infty} E[E[|\Delta X_j||\mathcal{F}_{j-1}] : T \ge j] \le C_1 \sum_{j=n+1}^{\infty} P(T\ge j) \stackrel{n\to\infty}{\longrightarrow} 0. \end{aligned}$$

The above estimate shows also that $X_T - X_{n\wedge T} \in L^1(P)$ for all $n \in \mathbb{N}$. By taking $n = 0$, we see that $X_T \in L^1(P)$. \(^□^)/

Example 4.6.8 Let $X = (X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ be adapted, integrable, and T be a stopping time w.r.t $(\mathcal{F}_n)_{n\in\mathbb{N}}$, such that $ET < \infty$.

a) Suppose that

$$\sup_{n\ge 1} E[|\Delta X_n||\mathcal{F}_{n-1}] \le C_1 \text{ for a constant } C_1 \in [0,\infty).$$

Let $(A_n)_{\in\mathbb{N}}$ be defined by (4.78). Then, $X_T, A_T \in L^1(P)$ and

$$EX_T = EX_0 + EA_T. \tag{4.88}$$

b) Suppose in addition that X is a martingale and that

$$\sup_{n\ge 1} E[|\Delta X_n|^2|\mathcal{F}_{n-1}] \le C_2 \text{ for a constant } C_2 \in [0,\infty).$$

Let $\langle X\rangle_n$, $n \in \mathbb{N}$ be given by (4.80). Then, $X_T^2, \langle X\rangle_T \in L^1(P)$ and

$$E[X_T^2] = E[X_0^2] + E\langle X\rangle_T. \tag{4.89}$$

Proof: a) Let $M_n \stackrel{\text{def}}{=} X_n - A_n$. Then, $(M_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ is a martingale (Proposition 4.6.2), and hence $(M_{n\wedge T}, \mathcal{F}_n)_{n\in\mathbb{N}}$ is a martingale (Corollary 4.4.4). This implies that

1) $EM_{n\wedge T} = EM_0, \quad \forall n \in \mathbb{N}$.

On the other hand, since

$$|\Delta A_n| \le E[|\Delta X_n||\mathcal{F}_{n-1}], \quad \Delta M_n = \Delta X_n - \Delta A_n,$$

we have

$$\begin{aligned}\sup_{n\ge 1} E[|\Delta A_n||\mathcal{F}_{n-1}] &= \sup_{n\ge 1} |\Delta A_n| \le \sup_{n\ge 1} E[|\Delta X_n||\mathcal{F}_{n-1}] \le C_1,\\ \sup_{n\ge 1} E[|\Delta M_n||\mathcal{F}_{n-1}] &\le 2C_1.\end{aligned}$$

Thus, by Lemma 4.6.7,

$$X_T, A_T, M_T \in L^1(P), \quad E|M_T - M_{n\wedge T}| \stackrel{n\to\infty}{\longrightarrow} 0.$$

In particular, by letting $n \to \infty$ in 1), we have $EM_T = EM_0$, and therefore,

$$EX_0 = EM_0 = EM_T = EX_T - EA_T,$$

which proves (4.88).

b) We will apply Proposition 4.6.5 to the stopped process $X^T \stackrel{\text{def}}{=} (X_{n\wedge T})_{n\in\mathbb{N}}$, which is a martingale (Corollary 4.4.4). We have

$$\langle X^T \rangle_\infty = \langle X \rangle_T = \sum_{n=1}^{T} E[|\Delta X_n|^2|\mathcal{F}_{n-1}] \le C_2 T \in L^1(P).$$

Therefore, $X_T = X^T_\infty \in L^2(P)$ and

$$E[(X^T_\infty)^2] \stackrel{(4.85)}{=} E[X_0^2] + E\langle X^T \rangle_\infty.$$

Since $X_T = X^T_\infty$, and $\langle X^T \rangle_\infty = \langle X \rangle_T$ by Exercise 4.6.1, we obtain (4.89). \\(^□^)/

Remarks: Referring to Example 4.6.8 a), suppose that (4.82) and $E[\Delta X_n] = m$, $n \ge 1$. Then, $A_n = mn$, $n \in \mathbb{N}$ (Corollary 4.6.4). Therefore, the equality (4.88) takes the following form

$$EX_T = EX_0 + mET \quad (\textbf{Wald's first equation}). \tag{4.90}$$

Let us now assume $m \neq 0$, but let not assume apriori that $ET < \infty$. Then, we have that

$$ET < \infty \iff \sup_{n\in\mathbb{N}} |EX_{n\wedge T}| < \infty. \tag{4.91}$$

In fact, by (4.90) applied to a bounded stopping time $n \wedge T$, we have that

$$EX_{n\wedge T} = EX_0 + mE[n \wedge T],$$

from which (4.91) follows immediately.

2) Referring to Example 4.6.8 b), suppose that (4.82), $E[\Delta X_n] = 0$, $E[(\Delta X_n)^2] = v$, $n \ge 1$. Then, $\langle X \rangle_n = vn$, $n \in \mathbb{N}$ (Corollary 4.6.4). Therefore, the equality (4.89) takes the following form

$$EX_T^2 = EX_0^2 + vET \quad (\textbf{Wald's second equation}). \tag{4.92}$$

\\(^□^)/

Exercise 4.6.1 Let $X = (X_n)_{n\in\mathbb{N}}$ be a process, $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be a filtration, and T be a stopping time. Let also $X^T \stackrel{\text{def}}{=} (X_{n\wedge T})_{n\in\mathbb{N}}$ be the stopped process. Prove the following. i) If X is predictable, then, so is X^T.
ii) Suppose that M and A are respectively the martingale part and the predictable part of an adapted, integrable process X. Then, M^T and A^T are respectively the martingale part and the predictable part of X^T.
iii) Suppose that X is a martingale such that $X_n \in L^2(P)$ for all $n \in \mathbb{N}$. Then $\langle X^T \rangle = \langle X \rangle^T$.

4.7 ($\star$) Structure of L^1-bounded martingales I

We have already seen the analogy between martingales and harmonic fnctions on the open unit disc $D \subset \mathbb{C}$. For a harmonic function u on D, it is known that the following conditions are equivalent, cf. [Dur84, p.160, (6)].

a) u is a difference of two nonnegative harmonic functions.

b) There exists a Borel signed measure μ on $[-\pi, \pi]$ such that

$$u(z) = \int_{-\pi}^{\pi} h(z, e^{\mathbf{i}\theta}) d\mu(\theta) \quad \text{for all } z \in D, \text{ where } h(z, w) = \tfrac{|w|^2 - |z|^2}{|w-z|^2}.$$

c) $\displaystyle\sup_{0<r<1} \int_{-\pi}^{\pi} |u(re^{\mathbf{i}\theta})| d\theta < \infty.$

Here is an analogue for martingales.

Proposition 4.7.1 *Suppose that the set $\mathbb{T}$ is unbounded from above, and that $X = (X_t, \mathcal{F}_t^X)_{t\in\mathbb{T}}$ is a martingale. Then, the following conditions are equivalent.*

a) *X is a difference of two nonnegative $(\mathcal{F}_t^X)$-martingales.*

b1) *There exists a signed measure Q on $(\Omega, \mathcal{F}_\infty^X)$ such that for all $t \in \mathbb{T}$, $|Q|_t \ll P_t$ and $dQ_t/dP_t = X_t$.*

b2) *There exists a signed measure Q on $(\Omega, \mathcal{F}_\infty^X)$ such that for all $t \in \mathbb{T}$, $Q_t \ll P$ and $dQ_t/dP_t = X_t$.*

c) $\displaystyle\sup_{t\in\mathbb{T}} E|X_t| < \infty.$

I am grateful to Francis Comets for bringing the following lemma into my interest.

Lemma 4.7.2 *Suppose that the set $\mathbb{T} \subset \mathbb{R}$ is unbounded from above and that $X = (X_t, \mathcal{F}_t)_{t\in\mathbb{T}}$ is a submartingale such that $\sup_{t\in\mathbb{T}} E[X_t^+] < \infty$.*

a) *There exists a martingale $Y = (Y_t, \mathcal{F}_t)_{t\in\mathbb{T}}$ such that $X_t^+ \le Y_t$ for all $t \in \mathbb{T}$.*

b) (**Krickeberg decomposition**) *There exists a nonnegative supermartingale $Z = (Z_t, \mathcal{F}_t)_{t\in\mathbb{T}}$ such that $X_t = Y_t - Z_t$ for all $t \in \mathbb{T}$. In particular, Z is a martingale if X is a martingale.*

Proof: a) We start by observing that

1) $t, u, v \in \mathbb{T}$, $t \le u < v \implies E[X_u^+|\mathcal{F}_t] \le E[X_v^+|\mathcal{F}_t]$, a.s.

Indeed, $(X_t^+, \mathcal{F}_t)_{t\in\mathbb{T}}$ is a submartingale by Lemma 4.3.3. Thus,

$$X_u^+ \le E[X_v^+|\mathcal{F}_u], \text{ a.s.}$$

We obtain 1) by taking the conditional expextations of the both hands sides of the above identity.
By 1), the limit $Y_t \stackrel{\text{def}}{=} \lim_{u\to\infty} E[X_u^+|\mathcal{F}_t] \in [0,\infty]$ exists and $X_t^+ \le Y_t$ for all $t \in \mathbb{T}$. We verify that

2) $Y = (Y_t, \mathcal{F}_t)_{t\in\mathbb{T}}$ is a martingale.

First, $Y_t \in L^1(P)$ for all $t \in \mathbb{T}$, since by 1) and the monotone convergence theorem,

$$EY_t = \lim_{u\to\infty} E[E[X_u^+|\mathcal{F}_t]] = \lim_{u\to\infty} E[X_u^+] < \infty.$$

Next, if $s, t \in \mathbb{T}$ and $s < t$, then, by the monotone convergence theorem for conditional expectations,

$$E[Y_t|\mathcal{F}_s] = \lim_{u\to\infty} E[E[X_u^+|\mathcal{F}_t]|\mathcal{F}_s] = \lim_{u\to\infty} E[X_u^+|\mathcal{F}_s] = Y_s, \text{ a.s.}$$

b) $Z_t \stackrel{\text{def}}{=} Y_t - X_t$, $t \in \mathbb{T}$ is a nonnegative supermartingale. In particular, Z is a martingale if X is a martingale. \\(^□^)/

Let $X = (X_t)_{t\in\mathbb{T}}$ be a process. We write $\mathcal{F}_t^X = \sigma(X_s \ ; \ s \in \mathbb{T}\cap[0,t])$ $t \in \mathbb{T}$, and $\mathcal{F}_\infty^X = \sigma(\mathcal{F}_t^X \ ; \ t \in \mathbb{T})$. For a signed measure Q on $(\Omega, \mathcal{F}_\infty^X)$, let $|Q|$ be its variation, $Q^\pm = (|Q| \pm Q)/2$ (Jordan decomposition) and $Q_t = Q|_{\mathcal{F}_t^X}$.

Lemma 4.7.3 *Let $Y = (Y_t, \mathcal{F}_t^X)_{t\in\mathbb{T}}$ be a nonnegative, mean-one martingale. Then, there exists a unique probability measure P^Y on $(\Omega, \mathcal{F}_\infty^X)$ such that*

$$P^Y(A) = E[Y_t : A] \text{ for all } t \in \mathbb{T} \text{ and } A \in \mathcal{F}_t^X.$$

Proof: For each $t \in \mathbb{T}$, let $\widetilde{P}_t(A) = E[Y_t : A]$ for $A \in \mathcal{F}_t^X$. Then, the family of measures $(\mathcal{F}_t^X, \widetilde{P}_t)$, $t \in \mathbb{T}$ are consistent in the sense that $\widetilde{P}_t|_{\mathcal{F}_s^X} = \widetilde{P}_s$ if $s, t \in \mathbb{T}$, $s < t$. Thus, by Kolmogorov's extension theorem, there exists a unique probability measure P^Y on $(\Omega, \mathcal{F}_\infty^X)$ such that $P^Y|_{\mathcal{F}_t^X} = \widetilde{P}_t$ for all $t \in \mathbb{T}$. \\(^□^)/

Proof of Proposition 4.7.1: a) $\Rightarrow$ b1): Suppose that X is a difference of two nonnegative $(\mathcal{F}_t^X)$-martingales Y_t and Z_t. Then, by Lemma 4.7.3, there exist finite measures Q^Y, Q^Z on $(\Omega, \mathcal{F}_\infty^X)$ such that for all $t \in \mathbb{T}$, $Q_t^Y \ll P_t$, $Q_t^Z \ll P_t$, $Y_t = dQ_t^Y/dP_t$, $Z_t = dQ_t^Z/dP_t$. Set $Q = Q^Y - Q^Z$. Then, $|Q| \le Q^Y + Q^Z$ and hence $|Q|_t \le (Q^Y + Q^Z)_t \ll P_t$. Moreover,

$$dQ_t/dP_t = d(Q_t^Y - dQ_t^Z)/dP_t = dQ_t^Y/dP_t - dQ_t^Z/dP_t = Y_t - Z_t = X_t.$$

b1) $\Rightarrow$ b2): This follows from the inequality $|Q_t| \le |Q|_t$.
b2) $\Rightarrow$ c): $E|X_t| = |Q_t|(\Omega) \le |Q|(\Omega) < \infty$.
c) $\Rightarrow$ a): This follows from Lemma 4.7.2. \\(^□^)/

5 Convergence Theorems for Martingales

5.1 Almost sure convergence

At the beginning of section 4.3, we have seen the analogy between martingales and harmonic fnctions on the unit disc $D \subset \mathbb{C}$. Suppose that a harmonic function u on D satisfies

$$\sup_{0<r<1} \int_{-\pi}^{\pi} |u(re^{\mathbf{i}\theta})| d\theta < \infty.$$

Then, it is known that there exists $f \in L^1([-\pi, \pi])$ such that

$$u(re^{\mathbf{i}\theta}) \stackrel{r \nearrow 1}{\longrightarrow} f(e^{\mathbf{i}\theta}) \text{ for almost all } \theta \in [-\pi, \pi].$$

cf. [Rud87, p.244,11.24].

The purpose of this subsection is to present the following analogue for the martingale.

Theorem 5.1.1 (**Martingale convergence theorem**) *Suppose that $X = (X_t, \mathcal{F}_t)_{t \in \mathbb{T}}$ is a submartingale or a supermartingale such that*

$$\textit{either } \mathbb{T} = \mathbb{N}, \textit{ or } \mathbb{T} = [0, \infty) \textit{ and } (X_t)_{t \geq 0} \textit{ is right-continuous,} \tag{5.1}$$

and that

$$\sup_{t \in \mathbb{T}} \|X_t\|_1 < \infty. \tag{5.2}$$

Then, there exists $X_\infty \in L^1(P)$ such that

$$X_t \stackrel{t \to \infty}{\longrightarrow} X_\infty \textit{ a.s.}$$

Remarks: **i)** Suppose that X in Theorem 5.1.1 is a martingale. Then, by the assumption (5.2), there exists a signed measure Q on $(\Omega, \mathcal{F}_\infty)$ such that $X_n = dQ_n/dP_n$, $n \in \mathbb{N}$, where $P_n = P|_{\mathcal{F}_n}$ and $Q_n = Q|_{\mathcal{F}_n}$, cf. Proposition 4.7.1. Moreover, the signed measure Q and the a.s. limit X_∞ in Theorem 5.1.1 are related as $dQ = X_\infty dP + \mathbf{1}_N dQ$, where $N \in \mathcal{F}$ and $P(N) = 0$, cf. Proposition 5.6.1 below. **ii)** Referring to Theorem 5.1.1, the condition (5.2) is not necessary for the conclusion of the theorem. An counterexample is provided as follows. Let S_n be the random walk considered in section 4.5 with $p_+ = p_- > 0$. Then, $X_n = S(n \wedge T_{-1})^2$ is a submartingale and $X_n \stackrel{n \to \infty}{\longrightarrow} S(T_{-1})^2 = 1$, a.s. However, since $S_n^2 - (1 - p_0)n$ is a martingale, so is $X_n - (1 - p_0)(n \wedge T_{-1})$ (Corollary 4.4.4). Hence,

$$EX_n = (1 - p_0)E[n \wedge T_{-1}] \stackrel{n \to \infty}{\longrightarrow} (1 - p_0)ET_{-1} \stackrel{(4.71)}{=} \infty.$$

We postpone the proof of Theorem 5.1.1 for a moment. As an immediate consequence of Theorem 5.1.1, we have

Corollary 5.1.2 *Suppose that $X = (X_t, \mathcal{F}_t)_{t\in\mathbb{T}}$ is a nonnegative supermartingale under assumption (5.1). Then, there exists $X_\infty \in L^1(P)$ such that $X_t \stackrel{t\to\infty}{\longrightarrow} X_\infty$ a.s. Moreover,*

a) *$X_t \geq E[X_\infty|\mathcal{F}_t]$ a.s. for all $t \in \mathbb{T}$.*

b) *The following conditions are equivalent.* **b1)** *X is a uniformly integrable martingale.* **b2)** *$EX_\infty = EX_0$.* **b3)** *$X_t = E[X_\infty|\mathcal{F}_t]$ a.s. for all $t \in \mathbb{T}$.*

Proof: Since $0 \leq EX_t \leq EX_0$ for all $t \in \mathbb{T}$, assmuption (5.2) is satisfied. Thus, by Theorem 5.1.1, there exists $X_\infty \in L^1(P)$ such that $X_t \stackrel{t\to\infty}{\longrightarrow} X_\infty$ a.s.
a) Since X_t is a supermartingale, $E[X_u|\mathcal{F}_t] \leq X_t$ for all $t, u \in \mathbb{T}$ with $t < u$. Hence by letting $u \to \infty$ and applying Fatou's lemma, we obtain the desired inequality.
b1) $\Rightarrow$ b2): Since X is a martingale, $EX_t = EX_0$ for all $t \in \mathbb{T}$. On the other hand, X is uniformly integrable and $X_t \stackrel{t\to\infty}{\longrightarrow} X_\infty$ a.s. Therefore, by Proposition 2.5.5, $X_t \stackrel{t\to\infty}{\longrightarrow} X_\infty$ in $L^1(P)$. THerefore, $EX_\infty = \lim_{t\to\infty} EX_t = EX_0$.
b2) $\Rightarrow$ b3): Suppose that $EX_\infty = EX_0$ and let $Y_t \stackrel{\text{def}}{=} E[X_\infty|\mathcal{F}_t]$. Then, for all $t \in \mathbb{T}$, $X_t \geq Y_t$ a.s. by a) and hence

$$EX_t \geq EY_t = EX_\infty = EX_0 \geq EX_t.$$

Thus, $X_t \geq Y_t$ a.s. and $EX_t = EY_t$, which, implies that $X_t = Y_t$.
b3) $\Rightarrow$ b1): This follows from Lemma 4.1.13. \(^□^)/

The following example is a simple application of Corollary 5.1.2. It shows also that the convergence of X_n in Theorem 5.1.1 and Corollary 5.1.2 does not necessarily take place in $L^1(P)$.

Example 5.1.3 Let $X_n = \prod_{j=0}^n \xi_j$, $n \in \mathbb{N}$, where $\xi_n \geq 0$, $n \in \mathbb{N}$ are independent r.v.'s such that $E\xi_n \leq 1$ for all $n \in \mathbb{N}$ and that $\displaystyle\prod_{j=0}^n E[\xi_j^\delta] \stackrel{n\to\infty}{\longrightarrow} 0$ for some $\delta \in (0, 1)$. Then,

a) $X_n \stackrel{n\to\infty}{\longrightarrow} 0$ a.s.

b) Suppose in particular that $E\xi_n = 1$ for all $n \in \mathbb{N}$. Then, X_n does not converge in $L^1(P)$.

Proof: a) X_n, $n \in \mathbb{N}\backslash\{0\}$ is a supermartingale by Example 4.3.6. Since $X_n \geq 0$, we see from Corollary 5.1.2 that there exists $X_\infty \in L^1(P)$ such that $X_n \stackrel{n\to\infty}{\longrightarrow} X_\infty$ a.s. On the other hand,

$$E[X_\infty^\delta] \stackrel{\text{Fatou}}{\leq} \varliminf_{n\to\infty} E[X_n^\delta] = \varliminf_{n\to\infty} \prod_{j=0}^n E[\xi_j^\delta] = 0.$$

Hence $X_\infty = 0$ a.s.
b) $X_n, n \in \mathbb{N}\backslash\{0\}$ is a martingale by Example 4.3.6. Suppose that $X_n \stackrel{n\to\infty}{\longrightarrow} Z$ in $L^1(P)$ for some $Z \in L^1(P)$. Then, $EZ = \lim_{n\to\infty} EX_n = 1$. On the other hand, there exists a subsequence $X_{n(k)}$ such that $X_{n(k)} \stackrel{k\to\infty}{\longrightarrow} Z$ a.s. This implies via a) that $Z = 0$ a.s., which is a contradiction.
\(^□^)/

Here is another example in which the convergence of X_n in Theorem 5.1.1 and Corollary 5.1.2 does not take place in $L^1(P)$.

Example 5.1.4 Let S_n, $n \in \mathbb{N}$ from section 4.5 with $p_+ = p_-$. Then, for $a \in \mathbb{N}\backslash\{0\}$, $X_n \stackrel{\rm def}{=} a + S(n \wedge T_{-a}) \geq 0$, $n \in \mathbb{N}$ is a martingale by Corollary 4.4.4. Since $T_{-a} < \infty$ a.s. by (4.69), we see that $X_n \stackrel{n\to\infty}{\longrightarrow} a + S(T_{-a}) = 0$ a.s. But the convergence does not take place in $L^1(P)$. Indeed, since $a + S(n \wedge T_{-a})$ is a martingale (Corollary 4.4.4),

$$EX_n = EX_0 = a > 0.$$

We now turn to the proof of Theorem 5.1.1. For a moment, we consider the case of $\mathbb{T} = \mathbb{N}$. Suppose that $X = (X_n)_{n\in\mathbb{N}}$ is a process. For $-\infty < a < b < \infty$ and $n \in \mathbb{N}$, we would like to formulate the number of upcrossing from a to b in the sequence $X_0, X_1, ..., X_n$. Let $T_0 \equiv 0$, and for $k \geq 1$, we set

$$\begin{array}{rcl} S_k &=& \inf\{n \geq T_{k-1}\ ;\ X_n \leq a\}, \\ T_k &=& \inf\{n \geq S_k\ ;\ X_n \geq b\}. \end{array}$$

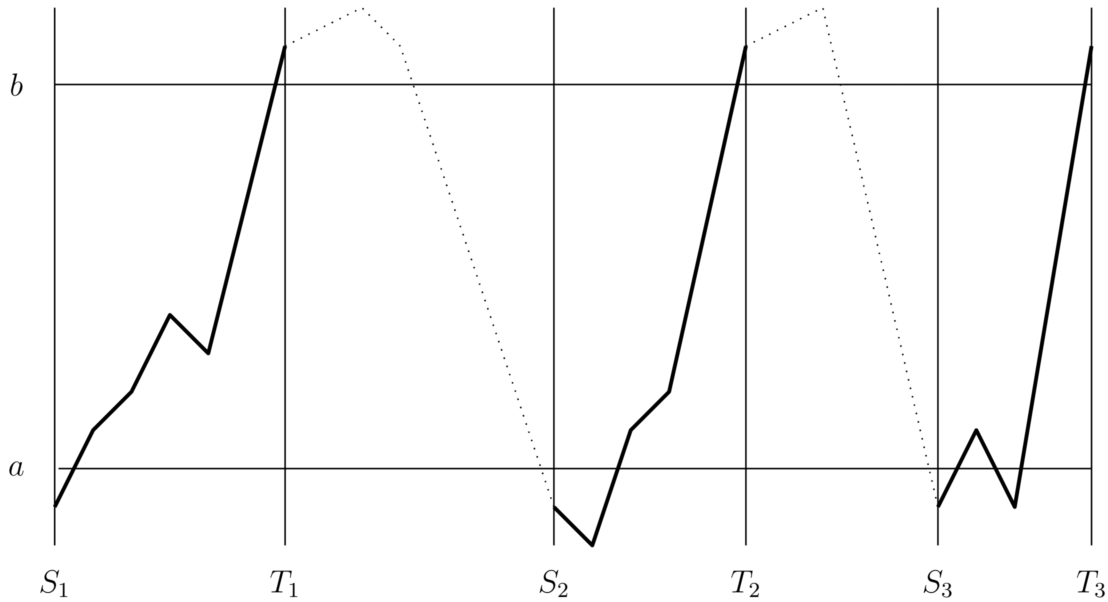
Then,

$$S_1 \leq T_1 \leq S_2 \leq T_2 \leq ...$$

If $T_k < \infty$, then the k-th upcrossing from a to b in the sequence $(X_n)_{n\in\mathbb{N}}$ starts at time S_k and is completed at time T_k. For $n \in \mathbb{N}$,

$$U_n \stackrel{\rm def}{=} \sup\{k \in \mathbb{N}\ ;\ T_k \leq n\},$$

which represents the number of completed upcrossing from a to b in the sequence $X_0, X_1, ..., X_n$. Noting that U_n is nondecreasing, we set $U_\infty = \lim_{n\to\infty} U_n \in [0, \infty]$.



Lemma 5.1.5 *Suppose that $U_\infty < \infty$ a.s. for any $-\infty < a < b < \infty$. Then:*

a) *The limit $X_\infty = \lim_{n\to\infty} X_n \in [-\infty, \infty]$ exists a.s.*

b) *Suppose in addition that (5.2) is satisfied. Then, $X_\infty \in L^1(P)$ and hence that $|X_\infty| < \infty$ a.s.*

Proof: a) It follows from the assumption that

1) $P\left(\underline{\lim}_{n\to\infty} X_n < a < b < \overline{\lim}_{n\to\infty} X_n\right) = 0$ for any $-\infty < a < b < \infty$.

On the other hand,

2) $$\left\{\underline{\lim}_{n\to\infty} X_n < \overline{\lim}_{n\to\infty} X_n\right\} = \bigcup_{\substack{a,b\in\mathbb{Q}\\ a<b}} \left\{\underline{\lim}_{n\to\infty} X_n < a < b < \overline{\lim}_{n\to\infty} X_n\right\},$$

We see from 1) and 2) that

$$\underline{\lim}_{n\to\infty} X_n = \overline{\lim}_{n\to\infty} X_n \text{ a.s.}$$

Hence the limit $X_\infty = \lim_{n\to\infty} X_n \in [-\infty, \infty]$ exists a.s.

b)

$$E|X_\infty| \stackrel{\text{Fatou}}{\leq} \underline{\lim}_{n\to\infty} E|X_n| \stackrel{(5.2)}{<} \infty.$$

Therefore, $X_\infty \in L^1(P)$ and hence that $|X_\infty| < \infty$ a.s. \(^□^)/

Lemma 5.1.6 **(The upcrossing inequality)** *If $X = (X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ is a submartingale, then,*

$$(b-a)EU_n \leq E[X_n \vee a] - E[X_0 \vee a].$$

Before going through the proof of Lemma 5.1.6, let us use the lemma to persent

Proof of Theorem 5.1.1 for $\mathbb{T} = \mathbb{N}$. By symmetry, we may focus on the case of submartingale. We first prove that $U_\infty < \infty$ a.s., which implies Theorem 5.1.1 by Lemma 5.1.5. Let $a, b \in \mathbb{R}$, $a < b$. We see from the monotone convergence theorem and Lemma 5.1.6 that

$$\begin{aligned}(b-a)EU_\infty &\stackrel{\text{MCT}}{=} (b-a)\lim_{n\to\infty} EU_n \\ &\stackrel{\text{Lemma 5.1.6}}{\leq} \sup_{n\in\mathbb{N}} E[X_n \vee a] - E[X_0 \vee a] \stackrel{(5.2)}{<} \infty.\end{aligned}$$

Therefore $U_\infty < \infty$ a.s.

\(^□^)/

Define $Y = (Y)_{n\in\mathbb{N}}$ by $Y_n = X_n \vee a$. Since $Y_n = X_n$ if $X_n \geq a$, S_k, T_k $(k \geq 1)$ are, and hence U_n is unchanged if we replace X by Y. We set

$$H_n = \begin{cases} 0 & \text{if } T_{k-1} < n \leq S_k \text{ for some } k \geq 1, \\ 1 & \text{if } S_k < n \leq T_k \text{ for some } k \geq 1. \end{cases} \tag{5.3}$$

We define $H \cdot Y$ by

$$(H \cdot Y)_n = \sum_{j=1}^{n} H_j (Y_j - Y_{j-1}).$$

We start by proving the following lemma [19]

Lemma 5.1.7 *$(b-a)U_n \leq (H \cdot Y)_n$ for $n \in \mathbb{N}$*

Proof: Note that

[19]The process X need not to be a submartingale or supermartingale for Lemma 5.1.7 to be true.

1) $T_k < \infty \implies Y(S_k) = a < b \le Y(T_k)$,

and that

2) $S_k < \infty \implies Y(S_k) = a \le Y(n)$ for all $n \in \mathbb{N}$.

(The inequality 2) is the reason for which we consider Y, instead of X.)

Now, let $U_n = \ell$, so that $T_\ell \le n < T_{\ell+1}$. Then, we will show that

3) $(H \cdot Y)(T_\ell) \ge (b-a)U_n$,

4) $(H \cdot Y)_n \ge (H \cdot Y)(T_\ell)$.

from which the lemma follows.

Indeed, 3) follows from the definition of H as follows.

$$\begin{aligned}(H \cdot Y)(T_\ell) &= \sum_{k=1}^{\ell}\left(\sum_{T_{k-1}<j\le S_k} + \sum_{S_k<j\le T_k}\right) H_j(Y_j - Y_{j-1}) \\ &\overset{(5.3)}{=} \sum_{k=1}^{\ell}\sum_{S_k<j\le T_k}(Y_j - Y_{j-1}) \\ &= \sum_{k=1}^{\ell}(Y(T_k) - Y(S_k)) \overset{1)}{\ge} (b-a)\ell = (b-a)U_n.\end{aligned}$$

Let us next show 4). Noting that $T_\ell < S_{\ell+1} \le T_{\ell+1}$, we consider the following two cases separately.

• Case 1: $T_\ell \le n \le S_{\ell+1}$. Since $H_j = 0$ for $T_\ell < j \le S_{\ell+1}$,

$$(H \cdot Y)_n - (H \cdot Y)(T_\ell) = \sum_{T_\ell<j\le n} H_j(Y_j - Y_{j-1}) \overset{(5.3)}{=} 0.$$

• Case 2:$S_{\ell+1} < n < T_{\ell+1}$. Then,

$$\begin{aligned}(H \cdot Y)_n - (H \cdot Y)(T_\ell) &= \left(\sum_{T_\ell<j\le S_{\ell+1}} + \sum_{S_{\ell+1}<j\le n}\right) H_j(Y_j - Y_{j-1}) \\ &\overset{(5.3)}{=} \sum_{S_{\ell+1}<j\le n}(Y_j - Y_{j-1}) = Y(n) - Y(S_{\ell+1}) \overset{3)}{\ge} 0.\end{aligned}$$

Proof of Lemma 5.1.6: We show that

1) $E(H \cdot Y)_n \le E[Y_n - Y_0]$ for $n \in \mathbb{N}$.

This, together with Lemma 5.1.7, implies Lemma 5.1.6. Note that Y is a submartingale (Lemma 4.3.3) and that S_k, T_k ($k \ge 1$) are stopping times. Note also that H_n, $n \ge 1$ is predictable, because for each $k \ge 1$,

$$\{S_k < n \le T_k\} = \{S_k \le n-1\} \backslash \{T_k \le n-1\} \in \mathcal{F}_{n-1}.$$

Since $H_n \le 1$, we see from Corollary 4.4.3 that

$$E(H \cdot Y)_n \le E(1 \cdot Y)_n = E[Y_n - Y_0].$$

This proves 1). \(^□^)/

Proof of Theorem 5.1.1 for $\mathbb{T} = [0,\infty)$: We assume that $\mathbb{T} = [0,\infty)$ and that $(X_t)_{t\ge 0}$ is right-continuous. By symmerty, we may focus on the case of submartingale. For $I \subset [0,\infty)$ and $-\infty < a < b < \infty$, Let

$$U(I) = \left\{ k \in \mathbb{N}\,;\ \begin{array}{l} \text{there exists a sequence } s_1 < t_1 < ... < s_k < t_k \text{ in } I \\ \text{such that } X_{s_j} \le a \text{ and } b \le X_{t_j} \text{ for all } j = 1, ..., k. \end{array} \right\}. \tag{5.4}$$

Let D be a dense subset of $[0,\infty)$, $t \in D$, and $(D_n)_{n\in\mathbb{N}}$ be a sequence of finite subsets of $D \cap [0,t]$ such that $0, t \in D_n$ for all $n \in \mathbb{N}$ and that $D_n \nearrow D \cap [0,t]$ as $n \nearrow \infty$. Then it follows from the proof of Lemma 5.1.5 that

$$(b-a)EU(D_n) \le E[X_- \vee a] - E[X_0 \vee a].$$

By the monotone convergence theorem in the limit $n \to \infty$,

$$(b-a)EU(D \cap [0,t]) \le E[X_- \vee a] - E[X_0 \vee a].$$

Then, by the monotone convergence theorem in the limit $t \to \infty$,

$$(b-a)EU(D) \le \sup_{t\ge 0} E[X_- \vee a] - E[X_0 \vee a] < \infty.$$

Hence $U(D) < \infty$, a.s., which implies, via the argument of Lemma 5.1.5 that the following limit exists a.s.

$$X_\infty = \lim_{\substack{t\to\infty \\ t\in D}} X_t \in [-\infty, \infty].$$

Moreover, by the right-continuity, we can remove the restriction $t \in D$ from the above limit. Finally, we see that $X_\infty \in L^1$, similarly as in Theorem 5.1.1. \(^□^)/

5.2 L^1 Convergence

Throughout this subsection, we assume that $X = (X_t, \mathcal{F}_t)_{t\in\mathbb{T}}$ is an adapted process. Here is the main result of this subsection.

Theorem 5.2.1 (**L^1 convergence theorem**) *Suppose that there exists a real r.v. X_∞ such that $X_t \stackrel{t\to\infty}{\longrightarrow} X_\infty$ a.s. Then, the following conditions are equivalent.*

$$X_\infty \in L^1(P) \text{ and } X_t = E[X_\infty|\mathcal{F}_t] \text{ a.s. for all } t \in \mathbb{T}. \tag{5.5}$$

$$\text{There exists a } Y \in L^1(P) \text{ such that } X_t = E[Y|\mathcal{F}_t] \text{ a.s. for all } t \in \mathbb{T}. \tag{5.6}$$

$$X \text{ is a uniformly integrable martingale.} \tag{5.7}$$

$$X \text{ is a martingale, } X_\infty \in L^1(P) \text{ and } X_t \stackrel{n\to\infty}{\longrightarrow} X_\infty \text{ in } L^1(P). \tag{5.8}$$

Moreover, it follows from (5.5) and (5.6) that the r.v's X_∞ and Y are related as

$$X_\infty = E[Y|\mathcal{F}_\infty] \text{ a.s. where } \mathcal{F}_\infty = \sigma\left[\bigcup_{t\in\mathbb{T}} \mathcal{F}_t\right]. \tag{5.9}$$

Proof: (5.5) ⇒ (5.6): Obvious.
(5.6) ⇒ (5.7): This follows from Lemma 4.1.13.
(5.7) ⇔ (5.8): This follows from Proposition 2.5.5.
(5.8) ⇒ (5.5): Since X is a martingale,

1) $$X_t = E[X_u|\mathcal{F}_t] \text{ a.s. for all } t, u \in \mathbb{T},\ t < u.$$

On the other hand, it follows from (5.8) and (4.13) that $E[X_u|\mathcal{F}_t] \stackrel{u\to\infty}{\longrightarrow} E[X_\infty|\mathcal{F}_t]$ in $L^1(P)$, which, together with 1), implies (5.5).

To prove (5.9), we take an arbitrary $t \in \mathbb{T}$ and $A \in \mathcal{F}_t$. Then, it follows from (5.5) and (5.6) that

$$E[X_\infty : A] = E[Y : A].$$

Since $t \in \mathbb{T}$ is arbitrary, the above equality is valid for all $A \in \bigcup_{t\in\mathbb{T}} \mathcal{F}_t$. Then, by Dynkin's Lemma (Lemma 1.3.1), the equality extends to all $A \in \mathcal{F}_\infty$, which implies (5.9). \(^□^)/

Remark: Suppose that X in Theorem 5.2.1 is bounded in $L^1(P)$. Then, there exists a signed measure Q on $(\Omega, \mathcal{F}_\infty)$ such that $X_n = dQ_n/dP_n$, $n \in \mathbb{N}$, where $P_n = P|_{\mathcal{F}_n}$ and $Q_n = Q|_{\mathcal{F}_n}$, cf. Proposition 4.7.1. Moreover, conditions (5.5)-(5.8) are equivalent to that $Q \ll P$, cf. Proposition 5.6.1 below.

As a direct consequence of Theorem 5.2.1, we obtain the following

Corollary 5.2.2 *Let $(\mathcal{F}_n)_{n\in\mathbb{T}}$ be a filtration and $Y \in L^1(P)$. Then,*

$$E[Y|\mathcal{F}_n] \stackrel{n\to\infty}{\longrightarrow} E[Y|\mathcal{F}_\infty] \text{ a.s. and in } L^1(P).$$

Proof: The martingale $X_n \stackrel{\text{def}}{=} E[Y|\mathcal{F}_n]$ satisfies (5.6). Therefore, by Theorem 5.1.1, there exists a real r.v. X_∞ such that $X_t \stackrel{t\to\infty}{\longrightarrow} X_\infty$ a.s. Moreover, by (5.8), $X_t \stackrel{t\to\infty}{\longrightarrow} X_\infty$ in $L^1(P)$. Finally, $X_\infty = E[Y|\mathcal{F}_\infty]$ by (5.9). \(^□^)/

Example 5.2.3 Let $X_n = \prod_{j=0}^n \xi_j$, where $(\xi_n)_{n\in\mathbb{N}}$ are mean-one nonnegative independent r.v's. Then, the following conditions are equivalent.

a) $\alpha \stackrel{\text{def}}{=} \prod_{n=1}^{\infty} E\sqrt{\xi_n} > 0$. **b)** $\sqrt{X_n} \stackrel{n\to\infty}{\longrightarrow} \sqrt{X_\infty}$ in $L^2(P)$. **c)** $(X_n)_{n\in\mathbb{N}}$ is uniformly integrable.

Proof: $X = (X_n)_{n\in\mathbb{N}}$ is a mean-one, nonnegative martingale by Example 4.3.6.
a) ⇒ b): It is enough to verify that $(\sqrt{X_n})_{n\in\mathbb{N}}$ is a Cauchy sequence, which can be done as follows. Let $m < n$. Then,

$$E\left[\sqrt{X_m}\sqrt{X_n}\right] = E\left[X_m\sqrt{\xi_{m+1}}\cdots\sqrt{\xi_n}\right] = \prod_{j=m+1}^{n} E\sqrt{\xi_j} \stackrel{m\to\infty}{\longrightarrow} 1,$$

and hence

$$\begin{aligned} E\left[|\sqrt{X_n} - \sqrt{X_m}|^2\right] &= EX_n + EX_m - 2E\left[\sqrt{X_n X_m}\right] \\ &= 2 - 2E\left[\sqrt{X_n X_m}\right] \stackrel{m\to\infty}{\longrightarrow} 0. \end{aligned}$$

b) ⇒ c):

$$EX_\infty = E[\sqrt{X_\infty}\sqrt{X_\infty}] \stackrel{\text{b)}}{=} \lim_{n\to\infty} E[\sqrt{X_n}\sqrt{X_n}] = \lim_{n\to\infty} E[X_n] = EX_0.$$

By Corollary 5.1.2, this implies c).
c) ⇒ a): To prove the contraposition, suppose $\alpha = 0$. Then, X does not converge in $L^1(P)$ by Example 5.1.3, hence X is not uniformly integrable, by the euivalence of (5.7) and (5.8). \(^□^)/

5.3 Optional Stopping Theorem

Throughout this subsection, we suppose that $X = (X_t, \mathcal{F}_t)_{t\in\mathbb{T}}$ is adapted process which satisfies (5.1). Now, suppose for a stopping time T that

$$X_t \text{ converges as } t \to \infty \text{ a.s. on the event } \{T = \infty\}. \tag{5.10}$$

If $T < \infty$ a.s., then nothing is assumed by (5.10). Let $S : \Omega \to [0, \infty]$ be a r.v. such that $\{S = \infty\} \subset \{T = \infty\}$. Then the r.v. X_S makes sense on the event $\{S < \infty\}$. Referring to (5.10),

$$X_S \stackrel{\text{def}}{=} \lim_{t\to\infty} X_t \text{ on the event } \{S = \infty\}. \tag{5.11}$$

The purpose of this subsection is to present the following theorem.

Theorem 5.3.1 (**Optional stopping theorem**) *Let $X = (X_t, \mathcal{F}_t)_{\in\mathbb{T}}$ be an adapted process and T be a stopping time for which we suppose (5.10) and adopt the convention (5.11). Then, the following conditions are equivalent.*

$$X_{S\wedge T} \in L^1(P) \text{ and } EX_T = EX_{S\wedge T} \text{ for any stopping time } S; \tag{5.12}$$

$$X_T \in L^1(P) \text{ and } E[X_T|\mathcal{F}_S] = X_{S\wedge T} \text{ a.s. for any stopping time } S; \tag{5.13}$$

$$X_T \in L^1(P) \text{ and } E[X_T|\mathcal{F}_t] = X_{t\wedge T} \text{ a.s. for all } t \in \mathbb{T}; \tag{5.14}$$

$$(X_{t\wedge T})_{t\in\mathbb{T}} \text{ is uniformly integrable martingale.} \tag{5.15}$$

Remark See Example 5.3.6 for typical examples for which condition (5.15) is valid.

We present the following Corollary to Theorem 5.3.1, which can easily be seen from the proof below.

Corollary 5.3.2 *Let $X = (X_t, \mathcal{F}_t)_{\in\mathbb{T}}$ be a uniformly integral submartingale (resp. supermartingale) and T be a stopping time for which we suppose (5.10) and adopt the convention (5.11). Then, (5.12)–(5.14) hold with the equalities replaced by $\geq$ (resp. $\leq$).*

For nonnegative supermartingales, (5.12)–(5.14) with the equalities replaced by $\leq$ are *always* true, even through they are not uniformly integrable in general. We note this fact as

Corollary 5.3.3 *(***Optional stopping theorem for nonnegative supermartigales***)* *Let $(X_t)_{t\in\mathbb{T}}$ be a nonnegative supermartingale and T be a stopping time for which we suppose (5.10) and adopt the convention (5.11). Then, (5.12)–(5.14) hold with the equalities replaced by $\le$. In particular, ,*

$$X_{T+t}\mathbf{1}\{X_T=0\}=0 \text{ for all } t\ge 0 \text{ a.s.} \tag{5.16}$$

Proof: We will prove (5.12) with the equality replaced by $\le$. Then, (5.12) and (5.14) with the equalities replaced by $\le$ follows from the proof of Theorem 5.3.1. We first observe that

1) $\quad E[X_{T\wedge t}|\mathcal{F}_S]\le X_{S\wedge T\wedge t}$ a.s. for arbitrarily fixed $t\in\mathbb{T}$.

This can be seen as follows. If $\mathbb{T}=\mathbb{N}$, then, $\{X_{s\wedge t\wedge T}\}_{s\in\mathbb{N}}\subset\{X_s\}_{s=0}^t$ is uniformly integrable. Thus, 1) follows from Corollary 5.3.2. If $T=[0,\infty)$ and $t\mapsto X_t$ is right-continuous, then, 1) follows from Lemma 5.3.4 below.

Note that $X_T=\lim_{t\to\infty}X_{t\wedge T}$. Then, by using Fatou's lemma for the conditional expectation given $\mathcal{F}_S$, we pass from 1) to (5.12) with the equality replaced by $\le$.

To see (5.16), we note that $E[X_{t+T}|\mathcal{F}_T]\le X_T$ a.s. and hence

$$E[X_{t+T}\mathbf{1}\{X_T=0\}|\mathcal{F}_T]=0 \text{ a.s.}$$

from which (5.16) follows. \(^□^)/

Proof of Theorem 5.3.1 for $\mathbb{T}=\mathbb{N}$:
(5.12) $\Leftrightarrow$ (5.13): It is enough to prove ($\Rightarrow$). By Lemma 5.3.8 below,

$$E[X_T|\mathcal{F}_{S\wedge T}]\stackrel{(5.20)}{=}E[X_T|\mathcal{F}_S].$$

Thus, it is enough to prove that

1) $E[X_T|\mathcal{F}_{S\wedge T}]=X_{S\wedge T}$ a.s.

To show this, we take arbitrary $A\in\mathcal{F}_{S\wedge T}$ and introduce

$$U=(S\wedge T)\mathbf{1}_A+T\mathbf{1}_{A^c},$$

which is a stopping time (Exercise 4.2.1) such that $U\le T$. Therefore, $X_U\in L^1(P)$ and

$$EX_T\stackrel{(5.12)}{=}EX_U=E[X_{S\wedge T}:A]+E[X_T:A^c],$$

i. e., $E[X_T:A]=E[X_{S\wedge T}:A]$, which implies 1).
(5.13) $\Leftrightarrow$ (5.14): It is enough to prove ($\Leftarrow$). Let $A\in\mathcal{F}_S$ be arbitrary. Then, $A\cap\{S=n\}\in\mathcal{F}_n$ for all $n\in\mathbb{N}$ and hence

3) $E[X_T:A\cap\{S=n\}]\stackrel{(5.14)}{=}E[X_{n\wedge T}:A\cap\{S=n\}]$.

Also, it is obvious that

4) $X_T=X_{S\wedge T}$ on the event $\{S=\infty\}$.

Therefore,

$$\begin{aligned} E[X_T : A] &= \sum_{n \in \mathbb{N}} E[X_T : A \cap \{S = n\}] + E[X_T : A \cap \{S = \infty\}] \\ &\stackrel{3),4)}{=} \sum_{n \in \mathbb{N}} E[X_{n \wedge T} : A \cap \{S = n\}] + E[X_{S \wedge T} : A \cap \{S = \infty\}] \\ &= E[X_{S \wedge T} : A], \end{aligned}$$

which implies (5.13).
(5.14) $\Leftrightarrow$ (5.15): This follows from Theorem 5.2.1 applied to $(X_{t \wedge T})_{t \in \mathbb{T}}$.

\(^□^)/

Remarks: 1) The condition (5.15) holds if $\sup_{t \in \mathbb{T}} |X_{t \wedge T}| \in L^1(P)$. This is in particular the case when $\mathbb{T} = \mathbb{N}$ and T is bounded.
2) Here is a well-known example for which a martingale does not satisfy (5.12) for a stopping time T, even with $S \equiv 0$. Let X be a simple random walk on $\mathbb{Z}$ such that $X_0 = 0$ and $T = \inf\{n \geq 1 \; ; \; X_n = x\}$ for $x \in \mathbb{Z}$. Since X is recurrent, we have $T < \infty$ a.s. and $X_T = x$ for all x. Thus, for $x \neq 0$, $EX_T = x \neq 0 = EX_0$.

We now turn to the proof of Theorem 5.3.1 for $\mathbb{T} = [0, \infty)$

- From here on, we assume that $\mathbb{T} = [0, \infty)$ and $(X_t)_{t \geq 0}$ is right-continuous.

The proofs of (5.15) $\Leftarrow$ (5.13) $\Leftrightarrow$ (5.12) are the same as those for the discrete-time case (Theorem 5.3.1). We will henceforth concentrate on the proof of (5.15) $\Rightarrow$ (5.13).

Let T_N, $N \in \mathbb{N}$ be a discrete approximation of T from the right defined by

$$T_N = \begin{cases} \frac{j}{2^N}, & \text{if } \frac{j-1}{2^N} < T \leq \frac{j}{2^N} \text{ for some } j \in \mathbb{N}, \\ \infty, & \text{if } T = \infty. \end{cases} \tag{5.17}$$

This approximation sequence is a subsequence of the one previously defined by (6.42). Thus, T_N are stopping times w.r.t. $(\mathcal{F}_t)_{t \geq 0}$ such that $0 \leq T_N - T \leq 2^{-N}$. Here, additionally, we have the monotonocity: $T_{N+1} \leq T_N$, $N \in \mathbb{N}$.

Lemma 5.3.4 *Suppose that $X = (X_t, \mathcal{F}_t)_{t \geq 0}$ is a right-continuous martingale and that T is a bounded stopping time. Then (5.13) is true. Moreover, if we suppose X is a right-continuous submartingale (resp. supermartingale) then, (5.13) holds with the equality replaced by $\geq$ (resp. $\leq$)*

Proof: We discuss only martingale case, adjustment needed for submartingale (supermartingale) cases being obvious. It is enough to prove that

$$E[X_T : A] = E[X_{S \wedge T} : A] \ \text{ for all } A \in \mathcal{F}_S. \tag{5.18}$$

For $N \in \mathbb{N}$ fixed, $X^{(N)} = (X_t, \mathcal{F}_t)_{t \in 2^{-N}\mathbb{N}}$ is a martingale, and S_N, T_N are stopping times w.r.t. the filtration $(\mathcal{F}_t)_{t \in 2^{-N}\mathbb{N}}$. Moreover, we have $A \in \mathcal{F}_S \subset \mathcal{F}_{S_N}$. Since T_N is bounded by assumption, it follows from Theorem 5.3.1 applied to the discrete-time martingale $X^{(N)}$ that

$$E[X(T_N) : A] = E[X(S_N \wedge T_N) : A].$$

Therefore, it only remains to prove that

$$X(T_N) \stackrel{N\to\infty}{\longrightarrow} X(T) \text{ and } X(S_N \wedge T_N) \stackrel{N\to\infty}{\longrightarrow} X(S \wedge T) \text{ in } L^1(P).$$

By right-continuity, the above convergences take place a.s. Hence it is enough to prove that

1) $\{X(T_N)\}_{N\in\mathbb{N}}$, $\{X(S_N \wedge T_N)\}_{N\in\mathbb{N}}$ are uniformly integrable.

Let U_N be either $S_N \wedge T_N$ or T_N. By assumption, there exists $m \in \mathbb{N}$ such that $T \leq m$ a.s., and hence $U_N \leq T_N \leq T_0 \leq T + 1 \leq m + 1$. Then, by Theorem 5.3.1 applied to the discrete-time submartingale $(|X_t|, \mathcal{F}_t)_{t\in 2^{-N}\mathbb{N}}$ and the bounded stopping times U_N, $m + 1$ w.r.t. $(\mathcal{F}_t)_{t\in 2^{-N}\mathbb{N}}$, we have

$$|X(U_N)| \leq E[|X_{m+1}||\mathcal{F}_{U_N}].$$

By Lemma 4.1.13, the right-hand side of the above inequality is uniformly integrable in N. Thus, $\{X(U_N)\}_{N\in\mathbb{N}}$ is uniformly integrable, which proves 1). \(^▫^)/

Lemma 5.3.5 *Suppose that $X = (X_t, \mathcal{F}_t)_{t\geq 0}$ is a right-continuous martingale (resp. submartingale, supermartingale) Then, for any stopping time R, $(X_{t\wedge R}, \mathcal{F}_t)_{t\geq 0}$ is a martingale (resp. submartingale, supermartingale).*

Proof: We discuss only martingale case, adjustment needed for submartingale (supermartingale) cases being obvious. By the right-continuity, the process $(X_{t\wedge R}, \mathcal{F}_t)_{t\geq 0}$ is adapted (Corollary 6.6.15). Let $0 \leq s < t$. Then, $t \wedge R$ is a bounded stopping time, and hence by Lemma 5.3.4,

$$X_{t\wedge R} \in L^1(P), \quad E[X_{t\wedge R}|\mathcal{F}_s] = X_{s\wedge R} \text{ a.s.}$$

This proves the lemma. \(^▫^)/

Proof of Theorem 5.3.1 for $\mathbb{T} = [0, \infty)$:
As is mentioned before, we have only to prove (5.15) ⇒ (5.13). For this purpose, it is enough to prove (5.18). By Lemma 5.3.5, $(X_{t\wedge T}, \mathcal{F}_t)_{t\geq 0}$ is a martingale and it is uniformly integrable by the assumption (5.15). Thus, for any $N \in \mathbb{N}$ fixed, $X^{(T,N)} = (X_{t\wedge T}, \mathcal{F}_t)_{t\in 2^{-N}\mathbb{N}}$ is a uniformly integrable martingale, and S_N, T_N are stopping times w.r.t. the filtration $(\mathcal{F}_t)_{t\in 2^{-N}\mathbb{N}}$. Moreover, we have $A \in \mathcal{F}_S \subset \mathcal{F}_{S_N}$. Thus, by Theorem 5.3.1 applied to the discrete-time martingale $X^{(T,N)}$, we have

$$E[X_T : A] \stackrel{T\leq T_N}{=} E[X(T_N \wedge T) : A] = E[X(S_N \wedge T) : A]$$

Therefore, it only remains to prove that

$$X(S_N \wedge T) \stackrel{N\to\infty}{\longrightarrow} X(S \wedge T) \text{ in } L^1(P).$$

By right-continuity, the above convergence takes place a.s. Hence it is enough to prove that

1) $\{X(S_N \wedge T)\}_{N\in\mathbb{N}}$ is uniformly integrable.

By assumption (5.15), the discrete-time submartingale $(|X_{t\wedge T}|, \mathcal{F}_t)_{t\in 2^{-N}\mathbb{N}}$ is uniformly integrable. Thus, by Theorem 5.3.1 applied to this submartingale, we see that $|X(S_0 \wedge T)| \in L^1(P)$ and that

$$|X(S_N \wedge T)| \leq E[|X(S_0 \wedge T)||\mathcal{F}_{S_N}].$$

By Lemma 4.1.13, the right-hand side of the above inequality is uniformly integrable in N, which proves 1). \(^▫^)/

Example 5.3.6 Let $X = (X_t, \mathcal{F}_t)_{t \in \mathbb{T}}$ be an adapted process. Here are typical examples for X and a stopping time T for which $(X_{t\wedge T})_{t\in\mathbb{T}}$ is uniformly integrable. Suppose that $|X_0| \le M$, a.s. for some $M \in (0,\infty)$ and let $T = \inf\{t \in \mathbb{T} \cap (0,\infty) \ ; \ |X_t| > M\}$. Suppose:
a) $C \stackrel{\text{def}}{=} \sup_{t\in\mathbb{T}} E|X_t| < \infty$,
b) Eihter the following b1) or b2) holds true.
b1) $\mathbb{T} = \mathbb{N}$ and there exists $R \in (0,\infty)$ such that $\sup_{n\ge 1} |X_n - X_{n-1}| \le R$.
b2) $\mathbb{T} = [0,\infty)$ and $t \mapsto X_t$ is continuous.
Then, $(X_{t\wedge T})_{t\in\mathbb{T}}$ is uniformly integrable.

Proof: Let $\lambda > 0$. Then,

$$E[|X_{t\wedge T}| : |X_{t\wedge T}| \ge \lambda] = I_t(\lambda) + J_t(\lambda),$$

where

$$I_t(\lambda) = E[|X_t| : |X_t| \ge \lambda, \ t < T], \quad J_t(\lambda) = E[|X_T| : |X_t| \ge \lambda, \ T \le t].$$

Since $\{t < T\} \subset \{|X_t| \le M\}$, we have

$$\sup_{t\in\mathbb{T}} I_t(\lambda) \le M \sup_{t\in\mathbb{T}} P(|X_t| \ge \lambda) \le MC/\lambda \stackrel{\lambda\to\infty}{\longrightarrow} 0.$$

As for $J_t(\lambda)$, let us first assume b1). Then, $|X_T| \le |X_{T-1}| + R \le M + R$ and hence

$$\sup_{t\in\mathbb{T}} J_t(\lambda) \le (M+R) \sup_{t\in\mathbb{T}} P(|X_t| \ge \lambda) \le (M+R)C/\lambda \stackrel{\lambda\to\infty}{\longrightarrow} 0.$$

If we assume b2), then, $|X_T| = M$. Thus, we have $\sup_{t\in\mathbb{T}} J_t(\lambda) \stackrel{\lambda\to\infty}{\longrightarrow} 0$ similarly as above. \(^□^)/

Example 5.3.7 Suppose that $X = (X_t, \mathcal{F}_t)_{t\ge 0}$ is a nonnegative martingale such that $t \mapsto X_t$ is continuous and that $X_\infty \equiv 0$ a.s. For a bounded stopping time S, we write $M_{[S,\infty)} = \sup_{t\ge S} X_t$. Then, for all $x \in (1,\infty)$

$$P\left(M_{[S,\infty)} > xX_S | \mathcal{F}_S\right) = x^{-1} \ \text{ a.s. on the set } \{X_S \neq 0\}.$$

In particular, if $P(X_S \neq 0) > 0$, then, conditionally on the event $X_S \neq 0$, the law of the r.v. $M_{[S,\infty)}/X_S$ is given by $x^{-2}\mathbf{1}_{\{x>1\}}dx$.

Proof: We will prove that for all $\mathcal{F}_S$-measurable, integrable r.v. $Z \ge 0$,

1) $P\left(M_{[S,\infty)} > Z \middle| \mathcal{F}_S\right) = 1 \wedge (X_S/Z)$ a.s. on the set $\{Z \neq 0\}$.

Then, the desired equality follows by setting $Z = xX_S$. It is easy to verify that

2) $P\left(M_{[S,\infty)} > Z \middle| \mathcal{F}_S\right) = 1$ a.s. on the set $\{X_S > Z\}$.

Indeed,

$$\begin{aligned} P\left(M_{[S,\infty)} > Z|\mathcal{F}_S\right)\mathbf{1}\{X_S > Z\} &= P\left(M_{[S,\infty)} > Z, X_S > Z|\mathcal{F}_S\right) \\ &= P\left(X_S > Z|\mathcal{F}_S\right) = \mathbf{1}\{X_S > Z\}. \end{aligned}$$

By 2), it is enough to prove that

3) $ZP\left(M_{[S,\infty)} > Z\middle|\, \mathcal{F}_S\right) = X_S$ a.s. on the set $\{X_S \le Z\}$.

For this purpose, we consider a stopping time $T = \inf\{t \ge S\ ;\ X_t > Z\}$. Then, for $n \in \mathbb{N}$, $(S+n) \wedge T$ is a bounded stopping time, and hence

4) $X_S \stackrel{\text{Lemma 5.3.4}}{=} E[X_{(S+n)\wedge T}|\mathcal{F}_S] = E[X_{S+n}\mathbf{1}\{T=\infty\}|\mathcal{F}_S] + E[X_{(S+n)\wedge T}\mathbf{1}\{T<\infty\}|\mathcal{F}_S]$.

Since $X_{S+n} \stackrel{n\to\infty}{\longrightarrow} 0$ a.s. and $\{T = \infty\} \subset \{X_{S+n} \le Z\}$ for all $n \in \mathbb{N}$, we have by DCT that

5) $E[X_{S+n}\mathbf{1}\{T=\infty\}|\mathcal{F}_S] \stackrel{n\to\infty}{\longrightarrow} 0$ a.s.

On the other hand, on the event $\{X_S \le Z, T < \infty\}$, $X_{(S+n)\wedge T} \stackrel{n\to\infty}{\longrightarrow} X_T = Z$ and $0 \le X_{(S+n)\wedge T} \le Z$ for all $n \in \mathbb{N}$. Therefore we have by DCT that

6) $E[X_{(S+n)\wedge T}\mathbf{1}\{T<\infty\}|\mathcal{F}_S] \stackrel{n\to\infty}{\longrightarrow} ZP(T<\infty|\mathcal{F}_S) = ZP(M_{[S,\infty)} > Z|\mathcal{F}_S)$ a.s.

Combining 4)–6), we obtain 3). \(^□^)/

(⋆) Complement

We present the following lemma, which was used in the proof of (5.13) ⇐ (5.12). This lemma is valid in the general setting of Definition 4.2.1.

Lemma 5.3.8 *Let S and T be stopping times and $X \in L^1(P)$. Then,*

$$E[X|\mathcal{F}_S] = E[X|\mathcal{F}_{S\wedge T}] \ \text{ a.s. on } \{S \le T\}. \tag{5.19}$$

Suppose in particular that X is $\mathcal{F}_T$-measurable. Then,

$$E[X|\mathcal{F}_S] = E[X|\mathcal{F}_{S\wedge T}] \ \text{ a.s.} \tag{5.20}$$

Proof: (5.19): The (5.19) is equivalent to

$$Y \stackrel{\text{def}}{=} E[X|\mathcal{F}_S]\mathbf{1}\{S \le T\} = E[X|\mathcal{F}_{S\wedge T}]\mathbf{1}\{S \le T\} \ \text{ a.s.},$$

which can be paraphrased as $Y = E[Y|\mathcal{F}_{S\wedge T}]$ a.s. Therefore, it is enough that Y is $\mathcal{F}_{S\wedge T}$-measurable.

On the other hand, $\{S \le T\} \in \mathcal{F}_{S\wedge T}$ by (4.40), and $E[X|\mathcal{F}_S]$ is $\mathcal{F}_S$-measurable. Therefore, Y is $\mathcal{F}_{S\wedge T}$-measurable by (4.41).

(5.20): By (5.19), (5.20) is equivalent to

$$Z \stackrel{\text{def}}{=} E[X|\mathcal{F}_S]\mathbf{1}\{T \le S\} = E[X|\mathcal{F}_{S\wedge T}]\mathbf{1}\{T \le S\} \ \text{ a.s.},$$

which can be paraphrased as $Z = E[Z|\mathcal{F}_{S\wedge T}]$ a.s. Therefore, it is enough that Z is $\mathcal{F}_{S\wedge T}$-measurable.

On the other hand, $X\mathbf{1}\{T \le S\}$ is $\mathcal{F}_{S\wedge T}$-measurable by (4.41), since X is $\mathcal{F}_T$-measurable. Hence

$$Z = E[X\mathbf{1}\{T \le S\}|\mathcal{F}_S] = X\mathbf{1}\{T \le S\}.$$

Therefore Z is $\mathcal{F}_{S\wedge T}$-measurable. \(^□^)/

Exercise 5.3.1 Let S nand T be stopping times and $X \in L^1(P)$. Prove then that $E[E[X|\mathcal{F}_T]|\mathcal{F}_S] = E[X|\mathcal{F}_{S\wedge T}]$ a.s. Hint: (5.20).

Exercise 5.3.2 Let $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be a filtration, $Y \in L^1(P)$, T be a stopping time, and $X_n = E[Y|\mathcal{F}_n]$, $n \in \mathbb{N}$. Then, prove that $X_T = E[Y|\mathcal{F}_T]$ a.s. on $\{T < \infty\}$. Hint: proof of (5.15) $\Rightarrow$ (5.13).

Exercise 5.3.3 Let $(X_t)_{t\in\mathbb{T}}$ be a nonnegative submartingale with assumption (5.1) and T be a stopping time. Then, prove the following. **i)** $EX_{t\wedge T} \leq EX_t$ for all $t \in \mathbb{T}$. **ii)** Suppose that $\sup_{t\in\mathbb{T}} EX_t < \infty$, so that $X_t \to X_\infty$, a.s. for some $X_\infty \in L^1(P)$ by the martingale convergence theorem (Theorem 5.1.1). Then, $EX_T \leq \sup_{t\in\mathbb{T}} EX_t$, where $X_T \stackrel{\text{def}}{=} X_\infty$ on the set $\{T = \infty\}$.

Exercise 5.3.4 Using the argument of Example 5.3.7, give an alternative proof of the equalities (4.76).

5.4 L^p Convergence

Throughout this subsection, we assume that $X = (X_t, \mathcal{F}_t)_{t\in\mathbb{T}}$ is an adapted process such that (5.1) holds. We set

$$Y_t = \sup_{s\in[0,t]\cap\mathbb{T}} X_s \ \text{and} \ \widetilde{Y}_t = \sup_{s\in[0,t]\cap\mathbb{T}} |X_s|. \tag{5.21}$$

We start by proving the following

Proposition 5.4.1 (**Doob's inequalities**) *Let $t \in \mathbb{T}$.*

a) (maximal inequality) *Suppose that $(X_s)_{s\in[0,t]\cap\mathbb{T}}$ is a submartingale. Then, for all $\lambda > 0$,*

$$\lambda P(Y_t \geq \lambda) \leq E[X_t : Y_t \geq \lambda]. \tag{5.22}$$

b) (L^p-maximal inequality) *Suppose that $(X_s)_{s\in[0,t]\cap\mathbb{T}}$ is a martingale, or a nonnegative submartingale. Then,*

$$\|\widetilde{Y}_t\|_p \leq \frac{p}{p-1}\|X_t\|_p \ \ \textit{if } p \in (1,\infty). \tag{5.23}$$

Remark The inequality (5.23) is no longer true for $p = 1$. In fact, we present an example of martingales for which there is no constant $c \in [0,\infty)$ such that

$$\|\widetilde{Y}_t\|_1 \leq c\|X_t\|_1 \ \text{ for all } t \in \mathbb{T} \tag{5.24}$$

cf. Example 5.4.3. In addition, the multiplicative constant $\frac{p}{p-1}$ on the RHS of (5.23) cannot be improved (Exercise 5.4.3).

Proof of Proposition 5.4.1 a): *Case 1*: $\mathbb{T} = \mathbb{N}$: Let $\lambda > 0$ be fixed and $T = \inf\{t \in \mathbb{T} \ ; \ X_t \geq \lambda\}$. Then,

1) $A \stackrel{\text{def}}{=} \{Y_t \geq \lambda\} = \{T \leq t\} \stackrel{\text{Lemma 4.2.3}}{\in} \mathcal{F}_{t\wedge T}$.

Moreover, for each fixed $t \in \mathbb{N}$, $(X_{s\wedge t})_{s\in\mathbb{N}}$ is clearly uniformly integrable submartingale, and hence by 1) and Corollary 5.3.2,

2) $E[X_{t\wedge T} : A] \leq E[X_t : A]$.

Finally, on the event A, we have that $\lambda \le X_T$ and $T = t \wedge T$. Therefore,

3) $A \subset \{\lambda \le X_{t\wedge T}\}$.

Combining these,

$$\lambda P(A) \stackrel{3)}{\le} E[X_{t\wedge T} : A] \stackrel{2)}{\le} E[X_t : A],$$

which proves (5.22).

Case 2: $\mathbb{T} = [0,\infty)$: Let $\lambda > 0$ and $t > 0$ be fixed. We approximate the interval $[0,t]$ by a finite subset set $I_N = \{2^{-N}kt\}_{k=0}^{2^N}$. We also take a strictly increasing positive sequence λ_n such that $\lambda_n \nearrow \lambda$, so that

1) $\quad (\lambda_n, \infty) \searrow [\lambda, \infty)$ as $n \to \infty$.

By the argument of Case 1, (5.22) is valid for the discrete-time submartingale $\{X_s\}_{s\in I_N}$. Therefore, we have for $m < n$ that

2)
$$\begin{cases} \lambda_n P\left(\max_{s\in I_N} X_s > \lambda_n\right) & \le \lambda_n P\left(\max_{s\in I_N} X_s \ge \lambda_n\right) \\ & \le E\left[X_t : \max_{s\in I_N} X_s \ge \lambda_n\right] \le E\left[X_t : \max_{s\in I_N} X_s > \lambda_m\right]. \end{cases}$$

Since X is right-continuous, $\max_{s\in I_N} X_s \nearrow Y_t$ as $N \to \infty$. Note also that the indicator function of an interval (a,∞) ($a \in \mathbb{R}$) is left-continuous, and hence

$$\mathbf{1}_{(a,\infty)}\left(\max_{s\in I_N} X_s\right) \stackrel{N\to\infty}{\longrightarrow} \mathbf{1}_{(a,\infty)}(Y_t).$$

Thus, by letting $N \to \infty$, it follows from 2) that

3) $$\lambda_n P(Y_t > \lambda_n) \le E[X_t : Y_t > \lambda_m].$$

By letting $n \to \infty$ first, and then letting $m \to \infty$, we obtain (5.22) from 1) and 3). \\(^□^)/

Proposition 5.4.1 b) will be proved via Lemma 5.4.2 below. The lemm has various applications beside the proof of Proposition 5.4.1 b), cf. Example 5.4.8. For this reason, we state the lemma in a setting which is more general than is necessary to prove Proposition 5.4.1 b). Here is the the settig for the lemma.

• Let $\varphi_1 : [0,\infty) \to [0,\infty)$ a right-continuous, nondecreasing function such that $\int_0^1 \frac{d\varphi_1(\lambda)}{\lambda} < \infty$. For $\varphi_1 \in \Phi$, we associate it with a function $\varphi_2 : [0,\infty) \to [0,\infty)$ defined by

$$\varphi_2(\lambda) = \int_0^\lambda \frac{d\varphi_1(t)}{t}, \quad \lambda \ge 0.$$

We denote the totality of such pairs (φ_1, φ_2) by Φ, of which two typical examples are

$$\varphi_1(\lambda) = \lambda^p \ (1 < p < \infty) \text{ and } \varphi_2(\lambda) = q\lambda^{p-1}, \text{ where } q = \tfrac{1}{1-p^{-1}}, \tag{5.25}$$

$$\varphi_1(\lambda) = (\lambda - 1)^+ \text{ and } \varphi_2(\lambda) = \log^+ \lambda \stackrel{\text{def}}{=} (\log \lambda) \vee 0. \tag{5.26}$$

Let $f, g \ge 0$ be measurable functions on a measure space $(S, \mathcal{B}, \mu)$. We consider the following conditions.

$$\mu(g \ge \lambda) \le \frac{1}{\lambda}\int_{g\ge\lambda} f d\mu \ \text{ if } \lambda > 0, \tag{5.27}$$

$$\int_S \varphi_1(g) d\mu \le \int_S f\varphi_2(g) d\mu \ \text{ if } (\varphi_1, \varphi_2) \in \Phi, \tag{5.28}$$

$$\int_S g^p d\mu \le \left(\frac{p}{p-1}\right)^p \int_S f^p d\mu \ \text{ if } p \in (1,\infty), \tag{5.29}$$

These conditions are related as follows.

Lemma 5.4.2 *(5.27) $\Longleftrightarrow$ (5.28) $\Longrightarrow$ (5.29).*

Proof: It follows from $\int_0^1 \frac{d\varphi_1(\lambda)}{\lambda} < \infty$ that $\varphi_1(0) = \varphi_2(0) = 0$, and hence $\varphi_1(\lambda) = \int_0^\lambda d\varphi_1(t)$ and $\varphi_2(\lambda) = \int_0^\lambda \frac{d\varphi_1(t)}{t}$ for all $\lambda \geq 0$. Therefore, for $j = 1, 2$,

$$\begin{aligned}\int_S f\varphi_j(g)d\mu &= \int_S f(x)d\mu(x)\int_0^\infty \mathbf{1}\{g(x) \geq \lambda\}d\varphi_j(\lambda) \\ &\stackrel{\text{Fubini}}{=} \int_0^\infty d\varphi_j(\lambda)\int_S f(x)\mathbf{1}\{g(x) \geq \lambda\}d\mu(x) \\ &= \int_0^\infty d\varphi_j(\lambda)\int_{g\geq\lambda} fd\mu. \end{aligned} \tag{5.30}$$

(5.27) $\Rightarrow$ (5.28):

$$\int_S \varphi_1(g)d\mu \stackrel{(5.30)}{=} \int_0^\infty d\varphi_1(\lambda)\mu(g \geq \lambda) \stackrel{(5.27)}{\leq} \int_0^\infty d\varphi_2(\lambda)\int_{g\geq\lambda} fd\mu \stackrel{(5.30)}{=} \int_S f\varphi_2(g)d\mu.$$

(5.27) $\Leftarrow$ (5.28): For fixed $\lambda > 0$, take $(\varphi_1, \varphi_2) \in \Phi$ defined by $d\varphi_1(t) = \delta_\lambda(dt)$ and $d\varphi_2(t) = \frac{1}{t}\delta_\lambda(dt) = \frac{1}{\lambda}\delta_\lambda(dt)$.
(5.28) $\Rightarrow$ (5.29): We may assume that $\int_S f^p d\mu < \infty$. We take $\varphi_1(\lambda) = \lambda^p$ and $\varphi_2(\lambda) = q\lambda^{p-1}$. As we have already seen, (5.28) implies (5.33). Thus, by applying (5.33) with $\beta = 2$, we see that

$$\int_S g^p d\mu \leq q2^q \int_S f^p d\mu < \infty.$$

Then applying (5.28),

$$\int_S g^p d\mu \stackrel{(5.28)}{\leq} q\int_S fg^{p-1}d\mu \stackrel{\text{Hölder}}{\leq} q\left(\int f^p d\mu\right)^{1/p}\left(\int g^p d\mu\right)^{1/q}$$

By dividing both sides by $\int g^p d\mu < \infty$, we obtain (5.29). \(^□^)/

Proof of Proposition 5.4.1 b): If X is a nonnegative submartingale, then $\widetilde{Y}_t = Y_t$. Hence, we conclude (5.23) from (5.22) and Lemma 5.4.2.

If X is a martingale, then, the desired inequality is obtained by applying (5.23) to the nonnegative submartingale $(|X_t|)_{t\in\mathbb{T}}$. \(^□^)/

Example 5.4.3 Here is an example of a martingale for which there is no constant $c \in [0, \infty)$ with property (5.24). Let $X = (X_t, \mathcal{F}_t)_{t\in\mathbb{T}}$ be a nonnegative martingale which is not uniformly integrable (for example the martingale in Example 5.2.3 which satisfies $\alpha = 0$) Then, by Theorem 5.2.1, X_t does not converge in $L^1(P)$ as $t \to \infty$. This implies, via Lemma 5.4.5 that $\|Y_t\|_1 \stackrel{t\to\infty}{\longrightarrow} \infty$. On the other hand, $\|X_t\|_1 = \|X_0\|_1$ for all $t \in \mathbb{T}$. In conclusion, there is no constant $c \in [0, \infty)$ with property (5.24).

The rest of this subsection is devoted to the proof of

Proposition 5.4.4 (**L^p convergence theorem**) *Let $X = (X_t, \mathcal{F}_t)_{t\in\mathbb{T}}$ be a martingale, or a nonnegative submartingale with assumption (5.1) in both cases. Suppose that $p \in (1,\infty)$ and that*

$$\sup_{t\in\mathbb{T}} \|X_t\|_p < \infty.$$

Then, there exists $X_\infty \in L^p(P)$ such that

$$X_t \stackrel{t\to\infty}{\longrightarrow} X_\infty \text{ a.s. and in } L^p(P).$$

To prove Proposition 5.4.4, we prepare the following

Lemma 5.4.5 *Let $\mathbb{T} \subset [0,\infty)$ be unbounded, $(X_t)_{t\in\mathbb{T}}$ be a sequence of r.v.'s and $\widetilde{Y}_t = \max_{s\in[0,t]\cap\mathbb{T}} |X_s|$. Suppose that there exists a r.v. X_∞ such that $X_t \stackrel{t\to\infty}{\longrightarrow} X_\infty$ a.s. and that*

$$\sup_{t\in\mathbb{T}} \|\widetilde{Y}_t\|_p < \infty \text{ for some } p \in [1,\infty), \tag{5.31}$$

Then, $X_t \stackrel{t\to\infty}{\longrightarrow} X_\infty$ in $L^p(P)$.

Proof: We let $\widetilde{Y}_\infty = \sup_{t\in\mathbb{T}} |X_t|$. Then,

$$\|\widetilde{Y}_\infty\|_p \stackrel{\text{Fatou}}{\leq} \varliminf_{t\to\infty} \|\widetilde{Y}_t\|_p < \infty.$$

Therefore, $\widetilde{Y}_\infty \in L^p(P)$ and hence

$$|X_t - X_\infty|^p \leq (2\widetilde{Y}_\infty)^p \in L^1(P).$$

We see from above considerations and the dominated convergence theorem that $X_t \stackrel{t\to\infty}{\longrightarrow} X_\infty$ in $L^p(P)$. \(^□^)/

Proof of Proposition 5.4.4: Note that

$$\sup_{t\in\mathbb{T}} \|X_t\|_1 \leq \sup_{t\in\mathbb{T}} \|X_t\|_p < \infty.$$

Then, it follows from the martingale convergence theorem (Theorem 5.1.1) that there exists $X_\infty \in L^1(P)$ such that $X_t \stackrel{t\to\infty}{\longrightarrow} X_\infty$, a.s. On the other hand, we let $\widetilde{Y}_t = \max_{s\in[0,t]\cap\mathbb{T}} |X_s|$. Then, by the L^p maximal inequality,

$$\sup_{t\in\mathbb{T}} \|\widetilde{Y}_t\|_p \stackrel{(5.23)}{\leq} q \sup_{t\in\mathbb{T}} \|X_t\|_p < \infty.$$

Therefore, we see from Lemma 5.4.5 that $X_n \stackrel{n\to\infty}{\longrightarrow} X_\infty$ in $L^p(P)$.

\(^□^)/

Complement In addition to the conditons (5.27) and (5.28), we consider the following conditions.

$$\mu(g/\beta \geq \lambda) \leq \frac{1}{(\beta-1)\lambda}\int_{f\geq\lambda} f d\mu \text{ if } \lambda > 0 \text{ and } \beta > 1, \tag{5.32}$$

$$\int_S \varphi_1(g/\beta)d\mu \leq \frac{1}{(\beta-1)}\int_S f\varphi_2(f)d\mu \text{ if } (\varphi_1,\varphi_2) \in \Phi \text{ and } \beta > 1, \tag{5.33}$$

$$\int_S (g-\beta)^+ d\mu \leq \alpha\int_S f\log^+ f d\mu \text{ if } \alpha,\beta \in (1,\infty),\ \tfrac{1}{\alpha}+\tfrac{1}{\beta}=1. \tag{5.34}$$

Remark Note that $x \leq (x-\beta)^+ + \beta$ for all $x, \beta \in \mathbb{R}$. Thus, if μ is a finite measure, then it follows from (5.34) that

$$\int_S g d\mu \leq \alpha \int_S f\log^+ f d\mu + \beta\mu(S). \tag{5.35}$$

We have the following lemma.

Lemma 5.4.6 ($\star$) *The conditions (5.27)–(5.34) are related as*

$$(5.27) \iff (5.28) \implies (5.32) \iff (5.33) \implies (5.34)$$

Proof:
(5.27) $\Rightarrow$ (5.32):

$$\begin{aligned}\beta\lambda\mu(g \geq \beta\lambda) &\overset{(5.27)}{\leq} \int_{g\geq\beta\lambda} f d\mu = \int_{\substack{g\geq\beta\lambda\\ f\geq\lambda}} f d\mu + \int_{\substack{g\geq\beta\lambda\\ f<\lambda}} f d\mu \\ &\leq \int_{f\geq\lambda} f d\mu + \lambda\mu(g \geq \beta\lambda).\end{aligned}$$

Subtracting $\lambda\mu(g \geq \beta\lambda)$ from the both-hand sides, we obtain (5.32).
(5.32) $\Leftrightarrow$ (5.33): This can be shown in the same way as (5.27) $\Leftrightarrow$ (5.28).
(5.33) $\Rightarrow$ (5.34): Apply (5.33) to $\varphi_1(\lambda) = (\lambda-1)^+$ and $\varphi_2(\lambda) = \log^+\lambda$. \\(^□^)/

Proposition 5.4.7 ($\star$) **(L^1-maximal inequality)** *Suppose that $t \in \mathbb{T}$ and that $(X_s)_{s\in[0,t]\cap\mathbb{T}}$ is a martingale, or a nonnegative submartingale. Then, for all $t \in \mathbb{T}$,*

$$\|\widetilde{Y}_t\|_1 \leq \alpha\||X_t|\log^+|X_t|\|_1 + \frac{\alpha}{\alpha-1} \text{ if } \alpha \in (1,\infty) \tag{5.36}$$

Proof: If X is a nonnegative submartingale, then $\widetilde{Y}_t = Y_t$. Hence, we conclude (5.36) from (5.22) and Lemma 5.4.6.

If X is a martingale, then, the desired inequality is obtained by applying (5.36) to the nonnegative submartingale $(|X_t|)_{t\in\mathbb{T}}$. \\(^□^)/

Remark The reverse inequality to (5.36) holds true in some cases, cf. Example 5.4.8.

Example 5.4.8 ($\star$) Here is an example of a martingale for which reverse inequality to (5.36) holds true. Suppose that $X = (X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ is a nonnegative supermartingale such that $X_0 = 1$ and that there exists $C \geq 1$ such that $X_{n+1} \leq CX_n$ for all $n \in \mathbb{N}$. Then,

$$E[X_\infty \log^+ X_\infty] \leq C(EY_\infty - 1).$$

Proof: Fix $\lambda > 1$ and set $T = \inf\{n \geq 1\ ;\ X_n > \lambda\}$. We observe that

1) $X_\infty > \lambda \Rightarrow Y_\infty > \lambda \iff T < \infty$,

2) $T < \infty \Rightarrow X_T \le CX_{T-1} \le C\lambda$.

Therefore,

$$\begin{aligned} E[X_\infty : X_\infty > \lambda] &\stackrel{1)}{\le} E[X_\infty : T < \infty] \stackrel{\text{Corollary 5.3.3}}{\le} E[X_T : T < \infty] \\ &\stackrel{2)}{\le} C\lambda P(T < \infty) \stackrel{1)}{=} C\lambda P(Y_\infty > \lambda). \end{aligned} \tag{5.37}$$

Noting that $Y_\infty \ge X_0 = 1$, we have

$$\begin{aligned} EY_\infty &= \int_0^\infty P(Y_\infty > \lambda) d\lambda = 1 + \int_1^\infty P(Y_\infty > \lambda) d\lambda \\ &\stackrel{(5.37)}{\ge} 1 + C^{-1} \int_1^\infty E[X_\infty : X_\infty > \lambda] \frac{d\lambda}{\lambda} \\ &\stackrel{(5.33)}{=} 1 + C^{-1} E[X_\infty \log^+ X_\infty]. \end{aligned}$$

Exercise 5.4.1 For $1 \le p < \infty$, let $\mathcal{M}^p$ be the totality of the martingales X such that $\|X\|_{\mathcal{M}^p} \stackrel{\text{def}}{=} \sup_{t \ge 0} \|X_t\|_p < \infty$. Also, let $\mathcal{M}^1_0$ be the totality of the uniformly integrable martingales in $\mathcal{M}^1$. Prove the following. **i)** The map $X \mapsto X_\infty$ defines a surjective isometry from $(\mathcal{M}^p, \|\cdot\|_{\mathcal{M}^p})$ to $L^p(\Omega, \mathcal{F}_\infty, P)$ for $1 < p < \infty$. The same map defines a surjective isometry from $(\mathcal{M}^1_0, \|\cdot\|_{\mathcal{M}^p})$ to $L^1(\Omega, \mathcal{F}_\infty, P)$. **ii)** For $1 < p < \infty$, the norms $\|X\|_{\mathcal{M}^p}$ and $\|\sup_{t\ge 0} |X_t|\|_p$ are equivalent.

Exercise 5.4.2 (**exponential maximal inequality**) Let $t \in \mathbb{T}$. Suppose that $(X_s)_{s \in [0,t] \cap \mathbb{T}}$ is a submartingale and that $E \exp X_t < \infty$. Then, prove that $E \exp Y_t \le eE \exp X_t$. **Hint**: Let $p \in (1, \infty)$. Then, it follows from the assumption and Lemma 4.3.3 that $\exp(X_s/p)$, $s \in [0,t] \cap \mathbb{T}$ is a nonnegative submartigale. Thus, applying (5.23) to this submartingale, we have

$$E \exp Y_t \le \left(\tfrac{p}{p-1}\right)^p E \exp X_t.$$

Then, we let $p \to \infty$.

Exercise 5.4.3 ($\star$) Let $\Omega = [0,1)$, $\mathcal{F} = \mathcal{B}(\Omega)$ and P be the Lebesgue measure on $(\Omega, \mathcal{B}(\Omega))$. We let $\|\cdot\|_p$ denote the norm of $L^p(P)$. For $f \in L^1(P)$ and $x \in \Omega$, define

$$Hf(x) = \frac{1}{x} \int_0^x f(y) dy.$$

The objective of this exercise is twofold. The first is to prove **Hardy's inequality**

$$\|Hf\|_p \le \frac{p}{p-1} \|f\|_p \ \text{ if } p \in (1, \infty). \tag{5.38}$$

as an application of Doob's L^p-maximal inequality (5.23). The second is to show that for both (5.23) and (5.38), the multiplicative constant $\frac{p}{p-1}$ cannot be improved.

For $t \in [0,1]$, $x \in \Omega$ and $f \in L^1(P)$, we set

$$\begin{aligned} \mathcal{F}_t &= \{A \in \mathcal{F} \ ; \ \text{either } A \subset [1-t, 1) \text{ or } \Omega \backslash A \subset [1-t, 1)\}, \\ f_t(x) &= (Hf)(1-t) \mathbf{1}_{[0,1-t)}(x) + f(x) \mathbf{1}_{[1-t,1)}(x). \end{aligned}$$

Then, prove the following. **i)** For fixed $x \in \Omega$, $t \mapsto f_t(x)$ is right-continuous, $f_1(x) = f(x)$ and $|Hf(x)| \le \sup_{x<t\le 1} |f_{1-t}(x)|$. **ii)** For fixed $t \in (0,1]$, $E[f|\mathcal{F}_t](x) = f_t(x)$, $P(dx)$-a.s. **iii)** For $p \in [1,\infty)$,

$$\|Hf\|_p \le \left\| \sup_{t\in[0,1]} f_t \right\|_p \le \text{the RHS of (5.38).}$$

iv) For (5.38), the multiplicative constant $\frac{p}{p-1}$ cannot be improved in the following sense. If $1 < p < \infty$ and $c < \frac{p}{p-1}$, there exists $f \in L^p(P)$ such that $\|Hf\|_p > c\|f\|_p$. **Hint**: Let $f(x) = x^{-\delta}$ $(0 < \delta < 1/p)$. Then, $\|Hf\|_p = (1-\delta)^{-1}(1-\delta p)^{-1/p}$, $\|f\|_p = (1-\delta p)^{-1/p}$ for $p \in (1,\infty)$ and $\|f\|_1 = \frac{\delta(2-\delta)}{(1-\delta)^2} \exp(-(1-\delta))$.
v) For (5.23), the multiplicative constant $\frac{p}{p-1}$ cannot be improved in the following sense. If $1 < p < \infty$ and $c < \frac{p}{p-1}$, there exists $f \in L^p(P)$ for which $\| \sup_{t\in[0,1]} f_t\|_p > c\|f_1\|_p$.

5.5 Backwards Martingales

Theorem 5.5.1 *Suppose that $X = (X_n)_{n\in -\mathbb{N}}$ is a submartingale (resp. supermartingale) and that*

$$-\infty < \inf_{n\in -\mathbb{N}} EX_n \quad (\textit{resp.} \ \sup_{n\in -\mathbb{N}} EX_n < \infty). \tag{5.39}$$

Then, there exists $X_{-\infty} \in L^1(P)$ such that

$$X_n \stackrel{n\to -\infty}{\longrightarrow} X_{-\infty} \ \textit{a.s. and in } L^1(P).$$

Proof: By symmerty, we may focus on the case of submartingale. We first prove that

1) $\exists X_{-\infty} = \lim_{n\to -\infty} X_n \in [-\infty, \infty]$ a.s.

Let $a, b \in \mathbb{R}$, $a < b$ and U_n $(n \in -\mathbb{N})$ be the number of upcrossing from a to b by the sequence $X_n, X_{n+1}, ..., X_0$. Noting that $U_{n-1} \ge U_n$ for $\forall n \in -\mathbb{N}$, we set $U_{-\infty} = \lim_{n\to -\infty} U_n \in [0,\infty]$. Then, we have by the argument of Lemma 5.1.6 that

$$EU_n \le E[(X_0 - a)^+] - E[(X_n - a)^+] \le E[(X_0 - a)^+].$$

This implies 1) by the argument in the proof of Theorem 5.1.1.
To prove that $X_n \stackrel{n\to -\infty}{\longrightarrow} X_{-\infty}$ in $L^1(P)$, it is enough to show that $(X_n)_{n\in -\mathbb{N}}$ is uniformly integrable. Noting (5.39) and that $EX_{n-1} \le EX_n$ for $\forall n \in -\mathbb{N}$, we set $m = \lim_{n\to -\infty} EX_n \in \mathbb{R}$. Then, for any $\varepsilon > 0$, there exists $k \in -\mathbb{N}$ such that $m \le EX_n \le m + \varepsilon$ for $\forall n \le k$. We claim for $n \le k$ and $\lambda > 0$ that

1) $\quad P(|X_n| > \lambda) \le (2E[X_0^+] - m)/\lambda$

2) $\quad E[|X_n| : |X_n| > \lambda] \le E[|X_k| : |X_n| > \lambda] + \varepsilon.$

These imply the desired uniform integrability. To prove 1), we note that $m \le EX_n$ and that X_n^+ is a submartingale (Lemma 4.3.3). Hence,

$$E|X_n| = 2E[X_n^+] - EX_n \le 2E[X_0^+] - m.$$

Then, 1) follows from Chebyshev's inequality. To prove 2), we note that

3) $\qquad E[X_n : X_n > \lambda] \le E[X_k : X_n > \lambda]$,

4) $$\begin{cases} E[X_n : X_n < -\lambda] &= EX_n - E[X_n : X_n \ge -\lambda] \\ &\stackrel{2)}{\ge} EX_k - \varepsilon - E[X_k : X_n \ge -\lambda] \\ &= E[X_k : X_n < -\lambda] - \varepsilon. \end{cases}$$

Putting these together,

$$\begin{aligned} E[|X_n| : |X_n| > \lambda] &= E[X_n : X_n > \lambda] - E[X_n : X_n < -\lambda] \\ &\stackrel{3),4)}{\le} E[X_k : X_n > \lambda] - E[X_k : X_n < -\lambda] + \varepsilon \\ &= E[|X_k| : |X_n| > \lambda] + \varepsilon. \end{aligned}$$

\(^□^)/

Remark: Supppose that $X = (X_n)_{n\in-\mathbb{N}}$ is a martingale. Then (5.39) is obviously true. Moreover, by Corollary 5.5.3 below, we have that

$$X_{-\infty} = E[X_0|\mathcal{F}_{-\infty}] \ a.s. \text{ with } \mathcal{F}_{-\infty} = \bigcap_{n\in-\mathbb{N}} \mathcal{F}_n.$$

Corollary 5.5.2 *Let $Y \in L^1(P)$ and $\mathcal{F}_{-\infty} = \bigcap_{n\in-\mathbb{N}} \mathcal{F}_n$. Then,*

$$E[Y|\mathcal{F}_n] \stackrel{n\to\infty}{\longrightarrow} E[Y|\mathcal{F}_{-\infty}] \ a.s. \text{ and in } L^1(P).$$

Proof: The process $X_n = E[Y|\mathcal{F}_n]$ $(n \in -\mathbb{N})$ is a martingale by Example 4.3.2. Thus, by Theorem 5.5.1, there exists an $X_{-\infty} \in L^1(P)$ such that $X_n \stackrel{n\to-\infty}{\longrightarrow} X_{-\infty}$ a.s. and in $L^1(P)$. Thus, it is enough to show that

1) $\qquad X_{-\infty} = E[Y|\mathcal{F}_{-\infty}]$ a.s.

To verify this, we take an arbitrary $A \in \mathcal{F}_{-\infty}$. Then, $A \in \mathcal{F}_n$ for all $n \in -\mathbb{N}$, and thus, $E[X_n : A] = E[Y : A]$. Letting $n \to -\infty$, we have

2) $\qquad E[X_{-\infty} : A] = E[Y : A]$,

which implies 1). \(^□^)/

Corollary 5.5.3 *Suppose that $X = (X_n)_{n\in-\mathbb{N}}$ is a submartingale (resp. supermartingale) and that $X_n \stackrel{n\to-\infty}{\longrightarrow} X_{-\infty}$ in $L^1(P)$. Then,*

$$X_{-\infty} \le E[X_0|\mathcal{F}_{-\infty}] \quad (\textit{resp. } X_{-\infty} \ge E[X_0|\mathcal{F}_{-\infty}]) \ a.s.,$$

where $\mathcal{F}_{-\infty} = \bigcap_{n\in-\mathbb{N}} \mathcal{F}_n$.

Proof: Suppose that $X = (X_n)_{n\in-\mathbb{N}}$ is a submartingale. Then, for all $n \in -\mathbb{N}$

$$X_n \le E[X_0|\mathcal{F}_n] \ a.s.$$

$X_n \stackrel{n\to-\infty}{\longrightarrow} X_{-\infty}$ in $L^1(P)$ by assumption. Moreover, $E[X_0|\mathcal{F}_n] \stackrel{n\to-\infty}{\longrightarrow} E[X_0|\mathcal{F}_{-\infty}]$ in $L^1(P)$ by Corollary 5.5.2. Threrefore, the result follows from Exercise 1.10.1. \(^□^)/

5.6 (⋆) Structure of L^1-bounded martingales II

Let u be a real harmonic function on the unit open disc D such that

$$\sup_{0<r<1} \int_{-\pi}^{\pi} |u(re^{\mathbf{i}\theta})| d\theta < \infty.$$

Then, there exists a unique Borel signed measure μ on $[-\pi, \pi]$ such that

$$u(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} h(z, e^{\mathbf{i}\theta}) d\mu(\theta) \ \text{ for all } z \in D,$$

where $h(z,w) = \frac{|w|^2-|z|^2}{|w-z|^2}$, cf. [Rud87, p.247,11.30]. Then, let $d\mu(\theta) = f(\theta)d\theta + \mathbf{1}_N(\theta)d\mu(\theta)$ be Lebesgue decomposition of μ with respect to the Lebesgue measure, where $f \in L^1([-\pi,\pi])$ and the signed measure and $N \subset [-\pi, \pi]$ is a Borel set with zero-Lebesgue measure.

$$u(re^{\mathbf{i}\theta}) \stackrel{r \nearrow \infty}{\longrightarrow} f(\theta) \ \text{ for almost all } \theta \in [-\pi, \pi],$$

cf. [Rud87, p.244,11.24].

We will explain that an L^1-bounded martingale X has an analogous properties.

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be a filtration such that $\mathcal{F} = \mathcal{F}_\infty \stackrel{\text{def}}{=} \sigma\left[\bigcup_{n\in\mathbb{N}} \mathcal{F}_n\right]$. Suppose that $X = (X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ is a martingale such that $\sup_{n\in\mathbb{N}} \|X_n\|_1 < \infty$. Then, there exists a signed measure Q on $(\Omega, \mathcal{F})$, such that $Q_n \ll P_n$ for all $n \in \mathbb{N}$, where $P_n = P|_{\mathcal{F}_n}$, $Q_n = Q|_{\mathcal{F}_n}$, and that $X_n = \frac{dQ_n}{dP_n}$, cf. Proposition 4.7.1. Moreover, by Theorem 5.1.1, there exists $X_\infty \in L^1(P)$ such that $X_n \stackrel{n\to\infty}{\longrightarrow} X_\infty$ P-a.s.

The signed measure Q and the r.v. X_∞ is related as follows.

Proposition 5.6.1 *Referring to the setting explained before the proposition, the following hold.*

a) *The conditions (5.5)–(5.8) in Theorem 5.2.1 for the martingale $X = (X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ are also equivalent to that $Q \ll P$. Moreover, if $Q \ll P$, then, $X_\infty = \frac{dQ}{dP}$.*

b) *There exists an $N \in \mathcal{F}$ such that $P(N) = 0$ and $dQ = X_\infty dP + \mathbf{1}_N dQ$.*

Proof: a) Suppose the condition (5.5) of Theorem 5.2.1 and let $\widetilde{Q}(A) = E[Y : A]$ $(A \in \mathcal{F})$. Then, for any $n \in \mathbb{N}$ and $A \in \mathcal{F}_n$,

$$\widetilde{Q}(A) = E[Y : A] = E[X_n : A] = Q(A).$$

Since n is arbitrary, it follows from Dynkin's Lemma (Lemma 1.3.1) that $\widetilde{Q} = Q$ on $\mathcal{F}_\infty = \mathcal{F}$. Thus, $Q \ll P$ and $dQ/dP = X_\infty$.

Suppose on the other hand that $Q \ll P$. Then, $Q(A) = E[\frac{dQ}{dP} : A]$ $(A \in \mathcal{F})$. Thus, by the definition of the conditional expectation (cf. (4.14)), we have that $E[\frac{dQ}{dP}|\mathcal{F}_n] = \frac{dQ_n}{dP_n} = X_n$ $(\forall n \in \mathbb{N})$. Moreover, since $\mathcal{F}_\infty = \mathcal{F}$, it follows from Corollary 5.2.2 that $X_n \stackrel{n\to\infty}{\longrightarrow} \frac{dQ}{dP}$ P-a.s. and in $L^1(P)$.

b) Let $Q^\pm$ be the positive and the negative parts of the Jordan decomposition of Q. Then, $X_n^\pm = dQ_n^\pm/dP_n$. Hence it is enough to prove the decomposition $dQ = X_\infty dP + \mathbf{1}_N dQ$ for

$Q^{\pm}$ separately. Therefore, we may assume that Q is a positive finite measure. If $Q = 0$, then $X_n \equiv 0$ and hence the decomposition $dQ = X_\infty dP + \mathbf{1}_N dQ$ holds with $N = \emptyset$. If $Q \neq 0$, then, by considering $Q(\cdot)/Q(\Omega)$ instead of Q, we may assume that $Q(\Omega) = 1$.

Let Q be a probability measure. Then, there exists an $N_1 \in \mathcal{F}$ such that $P(N_1) = 0$ and on $\Omega \backslash N_1$,

1) X_n $(n \in \mathbb{N} \cup \{\infty\})$ are well-defined and $X_n \stackrel{n\to\infty}{\longrightarrow} X_\infty$.

Let $R = \frac{P+Q}{2}$, $R_n = \frac{P_n+Q_n}{2}$, $n \in \mathbb{N}$. Note that $P \ll R$, $Q \ll R$, and $Q_n \ll P_n \ll R_n$, $n \in \mathbb{N}$. Let also $Y_n = \frac{dP_n}{dR_n}$ and $Z_n = \frac{dQ_n}{dR_n}$. Then, by Exercise 4.1.1, R-almost surely, X_n, Y_n, Z_n are well-defined and $Z_n = X_n Y_n$. Also, by part a), $Y_n \stackrel{n\to\infty}{\longrightarrow} \frac{dP}{dR}$, R-a.s. and $Z_n \stackrel{n\to\infty}{\longrightarrow} \frac{dQ}{dR}$, R-a.s. Therefore, there exists an $N_2 \in \mathcal{F}$ such that $R(N_2) = 0$ and on $\Omega \backslash N_2$,

2) X_n, Y_n, Z_n $(n \in \mathbb{N})$, $\frac{dP}{dR}$, $\frac{dQ}{dR}$ are well-defined, $Z_n = X_n Y_n$, $Y_n \stackrel{n\to\infty}{\longrightarrow} \frac{dP}{dR}$, and $Z_n \stackrel{n\to\infty}{\longrightarrow} \frac{dQ}{dR}$.

Let $N = N_1 \cup N_2$. Then, $P(N) = 0$ and on $\Omega \backslash N$, both 1) and 2) are true. Therefore, on $\Omega \backslash N$, we have that

3) $\frac{dQ}{dR} = X_\infty \frac{dP}{dR}$.

For $A \in \mathcal{F}$, we have that

$$Q(A \backslash N) = \int_{A\backslash N} \frac{dQ}{dR} dR \stackrel{3)}{=} \int_{A\backslash N} X_\infty \frac{dP}{dR} dR = E[X_\infty : A \backslash N] = E[X_\infty : A],$$

from which we conclude that $dQ = X_\infty dP + \mathbf{1}_N dQ$. \(^□^)/

Example 5.6.2 (**Kakutani's dichotomy**) Let $(S_n, \mathcal{B}_n)$, $n \in \mathbb{N}\backslash\{0\}$ be measurable spaces, $\mu_n, \nu_n \in \mathcal{P}(S_n, \mathcal{B}_n)$, $P = \otimes_{n=1}^\infty \mu_n$, and $Q = \otimes_{n=1}^\infty \nu_n$. Suppose that $\nu_n \ll \mu_n$ for all $n \in \mathbb{N}\backslash\{0\}$. Then,

$$\alpha \stackrel{\text{def}}{=} \prod_{n=1}^\infty \int \sqrt{\tfrac{d\nu_n}{d\mu_n}} d\mu_n \begin{cases} > 0 & \Rightarrow \quad Q \ll P, \\ = 0 & \Rightarrow \quad Q \perp P. \end{cases}$$

Proof: Let $(\Omega, \mathcal{F}) = \prod_{n=1}^\infty (S_n, \mathcal{B}_n)$ and $\xi_n(\omega) = \frac{d\nu_n}{d\mu_n}(\omega_n)$ for $\omega = (\omega_n)_{n=1}^\infty$. Then, $\xi_n \geq 0$, $n \in \mathbb{N}\backslash\{0\}$ are mean-one independent r.v.'s on $(\Omega, \mathcal{F}, P)$ and hence $X_n = \prod_{j=1}^n \xi_j$ is a nonnegative martingale. Moreover,

$$Q(A) = E[X_n : A] \quad \text{for all} \in \mathbb{N}\backslash\{0\} \text{ and } A \in \mathcal{F}_n = \sigma(\xi_1, \ldots, x_n).$$

Suppose first that $\alpha > 0$. Then, by Example 5.2.3, X_n converges in $L^1(P)$, which implies via Proposition 5.6.1 that $Q \ll P$.

Suppose on the other hand that $\alpha = 0$. Then, by Example 5.2.3, $X_\infty = 0$ a.s., which implies via Proposition 5.6.1 that $Q \perp P$. \(^□^)/

6 Brownian Motion and its Markov Property

6.1 Definition, and Some Basic Properties

The Brownian motion came into the history in 1827, when R. Brown, a British botanist, observed that pollen grains suspended in water perform a continual swarming motion. In 1905, A. Einstein derived (6.3) below from the moleculer physics point of view. A mathematically rigorous construction with a proof of the continuity (cf. B2) below) was given by N. Wiener (1923).

We fix a probability space $(\Omega, \mathcal{F}, P)$ in this subsection. In the sequel, we will repeatedly refer to a finite time series of the form

$$0 = t_0 < t_1 < ... < t_n, \quad n \geq 1. \tag{6.1}$$

Definition 6.1.1 (**Brownian motion**) Let $B = (B_t : \Omega \to \mathbb{R}^d)_{t \geq 0}$ be a family r.v.'s. We consider the following conditions.

B1) For any time series (6.1),

$$B(0), B(t_1) - B(0), \ldots, B(t_n) - B(t_{n-1}) \text{ are independent,} \tag{6.2}$$

$$B(t_j) - B(t_{j-1}) \approx N(0, (t_j - t_{j-1})I_d), \quad j = 1, \ldots, n, \tag{6.3}$$

where I_d is the identity matrix of degree d (cf. Example 1.2.4),

B2) There is an $\Omega_B \in \mathcal{F}$ such that $P(\Omega_B) = 1$ and $t \mapsto B_t(\omega)$ is continuous for all $\omega \in \Omega_B$.

B3) $B_0 = x$, for a nonrandom vector $x \in \mathbb{R}^d$.

► B is called a d-dimensional *Brownian motion* (BM^d for short) if the conditions B1), B2) are satisfied.

► B is called a d-dimensional *Brownian motion* started at x (BM^d_x for short), if the conditions B1)–B3) are satisfied.

► B is called a d-dimensional *pre-Brownian motion* (pre-BM^d for short), if the conditions B1) is satisfied. A d-dimensional *pre-Brownian motion* is said to be started at x, if it saitesfies B3) and is abbreviated by pre-BM^d_x.

Remark: 1) B2) does not follow from B1). In fact, there exists a pre-BM^1_0 $(B_t)_{t \geq 0}$ which is almost surely discontinuos at all $t \geq 0$ (Example 6.6.9). 2) If the condition B2) above is replaced by the following stronger one, B is called an *continuous modification* of a BM^d.

$$t \mapsto B_t(\omega) \text{ is continuous for all } \omega \in \Omega. \tag{6.4}$$

In some text books (e.g. [Bil95, p.503], [IkWa89, p.40], [KS91, p.47], [LeG16, p.27]), instead of B1)–B2) above, B1), B2) and (6.4) are adopted as the definition of the Brownian motion. However, there is no essential difference between B2) and (6.4). Suppose that B satisfies B1)–B2) and define $\widetilde{B}$ by

$$\widetilde{B}_t(\omega) = \begin{cases} B_t(\omega) & \text{if } \omega \in \Omega_B, \ t \geq 0, \\ B_0(\omega) & \text{if } \omega \notin \Omega_B, \ t \geq 0. \end{cases}$$

Then, $\widetilde{B}$ satisfies B1),B2) and (6.4).
3) In some text books (e.g. [Bil95, p.498], [KS91, p.47]), "$B_0 = x$" in the condition B4) above is replaced by "$B_0 = x$, a.s."

Lemma 6.1.2 *Suppose that $\widetilde{B}$ is a BM_0^d and that $X : \Omega \to \mathbb{R}^d$ is a r.v. independent of B Then, $B \stackrel{\text{def}}{=} (X + \widetilde{B}_t)_{t\geq 0}$ is a BM^d such that $B_0 = X$.*

Proof: Obvious from Definition 6.1.1. \(^□^)/

Recall that r.v.'s $\{X_j\}_{j=1}^m$ is called Gaussian r.v.'s if there exist i.i.d. $Z_1, \ldots, Z_n \approx N(0,1)$ such that each X_j $(j = 1, \ldots, m)$ is a linear combination of $Z_1, \ldots, Z_n$.

Lemma 6.1.3 *Referring to Definition 6.1.1, the condition B1) is equivalent to each of the following conditions*
B1') *For any time series (6.1), the r.v's*

$$X_j^\alpha \stackrel{\text{def}}{=} B^\alpha(t_j) - B^\alpha(t_{j-1}), \quad \alpha = 1, \ldots, d, \ j = 1, \ldots, n.$$

are independent and $X_j^\alpha \approx N(0, t_j - t_{j-1})$ for all $\alpha = 1, \ldots, d$ and $j = 1, \ldots, n$.
B1") *For any time series (6.1), $\{B^\alpha(t_k)\}_{\substack{1\leq\alpha\leq d\\1\leq k\leq n}}$ are mean-zero Gaussian r.v.'s such that*

$$\mathrm{cov}(B^\alpha(t_k), B^\beta(t_\ell)) = \delta_{\alpha,\beta} t_k \ \text{for all } \alpha, \beta = 1, ..., d \text{ and } 1 \leq k \leq \ell \leq n. \tag{6.5}$$

Proof: B1) $\Leftrightarrow$ B1'): This is because for each j,

$$B(t_j) - B(t_{j-1}) \approx N(0, (t_j - t_{j-1})I_d)$$

iff $\{X_j^\alpha\}_{1\leq\alpha\leq d}$ are independent and $X_j^\alpha \approx N(0, t_j - t_{j-1})$ for all $\alpha = 1, \ldots, d$.
B1') $\Rightarrow$ B1"): By B1'), $Z_j^\alpha \stackrel{\text{def}}{=} X_j^\alpha / \sqrt{t_j - t_{j-1}}$ $(1 \leq \alpha \leq d, 1 \leq j \leq n)$ are i.i.d., $\approx N(0,1)$. Thus, $\{B^\alpha(t_k)\}_{\substack{1\leq\alpha\leq d\\1\leq k\leq n}}$ are mean-zero Gaussian r.v.'s, since

$$B^\alpha(t_k) = \sum_{j=1}^k X_j^\alpha = \sum_{j=1}^k \sqrt{t_j - t_{j-1}} Z_j^\alpha,$$

for $\alpha = 1, \ldots, d$ and $k = 1, \ldots, n$. Moreover,

$$\mathrm{cov}(B^\alpha(t_k), B^\beta(t_\ell)) = \delta_{\alpha,\beta} \sum_{j=1}^k \mathrm{cov}(X_j^\alpha, X_j^\alpha) = \delta_{\alpha,\beta} \sum_{j=1}^k (t_j - t_{j-1}) = \delta_{\alpha,\beta} t_k$$

B1") $\Rightarrow$ B1'): Since $\{B^\alpha(t_k)\}_{\substack{1\leq\alpha\leq d\\1\leq k\leq n}}$ are mean-zero Gaussian r.v.'s, so are $\{X_j^\alpha\}_{\substack{1\leq\alpha\leq d\\1\leq j\leq n}}$. Moreover, for $\alpha, \beta = 1, ..., d$ and $1 \leq k \leq \ell \leq n$,

$$\begin{aligned}
&\mathrm{cov}(X_k^\alpha, X_\ell^\beta) \\
&= E[(B^\alpha(t_k) - B^\alpha(t_{k-1}))(B^\beta(t_\ell) - B^\beta(t_{\ell-1}))] \\
&= EB^\alpha(t_k)B^\beta(t_\ell) - EB^\alpha(t_k)B^\beta(t_{\ell-1}) - EB^\alpha(t_{k-1})B^\beta(t_\ell) + EB^\alpha(t_{k-1})B^\beta(t_{\ell-1}) \\
&\stackrel{\text{B1")}}{=} \delta_{\alpha,\beta}(t_k - t_k \wedge t_{\ell-1} - t_{k-1} + t_{k-1}) = \delta_{\alpha,\beta}\delta_{k,\ell}(t_k - t_{k-1}).
\end{aligned}$$

By Exercise 2.2.6, this implies B1'). \(^□^)/

We note that the Brownian motion can be defined in a different way.

Proposition 6.1.4 *Referring to Definition 6.1.1, let $B^\alpha = (B^\alpha_t)_{t\geq 0}$, $\alpha = 1, \ldots, d$ be the α-th coordinate of B. Then, the following conditions are equivalent.*
a) *B is a BM^d_0.*
b) *$B^1, ..., B^d$ are independent and each of them is a BM^1_0.*

Proof: The equivalence of a) and b) follows easily from that of B1) and B1') of Lemma 6.1.3. \(^□^)/

The following invariance property of the Brownian motion allows us to investigate its behavior as time $t \to \infty$ via that as time $t \to 0$, and vice versa.

Proposition 6.1.5 (Time inversion) *Let B be a BM^d. Define $\check{B} = (\check{B}_t)_{t\geq 0}$ by*

$$\check{B}_t = \begin{cases} B_0 + t(B_{1/t} - B_0), & \text{if } t > 0, \\ B_0, & \text{if } t = 0. \end{cases} \tag{6.6}$$

Then, $\check{B}$ is a BM^d such that $\check{B}_0 = B_0$.

Let us prove Proposition 6.1.5. Note that B_0 and $(B_{1/t} - B_0)_{t>0}$ are independent. Hence, by Lemma 6.1.2, it is enough to consider the case of $B_0 \equiv 0$. We first verify the following

Lemma 6.1.6 *Let B be a pre-BM^d_0. Then, so is the process $\check{B}$ defined by (6.6).*

Proof: We take arbitrary time sequence of the form (6.1). By Proposition 6.1.4, it is enough to show that

1) $(\check{B}^\alpha(t_j))_{\substack{1\leq\alpha\leq d\\1\leq j\leq n}}$ are Gaussian r.v.'s which satisfies (6.5).

We know that

2) $(B^\alpha(t_j))_{\substack{1\leq\alpha\leq d\\1\leq j\leq n}}$ are Gaussian r.v.'s which satisfies (6.5).

Since $0 < 1/t_n < 1/t_{n-1} < ... < 1/t_1$, $(B^\alpha(1/t_j))_{\substack{1\leq\alpha\leq d\\1\leq j\leq n}}$ is a mean-zero Gaussian r.v. by 2), and hence so is $(\check{B}^\alpha(t_j))_{\substack{1\leq\alpha\leq d\\1\leq j\leq n}} = (t_j B^\alpha(1/t_j))_{\substack{1\leq\alpha\leq d\\1\leq j\leq n}}$. Moreover, for $1 \leq k \leq \ell \leq n$ and $\alpha, \beta = 1, ..., d$,

$$\begin{aligned} \mathrm{cov}(\check{B}^\alpha(t_k), \check{B}^\beta(t_\ell)) &= t_k t_\ell E[B^\alpha(1/t_k)B^\beta(1/t_\ell)] \stackrel{2)}{=} \delta_{\alpha,\beta} t_k t_\ell \cdot t_\ell^{-1} \\ &= \delta_{\alpha,\beta} t_k. \end{aligned}$$

Thus, we have verified 1). \(^□^)/

To prove the continuity of $\check{B}(t)$ at $t = 0$, we prepare the following

Lemma 6.1.7 *For $f \in C((0,1) \to \mathbb{R})$,*

$$\overline{\lim_{t \to 0+}} f(t) = \overline{\lim_{\substack{r \to 0+ \\ r \in \mathbb{Q}}}} f(r), \quad \underline{\lim_{t \to 0+}} f(t) = \underline{\lim_{\substack{r \to 0+ \\ r \in \mathbb{Q}}}} f(r).$$

In particular, for $c \in [-\infty, \infty]$,

$$f(t) \stackrel{t \to 0+}{\longrightarrow} c \iff f(r) \stackrel{r \in \mathbb{Q},\, r \to 0+}{\longrightarrow} c.$$

Proof: Since the first and second equalities are equivalent, we only prove the first one. As for the first equality, note that

$$\text{LHS} = \lim_{\delta \to 0+} \sup_{t \in (0,\delta)} f(t), \quad \text{RHS} = \lim_{\delta \to 0+} \sup_{r \in (0,\delta) \cap \mathbb{Q}} f(r).$$

Thus, it is enough to verify that

1) $\sup_{t \in (0,\delta)} f(t) = \sup_{r \in (0,\delta) \cap \mathbb{Q}} f(r)$ for any $0 < \delta \leq 1$.

To prove 1), we have only to show that LHS $\leq$ RHS, since the opposite inequality is obvious. Let $c <$LHS, then, there exists $t \in (0,\delta)$ such that $c < f(t)$. Then, by the continuity, there exists $r \in (0,\delta) \cap \mathbb{Q}$ such that $c < f(r)$. Hence $c <$RHS of 1). Since c is arbitrary, we see that LHS $\leq$ RHS. \(^□^)/

Lemma 6.1.8 (Removability of isolated discontinuity) *Let $X = (X_t)_{t \geq 0}$ and $Y = (Y_t)_{t \geq 0}$ be two processes with values in $\mathbb{R}^d$ with the same law. Suppose that there exists $\Omega_X \in \mathcal{F}$ with $P(\Omega_X) = 1$ such that*
a) *$t \mapsto X_t(\omega)$ is continuous on $[0,\infty)$ for all $\omega \in \Omega_X$.*
b) *$t \mapsto Y_t(\omega)$ is continuous on $(0,\infty)$ for all $\omega \in \Omega_X$.*
Then, there exists $\Omega_Y \in \mathcal{F}$ with $P(\Omega_Y) = 1$ such that $t \mapsto Y_t(\omega)$ is continuous on $[0,\infty)$ for all $\omega \in \Omega_Y$.

Proof: Let

$$\begin{aligned} C_Y &= \left\{ Y_t - Y_0 \stackrel{t \to 0+}{\longrightarrow} 0 \right\}, \\ C_{X,\mathbb{Q}} &= \left\{ X_r - X_0 \stackrel{r \in \mathbb{Q},\, r \to 0+}{\longrightarrow} 0 \right\}, \quad C_{Y,\mathbb{Q}} = \left\{ Y_r - Y_0 \stackrel{r \in \mathbb{Q},\, r \to 0+}{\longrightarrow} 0 \right\}. \end{aligned}$$

It is enough to prove that

1) there exists $\Omega_Y \in \mathcal{F}$ with $P(\Omega_Y) = 1$ such that $\Omega_Y \subset C_Y$.

We will show this with $\Omega_Y \stackrel{\text{def}}{=} \Omega_X \cap C_{Y,\mathbb{Q}}$. We first verify that

2) $C_{X,\mathbb{Q}}, C_{Y,\mathbb{Q}} \in \mathcal{F}$.

Indeed,

$$C_{X,\mathbb{Q}} = \bigcap_{\substack{n \in \mathbb{N} \\ n \geq 1}} \bigcup_{\substack{m \in \mathbb{N} \\ m \geq 1}} \bigcap_{\substack{r \in (0,1/m) \\ r \in \mathbb{Q}}} \{|X_r - X_0| < 1/n\} \in \mathcal{F}.$$

Similarly, $C_{Y,\mathbb{Q}} \in \mathcal{F}$.

Now, $P(C_{X,\mathbb{Q}}) = 1$ by a) and hence $P(C_{Y,\mathbb{Q}}) = 1$, by 2) and $X \approx Y$. Therefore,

3) $\Omega_X \cap C_{Y,\mathbb{Q}} \in \mathcal{F}$, $P(\Omega_X \cap C_{Y,\mathbb{Q}}) = 1$.

On the other hand, b) and Lemma 6.1.7 implies that

4) $\Omega_X \cap C_{Y,\mathbb{Q}} = \Omega_X \cap C_Y \subset C_Y$.

3) and 4) implies 1) with $\Omega_Y \stackrel{\text{def}}{=} \Omega_X \cap C_{Y,\mathbb{Q}}$. \(^□^)/

Proof of Proposition 6.1.5: As is already explained, it is enough to consider the case of BM_0^d. Then, by Lemma 6.1.6, it is enough to verify the continuity of $\check{B}_t$ in $t \geq 0$. Recall that there exists an $\Omega_B \in \mathcal{F}$ such that $P(\Omega_B) = 1$ and $t \mapsto B_t(\omega)$ is continuous for all $\omega \in \Omega_B$ Then, for $\omega \in \Omega_B$, $\check{B}_t(\omega)$ is continuous at all $t > 0$. Therefore, the desired continuity follows from Lemma 6.1.8. \(^□^)/

For BM^d, we define the *canonical filtration* $(\mathcal{F}_t^0)_{t\geq 0}$ by

$$\mathcal{F}_t^0 = \sigma(B_s \; ; \; s \leq t). \tag{6.7}$$

The independence of the increments of the Brownian motion has the following consequence.

Proposition 6.1.9 (**Markov property I**) *Let B be a BM^d and $s \geq 0$. Define*

$$\widehat{B}^s = (\widehat{B}_t^s)_{t\geq 0} = (B_{s+t} - B_s)_{t\geq 0}. \tag{6.8}$$

Then,
a) *$\widehat{B}^s$ is a BM_0^d,*
b) *$\mathcal{F}_s^0$ and $\widehat{B}^s$ are independent.*

Proof: a) Clearly, $\widehat{B}_0^s = 0$, and $t \mapsto \widehat{B}_t^s$ is a.s. continuous. Let $0 \leq u < t$. Then,

$$\widehat{B}_t^s - \widehat{B}_u^s = B_{s+t} - B_{s+u}.$$

Hence, the increments of $\widehat{B}^s$ are independent and their laws are the same as those for B. Thus, $\widehat{B}^s$ is a BM_0^d.
b) We take $0 = r_0 < ... < r_m \leq s$ and $0 = t_0 < ... < t_n$. Then, it is enough to verify that

1) $(B(r_k))_{k=1}^m$ and $(\widehat{B}^s(t_\ell))_{\ell=1}^n$ are independent.

(cf. Lemma 1.6.5). Let

$$X \stackrel{\text{def}}{=} (B(r_j) - B(r_{j-1}))_{j=1}^m \quad \text{and} \quad Y \stackrel{\text{def}}{=} (B(s+t_j) - B(s+t_{j-1}))_{j=1}^n$$

We see from B1) in Definition 6.1.1 that

2) B_0, X and Y are independent.

Moreover, for $k = 1, ..., m$ and $\ell = 1, ..., n$,

$$B(r_k) = B_0 + \sum_{j=1}^{k}(B(r_j) - B(r_{j-1})), \quad \widehat{B}^s(t_\ell) = \sum_{j=1}^{\ell}(B(s+t_j) - B(s+t_{j-1})).$$

Thus,

3) $(B(r_k))_{k=1}^m$ is $\sigma(B_0, X)$-measurable, and $(\widehat{B}^s(t_\ell))_{\ell=1}^n$ is $\sigma(Y)$-measurable.

Now, 1) follows from 2) and 3). \(^□^)/

The Markov property implies that the past and the future are independent, given the present.

Corollary 6.1.10 *Let $s \geq 0$, $F \in \mathcal{F}_s^0$, and $G \in \mathcal{T}_s \stackrel{\text{def}}{=} \sigma(B_t \; ; \; t \geq s)$. Then,*

$$P(G|\mathcal{F}_s^0) = P(G|B_s), \quad a.s. \tag{6.9}$$

$$P(F \cap G|B_s) = P(F|B_s)P(G|B_s), \quad a.s. \tag{6.10}$$

Proof: Note that there exists $\Gamma \in \mathcal{B}((\mathbb{R}^d)^{[0,\infty)})$ such that

1) $G = \{(B_{s+t})_{t\geq 0} \in \Gamma\} = \{(B_s + B_t^s)_{t \geq 0} \in \Gamma\}$.

Note also that the following function is Borel measurable.

$$f(x) = P((x + B_t^s)_{t\geq 0} \in \Gamma), \quad x \in \mathbb{R}^d.$$

Since $\mathcal{F}_s^0$ and $(B_t^s)_{t\geq 0}$ is independent, we see from Exercise **??** that

$$P(G|\mathcal{F}_s^0) = f(B_s), \quad \text{a.s.}$$

In particular, $P(G|\mathcal{F}_s^0)$ is $\sigma(B_s)$-measurable, which implies (6.9). Then,

$$P(F \cap G|\mathcal{F}_s^0) = \mathbf{1}_F P(G|\mathcal{F}_s^0) \stackrel{(6.9)}{=} \mathbf{1}_F P(G|B_s), \quad \text{a.s.}$$

By taking the conditional expectations given $\sigma(B_s)$ of both hands sides of the above identity, we get (6.10). \(^□^)/

Let B be a BM^d and $s > 0$. The Markov property allows us to construct a new Brownian motion by replacing the path after the time s by an another Brownian motion β, which is independent of $\mathcal{F}_s$. More precisely, we have

Corollary 6.1.11 (**Concatenation of Brownian motions I**) *Let B be a BM^d, $s > 0$, and β be a BM_0^d which is independent of $\mathcal{F}_s^0$. Then the process $\widetilde{B} = (\widetilde{B}_t)_{t\geq 0}$ defined as follows is a BM^d.*

$$\widetilde{B}_t = \begin{cases} B_t, & \text{if } t \leq s, \\ B_s + \beta_{t-s}, & \text{if } t \geq s. \end{cases} \tag{6.11}$$

As a consequence, the Brownian motion β is expressed as

$$\beta_t = \widetilde{B}_{s+t} - \widetilde{B}_s, \quad t \geq 0.$$

Proof: Let $S = (\mathbb{R}^d)^{[0,\infty)}$ and define $F : S \times S \longrightarrow S$ by

$$F(x, y)(t) = \begin{cases} x(t), & \text{if } t \leq s, \\ x(s) + y(t-s), & \text{if } t \geq s. \end{cases}$$

Define also $X : \Omega \to S$ and $\widehat{B}^s : \Omega \to S$ by

$$X = (B_{t\wedge s})_{t\geq 0}, \quad \widehat{B}^s = (B_{t+s} - B_s)_{t\geq 0}.$$

Then,

1) $B = F(X, \widehat{B}^s)$, $\widetilde{B} = F(X, \beta)$.

Then, X is $\mathcal{F}_s^0$-measurable, and hence by assumption, β is a BM_0^d which is independent of X. On the other hand, we see from Proposition 6.1.9 that $\widehat{B}^s$ is a BM_0^d which is independent of X. As a consequence,

2) $(X, \widehat{B}^s) \approx (X, \beta)$.

This, together with 1), implies that $B \approx \widetilde{B}$. \(^□^)/

($\star$) Complement to section 6.1

We will prove that a BM_0^d exists on a suitable probability space $(\Omega, \mathcal{F}, P)$. Once we are given a BM_0^d, then, we can construct many other BM_0^d's (Exercise 6.1.2). However, "the law of BM_0^d is unique" in the following sense.

Proposition 6.1.12 (Uniqueness of the law of pre-BM^d) *Let $S = (\mathbb{R}^d)^{[0,\infty)}$ and let $\mathcal{B}(S)$ be its product σ-algebra (cf. Definition 1.5.1).*

a) *Suppose that B is a pre-BM_x^d. Then the map $\omega \mapsto (B_t(\omega))_{t \geq 0}$ $((\Omega, \mathcal{F}) \longrightarrow (S, \mathcal{B}(S))$ is measurable.*

b) *Suppose that B and $\widetilde{B}$ are pre-BM_x^d's. Then, their laws on $(S, \mathcal{B}(S))$ induced by the maps $\omega \mapsto (B_t(\omega))_{t \geq 0}$ and $\omega \mapsto (\widetilde{B}_t(\omega))_{t \geq 0}$ are the same;*

$$P((B_t)_{t \geq 0} \in A) = P((\widetilde{B}_t)_{t \geq 0} \in A) \quad \text{for all } A \in \mathcal{B}(S). \tag{6.12}$$

Proof: a): This follows from Lemma 1.5.2.
b): For time series of the form (6.1), the r.v.'s $(B(t_j))_{j=1}^n$ and $(\widetilde{B}(t_j))_{j=1}^n$ have the same law described in Proposition 6.1.4c). This proves (6.12) for all cylinder set $A \subset S$, and hence for all $A \in \mathcal{B}(S)$ (Lemma 1.5.4). \(^□^)/

Here is a variant of Proposition 6.1.12, which concerns a continuous modification of BM_x^d (cf. Definition 6.1.1).

Corollary 6.1.13 *Let $(S, \mathcal{B}(S))$ be as in Proposition 6.1.12 and let*

$$\begin{aligned} W &= \{w = (w_t)_{t \geq 0} \in S \ ; \ t \mapsto w_t \text{ is continuous}\}, \\ \mathcal{B}(W) &= \{A \cap W \ ; \ A \in \mathcal{B}(S)\}. \end{aligned}$$

a) *Suppose that B is a continuous modification of BM_x^d (cf. Definition 6.1.1). Then the map $\omega \mapsto (B_t(\omega))_{t \geq 0}$ from $(\Omega, \mathcal{F})$ to $(W, \mathcal{B}(W))$ is measurable.*

b) *Suppose that B and $\widetilde{B}$ are two continuous modifications of BM_x^d. Then, their laws on $(W, \mathcal{B}(W))$ induced by the maps $\omega \mapsto (B_t(\omega))_{t \geq 0}$ and $\omega \mapsto (\widetilde{B}_t(\omega))_{t \geq 0}$ are the same;*

$$P((B_t)_{t \geq 0} \in A) = P((\widetilde{B}_t)_{t \geq 0} \in A) \quad \text{for all } A \in \mathcal{B}(W). \tag{6.13}$$

Proof: a): This follows from Lemma 1.5.8.
b): This follows from the same argument as in Proposition 6.1.12, using Lemma 1.5.9 instead of Lemma 1.5.4. \(^□^)/

Remark: The unique law (6.13) on $(W, \mathcal{B}(W))$ of a continuous modification of a Brownian motion is called the *Wiener measure*. We note that $W \notin \mathcal{B}(S)$. In fact, suppose that $W \in \mathcal{B}(S)$, then, by Corollary 1.5.7, there exists an at most countable set $\Gamma \subset [0,\infty)$ with the following property.

1) $x \in S,\ y \in W,\ x_t = y_t$ for all $t \in \Gamma \implies x \in W$.

However, for any $y \in W$ and for any at most countable $\Gamma \subset [0,\infty)$, we can always find an $x \notin W$ (i.e., $t \mapsto x_t$ is discontinuous) such that $x_t = y_t$ for all $t \in \Gamma$. Therefore the set W does not have the property 1).

Lemma 6.1.14 *Let B be a* BM^d, $S = (\mathbb{R}^d)^{[0,\infty)}$ *and $\mathcal{B}(S)$ be the product σ-algebra of S. Then,*

a) *The map $(x,\omega) \mapsto x + B = (x + B_t(\omega))_{t\geq 0}$ is $(\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{F})/\mathcal{B}(S)$-measurable.*

b) *Let $F : S \to \mathbb{R}$ be bounded, $\mathcal{B}(S)$-measurable. Then, the function*

$$\mathbb{R}^d \ni x \mapsto EF(x+B)$$

is Borel measurable.

Proof: a) By Lemma 1.5.2, it is enough to verify that the map $(x,\omega) \mapsto x + B_t$ is $(\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{F})/\mathcal{B}(\mathbb{R}^d)$-measurable for each fixed $t \geq 0$. But this is obvious, since the map $(x,\omega) \mapsto x + B_t$ is a composition of

$$(x,\omega) \mapsto (x, B_t) \ \text{ and } \ (x,y) \mapsto x+y,$$

which are $(\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{F})/\mathcal{B}(\mathbb{R}^{2d})$-measurable and $\mathcal{B}(\mathbb{R}^{2d})/\mathcal{B}(\mathbb{R}^d)$-measurable, respectively.
b) It follows from a) that $(x,\omega) \mapsto F(x+B)$ is $\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{F}$-measurable. Thus, the measurability in question follows from a standard argument (Exercise 6.1.12). \(^□^)/

Exercise 6.1.1 Let B be a BM^d_x, and

$$h_t(x) = (2\pi t)^{-d/2}\exp\left(-\frac{|x|^2}{2t}\right), \quad t > 0, \quad x \in \mathbb{R}^d. \tag{6.14}$$

Then, prove that

$$\begin{aligned} &P(B_{t_1} \in A_1, \ldots, B_{t_n} \in A_n) \\ &= \int_{A_1} h_{t_1}(x_1 - x)dx_1 \int_{A_2} h_{t_2-t_1}(x_2-x_1)dx_2 \ldots \int_{A_n} h_{t_n-t_{n-1}}(x_n - x_{n-1})dx_n. \end{aligned} \tag{6.15}$$

for time series of the form (6.1) and $A_0, \ldots, A_n \in \mathcal{B}(\mathbb{R}^d)$.
Hint: Note that $\{B_{t_j} - B_{t_{j-1}}\}_{j=1}^n$ are independent and that $\{B_{t_1} \in A_1, \ldots, B_{t_n} \in A_n\} = \{(B_{t_j} - B_{t_{j-1}})_{j=1}^n \in D\}$, where $D = \bigcap_{j=1}^n \{y \in (\mathbb{R}^d)^n \ ; \ x + y_1 + \ldots + y_j \in A_j\}$. Therefore,

$$\text{LHS of (6.15)} = \int_D h_{t_1}(y_1)h_{t_2-t_1}(y_2)\cdots h_{t_n-t_{n-1}}(y_n)dy_1\cdots dy_n.$$

Exercise 6.1.2 Suppose that B is a BM^d_0. Then, prove that $(c^{-1/2}B_{ct})_{t\geq 0}$ is a BM^d_0 for all $c > 0$ and that $(UB_t)_{t\geq 0}$ is a BM^d_0 for any orthogonal $d \times d$ matrix U.

Exercise 6.1.3 Let B be a BM^d and $s > 0$. Then, prove that $(B_s - B_{s-t})_{0 \le t \le s} \approx (B_t - B_0)_{0 \le t \le s}$.

Exercise 6.1.4 Let B be a BM^d_0. Then, prove the following for $p > 0$ and $t > 0$. **i)** $E[|B_t|^{-p}] = t^{-p/2} C(p,d)$ where $C(p,d) < \infty$ if $p < d$ and $C(p,d) = \infty$ if otherwise. **ii)** $\int_0^t |B_s|^{-p} ds \in L^1(P)$ if $p < 2 \wedge d$.
Remark: For $d = 1$, it follows from ii) above that $\int_0^t |B_s|^{-p} ds < \infty$ a.s. for $p < 1$. On the other hand, it is known, as an application of Engelbert-Schmidt zero-one law that $\int_0^t |B_s|^{-1} ds = \infty$ a.s. cf. [KS91, p.217].

Exercise 6.1.5 Let B be BM^1_0 Prove the following. **i)** Suppose that $F : [0,t) \to \mathbb{R}$ be right-continuous and of bounded variation. Then, $B(F) \stackrel{\text{def}}{=} \int_0^t B_t dF(t)$ is a mean-zero Gaussian r.v. **Hint**: The step function $B^{(n)}_s = \sum_{j=1}^n B(tj/n) \mathbf{1}_{[(j-1)t/n, jt/n))}(s)$ $(0 \le s \le t)$ converges uniformly to B_s. **ii)** Suppose that $F_j : [0,t) \to \mathbb{R}$ $(j = 1,2)$ are continuous and of bounded variation. Then,

$$E\left[B(F_1)B(F_2)\right] = tF_1(t)F_2(t) + \int_0^t F_1(s)F_2(s)ds - F_1(t)\int_0^t F_2(s)ds - F_2(t)\int_0^t F_1(s)ds.$$

Exercise 6.1.6 Let B be a BM^q_x $(d \ge 2,\ x \in \mathbb{R}^d)$ and $f : [0,\infty) \to [0,\infty)$ be a measurable function. Let also $F_\nu(z)$ $(\nu, z \in \mathbb{C})$ be from (2.20). Then, prove that

$$E_x[f(|B_t|)] = \int_0^\infty k_t(|x|, r) f(r) dr,$$

where

$$k_t(r_0, r) = 2(2t)^{-\frac{d}{2}} r^{d-1} \exp\left(-\frac{r_0^2 + r^2}{2t}\right) F_{\frac{d}{2}-1}\left(\frac{r_0 r}{t}\right), \quad r_0, r \in [0,\infty).$$

Exercise 6.1.7 Let $X = (X_t : \Omega \to \mathbb{R})_{t \ge 0}$ be a process such that $t \mapsto X_t(\omega)$ is continuous for all $\omega \in \Omega$, and let $v : [0,\infty) \to [0,\infty)$ be continuous, strictly increasing, with $v(0) = 0$. Then, prove that the following conditions (a) and (b) are equivalent. **(a)** There exists a Brownian motion B such that $X_t - X_0 = B_{v(t)}$ $(\forall t \ge 0)$. **(b)** The process X is of independent increment and $X_t - X_s \approx N(0, v(t) - v(s))$ for all $0 \le s < t$.

Exercise 6.1.8 Let B be a BM^1_0 and $h : [0,\infty) \to \mathbb{R}$ be continuous, of bounded variation on any bounded interval. Then, prove the following. **(i)** The process

$$X_t = X_t(B) \stackrel{\text{def}}{=} B_t h(t) - \int_0^t B_u dh(u), \quad t \ge 0$$

is of independent increments and that $X_t - X_s \approx N(0, v(t) - v(s))$ for all $0 \le s < t$, where $v(t) = \int_0^t h(u)^2 du$. **Hint**: Take a sequence of partitions of $[0,t]$: $0 = t_{n,0} < t_{n,1} < \ldots < t_{n,p(n)} = t$ $(n \ge 1)$ such that $\max_{0 \le j \le p(n)-1}(t_{n,j+1} - t_{n,j}) \stackrel{n\to\infty}{\longrightarrow} 0$ and let $B^{(n)}_s = \sum_{j=1}^{p(n)-1} B(t_{n,j}) \mathbf{1}_{(t_{n,j}, t_{n,j+1}]}(s)$. Then, $X_t(B^{(n)}) \stackrel{n\to\infty}{\longrightarrow} X_t$ in $L^2(P)$. Moreover, by "summation by parts",

$$X_t(B^{(n)}) = \sum_{j=1}^{p(n)-1} (B(t_{n,j}) - B(t_{n,j-1})) h(t_{n,j}).$$

(ii) Suppose in addition that h vanishes on no open interval. Then, there exists a Brownian motion β such that $X_t = \beta_{v(t)}$ $(\forall t \ge 0)$. **Hint**: Exercise 6.1.7.

Exercise 6.1.9 Referring to Exercise 6.1.8, suppose in addition that h is strictly positive. Prove that, for $x \in \mathbb{R}$, $Y_t = h(t)^{-1}(h(0)x + X_t)$, $t \geq 0$ is the unique solution to the following integral equation.

$$(*) \qquad Y_t = x + B_t - \int_0^t Y_s \frac{dh(s)}{h(s)}.$$

Remark Let $\lambda > 0$. Then, with the choice $h(t) = \exp(\lambda t)$, the process $Y = (Y_t)_{t\geq 0}$ above is called the *Ornstein-Uhlenbeck process*, which is therefore defined by

$$Y_t = B_t + \exp(-\lambda t)\left(x - \lambda \int_0^t B_s \exp(\lambda s) ds\right), \quad t \geq 0.$$

By Exercise 6.1.8 (ii), there exists a Brownian motion β such that

$$Y_t = \exp(-\lambda t)\left(x + \beta\left(\frac{\exp(2\lambda t) - 1}{2\lambda}\right)\right), \quad t \geq 0.$$

In particular, for each $t > 0$, Y_t is a Gaussian r.v. with the mean $\exp(-\lambda t)x$ and the variance $\frac{1-\exp(-2\lambda t)}{2\lambda}$. By Exercise 6.1.9, $Y = (Y_t)_{t\geq 0}$ is the unique solution to the following integral equation.

$$Y_t = x + B_t - \lambda \int_0^t Y_s ds.$$

Exercise 6.1.10 (**Brownian bridge**) Let $a, b \in \mathbb{R}^d$, and $s > 0$. A process $X = (X_t : \Omega \to \mathbb{R}^d)_{0\leq t\leq s}$ is called a Brownian bridge from a to b ($\mathrm{BB}^d_{a,b,s}$ for short) if

$$X_t = B_t - \frac{t}{s}B_s + \left(1 - \frac{t}{s}\right)a + \frac{t}{s}b, \quad 0 \leq t \leq s,$$

where B is a BM^d_0. Prove the following. **(i)** If X is a $\mathrm{BB}^d_{a,b,s}$, then, $(X_{s-t})_{0\leq t\leq s}$ is a $\mathrm{BB}^d_{b,a,s}$. **Hint** Exercise 6.1.3. **(ii)** Suppose that two processes $X = (X_t : \Omega \to \mathbb{R}^d)_{0\leq t\leq s}$ and $\beta = (\beta_t : \Omega \to \mathbb{R}^d)_{t\geq 0}$ are related as

$$X_t = t\beta\left(\frac{1}{t} - \frac{1}{s}\right), \quad 0 < t \leq s,$$

or equivalently,

$$\beta_t = \left(t + \frac{1}{s}\right) X\left(\frac{1}{t + \frac{1}{s}}\right), \quad t \geq 0.$$

Then, X is a $\mathrm{BB}^d_{0,0,s}$ if and only if β is a BM^d_0. **Hint**: Suppose that β is a BM^d_0. Then, by Corollary 6.1.11, there exists a BM^d_0, say B, such that $\beta_t = B_{t+\frac{1}{s}} - B_{\frac{1}{s}}$. Then, use Proposition 6.1.5 to prove that X is a $\mathrm{BB}^d_{0,0,s}$. Suppose on the other hand that X is a $\mathrm{BB}^d_{0,0,s}$. Then, there exists a BM^d_0, say B, such that $X_t = B_t - \frac{t}{s}B_s$. Then, use Proposition 6.1.5 to prove that β is a BM^d_0.

Exercise 6.1.11 (**Markov porperty given the future**) Let B be a BM^d_0, $s > 0$, $b \in \mathbb{R}^d$, $\mathcal{T}_s = \sigma(B_t \,;\, t \geq s)$ and $X^b = (X^b_t)_{0<t\leq s} = (B_t - \frac{t}{s}B_s + \frac{t}{s}b)_{0<t\leq s}$. Prove then the following.
i) $\left(X^b, (B_t)_{t\geq s}\right) \approx \left((tB_{1/t} - tB_{1/s} + \frac{t}{s}b)_{0<t\leq s}, (tB_{1/t})_{t\geq s}\right)$. In particular, X^b is independent

of $\mathcal{T}_s$. [Hint:Proposition 6.1.5.]
ii) Suppose that $F : (\mathbb{R}^d)^{(0,s]} \to \mathbb{R}$ is bounded measurable and $A \in \mathcal{T}_s$. Then,

$$E[F((B_t)_{0<t\le s}) : A] = \int_A E[F(X^b)]|_{b=B_s(\omega)} P(d\omega).$$

Therefore,

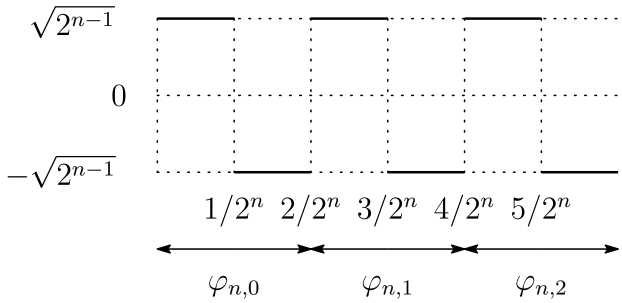
$$E[F((B_t)_{0<t\le s})|\mathcal{T}_s] = E[F(X^b)]|_{b=B_s}, \quad \text{a.s.}$$

Hint: For $0 < t \le s$, $B_t = X_t^0 + \frac{t}{s} B_s$.

Exercise 6.1.12 [20] Let $(S_1, \mathcal{A}_1)$ and $(S_2, \mathcal{A}_2)$ be measurable spaces and $\mu \in \mathcal{P}(S_2, \mathcal{A}_2)$. Then, for $F : S_1 \times S_2 \to \mathbb{R}$, bounded, $\mathcal{A}_1 \otimes \mathcal{A}_2$-measurable, prove that $f(x) = \int_{S_2} F(x, y)\mu(dy)$ is $\mathcal{A}_1$-measurable. [Hint: It is enough to consider the case where $F = 1_A$ for $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$. When $A = A_1 \times A_2$ $(A_j \in \mathcal{A}_j)$, $f = 1_{A_1}\mu(A_2)$ is clearly $\mathcal{A}_1$-measurable. Finally, use Dynkin's lemma.]

6.2 The Existence of the Brownian Motion

We present a construction of a BM_0^1 in this subsection. This is enough to prove the existence of BM_x^d for any $d \ge 1$ and $x \in \mathbb{R}^d$ (cf. Lemma 6.1.2, Corollary **??**). We begin by introducing Haar functions $\varphi_{n,k} : [0, \infty) \to \mathbb{R}$ $(n, k \in \mathbb{N})$ as follows.



$$\begin{aligned}\varphi_{0,k} &= 1_{(k,k+1]},\\ \varphi_{n,k} &= \sqrt{2^{n-1}}1_{(2k/2^n,(2k+1)/2^n]} - \sqrt{2^{n-1}}1_{(2k+1)/2^n,(2k+2)/2^n]}, \quad \text{for } n \ge 1.\end{aligned}$$

Let $X = (X_{n,k})_{n,k\in\mathbb{N}}$, where $X_{n,k}$ are iid $\approx N(0, 1)$, defined on a probability space $(\Omega, \mathcal{F}, P)$. We will prove the existence of BM_0^1 in the following form;

[20]This exercise is associated with Lemma 6.1.14 below.

Theorem 6.2.1 **a)** *The following series absolutely converges a.s.*

$$B_t = \sum_{n,k \geq 0} X_{n,k} \int_0^t \varphi_{n,k}, \quad t \geq 0. \tag{6.16}$$

More precisely, for any $\alpha \in [0, 1/2)$ and $T > 0$, there is an a.s. finite r.v. $M = M(\alpha, T) \geq 0$ such that

$$\sum_{n,k \geq 0} \left| X_{n,k} \int_s^t \varphi_{n,k} \right| \leq M|t-s|^\alpha \quad \textit{for all } 0 \leq s < t \leq T. \tag{6.17}$$

In particular,

$$|B_t - B_s| \leq M|t-s|^\alpha \quad \textit{for all } 0 \leq s < t \leq T. \tag{6.18}$$

b) $(B_t)_{t \geq 0}$ *defined above is a* BM_0^1.

Define

$$\langle f, g \rangle = \int_0^\infty fg, \quad f, g \in L^2[0, \infty).$$

We also introduce $\mathcal{X} \subset L^2([0, \infty))$ by:

$$\mathcal{X} = \text{finite linear combinations of } 1_{(0,t]} \ (t > 0).$$

Therefore, a function $h \in \mathcal{X}$ is expressed as

$$h = \sum_{i=1}^{\ell} c_i 1_{(0,t_i]}, \quad c_1, ..., c_\ell \in \mathbb{R}, \ t_1, \ldots, t_\ell \in (0, \infty) \tag{6.19}$$

for some $\ell \geq 1$. We will prove Theorem 6.2.1 in the following generalized form:

Lemma 6.2.2 *Then the following hold;*

a) *For $h \in \mathcal{X}$, the following series absolutely converges a.s.*

$$B(h) \stackrel{\rm def}{=} \sum_{n,k \geq 0} X_{n,k} \langle\, \varphi_{n,k}, h \,\rangle. \tag{6.20}$$

More precisely, there exists an a.s. finite r.v. $Z \geq 0$ for which the following holds true. Suppose that $h \in \mathcal{X}$ is of the form (6.19) with $t_1, \ldots, t_\ell \in (0,T]$ for some $T > 0$. Then, for any $q > 2$,

$$\sum_{n,k \geq 0} |X_{n,k} \langle\, \varphi_{n,k}, h \,\rangle| \leq C\ell Z \|h\|_q, \tag{6.21}$$

where $C = C(q,T) \in (0,\infty)$ is a constant and $\|\cdot\|_q = \|\cdot\|_{L^q[0,\infty)}$.

b) *$\{B(h)\}_{h \in \mathcal{X}}$ is a family of a mean-zero Gaussian r.v.'s such that*

$$E[B(h_1)B(h_2)] = \langle\, h_1, h_2 \,\rangle, \quad \textit{for all } h_1, h_2 \in \mathcal{X}. \tag{6.22}$$

c) *If $\{h_j\}_{j=1}^n \subset \mathcal{X}$ and $\langle\, h_i, h_j \,\rangle = 0$ for $i \neq j$, then $\{B(h_j)\}_{j=1}^n$ are independent.*

Remark: Note that $\mathcal{X}$ is dense in $L^2([0,\infty))$. Thus, by (6.22), the map $\mathcal{X} \ni h \mapsto B(h)$ extends to an isometry from $L^2([0,\infty))$ to $L^2(\Omega, \mathcal{F}, P)$.

We now finish the proof of Theorem 6.2.1 assuming Lemma 6.2.2.

Proof of Theorem 6.2.1: We see from (6.16) and (6.20) that for $0 \leq s \leq t < \infty$,

1) $\quad B_t - B_s = B(1_{(s,t]})$.

Since $\|1_{(s,t]}\|_q = |t-s|^{1/q}$, the bound (6.17) follows from (6.21) and 1) with $M(\alpha, T) = 2C(\alpha^{-1}, T)Z$. Let next us check **B0**)–**B2**) (with $d = 1$ and $x = 0$) for $\{B_t\}_{t \geq 0}$.
B0): This is obvious by the definition (6.16).
B1): If $n \geq 2$ and $0 = t_0 < t_1 < \ldots < t_n$, then for $i \neq j$, $\langle\, 1_{(t_{i-1}t_i]}, 1_{(t_{j-1}t_j]} \,\rangle = 0$. Therefore, $B_{t_j} - B_{t_{j-1}} = B(1_{(t_{j-1}t_j]})$ $(j = 1, \ldots n)$ are independent by Lemma 6.2.2 **c)**.
B2): $\langle\, 1_{(s,t]}, 1_{(s,t]} \,\rangle = t - s$ for $0 \leq s < t$. Hence it follows from Lemma 6.2.2 **b)** that $B_t - B_s = B(1_{(s,t]}) \approx N(0, t-s)$.
B2): This follows from (6.17).

\(^□^)/

We now turn to the proof of Lemma 6.2.2. We begin by proving the following

Lemma 6.2.3 $\{\varphi_{n,k}\}_{n,k\geq 0}$ *is a complete orthonormal system of* $L^2[0,\infty)$*, i.e.,*

$$\langle \varphi_{n,k}, \varphi_{n',k'} \rangle = \begin{cases} 1, & \text{if } (n,k) = (n',k') \\ 0, & \text{if otherwise.} \end{cases} \tag{6.23}$$

and

$$\bigcap_{n,k\geq 0} \{h \in L^2[0,\infty)\ ;\ \langle \varphi_{n,k}, h \rangle = 0\} = \{h \equiv 0\}. \tag{6.24}$$

Proof: The proof of (6.23) is easy and is left to the readers (cf. Exercise 6.2.1 below). To prove (6.24), we take a function h from the set on the left-hand side of (6.24) and show that

$$H(t) = H(0) \quad \text{for all } t \geq 0, \text{ where } H(t) \stackrel{\text{def.}}{=} \textstyle\int_0^t h.$$

Since diadic rationals are dense, it is enough to prove

1) $\quad H\left(\frac{2k+1}{2^n}\right) = H(0) \quad$ for all $n, k \geq 0$.

We will prove (1) by induction on n. We have

2) $$H(k+1) - H(k) = \int_k^{k+1} h = \langle \varphi_{0,k}, h \rangle = 0, \quad k = 0, 1, \ldots,$$

which proves 1) for $n = 0$. Suppose that 1) holds true with n replaced by $n-1$. Then, for $j, k \in \mathbb{N}$, $H\left(\frac{2k+2j}{2^n}\right) = H\left(\frac{k+j}{2^{n-1}}\right) = H(0)$. Therefore,

$$\begin{aligned} H\left(\tfrac{2k+1}{2^n}\right) - H(0) &= H\left(\tfrac{2k+1}{2^n}\right) - \tfrac{1}{2}H\left(\tfrac{2k}{2^n}\right) - \tfrac{1}{2}H\left(\tfrac{2k+2}{2^n}\right) \\ &= \tfrac{1}{2}\left(H\left(\tfrac{2k+1}{2^n}\right) - H\left(\tfrac{2k}{2^n}\right)\right) - \tfrac{1}{2}\left(H\left(\tfrac{2k+2}{2^n}\right) - H\left(\tfrac{2k+1}{2^n}\right)\right) \\ &= \tfrac{1}{2}\int_{\frac{2k}{2^n}}^{\frac{2k+1}{2^n}} h - \tfrac{1}{2}\int_{\frac{2k+1}{2^n}}^{\frac{2k+2}{2^n}} h = \tfrac{1}{2}2^{-\frac{n-1}{2}}\langle \varphi_{n,k}, h \rangle = 0. \end{aligned}$$

\\(^□^)/

Lemma 6.2.4

$$Z \stackrel{\text{def}}{=} \sup_{n,k\geq 0} |X_{n,k}|/\sqrt{\log(2+n+k)} < \infty, \quad a.s.$$

Proof: We will in fact prove that for $c > 2$,

$$P\left(|X_{n,k}| \leq c\sqrt{\log(2+n+k)} \quad \text{except finitely many } (n,k)\text{'s}\right) = 1.$$

We first compute for any $y > 0$ that

1) $$\begin{cases} P(|X_{n,k}| > y) &= \sqrt{2/\pi}\displaystyle\int_y^\infty \exp(-x^2/2)dx \\ &\leq \sqrt{2/\pi}\displaystyle\int_y^\infty (x/y)\exp(-x^2/2)dx = \sqrt{2/\pi}\exp(-y^2/2)/y. \end{cases}$$

We use this inequality as follows. (Note that $\sqrt{2/\pi} \leq 1$. Note also that $c\sqrt{\log(2+n+k)} > 1$, since $\sqrt{\log 2} = 0.83....$)

$$\begin{aligned} & E\left[\sum_{n,k\geq 0} 1\{|X_{n,k}| > c\sqrt{\log(2+n+k)}\}\right] \\ &= \sum_{n,k\geq 0} P(|X_{n,k}| > c\sqrt{\log(2+n+k)}) \stackrel{1)}{\leq} \sum_{n,k\geq 0} \exp\left(-\frac{c^2}{2}\log(2+n+k)\right) \\ &= \sum_{n,k\geq 0} (2+n+k)^{-\frac{c^2}{2}} < \infty. \end{aligned}$$

As a consequence, $\sum_{n,k\geq 0} 1\{|X_{n,k}| > c\sqrt{\log(2+n+k)}\} < \infty$, P-a.s., which is equivalent to what we wanted to prove. \(^□^)/

Proof of Lemma 6.2.2: **a**): Let It is enough to prove (6.21). We take $p \in (1,2)$ such that $\frac{1}{p} + \frac{1}{q} = 1$ and define $\varepsilon = \frac{1}{p} - \frac{1}{2} > 0$. We also introduce $K_n(h) \stackrel{\text{def}}{=} \{k \in \mathbb{N}\ ;\ \langle\, \varphi_{n,k}, h\,\rangle \neq 0\}$. We verify that

1) $\quad \|\varphi_{n,k}\|_p = 2^{\frac{n-1}{2} - \frac{n-1}{p}} = 2^{-(n-1)\varepsilon}$.

2) $\quad \max K_n(h) \leq 2^{n-1}T$,

3) $\quad |K_n(h)| \stackrel{\text{def}}{=} \sum_{k\in K_n(h)} 1 \leq (1+T)\ell$.

We also get 1) by a direct computation. To see 2), note that $h \equiv 0$ outside $(0,T]$ and that $\varphi_{n,k} \equiv 0$ outside $(\frac{2k}{2^n}, \frac{(2k+2)}{2^n}]$. If $k > 2^{n-1}T$, then $(0,T] \cap (\frac{2k}{2^n}, \frac{(2k+2)}{2^n}] = \emptyset$, and hence $\langle \varphi_{n,k}, h\rangle = 0$. The inequality 3) can be seen as follows. For any $t > 0$,

$$\langle\, \varphi_{n,k}, 1_{(0,t]}\,\rangle \neq 0 \implies \begin{cases} k \leq t, & \text{if } n = 0 \\ t \in [2k/2^n, (2k+2)/2^n), & \text{if } n \geq 1. \end{cases}$$

and hence,

$$|K_n(1_{(0,t]})| \leq \left\{ \begin{array}{ll} 1+t, & \text{if } n = 0 \\ 1, & \text{if } n \geq 1 \end{array} \right\} \leq 1+T.$$

Therefore,

$$|K_n(h)| \leq \sum_{j=1}^{\ell} |K_n(1_{(0,t_j]})| \leq \ell(1+T).$$

Let $c_{n,T} \stackrel{\text{def}}{=} \sqrt{\log(2+n+2^{n-1}T)}$. Then, for $k \in K_n(h)$,

4) $$\begin{cases} |X_{n,k}| & \stackrel{\text{Lemma 6.2.4}}{\leq} \quad Z\sqrt{\log(2+n+k)} \stackrel{2)}{\leq} c_{n,T}Z, \\ |\langle\, \varphi_{n,k}, h\,\rangle| & \stackrel{\text{Hölder}}{\leq} \quad \|\varphi_{n,k}\|_p \|h\|_q \stackrel{1)}{=} 2^{-(n-1)\varepsilon}\|h\|_q. \end{cases}$$

Therefore,

$$\begin{aligned}\sum_{n,k\ge 0}|X_{n,k}\langle\,\varphi_{n,k},h\,\rangle| &= \sum_{n\ge 0}\sum_{k\in K_n(h)}|X_{n,k}\langle\,\varphi_{n,k},h\,\rangle| \\ &\stackrel{4)}{\le} \|h\|_q Z\sum_{n\ge 0}c_{n,T}2^{-(n-1)\varepsilon}\sum_{k\in K_n(h)}1 \\ &\stackrel{3)}{\le} \|h\|_q Z(1+T)\ell\sum_{n\ge 0}c_{n,T}2^{-(n-1)\varepsilon}.\end{aligned}$$

The series in the third line converges and this proves (6.21).
b): Ingredients of the proof will be Lemma 6.2.3 and some basic properties of Gaussian r.v.'s listed in Exercise 2.2.4–Exercise 2.4.7. For $h\in\mathcal{X}$, we define $B(h)$ by (6.20) and $B_N(h)$ by the partial sum;

$$B_N(h)=\sum_{n=0}^{N}\sum_{k\ge 0}X_{n,k}\langle\,\varphi_{n,k},h\,\rangle.$$

Then,

- $B_N(h)$ for each $h\in\mathcal{X}$ is a mean-zero Gaussian r.v.

In fact, $B_N(h)$ is a finite summation of independent mean-zero Gaussian r.v.'s (cf. 3)) and hence is a mean-zero Gaussian r.v. by Exercise 2.2.4.
Next, as a consequence of part (a),

- $B_N(h)\stackrel{N\nearrow\infty}{\longrightarrow}B(h)$, P-a.s.

Moreover,

- $E[B_N(h_1)B_N(h_2)]\stackrel{N\nearrow\infty}{\longrightarrow}\langle\,h_1,h_2\,\rangle$ for $h_1,h_2\in\mathcal{X}$.

This can be seen as follows;

$$\begin{aligned}E[B_N(h_1)B_N(h_2)] &= \sum_{n,n'=0}^{N}\sum_{k,k'\ge 0}\langle\,\varphi_{n,k},h_1\,\rangle\langle\,\varphi_{n',k'},h_2\,\rangle E[X_{n,k}X_{n',k'}] \\ &= \sum_{n=0}^{N}\sum_{k\ge 0}\langle\,\varphi_{n,k},h_1\,\rangle\langle\,\varphi_{n,k},h_2\,\rangle \stackrel{N\nearrow\infty}{\longrightarrow}\sum_{n\ge 0}\sum_{k\ge 0}\langle\,\varphi_{n,k},h_1\,\rangle\langle\,\varphi_{n,k},h_2\,\rangle \\ &= \langle\,h_1,h_2\,\rangle,\quad \text{by Parseval's identity.}\end{aligned}$$

These, together with Exercise 2.4.7, prove that $B(h)$ for each $h\in\mathcal{X}$ is a Gaussian r.v. and that (6.22) holds for $h_1,h_2\in\mathcal{X}$.
c): By part **b**), $\sum_{j=1}^n c_jB(h_j)=B(\sum_{j=1}^n c_jh_j)$ is a Gaussian r.v. for $(c_j)_{j=1}^n\in\mathbb{R}^n$. Hence it follows from Exercise 2.2.5 that $(B(h_j))_{j=1}^n$ is an $\mathbb{R}^n$-valued Gaussian r.v. By this, (6.22) and Exercise 2.2.6, we see that $\{B(h_j)\}_{j=1}^n$ are independent. \(^□^)/

Exercise 6.2.1 Prove (6.23).

6.3 α-Hölder continuity for $\alpha < 1/2$

We start by proving the following estimate, which shows that the Brownian motion is α-Hölder continuous for any $\alpha < 1/2$.

Proposition 6.3.1 *If B is a BM_0^1, then for any $\alpha \in [0,1/2)$ and $T > 0$,*

$$\sup_{0 \le s < t \le T} \frac{|B_t - B_s|}{|t-s|^\alpha} < \infty, \quad a.s.$$

To prove Proposition 6.3.1, we prepare the following

Lemma 6.3.2 *For $f \in C([0,T] \to \mathbb{R})$ and $g \in C((0,T] \to (0,\infty))$,*

$$\sup_{0 \le s < t \le T} \frac{|f(t) - f(s)|}{g(t-s)} = \sup_{\substack{0 < s < t < T \\ s,t \in \mathbb{Q}}} \frac{|f(t) - f(s)|}{g(t-s)}.$$

Proof: We prove $\le$ only, since $\ge$ is obvious. Let M be the right-hand side of the equality to be proved. Then, we may assume that $M < \infty$. Let $0 \le s < t \le T$. We choose $s_n, t_n \in \mathbb{Q}$, $n \in \mathbb{N}$ such that $0 < s_n < t_n < T$, $s_n \to s$ and $t_n \to t$. We have that

$$|f(t_n) - f(s_n)| \le M g(t_n - s_n).$$

Letting $n \to \infty$, we obtain that $\frac{|f(t)-f(s)|}{g(t-s)} \le M$, as desired. \(^□^)/

Proof of Proposition 6.3.1: Let $\widetilde{B}$ be the BM_0^1 on a probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$, constructed by Theorem 6.2.1. Let

$$E = \left\{ \sup_{0 \le s < t \le T} \frac{|B_t - B_s|}{|t-s|^\alpha} < \infty \right\}, \quad F = \left\{ \sup_{\substack{0 < s < t < T \\ s,t \in \mathbb{Q}}} \frac{|B_t - B_s|}{|t-s|^\alpha} < \infty \right\}.$$

Note that $E \subset F$. Let also $\widetilde{E}$ and $\widetilde{F}$ be defined in the same way as above, with B replaced by $\widetilde{B}$. Then, we know from Theorem 6.2.1 that $\widetilde{E} \stackrel{\text{a.s.}}{=} \widetilde{\Omega}$. We want to conclude from this that $E \stackrel{\text{a.s.}}{=} \Omega$. Unfortunately, as is in the proof of Proposition 6.1.5, we can not do so directly, since $E \notin \sigma[B]$, as well as $\widetilde{E} \notin \sigma[\widetilde{B}]$. We will go around this bother by noting that

1) $F \in \sigma[B]$, $\widetilde{F} \in \sigma[\widetilde{B}]$ and $E \stackrel{\text{a.s.}}{=} F$.

Let us admit 1) for a moment to conclude the proof. By 1), it is enough to show that $P(F) = 1$. Since $B \approx \widetilde{B}$, $F \in \sigma[B]$, $\widetilde{F} \in \sigma[\widetilde{B}]$, we have that $P(F) = \widetilde{P}(\widetilde{F}) = 1$.
We now see 1) as follows. First,

$$F = \bigcup_{m \in \mathbb{N}} \bigcap_{\substack{0 < s < t < T \\ s,t \in \mathbb{Q}}} \left\{ \frac{|B_t - B_s|}{|t-s|^\alpha} \le m \right\} \in \sigma[B].$$

Similarly, $\widetilde{F} \in \sigma[\widetilde{B}]$. Now, recall that there exists an $\Omega_B \in \mathcal{F}$ such that $P(\Omega_B) = 1$ and $t \mapsto B_t(\omega)$ is continuous for all $\omega \in \Omega_B$. Thus, it follows from Lemma 6.3.2 that $E \cap \Omega_B = F \cap \Omega_B$, and hence $E \stackrel{\text{a.s.}}{=} F$. \(^□^)/

As an immediate conseqence of Proposition 6.3.1, we have the following

Corollary 6.3.3 *If B is a BM_0^1, then for any $\alpha \in [0,1/2)$ and $T > 0$,*

$$\lim_{h \searrow 0} \sup_{0<t\leq T} \frac{|B_{t\pm h} - B_t|}{h^\alpha} = 0, \quad a.s.$$

With Proposition 6.1.5 and Corollary 6.3.3, we obtain the following property of the Brownian motion as $t \to \infty$.

Corollary 6.3.4 (**The law of large numbers for the Brownian motion**) *Let B be a BM_0^d. Then, for any $\alpha > 1/2$,*

$$B_t/t^\alpha \stackrel{t\to\infty}{\longrightarrow} 0, \quad a.s.$$

Proof: Let $\check{B}$ be as in Proposition 6.1.5. Then,

$$B_t/t^\alpha \stackrel{t\to\infty}{\longrightarrow} 0 \iff t^{-(1-\alpha)}\check{B}_t \stackrel{t\to 0+}{\longrightarrow} 0.$$

Since $1-\alpha < 1/2$, we see from Corollary 6.3.3 that

$$t^{-(1-\alpha)}\check{B}_t \stackrel{t\to 0+}{\longrightarrow} 0, \quad \text{a.s.}$$

\(^□^)/

Remarks: 1) By Proposition 6.3.1, $t \mapsto B_t$ is α-Hölder continuous on any bounded interval for $\alpha < 1/2$. But this is no longer true for $\alpha = 1/2$ (Exercise 6.5.1).
2) Proposition 6.3.1 can be improved in the following way.

$$\sup_{0\leq s<t\leq T} \frac{|B_t - B_s|}{\sqrt{|t-s|\log(1/|t-s|)}} < \infty, \quad \text{a.s.} \tag{6.25}$$

See, e.g., [MP10, p.14, Theorem 1.12]. Moreover, this improvement is optimal, as can be seen from the following result, known as *Lévy's modulus of continuity* (P. Lévy (1937)).

$$\overline{\lim_{h\searrow 0}} \sup_{0\leq t\leq T} \frac{|B_{t+h} - B_t|}{\sqrt{h\log(1/h)}} = \sqrt{2}, \quad \text{a.s.} \tag{6.26}$$

See, e.g. [KS91, p.114, Theorem 9.25], [MP10, p.16, Theorem 1.14].
3) The following refinement of Corollary 6.3.4 is known as *the law of iterated logarithm* (A. Hincin (1933)).

$$\overline{\lim_{t\to\infty}} \frac{|B_t|}{\sqrt{t\log\log t}} = \sqrt{2}, \quad \text{a.s.} \tag{6.27}$$

See, e.g. [Dur95, p.434, (9.1)], [KS91, p.112, Theorem 9.22], [MP10, p.119, Theorem 5.1]. This, together with Proposition 6.1.9 and Proposition 6.1.5, implies that for any $t \geq 0$,

$$\overline{\lim_{h\searrow 0}} \frac{|B_{t+h} - B_t|}{\sqrt{h\log\log(1/h)}} = \sqrt{2}, \quad \text{a.s.} \tag{6.28}$$

Although the results (6.26) and (6.28) are of the similar kind, the functions on the denominators slightly differ, depending on whether the supremum of the time t is taken over an interval as in (6.26), or the time t is fixed as in (6.28).

Exercise 6.3.1 ($\star$) Let B be a pre-BM$^1_0$, U be a uniformly distributed r.v. on $(0,1)$, and $\varphi : [0,\infty) \to (0,\infty)$ be a nondecreasing function. We define $\widetilde{B} = (\widetilde{B}_t)_{t\geq 0}$ by

$$\widetilde{B}_t = \begin{cases} \varphi(n+1) & \text{if } t = n+U, \\ B_t & \text{if otherwise.} \end{cases}$$

Prove the following. i) $\widetilde{B}$ is a pre-BM$^1_0$. ii) $\overline{\lim}_{t\to\infty} |\widetilde{B}_t|/\varphi(t) \geq 1$, which shows that the conclusion of Corollary 6.3.4 is no longer true for pre-Brownian motions.

6.4 Nowhere α-Hölder continuity for $\alpha > 1/2$

One of the most striking property of the Brownian motion is the *nowhere differentiablity*[21]:

$$\text{With probability one, } t \mapsto B_t \text{ is not differentiable at any } t \geq 0. \tag{6.29}$$

Let us describe the above property in a more quantitative way. For a function $f : [0,\infty) \to \mathbb{R}$ and a exponent $\alpha \in (0,1]$, we define the right (resp. left) Hölder coefficients $C^+_{\alpha,f}(t)$, $t \geq 0$ (resp. $C^-_{\alpha,f}(t)$, $t > 0$) as follows.

$$C^{\pm}_{\alpha,f}(t) \stackrel{\text{def}}{=} \overline{\lim_{h\searrow 0}} \frac{|f(t\pm h) - f(t)|}{h^\alpha}. \tag{6.30}$$

If f is right (resp. left) differentiable at t, then, for all $\alpha \in (0,1]$,

$$C^+_{\alpha,f}(t) \leq C^+_{1,f}(t) < \infty \text{ (resp. } C^-_{\alpha,f}(t) \leq C^-_{1,f}(t) < \infty).$$

Thus, (6.29) is a consequence of the following

Proposition 6.4.1 *Let B be a* BM$^1_0$ *and $\alpha \in (1/2,1]$. Then, a.s.,*

$$C^+_{\alpha,B}(t) = \infty \textit{ for all } t \geq 0 \textit{ and } C^-_{\alpha,B}(t) = \infty \textit{ for all } t > 0. \tag{6.31}$$

Remark Davis, and independently, Perkins and Greenwood, proved in 1983 that

$$\inf_{t\in[0,1]} C^+_{1/2,B}(t) = 1, \quad \text{a.s.}$$

This shows that (6.31) is no longer true for $\alpha = 1/2$. See also Exercise 6.5.1 below.

We turn to the proof[22] of (6.31). We start with the following lemma, which has nothing to do with probability in itself. For $f : [0,\infty) \to \mathbb{R}$ and $\alpha \in (0,\infty)$, we define

$$S^+_{\alpha,f}(t) \stackrel{\text{def}}{=} \sup_{h\in(0,1]} \frac{|f(t+h) - f(t)|}{h^\alpha}. \tag{6.32}$$

[21] Due to R.E.A.C.Paley, N. Wiener and A. Zygmund (1933)

[22] We follow the line of argument by A. Dvoretsky, P. Erdös, S. Kakutani (1961).

Lemma 6.4.2 a) *Suppose that*

$$\inf_{t\in[0,T]} S^+_{\alpha,f}(t) < \ell \ \textit{for some}\ T, \ell \in (0,\infty).$$

Then, for any $\delta \in (0,1)$, *there exists* $i = 0, \dots, \lfloor T/\delta \rfloor$ *such that*

$$|f((i+j+1)\delta) - f((i+j)\delta)| \le 2\ell(j+1)^\alpha \delta^\alpha \ \textit{for all}\ j = 1, \dots, \lfloor 1/\delta \rfloor - 1.$$

b) *Suppose that* f *is bounded on* $[t, t+1]$ *for some* $t \ge 0$. *Then,*

$$S^+_{\alpha,f}(t) < \infty \iff C^+_{\alpha,f}(t) < \infty.$$

Proof: a) Take $t \in [0,T]$ such that $S^+_{\alpha,f}(t) < \ell$ and $i \in \mathbb{N}$ such that $i\delta \le t < (i+1)\delta$. Then, for $k = 0, 1$ and $j = 1, \dots, \lfloor \delta^{-1} \rfloor - 1$, we have

$$(i+j+k)\delta - t = \begin{cases} (j+k-1)\delta + (i+1)\delta - t & > \ (j+k-1)\delta \ \ge \ 0, \\ (j+k)\delta + i\delta - t & \le \ (j+k)\delta \ \ \ \ \ \ \ \ \le \ 1. \end{cases}$$

and hence,

$$\begin{aligned} |f((i+j+1)\delta) - f((i+j)\delta)| &\le \sum_{k=0,1} |f((i+j+k)\delta) - f(t)| \\ &\le S^+_{\alpha,f}(t) \sum_{k=0,1} ((i+j+k)\delta - t)^\alpha \\ &\le S^+_{\alpha,f}(t) \sum_{k=0,1} ((j+k)\delta)^\alpha \le 2\ell(j+1)^\alpha\delta^\alpha. \end{aligned}$$

a) $\Rightarrow$: Obvious, since $S^+_{\alpha,f}(t) \ge C^+_{\alpha,f}(t)$.
$\Leftarrow$ Since $\overline{\lim}_{h\searrow 0} = \lim_{\varepsilon\to 0} \sup_{h\in(0,\varepsilon]}$, there exists $0 < \varepsilon \le 1$ such that

1) $$\sup_{u\in(0,\varepsilon]} \frac{|f(t+h) - f(t)|}{h^\alpha} \le C^+_{\alpha,f}(t) + 1 < \infty.$$

On the other hand,

2) $$\sup_{h\in(\varepsilon,1]} \frac{|f(t+h) - f(t)|}{h^\alpha} \le \frac{1}{\varepsilon^\alpha} \sup_{h\in(\varepsilon,1]} |f(t+h) - f(t)| < \infty.$$

It follows from 1) and 2) that $S^+_{\alpha,f}(t) < \infty$. \(^□^)/

Proof of Proposition 6.4.1 *Step1* [23]: Referring to (6.32), we first prove that,

$$\text{a.s., } S^+_{\alpha,B}(t) = \infty \text{ for all } t \ge 0,$$

or equivalently that the following set F is a null set.

1) $F \stackrel{\text{def}}{=} \{S^+_{\alpha,B}(t) < \infty \text{ for some } t \ge 0\}$.

It is enough to prove that each $F_{T,\ell} = \{\inf_{t\in[0,T]} S^+_{\alpha,B}(t) < \ell\}$ $(T, \ell \in \mathbb{N}\backslash\{0\})$ is a null set, since $F = \bigcup_{T,\ell\in\mathbb{N}\backslash\{0\}} F_{T,\ell}$. For this purpose, take $m \in \mathbb{N}\backslash\{0\}$ such that

[23] The continuity of the path is not used here, so that the result is valid for pre-Brownian motion.

2) $\quad (\alpha - \frac{1}{2})m > 1$
and fix it. It follows from Lemma 6.4.2 a) that, on the set $F_{T,\ell}$, for any $\delta \in (0,1)$, there exists $i = 0, \ldots, \lfloor T/\delta \rfloor$ such that

$$X_{\delta,i,j} \stackrel{\text{def}}{=} |B((i+j+1)\delta) - B((i+j)\delta)| \le 2\ell(j+1)^\alpha \delta^\alpha \ \text{ for all } j = 1, \ldots, \lfloor 1/\delta \rfloor - 1.$$

Suppose from here on that $\delta \in (0, 1/(m+2))$ and hence $m \le \lfloor 1/\delta \rfloor - 1$. Then, the above inequality applied for $j = 1, \ldots, m$ yields

$$X_{\delta,i,j} \le L\delta^\alpha \ \text{ for } j = 1, \ldots, m, \text{ where } L \stackrel{\text{def}}{=} 2\ell(m+1)^\alpha.$$

From what we have dicussed so far, we obtain the following inclusion for any $\delta \in (0, 1/(m+2))$.

$$F_{T,\ell} \subset G_\delta \stackrel{\text{def}}{=} \bigcup_{i=0}^{\lfloor T/\delta \rfloor} \bigcap_{j=1}^{m} \{X_{\delta,i,j} \le L\delta^\alpha\}.$$

Thus, it is enough to prove that $P(G_\delta) \stackrel{\delta \to 0}{\longrightarrow} 0$. To see this, let us fix δ and i for a moment. Then, $((i+j)\delta,\ (i+j+1)\delta]$, $j \ge 1$ are disjoint intervals with the same length δ. Hence,
3) $\{X_{\delta,i,j}\}_{j=1}^m$ are i.i.d.$\approx \delta^{\frac{1}{2}}|Y|$ with $Y \approx N(0,1)$,
4) $P(X_{\delta,i,j} \le L\delta^\alpha) = P(\delta^{\frac{1}{2}}|Y| \le L\delta^\alpha) = P(|Y| \le L\delta^{\alpha - \frac{1}{2}}) \le L\delta^{\alpha-\frac{1}{2}}$,
where we have used the inequality $P(|Y| \le x) \le x$, which is easy to verify. Therefore,

$$\begin{aligned} P(G_\delta) &\le \sum_{i=0}^{\lfloor T/\delta \rfloor} P\left(\bigcap_{j=1}^{m} \{X_{\delta,i,j} \le L\delta^\alpha\}\right) \stackrel{3),4)}{\le} ((T/\delta)+1)\left(L\delta^{\alpha-\frac{1}{2}}\right)^m \\ &= (T+\delta)L^m \delta^{(\alpha-\frac{1}{2})m-1} \stackrel{\delta\to 0}{\longrightarrow} 0 \quad \text{(cf. 2)).} \end{aligned}$$

Step2: We prove (6.31). As for $C^+_{\alpha,B}(t)$, we have to prove that
5) $E \stackrel{\text{def}}{=} \{C^+_{\alpha,B}(t) < \infty \ \text{ for some } t \ge 0\}$ is a null set.
To show this, recall that there exists $\Omega_B \in \mathcal{F}$ with $P(\Omega_B) = 1$ on which $t \mapsto B_t$ is continuous, and hence $t \mapsto B_t$ is locally bounded. Thus, $\Omega_B \cap E \subset F$ (cf. 1)) by Lemma 6.4.2 b) and hence

$$E \subset (\Omega_B \cap E) \cup \Omega_B^{\mathsf{c}} \subset F \cup \Omega_B^{\mathsf{c}}.$$

Since F is a null set by Step 1, obtain 5).
To treat $C^-_{\alpha,B}(t)$, fix $T > 0$ and set $\beta(t) = B(T) - B(T-t)$ $(t \in [0,T])$. Then, $(\beta(t))_{t\in[0,T]}$ is a BM^1_0 and $C^-_{\alpha,B}(t) = C^+_{\alpha,\beta}(T-t)$ for $t \in (0,T]$. Thus, the assertion for $C^-_{\alpha,B}(t)$ follows from that for $C^+_{\alpha,B}(t)$. \(^□^)/

6.5 The Right-Continuous Enlargement of the Canonical Filtration

Let B be a BM^d. We define the *right-continuous enlargement* $(\mathcal{F}_t)_{t\ge0}$ of the canonical filtration $(\mathcal{F}^0_t)_{t\ge0}$ as follows;

$$\mathcal{F}^0_t = \sigma(B_s\ ;\ s \le t), \ \text{ and } \ \mathcal{F}_t = \bigcap_{\varepsilon>0} \mathcal{F}^0_{t+\varepsilon}. \tag{6.33}$$

In particular, $\mathcal{F}_0$ is called the *germ σ-algebra*. The technical advantage of introducing $\mathcal{F}_t$ ("an infinitesimal peeking in the future") is to enlarge $\mathcal{F}^0_t$ to get the right-continuity:

$$\bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon} = \mathcal{F}_t, \ \ \forall t \ge 0. \tag{6.34}$$

Indeed,

$$\bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon} = \bigcap_{\varepsilon>0}\bigcap_{\delta>0} \mathcal{F}^0_{t+\varepsilon+\delta} = \bigcap_{\varepsilon,\delta>0} \mathcal{F}^0_{t+\varepsilon+\delta} = \mathcal{F}_t.$$

Note that $\mathcal{F}_t$ is strictly larger than $\mathcal{F}^0_t$. For example, the r.v. $X = \overline{\lim_{n\to\infty}} B^1(t+\frac{1}{n})$ is $\mathcal{F}_t$-measurable, but not $\mathcal{F}^0_t$-measurable. Here, $X = B^1_t$ a.s. and hence X is $\mathcal{F}^0_t$-measurable up to a null function. In fact, $\mathcal{F}_t$ is larger than $\mathcal{F}^0_t$ only by the null sets in the following sense. Let $\mathcal{N}_t$ denote the totality of $\mathcal{F}_t$-measurable null sets. Then, $\mathcal{F}_t = \sigma(\mathcal{F}^0_t \cup \mathcal{N}_t)$ (Proposition 6.5.3).

Remark To avoid being confused in the future, we find it helpful to clarify the dependence of σ-algebra $\mathcal{F}_t$ on the value of B_0, particularly in the case of $B_0 \equiv x$. In this case, for any $t \geq 0$, the σ-algebra $\mathcal{F}_t$ does not depend on the starting point x. Indeed, $\mathcal{F}^0_t = \sigma(B_s \ ; \ 0 < s \leq t)$, since $B_0 \equiv x$, and hence neither $\mathcal{F}^0_t$ or $\mathcal{F}_t$ depends on x. However, an event A in $\mathcal{F}_t$ may depend on the value of x. For example, take $A = \{f(B_t) > f(B_0)\}$ for some Borel function $f : \mathbb{R}^d \to \mathbb{R}$.

Proposition 6.5.1 *Let B be a* BM^d*, $s \geq 0$, and $\widehat{B}^s = (B_{s+t} - B_s)_{t\geq 0}$ (cf. (6.8)). Then, $\mathcal{F}_s$ and $\widehat{B}^s$ is independent.*

Proof: We take arbitrary $A \in \mathcal{F}_s$, $m \in \mathbb{N}\backslash\{0\}$, $0 \leq t_1 < ... < t_m$ and verify that

$$A \text{ and } (\widehat{B}^s(t_j))_{j=1}^m \text{ are independent.}$$

(cf. Lemma 1.6.5) To do so, we take arbitrary $f \in C_{\mathrm{b}}((\mathbb{R}^d)^m)$ and write

$$F(\widehat{B}^s) = f(\widehat{B}^s(t_1), ..., \widehat{B}^s(t_m)).$$

It is enough to show that

1) $E[F(\widehat{B}^s) : A] = E[F(\widehat{B}^s)]P(A)$.

For $n \in \mathbb{N}\backslash\{0\}$, $A \in \mathcal{F}_s \subset \mathcal{F}^0_{s+\frac{1}{n}}$, and hence A and $\widehat{B}^{s+\frac{1}{n}}$ are independent by Proposition 6.1.9. Thus, we have that

2) $E[F(\widehat{B}^{s+\frac{1}{n}}) : A] = E[F(\widehat{B}^{s+\frac{1}{n}})]P(A)$.

Since $F(\widehat{B}^{s+\frac{1}{n}}) \stackrel{n\to\infty}{\longrightarrow} F(\widehat{B}^s)$ a.s., we obtain 1) from 2) by letting $n \to \infty$. \(^□^)/

By Proposition 6.5.1 and the proof of Corollary 6.5.2, we obtain the following

Corollary 6.5.2 *Let $s \geq 0$, $F \in \mathcal{F}_s$, and $G \in \mathcal{T}_s \stackrel{\text{def}}{=} \sigma(B_t \ ; \ t \geq s)$. Then,*

$$P(G|\mathcal{F}_s) = P(G|B_s), \quad a.s. \tag{6.35}$$

$$P(F \cap G|B_s) = P(F|B_s)P(G|B_s), \quad a.s. \tag{6.36}$$

Corollary 6.5.2 can be used to show that the right-continuous enlargement of $\mathcal{F}_t$ is larger than $\mathcal{F}^0_t$ by null sets:

Proposition 6.5.3 *Let B be a* BM^d, $t \geq 0$. *Then,*

a) $\mathcal{F}_t = \sigma(\mathcal{F}_t^0 \cup \mathcal{N}_t)$, *where* $\mathcal{N}_t$ *denotes the totality of* $\mathcal{F}_t$*-measurable null sets.*

b) (**germ triviality / Blumenthal zero-one law**) *If B is a* BM_x^d *for some* $x \in \mathbb{R}^d$ *and* $A \in \mathcal{F}_0$, *then,* $P(A) \in \{0, 1\}$.

Proof: a) It is clear that $\mathcal{F}_t \supset \sigma(\mathcal{F}_t^0 \cup \mathcal{N}_t)$. We will show the opposite inclusion. Let

$$G \in \mathcal{G}_t \stackrel{\text{def}}{=} \bigcap_{\varepsilon > 0} \sigma(B_{t+s}\ ;\ 0 \leq s \leq \varepsilon).$$

Since $\mathcal{G}_t \subset \mathcal{F}_t \cap \mathcal{T}_t$, we see from (6.35) that

$$\mathbf{1}_G = P(G|\mathcal{F}_t) \stackrel{(6.35)}{=} P(G|B_t), \quad \text{a.s.}$$

Thus, $\mathbf{1}_G$ is a.s. equals to an $\sigma(B_t)$-measurable function. This implies that

$$\mathcal{G}_t \subset \sigma(B_t) \vee \sigma(\mathcal{N}_t).$$

Hence

$$\mathcal{F}_t = \sigma(\mathcal{F}_t^0 \cup \mathcal{G}_t) \subset \sigma(\mathcal{F}_t^0 \cup \mathcal{N}_t).$$

b) Suppose in particular that B is a BM_x^d for some $x \in \mathbb{R}^d$. Then $\mathcal{F}_0^0 = \{\emptyset, \Omega\}$, and hence $\mathcal{F}_0 = \sigma(\mathcal{N}_0)$, which consists only of events A with $P(A) \in \{0, 1\}$. \(^□^)/

Remarks:
1) If B is a BM_x^d for some $x \in \mathbb{R}^d$ and $A \in \mathcal{F}_0$, the value $P(A) = 0, 1$ may differ depending on the choice of the starting point x. For example, let $A = \{B(1/n) \stackrel{n \to \infty}{\longrightarrow} 0\} \in \mathcal{F}_0$. Then, $P(A) = \delta_{0,x}$.
2) The germ triviality is not true in gereral for pre-Brownian motions. In fact, let B be BM_0^1, and U be a r.v. uniformly distributed on $(0, 1)$, which is independent of B. Now, define $\widetilde{B} = (\widetilde{B}_t)_{t \geq 0}$ by

$$\widetilde{B}_t = \begin{cases} B_t & \text{if } t \neq U/n \text{ for any } n \in \mathbb{N}, \\ U & \text{if } t = U/n \text{ for some } n \in \mathbb{N}. \end{cases}$$

Since $P(t = U/n \text{ for some } n \in \mathbb{N}) = 0$ for any fixed $t \geq 0$, B and $\widetilde{B}$ have the same law, and hence the latter is a pre-BM_0^1. However, the germ σ-algebra of $\widetilde{B}$ contains $\sigma(U)$.

Proposition 6.5.4 *Let B be a* BM^1, $t \geq 0$, *and* $h_1 > h_2 > \ldots > h_n \to 0$ *as* $n \to \infty$. *Then, a.s.,* $B(t + h_n) > B(t)$ *for infinitely many n, and* $B(t + h_n) < B(t)$ *for infinitely many n. In particular, the time t is an accumulation point of the set*

$$\{s > t\ ;\ B_s = B_t\}.$$

Proof: Let $\widehat{B}^t$ be defined as in Proposition 6.1.9. Then,

$$\{B(t + h_n) > B(t)\} = \{\widehat{B}^t(h_n) > 0\}, \quad \{B(t + h_n) < B(t)\} = \{\widehat{B}^t(h_n) < 0\}.$$

Since $\widehat{B}^t$ is a BM^1_0 by Proposition 6.1.9, it is enough to prove the proposition for BM^1_0 and for $t=0$. Let

$$A_m = \bigcup_{n \ge m} \{B(h_n) > 0\} \in \mathcal{F}_{h_m}, \quad \text{and} \quad A = \bigcap_{m \ge 1} A_m \in \mathcal{F}_0.$$

Then, $A_1 \supset A_2 \supset \dots$ and $P(A_m) \ge P(B(h_m) > 0) = 1/2$. Thus,

$$P(A) = \lim_{m \to \infty} P(A_m) \ge 1/2.$$

Therefore, $P(A) = 1$ by Proposition 6.5.3, which implies that $B(h_n) > 0$ for infinitely many n. Similarly, $B(h_n) < 0$ for infinitely many n. \(^□^)/

Proposition 6.5.5 *Let B be a* BM^d*. The σ-algebra $\mathcal{T}$ defined as follows is called the* **tail σ-algebra** *for the Brownian motion.*

$$\mathcal{T} \stackrel{\text{def}}{=} \bigcap_{t>0} \sigma(B_s \,;\, s \ge t). \tag{6.37}$$

Let $\check{B}$ be a BM^d *defined by*

$$\check{B}_t = \begin{cases} B_0 + t(B_{1/t} - B_0), & \text{if } t > 0, \\ B_0, & \text{if } t = 0. \end{cases}$$

(cf. Proposition 6.1.5) Then,

$$\check{\mathcal{F}}_0 = \sigma(B_0) \vee \mathcal{T}, \tag{6.38}$$

where $\check{\mathcal{F}}_0$ is the germ σ-algebra for $\check{B}$. In particular, if B is a BM^d_x *for some $x \in \mathbb{R}^d$, then,*

$$\check{\mathcal{F}}_0 = \mathcal{T}, \tag{6.39}$$

which implies that $P(A) \in \{0,1\}$ for all $A \in \mathcal{T}$ **(Tail triviality)**.

Proof: Note that the Brownian motion B is reconstructed from $\check{B}$ by

$$B_t = \begin{cases} \check{B}_0 + t(\check{B}_{1/t} - \check{B}_0), & \text{if } t > 0, \\ \check{B}_0, & \text{if } t = 0, \end{cases}$$

Thus,

$$\sigma(\check{B}_s \,;\, s \le t) = \sigma(B_0, B_{1/s} \,;\, s \le t) = \sigma(B_0, B_s \,;\, s \ge 1/t),$$

and hence

$$\check{\mathcal{F}}_0 = \bigcap_{t>0} \sigma(B_0, \check{B}_s \,;\, s \le t) = \bigcap_{t>0} \sigma(B_0, B_s \,;\, s \ge 1/t) = \sigma(B_0) \vee \mathcal{T}.$$

This proves (6.38), which implies (6.39) for BM^d_x. Finally, the tail triviality is a consequnce of the germ triviality (Proposition 6.5.3) for $\check{B}$. \(^□^)/

Remark: Referring to Proposition 6.5.5 in the case of BM^d_x, the value of $P(A)$ for $A \in \mathcal{T}$ does not depend on the starting point x. Moreover, the tail triviality is true for any BM^d (not only for BM^d_x for some $x \in \mathbb{R}^d$). See Example 6.7.3 below.

Exercise 6.5.1 Let B be a BM^1. Prove the following.
i) For $t \geq 0$, and a sequene $h_1 > h_2 > \ldots > h_n \to 0$, $\overline{\lim\limits_{n\to\infty}} \dfrac{B(t+h_n)-B(t)}{\sqrt{h_n}} = \infty$, a.s. Hint:By considering $\widehat{B}^t$ in Proposition 6.1.9, we may assume that B is a BM^1_0 and $t=0$. Then, prove that, for any $c > 0$, the event $\overline{\lim\limits_{n\to\infty}} \dfrac{B(h_n)}{\sqrt{h_n}} \geq c$ has positive probability.
ii) For a sequene $t_1 < t_2 < \ldots < t_n \to \infty$, $\overline{\lim\limits_{n\to\infty}} \dfrac{B(t_n)}{\sqrt{t_n}} = \infty$, a.s.

6.6 The Strong Markov Property

Throughout this subsection, we assume that $(\Omega, \mathcal{F}, P)$ is a probability space. We start with an abstract preparation. Let $\mathcal{G}$ be a sub σ-algebra of $\mathcal{F}$, $(S, \mathcal{B})$ be a measurable space and $\Omega_0 \subset \Omega$, without assuming that $\Omega_0 \in \mathcal{F}$. A map $\varphi : \Omega_0 \to S$ is said to be $\mathcal{G}/\mathcal{B}$-measurable on Ω_0 if

$$B \in \mathcal{B} \implies \exists A \in \mathcal{G}, \ \{\omega \in \Omega_0 \ ; \ \varphi(\omega) \in B\} = \Omega_0 \cap A. \tag{6.40}$$

If $\Omega_0 \in \mathcal{G}$, then, (6.40) is equivalent to that

$$B \in \mathcal{B} \implies \{\omega \in \Omega_0 \ ; \ \varphi(\omega) \in B\} \in \mathcal{G}. \tag{6.41}$$

In this subsection, we always assume (6.4) for BM^d, i.e. the map $t \mapsto B_t(\omega)$ is continuous for *all* $\omega \in \Omega$. We will denote by $(\mathcal{F}_t)_{t\geq 0}$ the right-continuous enlargement (6.33) of the canonical filtration.

Proposition 6.6.1 (**Strong Markov property I**) *Let B be a BM^d and T be a stopping time. Then,*

a) *the r.v. B_T is $\mathcal{F}_T$-measurable on $\{T < \infty\}$.*

Suppose in addition that $P(T < \infty) > 0$. Then, under $P(\,\cdot\,|T < \infty)$,

b) *the process $\widehat{B}^T$ defined as follows is a BM^d_0,*

$$\widehat{B}^T = (\widehat{B}^T_t)_{t\geq 0} = (B_{T+t} - B_T)_{t\geq 0}.$$

c) *$\mathcal{F}_T$ and $\widehat{B}^T$ are independent.*

Proof: a) This follows from Lemma 6.6.10 below.
b) and c) Let $m \geq 1$, $0 \leq t_1 < \ldots < t_m$, and $f \in C_{\mathrm{b}}((\mathbb{R}^d)^m \to \mathbb{R})$ be arbitrary. Let $\widehat{B}^s$ for $s \geq 0$ be defined by (6.8). We write

$$F(\widehat{B}^s) = f(\widehat{B}^s(t_1), \ldots, \widehat{B}^s(t_m))$$

We will prove the following equality for an arbitrary $A \in \mathcal{F}_T$.

1) $E[F(\widehat{B}^T)1_A|T < \infty] = E[F(\widehat{B}^0)]P(A|T < \infty)$.

Let us admit 1) for a moment to finish the proof. Setting $A = \Omega$, we have

2) $E[F(\widehat{B}^T)|T < \infty] = E[F(\widehat{B}^0)]$.

Plugging 2) into 1), we also have that

3) $E[F(\widehat{B}^T)1_A|T<\infty] = E[F(\widehat{B}^T)|T<\infty]P(A|T<\infty)$.

We see b) and c) respectively from 2) and 3) (cf. Lemma 1.6.5).
The equality 1) can be seen as follows. Let T_n, $n = 1, 2, ...$ be a discrete approximation of T from the right defined by

$$T_n = \begin{cases} \frac{j}{n}, & \text{if } \frac{j-1}{n} < T \leq \frac{j}{n} \text{ for some } j \in \mathbb{N}, \\ \infty, & \text{if } T = \infty. \end{cases} \tag{6.42}$$

If $T < \infty$, then $0 \leq T_n - T \leq \frac{1}{n}$, $n \geq 1$, and hence $T_n \stackrel{n\to\infty}{\longrightarrow} T$. Let $C_{n,j} \stackrel{\text{def}}{=} \{\frac{j-1}{n} < T \leq \frac{j}{n}\}$. Since $A \cap C_{n,j} \in \mathcal{F}_{j/n}$, we have by the Markov property I (Proposition 6.1.9) that

4) $E[F(\widehat{B}^{j/n}) : A \cap C_{n,j}] = E[F(\widehat{B}^0)]P(A \cap C_{n,j})$.

Therefore,

$$\begin{aligned} & E[F(\widehat{B}^{T_n}) : A \cap \{T < \infty\}] \\ &= \sum_{j\geq 0} E[F(\widehat{B}^{T_n}) : A \cap C_{n,j}] = \sum_{j \geq 0} E[F(\widehat{B}^{j/n}) : A \cap C_{n,j}] \\ &\stackrel{4)}{=} \sum_{j\geq 0} E[F(\widehat{B}^0)]P(A \cap C_{n,j}) = E[F(\widehat{B}^0)]P(A \cap \{T < \infty\}). \end{aligned}$$

Note that $\widehat{B}^{T_n}_t(\omega) \stackrel{n\to\infty}{\longrightarrow} \widehat{B}^T_t(\omega)$ for all $t \geq 0$ and $\omega \in \{T < \infty\}$. Thus, letting $n \to \infty$, and dividing the both hands sides by $P(T < \infty)$, we have 1). \(^□^)/

Remark T_n defined by (6.42) is a stopping time. Indeed, for $t \geq 0$,

$$\{T_n \leq t\} = \{T \leq \lfloor nt \rfloor /n\} \in \mathcal{F}_{\lfloor nt \rfloor /n} \in \mathcal{F}_t.$$

Let B be a BM^d, T be an a.s. finite stopping time for B. The strong Markov property allows us to construct a new Brownian motion by replacing the path after the time T by an another Brownian motion β, which is independent of $\mathcal{F}_T$. More precisely, we have

Corollary 6.6.2 (**Concatenation of Brownian motions II**) *Let B be a* BM^d, *T be an a.s. finite stopping time for B, and β be a* BM$^d_0$ *which is independent of $\mathcal{F}_T$. Then the process $\widetilde{B} = (\widetilde{B}_t)_{t\geq 0}$ defined as follows is a* BM^d *such that $\widetilde{B}_0 = B_0$.*

$$\widetilde{B}_t = \begin{cases} B_t, & \text{if } t \leq T, \\ B_T + \beta_{t-T}, & \text{if } t \geq T. \end{cases}$$

As a consequence, the Brownian motion β is expressed as

$$\beta_t = \widetilde{B}_{T+t} - \widetilde{B}_T, \ \ t \geq 0.$$

Proof: Let $S = (\mathbb{R}^d)^{[0,\infty)}$ and define $F : [0,\infty) \times S \times S \longrightarrow S$ by

$$F(s,x,y)(t) = \begin{cases} x(t), & \text{if } t \leq s, \\ x(s) + y(t-s), & \text{if } t \geq s. \end{cases}$$

Define also $X : \Omega \to S$ and $\widehat{B}^T : \Omega \to S$ by

$$X = (B_{t\wedge T})_{t\geq 0}, \quad \widehat{B}^T = (B_{t+T} - B_T)_{t\geq 0}.$$

Then,

1) $B = F(T, X, \widehat{B}^T), \quad \widetilde{B} = F(T, X, \beta).$

By (4.34) and Lemma 6.6.10, (T, X) is $\mathcal{F}_T$-measurable, and hence by assumption, β is a BM_0^d which is independent of (T, X). On the other hand, we see from Proposition 6.6.1 that $\widehat{B}^T$ is a BM_0^d which is independent of (T, X). As a consequence,

2) $(T, X, \widehat{B}^T) \approx (T, X, \beta).$

This, together with 1), implies that $B \approx \widetilde{B}$. \\(^□^)/

Let B be a BM^1 and

$$T_a = \inf\{t \geq 0 \; ; \; B_t = a\}, \quad a \in \mathbb{R}. \tag{6.43}$$

Recall that we assume (6.4). Thus, it follows from Lemma 6.6.11 below that T_a is a stopping time w.r.t. $(\mathcal{F}_t^0)_{t\geq 0}$, and hence w.r.t. $(\mathcal{F}_t)_{t\geq 0}$. Note also that

$$\varlimsup_{t\to\infty} B_t = \infty, \quad \varliminf_{t\to\infty} B_t = -\infty \;\text{ a.s.}$$

(cf. Exercise 6.5.1) Thus, $T_a < \infty$ a.s. for any $a \in \mathbb{R}$.

The following lemma (reflection principle) is the source of a couple of useful consequences (Proposition 6.6.4, Corollary 6.6.5). It will be useful to note in advance that for $a \in \mathbb{R}$, the map

$$x \mapsto 2a - x \;\; (\mathbb{R} \to \mathbb{R})$$

represents the reflection (mirror image) relative to the point a. The core of the reflection principle (which can be seen from the proof below) is that for BM_0^1,

$$(B_t)_{t\geq T_a} \approx (2a - B_t)_{t\geq T_a}.$$

Lemma 6.6.3 (Reflection principle). *Suppose that B is a BM_0^1, and that $a \in \mathbb{R}\backslash\{0\}$, $t > 0$, $J \in \mathcal{B}(\mathbb{R})$. Then,*

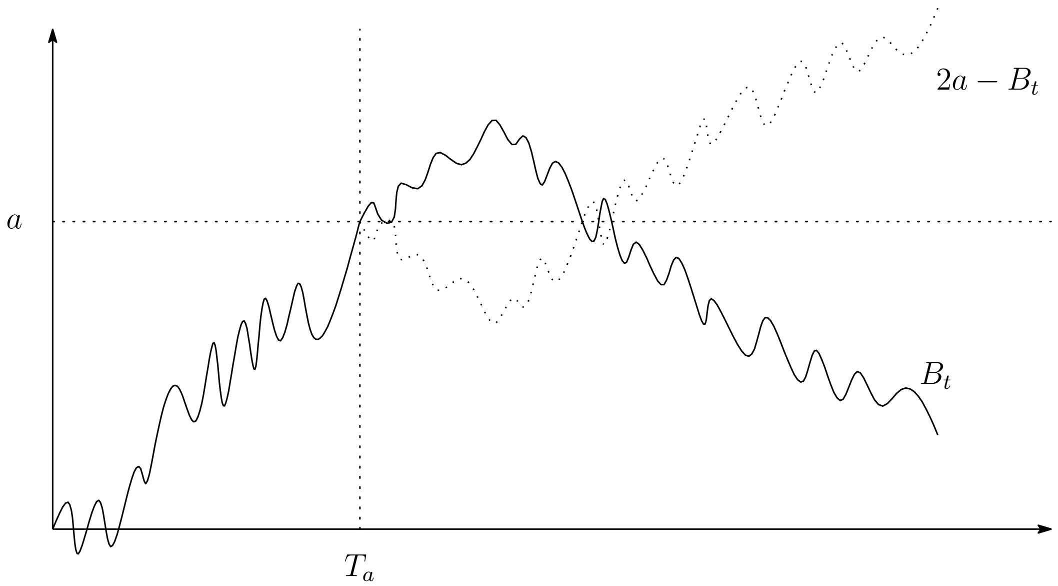
$$P(T_a \leq t, B_t \in J) = P(T_a \leq t, B_t \in 2a - J). \tag{6.44}$$

Let $J_a^+ = J \cap [a, \infty)$ and $J_a^- = J \cap (-\infty, a]$. Then, for $a > 0$,

$$P(T_a \leq t, B_t \in J) = P(B_t \in J_a^+) + P(B_t \in 2a - J_a^-) = \int_J h_t(x \vee (2a - x))dx, \tag{6.45}$$

where $h_t(x) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right)$. For $a < 0$,

$$P(T_a \leq t, B_t \in J) = P(B_t \in J_a^-) + P(B_t \in 2a - J_a^+) = \int_J h_t(x \wedge (2a - x))dx. \tag{6.46}$$



Proof: (6.44): Let

$$\widetilde{B}_t = \begin{cases} B_t, & \text{if } t \leq T_a, \\ 2a - B_t, & \text{if } t \geq T_a. \end{cases}$$

We first verify that

1) $(\widetilde{B}_t)_{t \geq 0}$ is a BM_0^1.

To do so, we define $\beta = (\beta_t)_{t \geq 0}$ as follows. If $T_a < \infty$, then

$$\beta_t \stackrel{\text{def}}{=} a - B(t + T_a) = -(B(t + T_a) - B(T_a)), \ \ \forall t \geq 0.$$

If $T_a = \infty$, then $\beta_t \stackrel{\text{def}}{=} 0$, $\forall t \geq 0$. Then, by the strong Markov property, β is a BM_0^1 which is independent of $\mathcal{F}_{T_a}$ Note that

$$t \geq T_a \implies B(T_a) + \beta(t - T_a) = a - (B_t - a) = 2a - B_t.$$

Thus, 1) follows from Corollary 6.6.2.
On the other hand, we have

2) $\begin{cases} \widetilde{T}_a \stackrel{\text{def}}{=} \inf\{t \geq 0 \ ; \ \widetilde{B}_t = a\} = T_a, \\ T_a \leq t \implies \widetilde{B}_t = 2a - B_t. \end{cases}$

Therefore,

$$P(T_a \leq t, B_t \in J) \stackrel{1)}{=} P(\widetilde{T}_a \leq t, \widetilde{B}_t \in J) \stackrel{2)}{=} P(T_a \leq t, B_t \in 2a - J).$$

This proves (6.44).
(6.45), (6.46): Since the proofs for (6.45) and (6.46) are similar, we present the proof only for (6.45). We have

3) $P(T_a \leq t, B_t \in J_a^-) \stackrel{(6.44)}{=} P(T_a \leq t, B_t \in 2a - J_a^-)$.

Moreover, for $a > 0$, $J_a^+ \cup (2a - J_a^-) \subset [a, \infty)$, and hence

4) $\{B_t \in J_a^+ \cup (2a - J_a^-)\} \subset \{T_a \leq t\}$.

Finally, note that

5) $x \vee (2a-x) = \begin{cases} x, & \text{if } x \geq a, \\ 2a - x, & \text{if } x \leq a. \end{cases}$

Therefore,

$$\begin{aligned}
P(T_a \leq t, B_t \in J) &= P(T_a \leq t, B_t \in J_a^+) + P(T_a \leq t, B_t \in J_a^-) \\
&\stackrel{3)}{=} P(T_a \leq t, B_t \in J_a^+) + P(T_a \leq t, B_t \in 2a - J_a^-) \\
&\stackrel{4)}{=} P(B_t \in J_a^+) + P(B_t \in 2a - J_a^-) \\
&= \int_{J_a^+} h_t(x)dx + \int_{J_a^-} h_t(2a-x)dx \stackrel{5)}{=} \int_J h_t(x \vee (2a-x))dx.
\end{aligned}$$

\(^□^)/

Remark: The equalities (6.45) and (6.46) can be used to prove the following. For $a > 0$ and $t > 0$,

$$P(T_a > t, B_t \in J) = P(B_t \in J_a^-) - P(B_t \in 2a - J_a^-) = \int_{J_a^-} (h_t(x) - h_t(2a-x))dx. \quad (6.47)$$

For $a < 0$ and $t > 0$,

$$P(T_a > t, B_t \in J) = P(B_t \in J_a^+) - P(B_t \in 2a - J_a^+) = \int_{J_a^+} (h_t(x) - h_t(2a-x))dx. \quad (6.48)$$

Indeed, for $a > 0$,

$$\begin{aligned}
P(T_a > t, B_t \in J) &= P(T_a > t, B_t \in J_a^-) = P(B_t \in J_a^-) - P(T_a \leq t, B_t \in J_a^-) \\
&\stackrel{(6.45)}{=} P(B_t \in J_a^-) - P(B_t \in 2a - J_a^-)
\end{aligned}$$

The proof for the case of $a < 0$ is similar.

For BM_0^1, the distribution of T_a can be computed as follows (See also Corollary 7.2.4).

Proposition 6.6.4 *For* BM_0^1 *and* $a \in \mathbb{R}\backslash\{0\}$,

$$T_a \approx a^2/B_1^2 \approx \frac{|a|}{\sqrt{2\pi t^3}} \exp\left(-\frac{a^2}{2t}\right) dt. \quad (6.49)$$

Proof: Since the proofs for the case of $a > 0$ and of $a < 0$ are similar, we present the proof only for the case $a > 0$.
(6.49): Let $t > 0$, For $J = \mathbb{R}$, $J_a^+ = 2a - J_a^- = [a, \infty)$. Thus, it follows from (6.45) for $J = \mathbb{R}$ that

$$P(T_a \leq t) \stackrel{(6.45)}{=} 2P(B_t \geq a) = P(a \leq |B_t|) = P(a^2/B_1^2 \leq t),$$

where we have used that $B_t \approx \sqrt{t} B_1$ to see the third equality. We see from Example 1.2.6 that $B_1^2/a^2 \approx \gamma(a^2/2, 1/2)$. Thus, we know the density of the r.v. a^2/B_1^2 from Exercise 1.2.8. This proves the last equality of (6.49). \(^□^)/

Remark: We have $T_a \approx a^2/B_1^2$ (Proposition 6.6.4) and $B_1^2/a^2 \approx \gamma(a^2/2, 1/2)$. Thus, by Example 2.3.5, we obtain the Laplace transform of T_a.

$$E \exp(-\lambda T_a) = \exp(-|a|\sqrt{2\lambda}), \quad \lambda > 0. \tag{6.50}$$

See also Proposition 7.2.3 for an alternative proof of (6.50).

Let B be a BM^1 and

$$S_t = \sup_{s \le t} B_s, \quad s_t = \inf_{s \le t} B_s, \quad t \ge 0. \tag{6.51}$$

Recall that we assume (6.4). Thus, S_t and s_t are $\mathcal{F}_t^0$-measurable, since the supremum/infimum over $s \le t$ can be replaced by that over $s \in \mathbb{Q} \cap [0, t]$.

Corollary 6.6.5 *Let*

$$Q_+ = \{(x, y) \in \mathbb{R} \times (0, \infty)\ ;\ x \le y\}, \quad Q_- = \{(x, y) \in \mathbb{R} \times (-\infty, 0)\ ;\ x \ge y\}.$$

Suppose that B is a BM_0^1 and that $t > 0$. Then,

a) $S_t \approx |B_t|$. *Moreover,*

$$(B_t, S_t) \approx (2y - x)\sqrt{\frac{2}{\pi t^3}} \exp\left(-\frac{(2y - x)^2}{2t}\right) dxdy \ \ \text{on } Q_+.$$

b) $s_t \approx -|B_t|$. *Moreover,*

$$(B_t, s_t) \approx (x - 2y)\sqrt{\frac{2}{\pi t^3}} \exp\left(-\frac{(x - 2y)^2}{2t}\right) dxdy \ \ \text{on } Q_-.$$

Proof: a) Since $S_t \ge a \iff T_a \le t$, we have for all $a > 0$,

$$P(S_t \ge a) = P(T_a \le t) \stackrel{(6.49)}{=} P(|B_t| \ge a).$$

This proves that $S_t \approx |B_t|$. On the other hand, we have

1) $P(B_t \in J, S_t \ge a) = P(T_a \le t, B_t \in J) \stackrel{(6.45)}{=} \displaystyle\int_{J_a^+} h_t(x)dx + \int_{J_a^-} h_t(2a - x)dx.$

On the other hand, let

$$k_t(x) \stackrel{\text{def}}{=} -h_t'(x) = \frac{x}{\sqrt{2\pi t^3}} \exp\left(-\frac{x^2}{2t}\right).$$

Then,

$$\textbf{2)} \quad \begin{cases} h_t(x) &= \displaystyle\int_x^\infty k_t(y)dy = 2\int_x^\infty k_t(2y - x)dy \\ h_t(2a - x) &= 2\displaystyle\int_a^\infty k_t(2y - x)dy. \end{cases}$$

Therefore,

$$P(B_t \in J, S_t \geq a) \stackrel{1),2)}{=} 2\int_{J_a^+} dx \int_x^\infty k_t(2y-x)dy + 2\int_{J_a^-} dx \int_a^\infty k_t(2y-x)dy$$
$$= 2\int_J dx \int_a^\infty k_t(2y-x)\mathbf{1}_{\{x\leq y\}}dy.$$

This shows that the r.v. (B_t, S_t) has the density $2k_t(2y-x)$ on the set Q_+, which proves a).
b) Similar to the above. (^□^)/

Our next objective is to prove

Proposition 6.6.6 *Let* B *be a* BM^1, *and*

$$\mathscr{Z}_a = \{t \geq 0 \ ; \ B_t = a\}, \ \ a \in \mathbb{R}.$$

Then, for any $a \in \mathbb{R}$, *a.s.,* $\mathscr{Z}_a$ *is a closed set with Lebesgue measure zero, without isolated points. In particular, a.s.,* $\mathscr{Z}_a$ *has the cardinality of continuity.*

We prepare two lemmas. Thanks to Proposition 6.6.1, Proposition 6.5.4 can be generalized in the following way.

Lemma 6.6.7 *Let* B *be a* BM^1, T *be a stopping time such that* $P(T < \infty) > 0$ *and* $h_1 > h_2 > \ldots > h_n \to 0$ *as* $n \to \infty$. *Then,* $P(\cdot|T < \infty)$*-a.s.,* $B(T+h_n) > B(T)$ *for infinitely many* n, *and* $B(T+h_n) < B(T)$ *for infinitely many* n. *In particular, the time* T *is an accumulation point of the set*

$$\{s > T \ ; \ B_s = B_T\}.$$

Proof: Let $\widehat{B}^T$ be defined as in Proposition 6.6.1. Then,

$$\{B(T+h_n) > B(T)\} = \{\widehat{B}^T(h_n) > 0\}, \ \ \{B(T+h_n) < B(T)\} = \{\widehat{B}^T(h_n) < 0\}.$$

By Proposition 6.6.1, $\widehat{B}^T$ is a BM$^1_0$ under $P(\cdot|T < \infty)$. Thus, it is enough to prove this proposition by replacing $\widehat{B}^T$ (under $P(\cdot|T < \infty)$) by BM$^1_0$. Therefore, we obtain the conclusion from Proposition 6.5.4. \(^□^)/

Lemma 6.6.8 *A complete metric space* $S \neq \emptyset$ *without isolated points has at least the cardinality of continuity.*

Proof: We construct an injection $f : \{0,1\}^{\mathbb{N}} \to S$ as follows. Choose an $x_0 \in S$ arbitrarily. Since x_0 is not isolated, there exists $x_1 \in S \backslash \{x_0\}$. We then take disjoint closed balls B_0, B_1 with radiuses ≤ 1, centered, respectively at x_0, x_1. Next, for $\alpha = 0, 1$, we take two different points $x_{\alpha 0}, x_{\alpha 1} \in B_\alpha$ and disjoint closed balls $B_{\alpha 0}, B_{\alpha 1} \subset B_\alpha$ with radiuses $\leq 1/2$, centered, respectively at $x_{\alpha 0}, x_{\alpha 1}$. By repeating this procedure, we obtain for any $\alpha = (\alpha_j)_{j=0}^\infty \in \{0,1\}^{\mathbb{N}}$ and $n \in \mathbb{N}$

$$f_n(\alpha) \stackrel{\text{def}}{=} x_{\alpha_0\alpha_1,\ldots,\alpha_n}, \ \ B_n(\alpha) \stackrel{\text{def}}{=} B_{\alpha_0\alpha_1,\ldots,\alpha_n}.$$

The sequence $f_n(\alpha)$ is a Cauchy sequence, since, if $m \leq n$, then, $f_n(\alpha) \in B_m(\alpha)$, and hence,

$$\mathrm{dist}(f_m(\alpha), f_n(\alpha)) \leq 1/m.$$

Consequently, the sequence $f_n(\alpha)$ converges a limit $f(\alpha)$ as $n \to \infty$. The map $f : \{0,1\}^{\mathbb{N}} \to S$ is injective, as is easily seen as follows. If $\alpha, \beta \in \{0,1\}^{\mathbb{N}}$, $\alpha \neq \beta$, then $\alpha_m \neq \beta_m$ for some $m \in \mathbb{N}$, and therefore,

$$B_m(\alpha) \cap B_m(\beta) = \emptyset, \quad f(\alpha) \in B_m(\alpha), \quad f(\beta) \in B_m(\beta).$$

Hence $f(\alpha) \neq f(\beta)$. \(^□^)/

Proof of Proposition 6.6.6: Clearly $\mathcal{Z}_a$ is closed, since it is the inverse image of a point a by the continuous function $t \mapsto B_t$. Denote by $|\mathcal{Z}_a|$ the Lebesgue measure of $\mathcal{Z}_a$. Since,

$$|\mathcal{Z}_a| = \int_0^\infty \mathbf{1}\{B_t = a\} dt,$$

We have

$$E|\mathcal{Z}_a| = \int_0^\infty P(B_t = a) dt = 0,$$

which implies that $|\mathcal{Z}_a| = 0$ a.s. Let

$$T_{a,r} = \inf\{t \geq r \ ; \ B_t = a\}, \quad r \geq 0.$$

Then, we see that $T_{a,r}$ is a stopping time, similarly as in Lemma 6.6.11. Therefore, by Lemma 6.6.7, and by the fact that $B(T_{a,r}) = a$ a.s., for any $r \geq 0$, there exists an event $A_r \in \mathcal{F}$ of probability one, on which $T_{a,r}$ is an accumulation point of the set

$$\{t > T_{a,r} \ ; \ B_t = a\} \subset \mathcal{Z}_a.$$

Let $\mathbb{Q}_+ = \mathbb{Q} \cap [0, \infty)$ and $A = \bigcap_{r \in \mathbb{Q}_+} A_r$. Then, $P(A) = 1$, and

1) on the event A, all $T_{a,r}$, $r \in \mathbb{Q}_+$ are accumulation points of $\mathcal{Z}_a$.

Thus, it is enough to prove that,

2) on the event A, all $t \in \mathcal{Z}_a \backslash \{T_{a,r} \ ; \ r \in \mathbb{Q}_+\}$ are accumulation points of $\mathcal{Z}_a$.

This can be seen as follows. For $t \in \mathcal{Z}_a \backslash \{T_{a,r} \ ; \ r \in \mathbb{Q}_+\}$, let $r(n) \in \mathbb{Q}_+ \cap [0, t)$ be such that $r(n) \nearrow t$. Then, $r(n) < t$ and $t \in \mathcal{Z}_a$. Thus, it follows from the definition of $T_{a,r(n)}$ that

$$r(n) \leq T_{a,r(n)} < t,$$

and hence $\mathcal{Z}_a \ni T_{a,r(n)} \stackrel{n \to \infty}{\longrightarrow} t$.
Since $\mathcal{Z}_a$ is a closed set $\neq \emptyset$ without isolated point, it has the cardinality of continuity by Lemma 6.6.8. \(^□^)/

Complement

Example 6.6.9 $(\star)$ Let B be BM_0^1 and U be a uniformly distributed r.v. on $(0,1)$. We define $\widetilde{B}$ by

$$\widetilde{B}(t) = \begin{cases} 0, & \text{if } t > 0,\ t \in U + \mathbb{Q} \text{ and } B(t) \neq 0, \\ 1, & \text{if } t > 0 \text{ and } B(t) = 0, \\ B(t), & \text{if otherwise.} \end{cases}$$

Then,
a) $\widetilde{B}$ is a pre-BM_0^1.
b) If $\omega \in \Omega_B$, then $t \mapsto \widetilde{B}(t)$ is discontinuous for all $t \geq 0$.

Proof: a) For any fixed $t > 0$, $P(\{t \in (U + \mathbb{Q})\} \cup \{B_t = 0\}) = 0$, and hence $B(t) = \widetilde{B}(t)$ a.s.
b) Let $\omega \in \Omega_B$ and $t_0 > 0$.
Case1: $t_0 = 0$ (Then,$\widetilde{B}(t_0) = \widetilde{B}(0) = 0$). Since there exists $t_n \in (0,\infty)$ such that $B(t_n) = 0$ and $t_n \stackrel{n\to\infty}{\longrightarrow} 0$,

$$\widetilde{B}(t_n) = 1 \stackrel{n\to\infty}{\longrightarrow} 1 \neq 0 = \widetilde{B}(0).$$

Thus $\widetilde{B}$ is discontinuous at 0.
Case2: $t_0 > 0$, $B(t_0) = 0$ ($\widetilde{B}(t_0) = 1$ in this case). Since $(U + \mathbb{Q}) \cap (0,\infty)$ is dense in $(0,\infty)$, there exists $r_n \in \mathbb{Q}$ such that $(0,\infty) \ni U + r_n \stackrel{n\to\infty}{\longrightarrow} t_0$. Then,

$$0 = \widetilde{B}(U + r_n) \stackrel{n\to\infty}{\longrightarrow} 0 \neq 1 = \widetilde{B}(t_0).$$

Thus $\widetilde{B}$ is discontinuous at t_0.
Case3: $t_0 > 0$, $B(t_0) \neq 0$ and $t_0 \in U + \mathbb{Q}$ ($\widetilde{B}(t_0) = 0$ in this case). Since $(0,\infty)\backslash((U+\mathbb{Q}) \cup \mathcal{Z}_0)$ is dense in $(0,\infty)$, there exists $t_n \in (0,\infty)\backslash((U+\mathbb{Q}) \cup \mathcal{Z}_0)$ such that $t_n \stackrel{n\to\infty}{\longrightarrow} t_0$. Then,

$$\widetilde{B}(t_n) = B(t_n) \stackrel{n\to\infty}{\longrightarrow} B(t_0) \neq 0 = \widetilde{B}(t_0).$$

Thus $\widetilde{B}$ is discontinuous at t_0.
Case4: $t_0 > 0$, $B(t_0) \neq 0$ and $t_0 \notin U + \mathbb{Q}$ ($\widetilde{B}(t_0) = B(t_0)$ in this case). Since $(U + \mathbb{Q}) \cap (0,\infty)$ is dense in $(0,\infty)$, there exists $r_n \in \mathbb{Q}$ such that $(0,\infty) \ni U + r_n \stackrel{n\to\infty}{\longrightarrow} t_0$. Then,

$$0 = \widetilde{B}(U + r_n) \stackrel{n\to\infty}{\longrightarrow} 0 \neq B(t_0) = \widetilde{B}(t_0).$$

Thus $\widetilde{B}$ is discontinuous at t_0. \(^□^)/

Exercise 6.6.1 Suppose that B is a BM_0^1. Then, prove that for $a \in \mathbb{R}\backslash\{0\}$ and $t > 0$,

$$P(T_a > t) = \sqrt{\frac{2}{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)2^n n!} \left(\frac{|a|}{\sqrt{t}}\right)^{2n+1}.$$

In particular, $P(T_a > t) = |a|\sqrt{\frac{2}{\pi t}} + O(t^{-3/2})$ as $t \to \infty$. [Hint: $T_a \approx a^2/B_1^2$]

Exercise 6.6.2 Suppose that B is a BM_0^1. Then, use Corollary 6.6.5 to prove the following. i) $S_t - B_t \approx |B_t|$. ii) $2S_t - B_t \approx |X_t|$, where X is a BM_0^3.

Exercise 6.6.3 Suppose that B is a BM_x^1 with $x > 0$ and that $J \in \mathcal{B}([0,\infty))$. Then, prove that

$$P(B_t \in J, T_0 > t) = \int_J (h_t(y - x) - h_t(y + x))dy.$$

[Hint: In terms of BM_0^1, the LHS $=P(x + B_t \in J, T_{-x} > t)$.]

Exercise 6.6.4 Suppose that B is a BM^1_0, $s > 0$, and X is a r.v. with the Cauchy distribution with paraameter 1. Then, prove the following.
i) Let $\widehat{B}^s$ be from Proposition 6.1.9 and let $T_a(\widehat{B}^s) = \inf\{t \geq 0\ ;\ \widehat{B}^s_t = a\}$, $a \in \mathbb{R}$. Then,

$$T_{\pm B_s}(\widehat{B}^s) \approx \left(B_s/\widehat{B}^s_1\right)^2 \approx sX^2.$$

[Hint: The first equality in law follows from Proposition 6.6.4, and the second from (1.70).]
ii) $T_{s,0} \stackrel{\rm def}{=} \inf\{t > s\ ;\ B_t = 0\} \approx (1 + X^2)s$. [Hint: $T_{s,0} = s + T_{-B_s}(\widehat{B}^s)$.]
iii) (**First Arcsin Law**) $T^-_{s,0} \stackrel{\rm def}{=} \sup\{t < s\ ;\ B_t = 0\} \approx s/(1 + X^2) \approx sY$, where Y is a r.v. with the arcsin law. [Hint: The first equality in law follows from the relation $T^-_{s,0} < t \Leftrightarrow s < T_{t,0}$, and the second from Exercise 1.2.14.]

($\star$) Complement to section 6.6

We prove Proposition 6.6.1a) in the following slightly generalized form.

Lemma 6.6.10 *Let S be a metric space, and $(X_t : \Omega \to S)_{t\geq 0}$ be a process adapted to a filtration $(\mathcal{F}_t)_{t\geq 0}$. Suppose that the function $t \mapsto X_t(\omega)$ is either right-continuous for all $\omega \in \Omega$, or left-continuous for all $\omega \in \Omega$. Then, for a stopping time T, the r.v. X_T is $\mathcal{F}_T$-measurable on $\{T < \infty\}$.*

Proof: Here, we assume that the function $t \mapsto X_t(\omega)$ is left-continuous for all $\omega \in \Omega$, since this is enough for Proposition 6.6.1a). See Corollary 6.6.15 below for the right-continuous case. Let T_n, $n = 1, 2, ...$ be a discrete approximation of T from the left defined by

$$T_n = \begin{cases} 0, & \text{if } T \leq \frac{1}{n}, \\ \frac{j}{n}, & \text{if } \frac{j}{n} < T \leq \frac{j+1}{n} \text{ for some } j = 1, 2, ..., \\ \infty, & \text{if } T = \infty. \end{cases}$$

If $T < \infty$, then $0 \leq T - T_n \leq \frac{1}{n}$, $n \geq 1$, and hence $T_n \stackrel{n\to\infty}{\longrightarrow} T$. Note that $\{T_n < \infty\} = \{T < \infty\}$ for all $n \geq 1$. By the left-continuity, $X(T_n) \stackrel{n\to\infty}{\longrightarrow} X(T)$ on $\{T < \infty\}$. Therefore, it is enough to prove that $X(T_n)$ is $\mathcal{F}_T$-measurable on $\{T < \infty\}$ for all $n \geq 1$. (We need to approximate T from the left, rather than the right, so that the following argument goes through.) Now, for $B \in \mathcal{B}(S)$, let

$$C_{n,0} = \left\{T \leq \tfrac{1}{n},\ X_0 \in B\right\}, \quad C_{n,j} = \left\{\tfrac{j}{n} < T \leq \tfrac{j+1}{n},\ X_{j/n} \in B\right\}, \ \ j \geq 1.$$

Then,

$$\{T < \infty, X(T_n) \in B\} = \bigcup_{j\in\mathbb{N}} C_{n,j}.$$

Thus, in view of (6.41), it is enough to show that

1) $C_{n,j} \in \mathcal{F}_T$ for all $j \in \mathbb{N}$.

This can be seen as follows. For $t \geq 0$,

$$\begin{aligned} C_{n,0} \cap \{T \leq t\} &= \left\{T \leq \tfrac{1}{n} \wedge t,\ X_0 \in B\right\} \in \mathcal{F}_t, \\ C_{n,j} \cap \{T \leq t\} &= C_{n,j} = \left\{\tfrac{j}{n} < T \leq \tfrac{j+1}{n} \wedge t,\ X_{j/n} \in B\right\} \in \mathcal{F}_t, \ \ j \geq 1. \end{aligned}$$

These imply 1). \\(^□^)/

Lemma 6.6.11 *Let S be a metric space, $X = (X_t : \Omega \to S)_{t \geq 0}$ be a process, T_A, and $(\mathcal{F}_t^0)_{t \geq 0}$ be defined as Example 4.2.2. Suppose that the function $t \mapsto X_t(\omega)$ is continuous for all $\omega \in \Omega$ and that $A \subset S$ is closed. Then, T_A is a stopping time w.r.t. $(\mathcal{F}_t^0)_{t \geq 0}$.*

Proof: We introduce a process $Y_t \stackrel{\text{def}}{=} \text{dist}(X_t, A)$, and observe that the following are equivalent.

1) $T_A \leq t$,

2) $\exists s \in [0, t]$, $X_s \in A$.

3) $\exists s \in [0, t]$, $Y_s = 0$.

4) $\inf_{r \in [0,t] \cap \mathbb{Q}} Y_r = 0$.

1) $\Leftrightarrow$ 2): Since A is closed, the set $\{t \geq 0 \; ; \; X_t \in A\} \subset [0, \infty)$ is also closed, and hence has a minimum, which is T_A. This explains 1) $\Rightarrow$ 2), while the converse is obvious.
2) $\Leftrightarrow$ 3): Since A is closed, $X_s \in A$ if and only if $Y_s = 0$.
3) $\Rightarrow$ 4): Assume 3) and let $r_n \in \mathbb{Q} \cap [0, t]$ be such that $r_n \to s$. Then, by the continuity of $t \mapsto Y_t$, $Y(r_n) \to Y(s) = 0$, and hence 4) holds.
3) $\Leftarrow$ 4): Let $r_n \in \mathbb{Q} \cap [0, t]$ be such that $Y(r_n) \to 0$. Then, there exist $s \in [0, t]$ and a subsequence $r_{n(k)} \to s$. By the continuity of $t \mapsto Y_t$, $Y(r_{n(k)}) \to Y(s) = 0$.
The equivalence of 1) and 4) implies that

$$\{T_A \leq t\} = \left\{ \inf_{r \in [0,t] \cap \mathbb{Q}} Y_r = 0 \right\} \in \mathcal{F}_t^0.$$

Thus, T_A is a stopping time w.r.t. $(\mathcal{F}_t^0)_{t \geq 0}$. \\(^□^)/

In what follows, we give a more complete account to Lemma 6.6.10 including the right-continuous case. We assume that a filtration $(\mathcal{F}_t)_{t \geq 0}$ is given and that stopping times are associated with this filtration.

Definition 6.6.12 (**Adaptedness, progressive measurability**) Suppose that $(S, \mathcal{B})$ is a measurable space and that $X = (X_t)_{t \in \mathbb{T}}$ is a process with values in S.
► X is said to be **adapted** if the map $X_t : \Omega \to S$ is $\mathcal{F}_t/\mathcal{B}$-measurable for all $t \geq 0$.
► X is said to be **progressively measurable** if the following map is $(\mathcal{B}(\mathbb{T} \cap [0, t]) \otimes \mathcal{F}_t)/\mathcal{B}$-measurable for all $t \geq 0$.

$$(s, \omega) \mapsto X_s(\omega) \quad ((\mathbb{T} \cap [0, t]) \times \Omega \longrightarrow S)$$

Clearly, a progressively measurable process is adapted. In the following proposition, we will see two basic conditions under which the converse is also true.

Proposition 6.6.13 *Let the process X in Definition 6.6.12 be adapted. Then, under either of the following conditions a),b), X is progressively measurable.*

a) *$\mathbb{T}$ is at most countable.*

b) *$\mathbb{T} = [0, \infty)$, S is a metric space, and that the function $t \mapsto X_t(\omega)$ is right-continuous for all $\omega \in \Omega$, or left-continuous for all $\omega \in \Omega$.*

Proof: a) Let $t \geq 0$ and $B \in \mathcal{B}$. Since X is adapted, we have for $s \in \mathbb{T} \cap [0, t]$,

$$\{s\} \times \{\omega \in \Omega \; ; \; X_s(\omega) \in B\} \in \mathcal{B}(\mathbb{T} \cap [0, t]) \otimes \mathcal{F}_t.$$

Thus,

$$\begin{aligned} &\{(s, \omega) \in \mathbb{T} \cap [0, t] \times \Omega \; ; \; X_s(\omega) \in B\} \\ &= \bigcup_{s \in \mathbb{T} \cap [0,t]} \{s\} \times \{\omega \in \Omega \; ; \; X_s(\omega) \in B\} \in \mathcal{B}(\mathbb{T} \cap [0, t]) \otimes \mathcal{F}_t. \end{aligned}$$

Thus, X is progressively measurable.
b) Suppose that the function $t \mapsto X_t(\omega)$ is right-continuous for all $\omega \in \Omega$ (The proof is similar if we suppose the left-continuity). For $n \in \mathbb{N}$, let

$$X^{(n)}(s, \omega) = \sum_{j=0}^{\infty} X((j+1)/2^n, \omega) \mathbf{1}\{s \in [j/2^n, (j+1)/2^n)\}, \quad s \geq 0.$$

Then, for $t \geq 0$ and $B \in \mathcal{B}$,

$$\begin{aligned} &\{(s, \omega) \in [0, t] \times \Omega \; ; \; X^{(n)}(s, \omega) \in B\} \\ &= \bigcup_{\substack{j \in \mathbb{N} \\ (j+1)/2^n \leq t}} [j/2^n, (j+1)/2^n) \times \{\omega \in \Omega \; ; \; X((j+1)/2^n, \omega) \in B\} \in \mathcal{B}([0, t]) \otimes \mathcal{F}_t. \end{aligned}$$

Thus, $X^{(n)}$ is progressively measurable for all $n \in \mathbb{N}$. Moreover, by $X^{(n)}(s, \omega) \stackrel{n \to \infty}{\longrightarrow} X(s, \omega)$ by the right-continuity. Therefore, X is progressively measurable. \(^□^)/

Proposition 6.6.14 *Let everything be as in Definition 6.6.12, and let T be a stopping time.*

a) *The process $(X_{t \wedge T})_{t \in \mathbb{T}}$ is adapted. $\iff$ The r.v. X_T is $\mathcal{F}_T$-measurable on $\{T < \infty\}$.*

b) *Suppose that the process $(X_t)_{t \geq 0}$ is progressively measurable. Then, the process $(X_{t \wedge T})_{t \geq 0}$ is again progressively measurable, hence is adapted. As a consequence, X_T is $\mathcal{F}_T$-measurable on $\{T < \infty\}$.*

Proof: a) ($\Rightarrow$) Let $B \in \mathcal{B}$ and $t \geq 0$. Then, $\{X_{t \wedge T} \in B\} \in \mathcal{F}_t$ by the assumption. Therefore,

$$\{X_T \in B\} \cap \{T \leq t\} = \{X_{t \wedge T} \in B\} \cap \{T \leq t\} \in \mathcal{F}_t.$$

($\Leftarrow$) Let $B \in \mathcal{B}$ and $t \geq 0$. Then,

1) $\{X_{t \wedge T} \in B\} = \{t < T, X_t \in B\} \cup \{T \leq t, X_T \in B\}$.

Clearly,

2) $\{t < T, X_t \in B\} \in \mathcal{F}_t$.

On the other hand, by the assumption, $\{T < \infty, X_T \in B\} = A \cap \{T < \infty\}$ for some $A \in \mathcal{F}_T$, and hence

3) $\{T \leq t, X_T \in B\} = A \cap \{T \leq t\} \in \mathcal{F}_t$.

It follows from 1)–3) that $\{X_{t\wedge T} \in B\} \in \mathcal{F}_t$.
b) For notational simplicity, we consider the case of $\mathbb{T} = [0, \infty)$. It is easy to see that the function $(s, \omega) \mapsto s \wedge T$ is $(\mathcal{B}([0, t]) \otimes \mathcal{F}_t)/\mathcal{B}([0, t])$-measurable. In fact, for any $u \in [0, t]$,

$$\begin{aligned}\{(s, \omega) \; ; \; s \wedge T \leq u\} &= \{(s, \omega) \; ; \; s \leq u\} \cup \{(s, \omega) \; ; \; T \leq u\} \\ &= ([0, u] \times \Omega) \cup ([0, t] \times \{\omega \; ; \; T \leq u\}) \in \mathcal{B}([0, t]) \otimes \mathcal{F}_t.\end{aligned}$$

Hence

1) the map $(s, \omega) \mapsto (s \wedge T, \omega)$ is $(\mathcal{B}([0, t]) \otimes \mathcal{F}_t)/(\mathcal{B}([0, t]) \otimes \mathcal{F}_t)$-measurable.

On the other hand, by assumption,

2) the map $(s, \omega) \mapsto X_s(\omega)$ is $(\mathcal{B}([0, t]) \otimes \mathcal{F}_t)/\mathcal{B}$-measurable.

Since the map $(s, \omega) \mapsto X_{s\wedge T}(\omega)$ is the composition of those of 1) and 2), it is $(\mathcal{B}([0, t]) \otimes \mathcal{F}_t)/\mathcal{B}$-measurable. \(^□^)/

Combinning Proposition 6.6.13 and Proposition 6.6.14, we obtain the following

Corollary 6.6.15 *Let the process X in Definition 6.6.12 be adapted and T be a stopping time. Then, under either of the conditions a),b) in Proposition 6.6.13, $(X_{t\wedge T})_{t\in\mathbb{T}}$ is adapted, and X_T is $\mathcal{F}_T$-measurable on $\{T < \infty\}$.*

6.7 Alternative Formulations of Markov Properties

For the rest of section 6, we will work on a special measurable space $(\Omega, \mathcal{F})$ defined by

$$\Omega = \{\omega = (\omega_t)_{t\geq 0} \in (\mathbb{R}^d)^{[0,\infty)} \; ; \; t \mapsto \omega_t \text{ is continuous.}\}, \tag{6.52}$$

$$\mathcal{F} = \sigma[\omega_t \; ; \; t \geq 0]. \tag{6.53}$$

For $\omega = (\omega_t)_{t\geq 0} \in \Omega$, we write $B_t = B_t(\omega) = \omega_t$. Then, we consider the filtration $(\mathcal{F}_t)_{t\geq 0}$ defined by (6.33). For $x \in \mathbb{R}^d$, we let P_x denote a unique probability measure on $(\Omega, \mathcal{F})$ under which $(B_t)_{t\geq 0}$ is a BM^d_x. (cf. Proposition 6.1.12). We denote by E_x the expectation w.r.t. P_x. For $x \in \mathbb{R}^d$, let

$$x + B \stackrel{\text{def}}{=} (x + B_t)_{t\geq 0} \in \Omega. \tag{6.54}$$

For $s \geq 0$ and $\omega \in \Omega$, we define

$$\theta_s\omega = (B_{s+t}(\omega))_{t\geq 0} \tag{6.55}$$

Lemma 6.7.1 *The map $(s, \omega) \mapsto \theta_s\omega$, $([0, \infty)\times\Omega \longrightarrow \Omega)$ is $\mathcal{B}([0, \infty))\otimes\mathcal{F}/\mathcal{F}$-measurable.*

Proof: By Lemma 1.5.2, it is enough to verify that the map $(s, \omega) \mapsto \omega_{s+t}$ is $\mathcal{B}([0, \infty)) \otimes \mathcal{F}/\mathcal{B}(\mathbb{R}^d)$-measurable for each fixed $t \geq 0$. The map $\omega \mapsto \omega_{s+t}$ is clearly $\mathcal{F}/\mathcal{B}(\mathbb{R}^d)$-measurable. This, together with the continuity of $s \mapsto \omega_{s+t}$, implies that the map $(s, \omega) \mapsto \omega_{s+t}$ is $\mathcal{B}([0, \infty)) \otimes \mathcal{F}/\mathcal{B}(\mathbb{R}^d)$-measurable. \(^□^)/

Let $\mu \in \mathcal{P}(\mathbb{R}^d)$ be arbitrary. For $A \in \mathcal{F}$, the function $x \mapsto P_x(A)$ is Borel measurable by Lemma 6.1.14. Therefore, we can define

$$P(A) = \int_{\mathbb{R}^d} P_x(A) d\mu(x). \tag{6.56}$$

It follows from the bounded convergence theorem that $A \mapsto P(A)$ is a probability measure on $(\Omega, \mathcal{F})$.

- For the rest of this section, P denotes the probability measure (6.56) on $(\Omega, \mathcal{F})$, and the associated expectation will be denoted by E.

Theorem 6.7.2 (**Markov property II**) *Let $F : \Omega \to \mathbb{R}$ be bounded, $\mathcal{F}$-measurable, and $G : \Omega \to \mathbb{R}$ be bounded, $\mathcal{F}_s$-measurable for $s \geq 0$. Then,*

$$E[G \cdot F \circ \theta_s] = E[G E_{B(s)} F]. \tag{6.57}$$

Remark: Since F is $\mathcal{F}$-measurable, and $\theta_s$is $\mathcal{F}/\mathcal{F}$-measurable (Lemma 6.7.1), $F \circ \theta_s$ is $\mathcal{F}$-measurable. Thus, the left-hand side of (6.57) is well defined. On the other hand, the quantity $E_{B(s)}F$ on the right-hand side of (6.57) should be understood as the value of the function $f(x) \stackrel{\text{def}}{=} E_x F$ evaluated at $x = B_s$. Since f is Borel measurable (Lemma 6.1.14), $f(B_s)$ is $\sigma[B_s]$-measurable.

Proof: We see from Proposition 6.1.9 that

1) $\mathcal{F}_s$ and $(\widehat{B}^s_t)_{t \geq 0}$ are independent,

2) $E[F((y + \widehat{B}^s_t)_{t \geq 0})] = E_0[F((y + B_t)_{t \geq 0})] = E_y F$ for $y \in \mathbb{R}^d$.

Let us consider the product space $(\Omega^2, \mathcal{F} \otimes \mathcal{F}, P \otimes P)$ and denote an element of Ω^2 by $(\omega, \widehat{\omega})$. Then, by 1),

3) the law of the r.v. $G(\omega)F((B_s(\omega) + \widehat{B}^s_t(\omega))_{t \geq 0})$ under $P(d\omega)$ is the same as the law of $G(\omega)F((B_s(\omega) + \widehat{B}^s_t(\widehat{\omega}))_{t \geq 0})$ under $(P \otimes P)(d\omega d\widehat{\omega})$.

Since $B_t \circ \theta_s = B_s + \widehat{B}^s_t$, we have that

4)
$$\begin{cases} E[G \cdot F \circ \theta_s] & = & E[G \cdot F((B_s + \widehat{B}^s_t)_{t \geq 0})] \\ & \stackrel{3)}{=} & \int_{\Omega^2} (P \otimes P)(d\omega d\widehat{\omega}) G(\omega) F((B_s(\omega) + \widehat{B}^s_t(\widehat{\omega}))_{t \geq 0}) \\ & \stackrel{\text{Fubini}}{=} & \int_{\Omega} G(\omega) P(d\omega) \int_{\Omega} P(d\widehat{\omega}) F((B_s(\omega) + \widehat{B}^s_t(\widehat{\omega}))_{t \geq 0}) \end{cases}$$

On the other hand,

5) $\displaystyle\int_{\Omega} P(d\widehat{\omega}) F((B_s(\omega) + \widehat{B}^s_t(\widehat{\omega}))_{t \geq 0}) \stackrel{2)}{=} E_{B_s(\omega)} F.$

Putting 4) and 5) together, we obtain

$$E[G \cdot F \circ \theta_s] = \int_\Omega G(\omega)\, E_{B_s(\omega)} F\, P(d\omega) = E[G E_{B(s)} F].$$

\(^□^)/

We present a couple of applications of Theorem 6.7.2.

Example 6.7.3 Let $t > 0$, $\mathcal{T}_t = \sigma(B_{t+s}\ ;\ s \geq 0)$ and $\mathcal{T} = \bigcap_{t>0} \mathcal{T}_t$ ($\mathcal{T}$ is the tail σ-algebra for Brownian motion, cf. Proposition 6.5.5).

a) For any $t > 0$, $x, y \in \mathbb{R}^d$, the measures P_x and P_y are mutually absolutely continuous on $\mathcal{T}_t$.

b) $P(A) = P_0(A) \in \{0, 1\}$ for any $A \in \mathcal{T}$ and $\mu \in \mathcal{P}(\mathbb{R}^d)$, where the measure P is defined by (6.56).

Proof: a) Note that

$$\mathcal{T}_t = \sigma(B_s \circ \theta_t\ ;\ s \geq 0).$$

Thus, if $A \in \mathcal{T}_t$, then, $A = \theta_t^{-1} C$ for some $C \in \mathcal{F}$. Therefore, for all $x \in \mathbb{R}^d$,

1) $P_x(A) \stackrel{(6.57)}{=} E_x[P_{B(t)}(C)] = \int_{\mathbb{R}^d} h_t(y - x) P_y(C)$, cf. (6.14).

Suppose that $P_x(A) = 0$ for some $x \in \mathbb{R}^d$. Then, it follows from 1) that $P_y(C) = 0$ for almost all $y \in \mathbb{R}^d$, which implies again by 1), that $P_x(A) = 0$ for all $x \in \mathbb{R}^d$.
b) It follows from Proposition 6.5.5 and a) above that $P_x(A) = P_0(A) \in \{0, 1\}$ for all $x \in \mathbb{R}^d$. Thus,

$$P(A) = \int_{\mathbb{R}^d} P_x(A) d\mu(x) = P_0(A) \in \{0, 1\}.$$

\(^□^)/

Example 6.7.4 Let $A \subset \mathbb{R}^d$ be either closed or open, and let $T_A = \inf\{t \geq 0\ ;\ B_t \in A\}$. Suppose that

$$M \stackrel{\text{def}}{=} \sup_{x \in A^c} E_x T_A < \infty.$$

Then, for any $\lambda \in (0, 1/M)$,

$$\sup_{x \in A^c} E_x \exp(\lambda T_A) \leq 1/(1 - \lambda M) < \infty.$$

Proof: We write $T = T_A$ for simplicity. By the power series expansion of the exponential, it is enough to show that

1) $\sup_{x \in A^c} E_x[T^n] \leq n! M^n$ for all $n \in \mathbb{N} \backslash \{0\}$.

We prove this by induction on n. By assumption, 1) is true for $n = 1$. Suppose that $n \geq 2$ and 1) is true for $n-1$. For $t \geq 0$, note that

$$t^n = n! \int_{0<s_1<s_2<\ldots<s_n<t} ds_1 ds_2 \cdots ds_n$$

and that $T = t + T \circ \theta_t$ on the event $\{T \geq t\}$. Hence

$$\begin{aligned}
E_x[T^n] &= n! E_x \int_{0<s_1<s_2<\ldots<s_n<T} ds_1 ds_2 \cdots ds_n \\
&= n! \int_0^\infty ds_1 E_x \left[\mathbf{1}\{T > s_1\} E_{B(s_1)} \int_{s_1<s_2<\ldots<s_n<s_1+T} ds_2 \cdots ds_n \right] \\
&= n! \int_0^\infty ds_1 E_x \left[\mathbf{1}\{T > s_1\} E_{B(s_1)} \int_{0<s_2<\ldots<s_n<T} ds_2 \cdots ds_n \right] \\
&= n \int_0^\infty ds_1 E_x \left[\mathbf{1}\{T > s_1\} E_{B(s_1)}[T^{n-1}] \right] \\
&\leq n \cdot (n-1)! M^{n-1} \int_0^\infty P_x(T > s_1) ds_1 = n! M^n.
\end{aligned}$$

\(^□^)/

Lemma 6.7.5 *Let T be a stopping time. Then, the map $\omega \mapsto (B_{T(\omega)+t}(\omega))_{t \geq 0}$ is $\mathcal{F}/\mathcal{F}$-measurable on $\{T < \infty\}$, cf. (6.40).*

Proof: By Lemma 1.5.2, it is enough to verify that the map $\omega \mapsto B_{T(\omega)+t}(\omega)$ is $\mathcal{F}/\mathcal{B}(\mathbb{R}^d)$-measurable on $\{T < \infty\}$ for each fixed $t \geq 0$. Since $T + t$ is a stopping time, it follows from Lemma 6.6.10 that the map $\omega \mapsto B_{T(\omega)+t}(\omega)$ is $\mathcal{F}_{T+t}/\mathcal{B}(\mathbb{R}^d)$-measurable on $\{T < \infty\}$, and hence is $\mathcal{F}/\mathcal{B}(\mathbb{R}^d)$-measurable on $\{T < \infty\}$ \(^□^)/

Let T be a stopping time. For an $\omega \in \Omega$ with $T(\omega) < \infty$, we define

$$\theta_T \omega = (B_{T(\omega)+t}(\omega))_{t \geq 0} \tag{6.58}$$

By Lemma 6.7.5, the map $\omega \mapsto \theta_T \omega$ is $\mathcal{F}/\mathcal{F}$-measurable on $\{T < \infty\}$.

Theorem 6.7.6 (**Strong Markov property II**) *Let T be a stopping time. Suppose that $F : \Omega \to \mathbb{R}$ is a bounded, $\mathcal{F}$-measurable, and that $G : \Omega \to \mathbb{R}$ is bounded, $\mathcal{F}_T$-measurable. Then,*

$$E[G \cdot F \circ \theta_T : T < \infty] = E[G E_{B(T)} F : T < \infty]. \tag{6.59}$$

Remark: Since F is $\mathcal{F}$-measurable, and $\theta_T$is $\mathcal{F}/\mathcal{F}$-measurable on $\{T < \infty\}$, $F \circ \theta_T$ is $\mathcal{F}$-measurable on $\{T < \infty\}$. Thus, the left-hand sides of (6.59) is well defined. On the other hand, the quantity $E_{B(T)}F$ on the right-hand side of (6.59) should be understood as the value of the function $f(x) \stackrel{\text{def}}{=} E_x F$ evaluated at $x = B_T$. Since f is Borel measurable (Lemma 6.1.14), and B_T is $\mathcal{F}_T$-measurable on $\{T < \infty\}$ (Lemma 6.6.10), $f(B_T)$ is $\mathcal{F}_T$-measurable on $\{T < \infty\}$.

Proof: We may assume that $P(T < \infty) > 0$. We write $P' = P(\,\cdot\,|T < \infty)$ and $E' = E[\,\cdot\,|T < \infty]$. Then, we see from Proposition 6.6.1 that

1) $\mathcal{F}_T$ and $(\widehat{B}^T_t)_{t\ge 0}$ are independent under P',

2) $E'[F((y+\widehat{B}^T_t)_{t\ge 0})] = E_0[F((y+B_t)_{t\ge 0})] = E_y F$ for $y \in \mathbb{R}^d$.

Let us consider the product space $(\Omega^2, \mathcal{F}\otimes\mathcal{F}, P'\otimes P')$ and denote an element of Ω^2 by $(\omega, \widehat{\omega})$. Since B_T is $\mathcal{F}_T$-measurable on $\{T<\infty\}$ (Lemma 6.6.10), it follows from 1) that

3) the law of the r.v. $G(\omega)F((B_T(\omega)+\widehat{B}^T_t(\omega))_{t\ge 0})$ under $P'(d\omega)$ is the same as the law of $G(\omega)F((B_T(\omega)+\widehat{B}^T_t(\widehat{\omega}))_{t\ge 0})$ under $(P'\otimes P')(d\omega d\widehat{\omega})$.

Since $B_t\circ\theta_T = B_T + \widehat{B}^T_t$ on $\{T<\infty\}$, we have that

$$
\text{4)}\ \begin{cases} E'[G\cdot F\circ\theta_T] &= E'[G\cdot F((B_T+\widehat{B}^T_t)_{t\ge 0})] \\ &\stackrel{3)}{=} \displaystyle\int_{\Omega^2}(P'\otimes P')(d\omega d\widehat{\omega})G(\omega)F((B_T(\omega)+\widehat{B}^T_t(\widehat{\omega}))_{t\ge 0}) \\ &\stackrel{\text{Fubini}}{=} \displaystyle\int_{\Omega}G(\omega)P'(d\omega)\int_{\Omega}P'(d\widehat{\omega})F((B_T(\omega)+\widehat{B}^T_t(\widehat{\omega}))_{t\ge 0}) \end{cases}
$$

On the other hand,

5) $\displaystyle\int_{\Omega}P'(d\widehat{\omega})F((B_T(\omega)+\widehat{B}^T_t(\widehat{\omega}))_{t\ge 0}) \stackrel{2)}{=} E_{B_T(\omega)}F.$

Putting 4) and 5) together, we obtain

$$
E'[G\cdot F\circ\theta_T] = \int_{\Omega} G(\omega)\, E_{B_T(\omega)}F\, P'(d\omega) = E'[GE_{B(T)}F].
$$

Multiplying the both hands sides by $P(T<\infty)$, we obtain (6.59). \\(^□^)/

Exercise 6.7.1 (**Khasmin'skii's lemma**) Suppose that $f:\mathbb{R}^d\to[0,\infty)$ is Borel measurable, $0<t\le\infty$ and that

$$
M \stackrel{\text{def}}{=} \sup_{x\in\mathbb{R}^d} E_x\int_0^t f(B_s)ds < 1.
$$

Then, prove that

$$
\sup_{x\in\mathbb{R}^d} E_x \exp\left(\int_0^t f(B_s)ds\right) \le 1/(1-M) < \infty.
$$

[Hint:Example 6.7.4]

6.8 (⋆) The Second Arcsin Law

Throught this subsection, we denote by $\mathbf{M}_{\mathrm{b}}(\mathbb{R}^d)$ the set of bounded Borel measurable functions on $\mathbb{R}^d$. For $V\in\mathbf{M}_{\mathrm{b}}(\mathbb{R}^d)$, with $\inf V>0$, we define the **resolvent operator** $G_V:\mathbf{M}_{\mathrm{b}}(\mathbb{R}^d)\to\mathbf{M}_{\mathrm{b}}(\mathbb{R}^d)$ by

$$
G_V f(x) = E_x\int_0^\infty \exp\left(-\int_0^t V(B_s)ds\right) f(B_t)dt,\quad f\in\mathbf{M}_{\mathrm{b}}(\mathbb{R}^d),\ x\in\mathbb{R}^d. \tag{6.60}
$$

Lemma 6.8.1 *For* $U \in \mathbf{M}_{\mathrm{b}}(\mathbb{R}^d)$*, with* $\inf U > 0$*, and* $V \in \mathbf{M}_{\mathrm{b}}(\mathbb{R}^d \to [0,\infty))$*, the operators* G_U *and* G_{U+V} *satisfy the* **resolvent equation***:*

$$G_U - G_{U+V} = G_U V G_{U+V}.$$

Proof: To simplify the notation, we introduce $A_t^U \stackrel{\text{def}}{=} \int_0^t U(B_s)ds$, and similarly, A_t^V and A_t^{U+V}. Note that

1) $1 - \exp(-A_t^V) = \exp(-A_t^V)(\exp(A_t^V) - 1) = \displaystyle\int_0^t V(B_s)\exp(-(A_t^V - A_s^V))ds.$

and that

2) $$\left\{\begin{array}{rl} & E_x\left[\displaystyle\int_s^\infty \exp(-(A_t^{U+V} - A_s^{U+V}))f(B_t)dt \,\middle|\, \mathcal{F}_s\right] \\ = & E_x\left[\displaystyle\int_0^\infty \exp(-(A_{s+t}^{U+V} - A_s^{U+V}))f(B_{s+t})dt \,\middle|\, \mathcal{F}_s\right] \\ = & E_{B_s}\left[\displaystyle\int_0^\infty \exp(-A_t^{U+V})f(B_t)dt\right] = G_{U+V}f(B_s). \end{array}\right.$$

Therefore,

$$\begin{aligned} & G_U f(x) - G_{U+V} f(x) \\ &= E_x \int_0^\infty \exp\left(-A_t^U\right)\left(1 - \exp\left(-A_t^U\right)\right) f(B_t)dt \\ &\stackrel{1)}{=} E_x \int_0^\infty \exp\left(-A_t^U\right) dt \int_0^t V(B_s)\exp(-(A_t^V - A_s^V))f(B_t)ds \\ &= \int_0^\infty ds E_x\left[\exp\left(-A_s^U\right) V(B_s) \int_s^\infty \exp(-(A_t^{U+V} - A_s^{U+V}))f(B_t)dt\right] \\ &\stackrel{2)}{=} \int_0^\infty ds E_x\left[\exp\left(-A_s^U\right) V(B_s) G_{U+V} f(B_s)\right] = G_U(VG_{U+V}f)(x). \end{aligned}$$

\\(^□^)/

From here on, we focus on the case of $d = 1$.

Lemma 6.8.2 *For* $V \in \mathbf{M}_{\mathrm{b}}(\mathbb{R})$ *with* $\inf V > 0$ *and* $f \in \mathbf{M}_{\mathrm{b}}(\mathbb{R})$*,*

$$u \stackrel{\text{def}}{=} G_V f \in C_{\mathrm{b}}(\mathbb{R}).$$

Suppose in addition that V *and* f *are piecewise continuous, with the respective sets of discontinuities* D_V *and* D_f*. Then,* $u \in C^1(\mathbb{R}) \cap C^2(\mathbb{R}\backslash(D_V \cup D_f))$ *and*

$$\tfrac{1}{2}u'' = Vu - f, \quad \text{on } \mathbb{R}\backslash(D_V \cup D_f). \tag{6.61}$$

Proof: Let $\lambda \stackrel{\text{def}}{=} \inf V > 0$ and $\widetilde{V} \stackrel{\text{def}}{=} V - \lambda \in \mathbf{M}_{\mathrm{b}}(\mathbb{R} \to [0,\infty))$. We then have by the resolvent equation that

1) $u = G_\lambda f - G_\lambda(\widetilde{V}u)$.

Let $h_t(x) \stackrel{\text{def}}{=} \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right)$, $t > 0$, $x \in \mathbb{R}$. We see from Lemma 2.3.4 that

2) $\displaystyle\int_0^\infty e^{-\lambda t} h_t(x) dt = \frac{1}{\sqrt{2\lambda}} e^{-|x|\sqrt{2\lambda}}.$

Thus,

3)
$$\left\{ \begin{array}{rcl} G_\lambda f(x) & = & \displaystyle\int_0^\infty e^{-\lambda t} E_x f(B_t) dt = \int_0^\infty e^{-\lambda t} dt \int_{-\infty}^\infty h_t(x-y) f(y) dy \\ & \stackrel{1)}{=} & \displaystyle\frac{1}{\sqrt{2\lambda}} \int_{-\infty}^\infty e^{-|x-y|\sqrt{2\lambda}} f(y) dy \\ & = & \displaystyle\frac{e^{-x\sqrt{2\lambda}}}{\sqrt{2\lambda}} \int_{-\infty}^x e^{y\sqrt{2\lambda}} f(y) dy + \frac{e^{x\sqrt{2\lambda}}}{\sqrt{2\lambda}} \int_x^\infty e^{-y\sqrt{2\lambda}} f(y) dy. \end{array} \right.$$

We see from 3) that $G_\lambda f \in C_{\mathrm{b}}(\mathbb{R})$. Similarly, $G_\lambda(\widetilde{V}u) \in C_{\mathrm{b}}(\mathbb{R})$. Hence $u \in C_{\mathrm{b}}(\mathbb{R})$ by 1). We suppose from here on that V and f are piecewise continuous. Then, we see from 3) that $G_\lambda f \in C^1(\mathbb{R}\backslash D_f)$. Similarly, $G_\lambda(\widetilde{V}u) \in C^1(\mathbb{R}\backslash D_V)$ (Note that $D_{\widetilde{V}u} \subset D_V$). Hence $u \in C^1(\mathbb{R}\backslash(D_V \cup D_f))$ by 1). Moreover, for $x \in \mathbb{R}\backslash D_f$,

$$(G_\lambda f)'(x) = -\frac{e^{-x\sqrt{2\lambda}}}{\sqrt{2\lambda}} \int_{-\infty}^x e^{y\sqrt{2\lambda}} f(y) dy + \frac{e^{x\sqrt{2\lambda}}}{\sqrt{2\lambda}} \int_x^\infty e^{-y\sqrt{2\lambda}} f(y) dy.$$

In particular, we have $(G_\lambda f)'(y-) = (G_\lambda f)'(y+)$ for each $y \in D_f$. Therefore, we have $G_\lambda f \in C^1(\mathbb{R})$. Similarly, $G_\lambda(\widetilde{V}u) \in C^1(\mathbb{R})$. Hence $u \in C^1(\mathbb{R})$ by 1). Moreover, we see from 3) that

4) $\frac{1}{2}(G_\lambda f)'' = \lambda G_\lambda f - f$ on $\mathbb{R}\backslash D_f$.

Similarly,

5) $\frac{1}{2}(G_\lambda(\widetilde{V}u))'' = \lambda G_\lambda(\widetilde{V}u) - \widetilde{V}u$ on $\mathbb{R}\backslash D_V$.

We see from 1),4),5) that $u \in C^2(\mathbb{R}\backslash(D_V \cup D_f))$ and (6.61). \(^□^)/

Lemma 6.8.3 *Let $\alpha, \beta > 0$ and $\gamma \in \mathbb{R}$. Suppose that $u \in C^1(\mathbb{R}) \cap C^2(\mathbb{R}\backslash\{0\})$ is a bounded solution to the following differerential equation.*

$$\frac{1}{2}u''(x) = \left\{ \begin{array}{ll} \alpha u(x) - \gamma, & \text{if } x < 0, \\ \beta u(x) - \gamma, & \text{if } x > 0. \end{array} \right.$$

Then,

$$u(x) = \left\{ \begin{array}{ll} \frac{\gamma}{\alpha}\left(\frac{\sqrt{\alpha}-\sqrt{\beta}}{\sqrt{\beta}} \exp(x\sqrt{2\alpha}) + 1\right), & \text{if } x < 0, \\ \frac{\gamma}{\beta}\left(\frac{\sqrt{\beta}-\sqrt{\alpha}}{\sqrt{\alpha}} \exp(-x\sqrt{2\beta}) + 1\right), & \text{if } x > 0. \end{array} \right.$$

In particular, $u(0) = \gamma/\sqrt{\alpha\beta}$.

Proof: The solution to the differerential equation in question must be of the form:

$$u(x) = \left\{ \begin{array}{ll} A_+ \exp(x\sqrt{2\alpha}) + A_- \exp(-x\sqrt{2\alpha}) + \frac{\gamma}{\alpha}, & \text{if } x < 0, \\ B_+ \exp(x\sqrt{2\beta}) + B_- \exp(-x\sqrt{2\beta}) + \frac{\gamma}{\beta}, & \text{if } x > 0. \end{array} \right.$$

Since u is bounded, we have $A_- = B_+ = 0$. Then,

$$\begin{aligned} u(0-) &= A_+ + (\gamma/\alpha), \ \ u'(0-) = \sqrt{2\alpha}A_+, \\ u(0+) &= B_- + (\gamma/\beta), \ \ u'(0+) = -\sqrt{2\beta}B_-. \end{aligned}$$

These, together with $u(0-) = u(0+)$, and $u'(0-) = u'(0+)$ imply that $A_+ = \frac{\gamma}{\alpha}\frac{\sqrt{\alpha}-\sqrt{\beta}}{\sqrt{\beta}}$ and $B_- = \frac{\gamma}{\beta}\frac{\sqrt{\beta}-\sqrt{\alpha}}{\sqrt{\alpha}}$. \(^□^)/

Proposition 6.8.4 (**The Second Arcsin Law**) *Let B be a* BM_0^1, $t > 0$, *and*

$$A_t = \int_0^t \mathbf{1}_{\{B_s>0\}} ds.$$

Then, the r.v. A_t/t has the arcsin law, i.e., $A_t/t \approx \frac{dx}{\pi\sqrt{x(1-x)}}$ on $(0,1)$.

Proof: Let $\alpha, \beta > 0$ and $V(x) = \alpha + \beta\mathbf{1}_{\{x>0\}}$, $x \in \mathbb{R}$. Then, by Lemma 6.8.2,

$$u \stackrel{\text{def}}{=} G_V 1 \in C^1(\mathbb{R}) \cap C^2(\mathbb{R}\backslash\{0\}),$$

and

$$\tfrac{1}{2}u''(x) = \begin{cases} \alpha u(x) - 1, & \text{if } x < 0, \\ (\alpha+\beta)u(x) - 1, & \text{if } x > 0. \end{cases}$$

Thus, by Lemma 6.8.3, we have $u(0) = 1/\sqrt{\alpha(\alpha+\beta)}$, i.e.,

1) $\displaystyle\int_0^\infty e^{-\alpha t} E\exp\left(-\beta A_t\right) dt = \frac{1}{\sqrt{\alpha(\alpha+\beta)}}.$

We have on the other hand that

2) $\displaystyle\int_0^\infty e^{-\alpha t} dt \int_0^t \frac{e^{-\beta y} dy}{\pi\sqrt{y(t-y)}} = \frac{1}{\sqrt{\alpha(\alpha+\beta)}}.$

To prove 2), we note that

3) $\displaystyle\int_0^\infty \frac{e^{-\alpha t} dt}{\sqrt{t}} = \sqrt{\frac{\pi}{\alpha}}.$

Then,

$$\begin{aligned} \text{LHS of 2)} &= \frac{1}{\pi}\int_0^\infty \frac{e^{-\beta y} dy}{\sqrt{y}} \int_y^\infty \frac{e^{-\alpha t} dt}{\sqrt{t-y}} = \frac{1}{\pi}\int_0^\infty \frac{e^{-(\alpha+\beta)y} dy}{\pi\sqrt{y}} \int_0^\infty \frac{e^{-\alpha y} dt}{\sqrt{t}} \\ &\stackrel{3)}{=} \frac{1}{\sqrt{\alpha(\alpha+\beta)}}. \end{aligned}$$

By 1),2) and the uniqueness of the Laplace transform (Example 1.8.3) in the variable α, we have that

$$E\exp\left(-\beta A_t\right) = \int_0^t \frac{e^{-\beta y} dy}{\pi\sqrt{y(t-y)}},$$

and hence

$$E\exp\left(-\beta A_t/t\right) = \int_0^1 \frac{e^{-\beta y}dy}{\pi\sqrt{y(1-y)}}.$$

Then, by the uniqueness of the Laplace transform in the variable β, we arrive at the conclusion. \(^□^)/

6.9 Filtrations and Stopping Times II

Throughout this subsection, we assume that $(\Omega, \mathcal{F}, P)$ is a probability space.

Definition 6.9.1 Let $(\mathcal{F}_t)_{t\geq 0}$ be a filtration, and $T : \Omega \to [0, \infty]$ be a r.v.

► $(\mathcal{F}_t)_{t\geq 0}$ is said to be **right-continuous** if

$$\bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon} = \mathcal{F}_t, \quad \forall t \geq 0. \tag{6.62}$$

► T is said to be an **optional time** if

$$\{T < t\} \in \mathcal{F}_t \text{ for all } t > 0. \tag{6.63}$$

Lemma 6.9.2 *Let everything be as in Definition 6.9.1.*

a) *Then, for all $t \geq 0$ and $A \in \mathcal{F}$,*

$$A \cap \{T \leq t\} \in \mathcal{F}_t \text{ for all } t \geq 0 \implies A \cap \{T < t\} \in \mathcal{F}_t \text{ for all } t \geq 0. \tag{6.64}$$

In particular,

$$T \text{ is a stopping time} \implies T \text{ is an optional time.} \tag{6.65}$$

b) *Suppose that $(\mathcal{F}_t)_{t\geq 0}$ is right-continuous. Then the converse to (6.64) and (6.65) are also true.*

Proof: a) It is enough to show (6.64), which can be seen as follows.

$$A \cap \{T < t\} = \bigcup_{n\geq 1} (A \cap \{T \leq t - \tfrac{1}{n}\}) \in \mathcal{F}_t.$$

b) It is enough to show the converse to (6.64), which can be seen as follows.

$$A \cap \{T \leq t\} = \bigcap_{n\geq 1} (A \cap \{T < t + \tfrac{1}{n}\}) \in \bigcap_{n\geq 1} \mathcal{F}_{t+\frac{1}{n}} \stackrel{(6.62)}{=} \mathcal{F}_t.$$

\(^□^)/

Proposition 6.9.3 *Let S be a metric space, $X = (X_t : \Omega \to S)_{t\geq 0}$ be a process, T_A, T_A^+ and $(\mathcal{F}_t^0)_{t\geq 0}$ be defined as Example 4.2.2. Then, under the one of the following assumptions a),b), T_A and T_A^+ are optional times w.r.t. $(\mathcal{F}_t^0)_{t\geq 0}$. Moreover, under the assumption b), T_A is a stopping time w.r.t. $(\mathcal{F}_t^0)_{t\geq 0}$.*

a) *The function $t \mapsto X_t(\omega)$ is right-continuous for all $\omega \in \Omega$ and that A is open.*

b) *The function $t \mapsto X_t(\omega)$ is continuous for all $\omega \in \Omega$ and that A is closed.*

Proof: a) We concentrate on the case of the first entry time, since the proof for the the first hitting time is similar. We start by observing that the following are equivalent.

1) $T_A < t$,

2) $\exists s \in [0, t)$, $X_s \in A$.

3) $\exists r \in [0, t) \cap \mathbb{Q}$, $X_r \in A$.

1) $\Leftrightarrow$ 2): This follows from the definition of T_A, and is valid for *any* $A \subset \mathbb{R}^d$.
2) $\Rightarrow$ 3): Since $s \mapsto X_s$ is right-continuous and A is open, $s < \exists u < t$ such that $X_r \in A$ for all $r \in [s, u]$. Thus, we can find $r \in [s, u] \cap \mathbb{Q}$ such that $X_r \in A$, and hence 3) holds.
2) $\Leftarrow$ 3): Obvious.
The equivalence of 1) and 3) implies that

$$\{T_A < t\} = \bigcup_{r \in [0,t) \cap \mathbb{Q}} \{X_r \in A\} \in \sigma[(X_r)_{r \in [0,t) \cap \mathbb{Q}}] \subset \mathcal{F}_t^0.$$

Thus, T_A is an optional time w.r.t. $(\mathcal{F}_t^0)_{t \geq 0}$.
By Lemma 6.6.11, T_A is a stopping time w.r.t. $(\mathcal{F}_t^0)_{t \geq 0}$. Next, for $r \geq 0$, define

$$T_{A,r} = \inf\{t \geq r \ ; \ X_t \in A\}.$$

Then, by the same argument as above, we see that $T_{A,r}$ is a stopping time w.r.t. $(\mathcal{F}_t^0)_{t \geq 0}$, hence by Lemma 6.9.2,

4) $\{T_{A,r} < t\} \in \mathcal{F}_t^0$.

Note also that

$$\{t > 0 \ ; \ X_t \in A\} = \bigcup_{\substack{r > 0 \\ r \in \mathbb{Q}}} \{t \geq r \ ; \ X_t \in A\},$$

and hence that $T_A^+ = \inf_{\substack{r > 0 \\ r \in \mathbb{Q}}} T_{A,r}$. Therefore,

$$\{T_A^+ < t\} = \bigcup_{\substack{r > 0 \\ r \in \mathbb{Q}}} \{T_{A,r} < t\} \overset{4)}{\in} \mathcal{F}_t^0.$$

Thus, T_A^+ is an optional time w.r.t. $(\mathcal{F}_t^0)_{t \geq 0}$. \(^□^)/

Lemma 6.9.2 can be used to prove

Corollary 6.9.4 *In Proposition 6.9.3, suppose that X is adapted to a right-continuous filtration $(\mathcal{F}_t)_{t \geq 0}$. Then, under the one of the following assumptions a),b), T_A and T_A^+ are stopping times w.r.t. $(\mathcal{F}_t)_{t \geq 0}$.*

a) *The function $t \mapsto X_t(\omega)$ is right-continuous for all $\omega \in \Omega$ and that A is open.*

b) *The function $t \mapsto X_t(\omega)$ is continuous for all $\omega \in \Omega$ and that A is closed.*

Proof Since X is adapted to $(\mathcal{F}_t)_{t \geq 0}$, we have $\mathcal{F}_t^0 \subset \mathcal{F}_t$ for all $t \geq 0$. By Proposition 6.9.3, T_A and T_A^+ are optional times w.r.t. $(\mathcal{F}_t^0)_{t \geq 0}$ and hence w.r.t. $(\mathcal{F}_t)_{t \geq 0}$. This, together with Lemma 6.9.2 and the right-continuity of $(\mathcal{F}_t)_{t \geq 0}$, we see that T_A and T_A^+ are stopping times w.r.t. $(\mathcal{F}_t)_{t \geq 0}$. \(^□^)/

Example 6.9.5 Referring to Proposition 6.9.3, we suppose that $t \mapsto X_t(\omega)$ is continuous for all $\omega \in \Omega$. We show by an example that T_A and T_A^+ for an open A may not be a stopping time w.r.t. $(\mathcal{F}_t^0)_{t \geq 0}$. This, together with Proposition 6.9.3, shows that the filtration $(\mathcal{F}_t^0)_{t \geq 0}$ is not right-continuous. Suppose that $X_0 = 0$ and $A = (1, \infty) \times \mathbb{R}^{d-1}$. Then, $T_A = T_A^+$. Let us consider an event

$$E = \{X_s = se_1, \ \forall s \in [0,1]\} \in \mathcal{F}_1^0,$$

where $e_1 = (1, 0, \ldots, 0)$. Since all the coordinates X_s, $s \in [0,1]$ are already fixed on E, the set E does not contain any nonempty proper subset which belong to $\mathcal{F}_1^0$. On the other hand,

$$\begin{aligned} E \cap \{T_A \leq 1\} &= \{X_s = se_1, \ \forall s \in [0,1], T_A = 1\} \neq \emptyset, \\ E \backslash \{T_A \leq 1\} &= \{X_s = se_1, \ \forall s \in [0,1], T_A > 1\} \neq \emptyset. \end{aligned}$$

If we had that $\{T_A \leq 1\} \in \mathcal{F}_1^0$, then, the above two sets would belong to $\mathcal{F}_1^0$, which is a contradiction.

This example can also be used to construct a sequence of stopping times, whose infimum is not a stopping time. Let A as above and let $A_n = [\frac{n+2}{n+1}, \infty) \times \mathbb{R}^{d-1}$, $n \in \mathbb{N}$, so that $A = \bigcup_{n \in \mathbb{N}} A_n$. Then, we have $T_A = \inf_{n \in \mathbb{N}} T_{A_n}$ (Exercise 4.2.3). T_{A_n} are stopping times w.r.t. $(\mathcal{F}_t^0)_{t \geq 0}$, since A_n are closed (Proposition 6.9.3). However, $T_A = \inf_{n \in \mathbb{N}} T_{A_n}$ is not a stopping time w.r.t. $(\mathcal{F}_t^0)_{t \geq 0}$ as we have already seen.

Exercise 6.9.1 Prove that, if T_n, $n \in \mathbb{N}$ are optional times, then, so is $T \stackrel{\text{def}}{=} \inf_{n \in \mathbb{N}} T_n$.

Exercise 6.9.2 Suppose that a filtration $(\mathcal{F}_t)_{t \geq 0}$ is right-continuous. Prove the following.
i) For a stopping time T, $A \in \mathcal{F}_T \iff A \cap \{T < t\} \in \mathcal{F}_t$ for all $t \geq 0$. Hint: (6.64).
ii) If T_n, $n \in \mathbb{N}$ are stopping times, then so is $T \stackrel{\text{def}}{=} \inf_{n \in \mathbb{N}} T_n$, and $\mathcal{F}_T = \bigcap_{n \in \mathbb{N}} \mathcal{F}_{T_n}$.

Exercise 6.9.3 Suppose that a filtration $(\mathcal{F}_t)_{t \geq 0}$ is right-continuous. Then, give an alternative proof of Lemma 4.2.4 in the case of $\mathbb{T} = [0, \infty)$, by approximating $S + T$ by $S_N + T_N$, where S_N and T_N are defined by (5.17).

7 Brownian Motion and the Related Martingales

7.1 Martingales Related to the Brownian Motion

Definition 7.1.1 Suppose that $(\mathcal{F}_t)_{t\ge 0}$ is a filtration and that B is a continuous, adapted process with values in $\mathbb{R}^d$. B is called a *Brownian motion (or* BM^d*) w.r.t.* $(\mathcal{F}_t)_{t\ge 0}$ if for any $0 \le s < t$, $B_t - B_s$ is a mean-zero Gaussian r.v. with covaiance matrix $(t-s)(\delta_{\alpha\,\beta})_{\alpha,\beta=1}^d$, and is independent of $\mathcal{F}_s$.

Remark Suppose that B is a BM^d w.r.t. $(\mathcal{F}_t)_{t\ge 0}$. Then, it follows from the above definition that for any $s \ge 0$, the process $(B_{t+s} - B_s)_{t\ge 0}$ is independent of $\mathcal{F}_s$.

The notion of "Brownian motion w.r.t. a filtration" introduced above gives to a Brownian motion a certain amount of flexibility for the choice of the filtration to be associated with. Suppose that B is a BM^d and that a filtration $(\mathcal{G}_t)_{t\ge 0}$ satisfies $\mathcal{F}_t^0 \subset \mathcal{G}_t \subset \mathcal{F}_t$ for all $t \ge 0$, where $(\mathcal{F}_t^0)_{t\ge 0}$ is the canonical filtration, and $(\mathcal{F}_t)_{t\ge 0}$ is its right-continuous enlargement, cf. (6.33). Then, by Proposition 6.5.1, B is a BM^d w.r.t. $(\mathcal{G}_t)_{t\ge 0}$. Moreover, for any $\alpha = 1, \ldots, d$, the α-th coordinate process B^α is a BM^1 w.r.t. $(\mathcal{G}_t)_{t\ge 0}$

We first present the following simple, but useful characteriztion of the Brownian motion. This proposition is applied later to Proposition 7.1.3, Proposition 7.9.6 and Theorem 7.8.1.

Proposition 7.1.2 *Suppose that* $X = (X_t)_{t\ge 0}$ *is a continuous process with values in* $\mathbb{R}^d$*, adapted to a filtration* $(\mathcal{F}_t)_{t\ge 0}$*, such that* $X_0 = 0$*. Then, the following conditions are equivalent.*
a) X *is a* BM_0^d *w.r.t.* $(\mathcal{F}_t)_{t\ge 0}$*;*
b) $\exp(\mathbf{i}\theta \cdot X_t + t|\theta|^2/2)$*,* $t \ge 0$ *is a martingale for all* $\theta \in \mathbb{R}^d$*;*
c) $\exp(\theta \cdot X_t - t|\theta|^2/2)$*,* $t \ge 0$ *is a martingale for all* $\theta \in \mathbb{R}^d$*.*

Proof: a) $\Leftrightarrow$ b): a) is equivalent to that
$$E[\exp(\mathbf{i}\theta \cdot (X_t - X_s))\,|\mathcal{F}_s] = \exp(-(t-s)|\theta|^2/2) \text{ a.s. for all } \theta \in \mathbb{R}^d.$$
Multiplying the both-hand sides by $\exp(\mathbf{i}\theta \cdot X_s + t|\theta|^2/2)$, we see that this is equivalent to
$$E[\exp(\mathbf{i}\theta \cdot X_t + t|\theta|^2/2)|\mathcal{F}_s] = \exp(\mathbf{i}\theta \cdot X_s + s|\theta|^2/2), \text{ a.s. for all } \theta \in \mathbb{R}^d,$$
which is equivalent to b). The equivalence of a) $\Leftrightarrow$ c) is obtained in the same way. \(^□^)/

We define the Hermite polynomials $H_n : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$, $n \in \mathbb{N}^d$ inductively by
$$H_0(x,t) = 1, \quad H_{n+e_\alpha}(x,t) = x_\alpha H_n(x,t) - t\frac{\partial H_n}{\partial x_\alpha}(x,t), \quad n \in \mathbb{N}^d, \tag{7.1}$$
where $e_\alpha = (\delta_{\alpha\,\beta})_{\beta=1}^d$. For example,
$$H_{e_\alpha}(x,t) = x_\alpha, \quad H_{e_\alpha+e_\beta}(x,t) = x_\alpha x_\beta - t\delta_{\alpha\,\beta}. \tag{7.2}$$
On the other hand, we define, for $\theta \in \mathbb{R}^d$ and $(x,t) \in \mathbb{R}^d \times \mathbb{R}$,
$$g_\theta(x,t) \stackrel{\text{def}}{=} \exp\left(\theta \cdot x - \frac{t|\theta|^2}{2}\right). \tag{7.3}$$
For $n = (n_\alpha)_{\alpha=1}^d$, we write $\left(\frac{\partial}{\partial\theta}\right)^n = \left(\frac{\partial}{\partial\theta_1}\right)^{n_1} \cdots \left(\frac{\partial}{\partial\theta_d}\right)^{n_d}$. Then, the functions g_θ and H_n are related as
$$\left(\frac{\partial}{\partial\theta}\right)^n g_\theta(x,t) = H_n(x - t\theta, t) g_\theta(x,t) \tag{7.4}$$

for all $\theta \in \mathbb{R}^d$ and $(x,t) \in \mathbb{R}^d \times \mathbb{R}$. In particular,

$$\left(\frac{\partial}{\partial\theta}\right)^n g_\theta(x,t)\Big|_{\theta=0} = H_n(x,t). \tag{7.5}$$

Let B be a BM_0^d w.r.t. a right-continuous filtration $(\mathcal{F}_t)_{t\geq 0}$.

Proposition 7.1.3 *Let B be a BM_0^d w.r.t. a filtration $(\mathcal{F}_t)_{t\geq 0}$ Then, referring to (7.3) and (7.1), the following processes are martingales w.r.t. $(\mathcal{F}_t)_{t\geq 0}$ for any $\theta \in \mathbb{R}^d$ and $n \in \mathbb{N}^d$.*

$$(H_n(B_t - \theta t, t) g_\theta(B_t, t))_{t\geq 0}, \quad (g_\theta(B_t, t))_{t\geq 0}, \quad (H_n(B_t, t))_{t\geq 0} \tag{7.6}$$

Proof: Among the three processes in question, the second and the third one are special cases of the first one ($n = 0$ and $\theta = 0$). Therefore, we may forcus on the first one. In what follows, we consider the case of $d = 1$ for notational simplicity. Let $0 \leq s < t < \infty$. Then, by Proposition 7.1.2,

$$\begin{aligned} E[\exp(\theta \cdot (B_t - B_s)) | \mathcal{F}_s] &= E\exp(\theta \cdot (B_t - B_s)), \quad \text{a.s.} \\ &= \exp\left((t-s)|\theta|^2/2\right). \end{aligned}$$

Multiplying the both-hand sides by $\exp(\theta \cdot B_s - t|\theta|^2/2)$, we see that

1) $E[g_\theta(B_t, t)|\mathcal{F}_s] = g_\theta(B_s, s)$, a.s.

We see from 1), (7.4) and the dominated convergence theorem for the conditional expectation (Proposition 4.1.12) that

$$E\left[\left(\tfrac{\partial}{\partial\theta}\right)^n g_\theta(B_t, t)|\mathcal{F}_s\right] = \left(\tfrac{\partial}{\partial\theta}\right)^n E[g_\theta(B_t, t)|\mathcal{F}_s] \text{ a.s.}$$

This, together with 1), implies that

$$E\left[\left(\tfrac{\partial}{\partial\theta}\right)^n g_\theta(B_t, t)|\mathcal{F}_s\right] = \left(\tfrac{\partial}{\partial\theta}\right)^n g_\theta(B_s, s), \quad \text{a.s.}$$

By (7.4), this proves the desired martingale property. \(^□^)/

Remark See Example 7.6.2 for a representation of the martingales in Proposition 7.1.3 in terms of the stochastic integral.

Example 7.1.4 (**Exit time from a bounded set**) Let B be BM_x^d. We adopt the notation introduced at the beginning of section 6.7. Suppose that $A \subset \mathbb{R}^d$ is bounded, either closed or open, and let

$$T = T_{A^c} = \inf\{t \geq 0 \ ; \ B_t \in A^c\}.$$

Then, there is $\lambda > 0$ such that

$$\sup_{x \in A} E_x \exp(\lambda T) < \infty. \tag{7.7}$$

Proof: By Example 6.7.4, it is enough to prove that

1) $\sup_{x \in A} E_x T < \infty$

Since $(B^1_t - x^1)^2 - t$ is a martingale by Proposition 7.1.3, we have by Theorem 5.3.1 that

$$E_x[t \wedge T] = E_x\left[(B^1_{t\wedge T} - x^1)^2\right] \le \sup_{y \in \overline{A}} |y - x|^2,$$

from which we obtain 1) by letting $t \nearrow \infty$. \(^□^)/

Exercise 7.1.1 Let $g_\lambda : \mathbb{R}^2 \to \mathbb{R}$ $(\lambda \in \mathbb{R})$ and $H_n : \mathbb{R}^2 \to \mathbb{R}$ $(n \in \mathbb{N})$ be from Proposition 7.1.3. Then, prove the following. i) $g_\lambda(x,t) = \sum_{n=0}^\infty \frac{\lambda^n}{n!} H_n(x,t)$. [Hint: (7.5).] ii) $\frac{\partial}{\partial x} H_n(x,t) = nH_{n-1}(x,t)$. [Hint: $\frac{\partial g_\lambda}{\partial x}(x,t) = \lambda g_\lambda(x,t)$.] iii) $\frac{\partial H_n}{\partial t}(x,t) + \frac{1}{2}\frac{\partial^2 H_n}{\partial x^2}(x,t) = 0$. [Hint: $\frac{\partial g_\lambda}{\partial t}(x,t) + \frac{1}{2}\frac{\partial^2 g_\lambda}{\partial x^2}(x,t) = 0$.]

7.2 Hitting Times for One-dimensional Brownian Motions with Drift

Let B be BM^1_0. We will denote by $(\mathcal{F}_t)_{t\ge 0}$ the right-continuous enlargement of the canonical filtration defined by (6.33). For $c > 0$, we define $(X_t)_{t\ge0}$ by

$$X_t = B_t - ct.$$

Let also

$$g(\lambda, \mu) = \mu^2 - 2c\mu - 2\lambda, \quad \text{for } \lambda, \mu \in \mathbb{R}. \tag{7.8}$$

For any fixed $\lambda \ge -c^2/2$, the equation $g(\lambda,\mu) = 0$ has real solutions $\mu = f_+(\lambda)$, and $\mu = -f_-(\lambda)$, where

$$f_\pm(\lambda) \stackrel{\text{def}}{=} \sqrt{c^2 + 2\lambda} \pm c. \tag{7.9}$$

In particular, for $\lambda > 0$, we have

$$f_+(\lambda) > f_+(0) = 2c, \quad f_-(\lambda) > f_-(0) = 0. \tag{7.10}$$

Lemma 7.2.1 *Let $\lambda \ge -c^2/2$, $\mu \in \{f_+(\lambda), -f_-(\lambda)\}$, and $M_t = \exp(\mu X_t - \lambda t)$, $t \ge 0$ Then, $M = (M_t, \mathcal{F}_t)_{t\ge0}$ is a martingale.*

Proof: Since

$$\mu X_t - \lambda t = \mu B_t - (c\mu + \lambda)t = \mu B_t - \mu^2 t/2.$$

Therefore, M is a martigale by Proposition 7.1.3. \(^□^)/

Corollary 7.2.2 *Let $\varphi : \mathbb{R} \to \mathbb{R}$ be defined by $\varphi(x) = x$ if $c = 0$ and $\varphi(x) = \exp(2cx)$ if $c \ne 0$. Then, $(\varphi(X_t), \mathcal{F}_t)_{t\ge0}$ is a martingale.*

Proof: If $c = 0$, then $\varphi(X_t) = B_t$ is a martingale by Proposition 7.1.3. If $c > 0$, then $\varphi(X_t) = \exp(2cX_t) = \exp(f_+(0)X_t)$ is a martingale by Lemma 7.2.1. \(^□^)/

For $a \in \mathbb{R}$, let

$$T_a = \inf\{t \ge 0\ ;\ X_t = a\}.$$

Proposition 7.2.3 *For $a > 0$ and $\lambda > 0$,*

$$E\exp(-\lambda T_{-a}) = \exp(-af_-(\lambda)), \quad E\exp(-\lambda T_a) = \exp(-af_+(\lambda)), \tag{7.11}$$
$$P(T_{-a} < \infty) = 1, \quad P(T_a < \infty) = \exp(-2ac). \tag{7.12}$$

with the convention that $\exp(-\infty) = 0$. *Moreover, if $c > 0$, then*

$$ET_{-a} = E[T_a | T_a < \infty] = a/c. \tag{7.13}$$

On the other hand, if $c = 0$, then

$$ET_{-a} = ET_a = \infty. \tag{7.14}$$

Proof: Let M be as in Lemma 7.2.1. By Theorem 5.3.1, we have for any stopping time T and $t \geq 0$ that,

$$1 = M_0 \stackrel{(5.12)}{=} EM_{t \wedge T}. \tag{7.15}$$

(7.11): To prove the equality for T_{-a}, we apply (7.15) for $\mu = -f_-(\lambda) < 0$ and $T = T_{-a}$. Note that $-a \leq X(t \wedge T_{-a})$, and hence

1) $0 \leq M(t \wedge T_{-a}) = \exp(\mu X(t \wedge T_{-a}) - \lambda t \wedge T_{-a}) \leq \exp(-\mu a - \lambda t \wedge T_{-a}) \leq \exp(-\mu a)$.

On the other hand, we have

2) $M(t \wedge T_{-a}) \stackrel{t \to \infty}{\longrightarrow} \exp(-\mu a - \lambda T_{-a})$.

Indeed, if $T_{-a} < \infty$, then, $X(t \wedge T_{-a}) \stackrel{t \to \infty}{\longrightarrow} X(T_{-a}) = -a$, and hence,

$$M(t \wedge T_{-a}) = \exp(\mu X(t \wedge T_{-a}) - \lambda t \wedge T_{-a}) \stackrel{t \to \infty}{\longrightarrow} \exp(-\mu a - \lambda T_{-a}).$$

If $T_{-a} = \infty$, then, $0 \leq M_t \stackrel{1)}{\leq} \exp(-\mu a - \lambda t)$, $\forall t \geq 0$, and hence

$$M(t \wedge T_{-a}) = M_t \stackrel{t \to \infty}{\longrightarrow} 0 = \exp(-\mu a - \lambda T_{-a}).$$

By 1) and 2), we can use BCT in the limit $t \to \infty$ to conclude from (7.15) that

$$1 = \exp(-\mu a) E\exp(-\lambda T_{-a})$$

This proves the equality for T_{-a}. The other equality is obtained in the same way.
(7.12): We have for any r.v. $T : \Omega \to [0, \infty]$ that

$$\lim_{\substack{\lambda \to 0 \\ \lambda > 0}} E\exp(-\lambda T) = P(T < \infty).$$

Thus, we see (7.12) from (7.10) and (7.11).
(7.13), (7.14): By Exercise 1.1.6, we have for any r.v. $T : \Omega \to [0, \infty]$ that

3) $\qquad E[T : T < \infty] = -\lim_{\substack{\lambda \to 0 \\ \lambda > 0}} \frac{d}{d\lambda} E\exp(-\lambda T)$

On the other hand, the function $f_\pm$ is differentiable on $(-c^2/2, \infty)$ and

$$f'_\pm(\lambda) = 1/\sqrt{c^2+2\lambda}, \quad \lambda > -c^2/2.$$

Thus,

4) $$f'_\pm(0+) \stackrel{\text{def}}{=} \lim_{\substack{\lambda\to 0\\ \lambda>0}} f'_\pm(\lambda) = \begin{cases} 1/c, & \text{if } c>0, \\ \infty, & \text{if } c=0. \end{cases}$$

Then, it follows from (7.11) and that

$$\begin{aligned} E[T_a : T_a < \infty] &\stackrel{3)}{=} -\lim_{\substack{\lambda\to 0\\ \lambda>0}} \frac{d}{d\lambda} E\exp(-\lambda T_{-a}) \stackrel{(7.11)}{=} -\lim_{\substack{\lambda\to 0\\ \lambda>0}} \frac{d}{d\lambda}\exp(-af_+(\lambda)) \\ &= a\exp(-af_+(0))f'_+(0+) \stackrel{(7.10),\,4)}{=} (a/c)\exp(-2ac). \end{aligned}$$

Since $P(T_a<\infty) = \exp(-2ac)$ by (7.12), we obtain the second equality of (7.13). The other equalities can be obtained in the same way. \(^□^)/

Remark 1) If $c>0$, then, $Y \stackrel{\text{def}}{=} \sup_{t\geq 0} X_t < \infty$ a.s. Moreover, we see from the equality (7.12) that the r.v. Y is exponentialy distributed.

$$P(Y\geq a) = P(T_a<\infty) \stackrel{(7.12)}{=} \exp(-2ac). \tag{7.16}$$

2) If $c>0$ again, the validity of the first identity of (7.11) extends to all $\lambda \geq -c^2/2$. To see this, we note that $\exp(cX_t + c^2t/2)$ is a martingale by Lemma 7.2.1. Thus $E\exp(c^2T_{-a}/2) \leq e^{ca}$ by Corollary 5.3.3. This implies that $E\exp(-\lambda T_{-a})$ for $\lambda \in \mathbb{C}$, $\mathrm{Re}\,\lambda > -c^2/2$ is holomorphic. Therefore, by the unicity theorem, the first identity of (7.11) extends to all $\lambda > -c^2/2$. Finally, the case of $\lambda = -c^2/2$ is obtained by the monotone convergence theorem.

By (7.11) and the uniqueness of the Laplace transform (Example 1.8.3), we can identify the density of the r.v. T_a for all $a \in \mathbb{R}\backslash\{0\}$ (See also Proposition 6.6.4 for the case of $c=0$).

Corollary 7.2.4 *For $c\geq 0$ and $a \in \mathbb{R}\backslash\{0\}$, $T_a \approx k_t(a,c)dt$, where*

$$k_t(a,c) = \frac{|a|}{\sqrt{2\pi t^3}}\exp\left(-\frac{(a+ct)^2}{2t}\right).$$

Proof: By (7.11) and the uniqueness of the Laplace transform (Example 1.8.3), it is enough to verify for all $\lambda>0$ that

1) $$I \stackrel{\text{def}}{=} \int_0^\infty \exp(-\lambda t)k_t(a,c)dt = \begin{cases} \exp(-af_+(\lambda)), & a>0, \\ \exp(af_-(\lambda)), & a<0. \end{cases}$$

We first consider the case of $a>0$. Note that

2) $$\lambda t + \frac{(a+ct)^2}{2t} = ac + \frac{c^2+2\lambda}{2}t + \frac{a^2}{2t}.$$

We also recall from Lemma 2.3.4 with $n=1$ that

3) $\dfrac{a}{\sqrt{2\pi}}\displaystyle\int_0^\infty t^{-3/2}\exp\left(-\frac{b^2 t}{2}-\frac{a^2}{2t}\right)dt=\exp(-ab),\quad a,b>0.$

Therefore,

$$
\begin{aligned}
I &\stackrel{2)}{=} \exp(-ac)\frac{a}{\sqrt{2\pi}}\int_0^\infty t^{-3/2}\exp\left(-\frac{c^2+2\lambda}{2}t-\frac{a^2}{2t}\right)dt \\
&\stackrel{3)}{=} \exp(-ac)\exp\left(-a\sqrt{c^2+2\lambda}\right)=\exp(-af_+(\lambda)),
\end{aligned}
$$

which proves 1). The proof for the case of $a<0$ is similar. \(^□^)/

Remark Corollary 7.2.4 can also be derived as an application of the Cameron-Martin formula [LeG16, pp. 140–141].

Proposition 7.2.5 *For $a,b>0$ and $\lambda>0$,*

$$
\begin{aligned}
E[\exp(-\lambda T_{-a}):T_{-a}<T_b] &= \frac{e^{ac}\sinh(b\sqrt{c^2+2\lambda})}{\sinh((a+b)\sqrt{c^2+2\lambda})}, &(7.17)\\
E[\exp(-\lambda T_b):T_b<T_{-a}] &= \frac{e^{-bc}\sinh(a\sqrt{c^2+2\lambda})}{\sinh((a+b)\sqrt{c^2+2\lambda})}, &(7.18)
\end{aligned}
$$

with the convention that $\exp(-\infty)=0$. *Moreover, if $c>0$, then*

$$P(T_{-a}<T_b)=\frac{e^{2bc}-1}{e^{2bc}-e^{-2ac}},\quad P(T_b<T_{-a})=\frac{1-e^{-2ac}}{e^{2bc}-e^{-2ac}}. \tag{7.19}$$

On the other hand, if $c=0$, then

$$P(T_{-a}<T_b)=\frac{b}{a+b},\quad P(T_b<T_{-a})=\frac{a}{a+b}. \tag{7.20}$$

Proof: (7.17), (7.18): Let M be as in Lemma 7.2.1. We write $M=M_+$ if $\mu=f_+(\lambda)$, and $M=M_-$ if $\mu=-f_-(\lambda)$. We take $T=T_{-a}\wedge T_b$. Then, we see from (7.15) that

1) $1=EM_\pm(t\wedge T)$.

On the other hand,

2) $0\le M_+(t\wedge T)\le\exp(\mu b)$, $0\le M_-(t\wedge T)\le\exp(-\mu a)$.

We now note that

3) $T_{-a}\neq T_b$ a.s.

This can be seen as follows. If $T_{-a}=T_b<\infty$, then, $-a=X(T_{-a})=X(T_b)=b$, which is impossible. Hence, $\{T_{-a}=T_b<\infty\}=\emptyset$. On the other hand, $T_b<\infty$ a.s. by (7.12). Thus, $P(T_{-a}=T_b=\infty)=0$.
It follows from 3) that almost surely,

4) $$\left\{\begin{array}{rcl} M_\pm(t\wedge T) &=& M_\pm(t\wedge T_{-a})\mathbf{1}\{T_{-a}<T_b\}+M_\pm(t\wedge T_b)\mathbf{1}\{T_b<T_{-a}\} \\ &\stackrel{t\to\infty}{\longrightarrow}& \exp(\mp af_\pm(\lambda)-\lambda T_{-a})\mathbf{1}\{T_{-a}<T_b\}+\exp(\pm bf_\pm(\lambda)-\lambda T_b)\mathbf{1}\{T_b<T_{-a}\}.\end{array}\right.$$

Let E_1 and E_2 be the LHS's of (7.17) and (7.18), respectively. Then, by 2) and 4), we can apply BCT for 1) in the limit $t \to \infty$ to conclude that

$$1 = \exp(\mp a f_\pm(\lambda))E_1 + \exp(\pm b f_\pm(\lambda))E_2.$$

By solving the above equation, we have

$$\begin{aligned} E_1 &= \frac{\exp(bf_+(\lambda)) - \exp(-bf_-(\lambda))}{\exp(bf_+(\lambda) + af_-(\lambda)) - \exp(-af_+(\lambda) - bf_-(\lambda))}, \\ E_2 &= \frac{\exp(af_-(\lambda)) - \exp(-af_+(\lambda))}{\exp(bf_+(\lambda) + af_-(\lambda)) - \exp(-af_+(\lambda) - bf_-(\lambda))}, \end{aligned}$$

from which we obtain (7.17) and (7.18).
(7.19),(7.20): These follow from (7.17) and (7.18) by letting $\lambda \searrow 0$, cf. (7.10). \(^□^)/

Remark Using the function φ, introduced in Corollary 7.2.2, the equalities (7.19) and (7.20) can be written at the same time as:

$$P(T_{-a} < T_b) = \frac{\varphi(b) - \varphi(0)}{\varphi(b) - \varphi(-a)}, \quad P(T_b < T_{-a}) = \frac{\varphi(0) - \varphi(-a)}{\varphi(b) - \varphi(-a)}.$$

The equalities (7.19) and (7.20) tell us the distribution of the r.v. $Z \stackrel{\text{def}}{=} \sup_{t \le T_{-a}} X_t$ (Note that $T_{-a} < \infty$ a.s. by (7.12)).

$$P(Z \ge b) = P(T_b < T_{-a}) \stackrel{(7.19),(7.20)}{=} \begin{cases} (1 - e^{-2ac}/(e^{2bc} - e^{-2ac}) & \text{if } c > 0, \\ a/(a+b) & \text{if } c = 0. \end{cases} \tag{7.21}$$

In particular,

$$EZ = \int_0^\infty P(Z \ge b)\, db \begin{cases} < \infty & \text{if } c > 0, \\ = \infty & \text{if } c = 0. \end{cases}$$

Exercise 7.2.1 Prove that

$$E[T_{-a} \wedge T_b] = \begin{cases} (a+b)(e^{bc}/c)\sinh(ac)/\sinh((a+b)c) & \text{if } c > 0, \\ ab & \text{if } c = 0. \end{cases}$$

[Hint: For $c > 0$, use the martingale $X_t + ct = B_t$, and for $c = 0$,use the martingale $B_t^2 - t$.]

Remark If we consider BM^1_x instead of BM^1_0. Then, for $c = 0$, it follows from Exercise 7.2.1 that $m_x \stackrel{\text{def}}{=} E[T_{-a} \wedge T_b] = (a+x)(b-x) \le (a+b)^2/4$ if $x \in [-a, b]$, and $m_x = 0$ if $x \notin [-a, b]$. Thus, we have by Example 6.7.4 that $E\exp(\lambda(T_{-a} \wedge T_b)) < \infty$ for any $\lambda \in (-\infty, 4/(a+b)^2)$.

Exercise 7.2.2 For $c = 0$, prove that

$$E\exp(-\lambda(T_{-a} \wedge T_b)) = \frac{\cosh\left((a-b)\sqrt{\lambda/2}\right)}{\cosh\left((a+b)\sqrt{\lambda/2}\right)}.$$

[Hint: For $x, y \in \mathbb{R}$, $\sinh x + \sinh y = 2\sinh\left(\frac{x+y}{2}\right)\cosh\left(\frac{x-y}{2}\right)$, $\sinh(x+y) = 2\cosh\left(\frac{x+y}{2}\right)\sinh\left(\frac{x+y}{2}\right)$.]

Remark By the remark after Exercise 7.2.1, we see from Exercise 7.2.2 and the analytic continuation that for any $\lambda \in (-\infty, 4/(a+b)^2)$,

$$E\exp(\lambda(T_{-a} \wedge T_b)) = \frac{\cos\left((a-b)\sqrt{\lambda/2}\right)}{\cos\left((a+b)\sqrt{\lambda/2}\right)}.$$

Exercise 7.2.3 Let $k_t(x) = \frac{x}{\sqrt{2\pi t^3}} \exp\left(-\frac{x^2}{2t}\right)$, $x \in \mathbb{R}$, $t > 0$. Then, for $c = 0$, prove that $T_{-a} \wedge T_b \approx k_t(a,b)dt$, where

$$k_t(a,b) = \sum_{j=0}^{\infty} (-1)^j (k_t(a + (a+b)j) + k_t(b + (a+b)j)).$$

[Hint: Compute the Laplace transform $\int_0^\infty \exp(-\lambda t) k_t(a,b)dt$, $\lambda > 0$ and compare it with Exercise 7.2.2.]

7.3 Stochastic Integrals

Let B be BM^1 w.r.t. a filtration $(\mathcal{F}_t)_{t\ge 0}$, cf. Definition 7.1.1. For a suitable class of processes $H = (H_t)_{t\ge 0}$, we will define the integral of the form

$$\int_0^t H_s dB_s, \quad t \ge 0, \tag{7.22}$$

which is called the stochasic integral with respect to the Brownian motion. The function $s \mapsto B_s$ is not of bounded variation in any interval. Therefore, the above integral cannot be defined as a Lebesgue-Stieltjes integral.

We start by introducing some classes of integrands for the stochastic integral.

Definition 7.3.1 (**Integrands for stochastic integral**)
▶ We denote by $\mathcal{L}$ the totality of progressive real processes w.r.t. $(\mathcal{F}_t)_{t\ge 0}$ (cf. Definition 6.6.12).
▶ We define

$$\begin{aligned} \mathcal{L}^2_\infty &= \left\{ H \in \mathcal{L} \; ; \; E\int_0^\infty H_s^2 ds < \infty \right\}, \\ \mathcal{L}^2 &= \left\{ H \in \mathcal{L} \; ; \; E\int_0^t H_s^2 ds < \infty \text{ for all } t \in (0,\infty) \right\}, \\ \mathcal{L}^2_{\text{a.s.}} &= \left\{ H \in \mathcal{L} \; ; \; \int_0^t H_s^2 ds < \infty, \; P\text{-a.s. for all } t > 0 \right\}. \end{aligned}$$

▶ A process $H \in \mathcal{L}$ is said to be **elementary**, if it is a finite linear combimations of the processes of the following form

$$(\mathbf{1}_{(a,b]} \otimes h)_t(\omega) = h(\omega)\mathbf{1}_{(a,b]}(t), \quad (t,\omega) \in [0,\infty) \times \Omega. \tag{7.23}$$

for some $0 \le a < b < \infty$ and $h \in L^2(\Omega, \mathcal{F}_a, P)$. The totality of elementary processes is denoted by $\mathcal{E}$.

Remark: Clearly, $\mathcal{E} \subset \mathcal{L}^2_\infty \subset \mathcal{L}^2 \subset \mathcal{L}^2_{\text{a.s.}} \subset \mathcal{L}$.

Definition 7.3.2 (**Spaces of continuous (local) martingales**)
▶ We denote by $\mathcal{M}_{\text{c}}$ the totality of martingales $M = (M_t, \mathcal{F}_t)_{t\ge 0}$ such that $M_0 = 0$ and $t \mapsto M_t$ is a.s. continuous.
▶ We define

$$\begin{aligned} \mathcal{M}^2_{\text{c},\infty} &= \left\{ M \in \mathcal{M}^{\text{c}} \; ; \; \sup_{t\ge 0} E[M_t^2] < \infty \right\}. \\ \mathcal{M}^2_{\text{c}} &= \left\{ M \in \mathcal{M}^{\text{c}} \; ; \; E[M_t^2] < \infty \text{ for all } t \in (0,\infty) \right\}. \end{aligned}$$

▶ An adapted process $M = (M_t, \mathcal{F}_t)_{t\geq 0}$ is called a **local martingale**, if there exists a nondecreasing sequence of finite stopping times $(T_n)_{n\geq 1}$ such that $T_n \stackrel{n\to\infty}{\longrightarrow} \infty$ a.s. and for any $n \geq 1$, $(M_{t\wedge T_n})_{t\geq 0}$ is uniformly integrable martingale. The above sequence $(T_n)_{n\geq 1}$ of stopping times is then called a **reduction sequence**.
▶ We denote by $\mathcal{M}^2_{\text{c,loc}}$ the totality of local martingales $M = (M_t, \mathcal{F}_t)_{t\geq 0}$ such that $M_0 = 0$ and $t \mapsto M_t$ is a.s. continuous, and there exists reduction a sequence $(T_n)_{n\geq 1}$ such that $M_{\cdot\wedge T_n} \in \mathcal{M}^2_{\text{c},\infty}$ for all $n \geq 1$. We identify two elements $M, \widetilde{M}$ in $\mathcal{M}^2_{\text{c,loc}}$, if $M_t = \widetilde{M}_t$ a.s. for all $t \geq 0$.

Remark: If $M \in \mathcal{M}^2_{\text{c},\infty}$, then $(M_t)_{t\geq 0}$ is uniformly integrable. Indeed, by L^2-maximal inequality (5.23),

$$E\left[\sup_{t\geq 0}(M_t)^2\right] \leq 4\sup_{t\geq 0} E\left[(M_t)^2\right] < \infty.$$

Theorem 7.3.3 *There exists a unique map $H \mapsto H \cdot B$ from $\mathcal{L}^2_{\text{a.s.}}$ to $\mathcal{M}^2_{\text{c,loc}}$ which satisfies the following properties.*
a) *For all $H, K \in \mathcal{L}^2_{\text{a.s.}}$, $\alpha, \beta \in L^\infty(\Omega, \mathcal{F}_0, P)$, and $t \geq 0$,*

$$((\alpha H + \beta K) \cdot B)_t = \alpha(H \cdot B)_t + \beta(K \cdot B)_t. \tag{7.24}$$

b) *Referring to (7.29), for all $0 \leq a < b < \infty$ and $h \in L^2(\Omega, \mathcal{F}_a, P)$,*

$$((\mathbf{1}_{(a,b]} \otimes h) \cdot B)_t(\omega) = h(\omega)(B_{t\wedge b}(\omega) - B_{t\wedge a}(\omega)), \quad (t, \omega) \in [0, \infty) \times \Omega. \tag{7.25}$$

c) *For all $H, K \in \mathcal{L}^2$, the following processes are martingales w.r.t. $(\mathcal{F}_t)_{t\geq 0}$.*

$$(H \cdot B)_t \;\; \textit{and} \;\; Q_t(H, K) \stackrel{\text{def}}{=} (H \cdot B)_t(K \cdot B)_t - \int_0^t H_s K_s ds, \quad t \geq 0. \tag{7.26}$$

d) *For all $H \in \mathcal{L}^2_{\text{a.s.}}$ and stopping time T,*

$$(H\chi_T \cdot B)_t = (H \cdot B)_{t\wedge T}, , \quad t \geq 0. \tag{7.27}$$

where $\chi_T(t, \omega) = \mathbf{1}_{[0,T(\omega)]}(t)$, $(t, \omega) \in [0, \infty) \times \Omega$.

The process $H \cdot B \in \mathcal{M}^2_{\text{c,loc}}$ stated in Theorem 7.3.3 is called the **stochastic integral** of $H \in \mathcal{L}^2_{\text{a.s.}}$ w.r.t. the Brownian motion B and is also denoted also by the integral notation (7.22). It follows from Theorem 7.3.3 b) that, for $H \in \mathcal{L}^2$ and $t \geq 0$,

$$E[(H \cdot B)_t] = 0, \quad E[(H \cdot B)_t^2] = E\int_0^t H_s^2 ds. \tag{7.28}$$

The second equality of (7.28) is called **Itô's isometry**. We now prove Theorem 7.3.3 in three successive steps.

Step1 (**The case of** $H \in \mathcal{E}$) We first consider the stochastic integral of an elementary process. Suppose that $H \in \mathcal{E}$ is expressed as

$$H = \sum_{j=0}^{N-1} \mathbf{1}_{(c_j, c_{j+1}]} \otimes h_j \tag{7.29}$$

with a strictly increasing sequence $(c_j)_{j=0}^N$, $c_0 = 0$ and $h_j \in L^2(\Omega, \mathcal{F}_{c_j}, P)$ $(0 \le j \le N-1)$. Then, we define $H \cdot B$ by

$$(H \cdot B)_t = \sum_{j=0}^{N-1} h_j (B_{t \wedge c_{j+1}} - B_{t \wedge c_j}), \quad t \ge 0. \tag{7.30}$$

Lemma 7.3.4 **a)** *The definition (7.30) is well defined, i.e., it does not depend on the way in which H is expressed as on the right-hand side of (7.29).*

b) *The properties (7.24) and (7.27) hold for $H, K \in \mathcal{E}$*

Proof: Let $H, K \in \mathcal{E}$ be such that

$$H_t = \sum_{\ell=0}^{L-1} \mathbf{1}_{(a_\ell, a_{\ell+1}]} \otimes h_\ell, \quad K_t = \sum_{m=0}^{M-1} \mathbf{1}_{(b_m, b_{m+1}]} \otimes k_m,$$

where $(a_\ell)_{\ell=0}^L$ and $(b_m)_{m=0}^M$ are strictly increasing sequence, $a_0 = b_0 = 0$ and $h_\ell \in L^2(\Omega, \mathcal{F}_{a_\ell}, P)$, $k_m \in L^2(\Omega, \mathcal{F}_{b_m}, P)$ $(0 \le \ell < L,\ 0 \le m < M)$. We define a sequence $(c_j)_{j=0}^N$ by

$$\{c_1 < \ldots < c_N\} = \{a_\ell\}_{\ell=1}^L \cup \{b_m\}_{m=1}^M, \quad c_0 = 0.$$

As a consequence, there exist $0 = p(0) < p(1) < \ldots < p(L) \le N$ and $0 = q(0) < p(1) < \ldots < q(M) \le N$ such that

$$a_\ell = c_{p(\ell)} \ (1 \le \ell \le L), \text{ and } b_m = c_{q(m)} \ (1 \le m \le M).$$

We then define r.v.'s $\{\widetilde{h}_j\}_{j=1}^N$, $\{\widetilde{k}_j\}_{j=1}^N$ by

$$\widetilde{h}_j = h_\ell \text{ for } p(\ell) \le j < p(\ell+1) \text{ and } \widetilde{k}_j = k_m \text{ for } q(m) \le j < q(m+1).$$

Then,

1) $$H = \sum_{j=0}^{N-1} \mathbf{1}_{(c_j, c_{j+1}]} \otimes \widetilde{h}_j, \quad K = \sum_{j=0}^{N-1} \mathbf{1}_{(c_j, c_{j+1}]} \otimes \widetilde{k}_j,$$

On the other hand,

2) $I_t = \widetilde{I}_t$ and $J_t = \widetilde{J}_t$,

where

$$I_t = \sum_{\ell=0}^{L-1} h_\ell (B_{t \wedge a_{\ell+1}} - B_{t \wedge a_\ell}), \quad J_t = \sum_{m=0}^{M-1} k_m (B_{t \wedge b_{m+1}} - B_{t \wedge b_m}),$$
$$\widetilde{I}_t = \sum_{j=1}^{N} \widetilde{h}_j (B_{t \wedge c_j} - B_{t \wedge c_{j-1}}), \quad \widetilde{J}_t = \sum_{j=1}^{N} \widetilde{k}_j (B_{t \wedge c_j} - B_{t \wedge c_{j-1}}).$$

Indeed,

$$\begin{aligned} \widetilde{I}_t &= \sum_{j=0}^{N-1} \widetilde{h}_j (B_{t \wedge c_{j+1}} - B_{t \wedge c_j}) = \sum_{\ell=0}^{L-1} h_\ell \sum_{k(\ell) \le j < k(\ell+1)} (B_{t \wedge c_{j+1}} - B_{t \wedge c_j}) \\ &= \sum_{\ell=0}^{L-1} h_\ell (B_{t \wedge a_{\ell+1}} - B_{t \wedge a_\ell}) = I_t. \end{aligned}$$

Similarly $J_t = \widetilde{J}_t$.
a): To ses (7.24), suppose that $H = K$. Then, it follows from 1) that $\widetilde{h}_j = \widetilde{k}_j$ for all $j = 1, ..., N$ and hence $\widetilde{I}_t = \widetilde{J}_t$. Thus, we have $I_t = J_t$ by 2). Therefore, the definition (7.30) does not depend on the way in which H is expressed as on the right-hand side of (7.29).
b): Let $H, K \in \mathcal{E}$ be as at the beginning of the proof and $\alpha, \beta \in L^\infty(\Omega, \mathcal{F}_0, P)$, then,

$$\alpha H + \beta K \stackrel{1)}{=} \sum_{j=0}^{N-1} \mathbf{1}_{(c_j, c_{j+1}]} \otimes (\alpha \widetilde{h}_j + \beta \widetilde{k}_j)$$

Hence

$$\begin{aligned} ((\alpha H + \beta K) \cdot B)_t &\stackrel{(7.30)}{=} \sum_{j=0}^{N-1} (\alpha \widetilde{h}_j + \beta \widetilde{k}_j)(B_{t \wedge c_{j+1}} - B_{t \wedge c_j}) \\ &= \alpha \sum_{j=0}^{N-1} \widetilde{h}_j (B_{t \wedge c_{j+1}} - B_{t \wedge c_j}) + \beta \sum_{j=0}^{N-1} \widetilde{k}_j (B_{t \wedge c_{j+1}} - B_{t \wedge c_j}) \\ &\stackrel{2)}{=} \alpha (H \cdot B)_t + \beta (K \cdot B)_t. \end{aligned}$$

To see (7.27), suppose that H is expressed as (7.29) and that T is a stopping time. Then,

$$H \chi_T = \sum_{j=1}^{N} \mathbf{1}_{(c_j, c_{j+1}]} h_j \chi_T.$$

It follows from (4.41) that $h_j \chi_T$ is $\mathcal{F}_{c_j}$-measurable, and hence

$$\begin{aligned} (H \chi_T \cdot B)_t &= \sum_{j=1}^{N} \mathbf{1}_{(c_j, c_{j+1}]} h_j \chi_T (B_{t \wedge c_{j+1}} - B_{t \wedge c_j}) \\ &= \sum_{j=1}^{N} \mathbf{1}_{(c_j, c_{j+1}]} h_j (B_{t \wedge T \wedge c_{j+1}} - B_{t \wedge T \wedge c_j}) = (H \cdot B)_{t \wedge T}. \end{aligned}$$

\(^□^)/

Next, in order to verify that the processes (7.26) are martingales, we prepare the following lemma.

Lemma 7.3.5 a) *Let $0 \le a < b \le \infty$, $h \in L^1(P)$ be $\mathcal{F}_a$-measurable. Then, the following processes are martingales.*

$$U_t = h(B_{t \wedge b} - B_{t \wedge a}), \quad V_t = h\left((B_{t \wedge b} - B_{t \wedge a})^2 - (t \wedge b - t \wedge a)\right).$$

b) *Let $0 \le a_1 < b_1 \le a_2 < b_2 \le \infty$, $h_j \in L^2(P)$ be $\mathcal{F}_{a_j}$-measurable, $j = 1, 2$. Then, the following process is a martingale.*

$$W_t = h_1 h_2 (B_{t \wedge b_1} - B_{t \wedge a_1})(B_{t \wedge b_2} - B_{t \wedge a_2})$$

Proof: a) We first check that $U_t, V_t \in L^1(P)$ for all $t \ge 0$. Let $L_t = B_{t \wedge b} - B_{t \wedge a}$. Then, for $t \le a$, $L_t = 0$ and hence $U_t = 0$. For $t \ge a$, $h \in L^1(P)$ and $L_t \in L^1(P)$ are independent, and hence $U_t = hL_t \in L^1(P)$. Similarly, $V_t \in L^1(P)$. We next prove that U_t, V_t are martingales. Since h is $\mathcal{F}_a$-measurable and L_t is a martingale such that $L_t = 0$ if $t \le a$, it follows from Exercise 4.3.3 that $U_t = hL_t$ is a martingale. On the other hand, it is not difficult to see that the process

$$M_t \stackrel{\text{def}}{=} (B_{t\wedge b} - B_{t\wedge a})^2 - (t\wedge b - t\wedge a), \quad t \geq 0$$

is a martingale such that $M_t = 0$ if $t \leq a$. Thus, it follows from Exercise 4.3.3 that $V_t = hM_t$ is a martingale.

b) We first check that $W_t \in L^1(P)$ for all $t \geq 0$. $N_t = B_{t\wedge b_2} - B_{t\wedge a_2}$ and $Z_t \stackrel{\text{def}}{=} h_1 h_2 (B_{t\wedge b_1} - B_{t\wedge a_1})$. Then, $W_t = Z_t N_t$. For $t \leq a_2$, $N_t = 0$, and hence $W_t = 0$. Thus, we suppose that $t \geq a_2$. Since $h_1 \in L^2(P)$ and $B_{b_1} - B_{a_1} \in L^2(P)$ are independent, $h_1(B_{b_1} - B_{a_1}) \in L^2(P)$, and hence $h_1 h_2 (B_{b_1} - B_{a_1}) \in L^1(P)$. Moreover, $h_1 h_2 (B_{b_1} - B_{a_1}) \in L^1(P)$ and $N_t = B_{t\wedge b_2} - B_{a_2} \in L^1(P)$ are independent, and hence $W_t \in L^1(P)$. Next, we prove that W_t is a martingale. Since N_t is a martingale such that $N_t = 0$ if $t \leq a_2$, and $Z_t = Z_{t\wedge a_2}$ is $\mathcal{F}_{a_2}$-measurable for all $t \geq 0$, it follows from Exercise 4.3.3 that $W_t = Z_t N_t$ is a martingale. \(^□^)/

Now, it is easy to prove

Lemma 7.3.6 *Suppose that $H, K \in \mathcal{E}$. Then, $H \cdot B \in \mathcal{M}^2_{\mathrm{c}}$ and the processes (7.26) are martingales.*

Proof: It is clear that the process $H \cdot B$ defined by (7.30) is a.s. continuous and that $E[(H \cdot B)^2_t] < \infty$ for all $t > 0$. Moreover,

$$Q_t(H, K) = \sum_{i,j=1}^{N} h_i k_j \left((B_{t\wedge c_i} - B_{t\wedge c_{i-1}})(B_{t\wedge c_j} - B_{t\wedge c_{j-1}}) - \delta_{i,j}((t\wedge c_i) - (t\wedge c_{i-1})) \right).$$

We see from Lemma 7.3.5, that all the terms on the RHS of (7.30) and that of the above display are martingales. Hence $H \cdot B \in \mathcal{M}^2_{\mathrm{c}}$ and the process $Q_t(H, K)$ is a martingale. \(^□^)/

Step2: (**The case of $H \in \mathcal{L}^2_\infty$**) It is convenient to organize the construction in the abstract framework, concerning the isometry between two Hilbert spaces. For $H, K \in \mathcal{L}^2_\infty$, we define their inner product by

$$\langle H, K \rangle_{\mathcal{L}^2_\infty} = E \int_0^\infty H_s K_s ds. \tag{7.31}$$

We identify two elements $H, \widetilde{H}$ in $\mathcal{L}^2_\infty$, if $H_t(\omega) = \widetilde{H}_t(\omega)$, $dt \otimes P(d\omega)$-a.s. on $[0, \infty) \times \Omega$. Then, it is easy to show that $\mathcal{L}^2_\infty$ is a Hilbert space. We have the following lemma.

Lemma 7.3.7 *$\mathcal{E}$ is dense in $\mathcal{L}^2_\infty$.*

Proof: It is enough to show that the ortogonal complement $\mathcal{E}^\perp$ contains only of the null function. For this purpose, suppose that $H \in \mathcal{E}^\perp$. Then, considering $\mathbf{1}_{(a,b]} \otimes \mathbf{1}_A \in \mathcal{E}$, with $0 \leq a < b < \infty$ and $A \in \mathcal{F}_a$, we have

$$E\left[\int_a^b H_s ds : A\right] = 0.$$

This implies that the process $M_t \stackrel{\text{def}}{=} \int_0^t H_s ds$, $t \geq 0$ is a continuous martingale, and hence $M \equiv 0$, a.s., since M is at the same time of bounded variation (cf. Lemma 7.3.12 below). Consequently, $H \equiv 0$, a.s. \(^□^)/

For $M \in \mathcal{M}^2_{c,\infty}$, we define two norms

$$\rho_2(X) = \sup_{t\geq 0} E[X_t^2]^{1/2}, \quad \overline{\rho}_2(M) = E\left[\sup_{t\geq 0} M_t^2\right]^{1/2}. \tag{7.32}$$

By L^2-maximal inequality (5.23),

$$\rho_2(M) \leq \overline{\rho}_2(M) \leq 2\rho_2(M) \text{ for } M \in \mathcal{M}^2_{c,\infty},$$

and hence the norms ρ_2 and $\overline{\rho}_2$ are equivalent on $\mathcal{M}^2_{c,\infty}$.

Lemma 7.3.8 $(\mathcal{M}^2_{c,\infty}, \rho_2)$ *is a Hilbert space.*

Proof: It is enough to show that $(\mathcal{M}^2_{c,\infty}, \overline{\rho}_2)$ is a Hilbert space. Suppose that $(M^{(k)})_{k\in\mathbb{N}}$ is a Cauchy sequence in $(\mathcal{M}^2_{c,\infty}, \overline{\rho}_2)$. To prove that $(M^{(k)})_{k\in\mathbb{N}}$ converges in $\mathcal{M}^2_{c,\infty}$, it is enough to find a subsequence which converges in $\mathcal{M}^2_{c,\infty}$. Then, by taking a subsequence, we may assume that $\overline{\rho}_2(M^{(k+1)}, M^{(k)}) \leq 2^{-k}$, so that the following series converges w.r.t. $\overline{\rho}_2$:

$$M_t = M_t^{(0)} + \sum_{k=0}^{\infty} (M_t^{(k+1)} - M_t^{(k)}).$$

Moreover, we have

$$\overline{\rho}_2(M, M^{(k)}) \leq \sum_{j=k+1}^{\infty} 2^{-j} \stackrel{k\to\infty}{\longrightarrow} 0.$$

In particular, for each $n \geq 1$,

$$\sup_{t\geq 0} |M_t - M_t^{(k)}| \stackrel{k\to\infty}{\longrightarrow} 0 \text{ in } L^2(P).$$

By taking subsequence again, we may assume that the above convergence takes place a.s., and hence a.s., $M_t^{(k)}$ converges to M_t uniformly in $t \geq 0$. This implies that $M \in \mathcal{M}^2_{c,\infty}$. \(^□^)/

Lemma 7.3.9 *The map $H \mapsto H \cdot B$ ($\mathcal{E} \to \mathcal{M}^2_{c,\infty}$) defined by Step1 is uniquely extended to a linear isometry*

$$H \mapsto H \cdot B \ (\mathcal{L}^2_\infty, \|\cdot\|_{\mathcal{L}^2_\infty}) \longrightarrow (\mathcal{M}^2_{c,\infty}, \rho_2). \tag{7.33}$$

Moreover for $H \in \mathcal{L}^2_\infty$, the process $H \cdot B$ defined this way satisfies the equality (7.27) for all stopping time T.

Proof: The map $H \mapsto H \cdot B$ ($\mathcal{E} \to \mathcal{M}^2_{c,\infty}$) defined by Step1 is a linear operator by (7.24). Moreover, it follows from the Itô's isometry (7.28) for $\mathcal{E}$ that

$$\rho_2(H \cdot B) = \|H\|_{\mathcal{L}^2_\infty}, \quad H \in \mathcal{E}.$$

Therefore, by Lemma 7.3.7, the map $H \mapsto H \cdot B$ ($\mathcal{E} \longrightarrow \mathcal{M}^2_{c,\infty}$) can uniquely be extended to a linear isometry from $\mathcal{L}^2_\infty$ to $\mathcal{M}^2_{c,\infty}$.
To show the equality (7.27) for $H \in \mathcal{L}^2_\infty$, take a sequence $H^{(n)} \in \mathcal{E}$ which converges in $\mathcal{L}^2_\infty$ to H. Then, $H^{(n)}\chi_T$ converges in $\mathcal{L}^2_\infty$ to $H\chi_T$. These imply via Lemma 7.3.4 that

1) $H^{(n)} \cdot B \stackrel{n\to\infty}{\longrightarrow} H \cdot B$ and $H^{(n)}\chi_T \cdot B \stackrel{n\to\infty}{\longrightarrow} H\chi_T \cdot B$ in $\mathcal{M}_{c,\infty}$.

2) $$(H^{(n)}\chi_T \cdot B)_t = (H^{(n)} \cdot B)_{t\wedge T}$$

Therefore,

$$
\begin{aligned}
E\left[\sup_{t\geq 0}|(H\cdot B)_{t\wedge T}-(H\chi_T\cdot B)_t|^2\right] &\overset{1)}{=} \lim_{n\to\infty} E\left[\sup_{t\geq 0}|(H\cdot B)_{t\wedge T}-(H^{(n)}\chi_T\cdot B)_t|^2\right]\\
&\overset{2)}{=} \lim_{n\to\infty} E\left[\sup_{t\geq 0}|(H\cdot B)_{t\wedge T}-(H^{(n)}\cdot B)_{t\wedge T}|^2\right]\\
&\leq \lim_{n\to\infty} E\left[\sup_{t\geq 0}|(H\cdot B)_{t}-(H^{(n)}\cdot B)_{t}|^2\right] \overset{1)}{=} 0.
\end{aligned}
$$

\(^□^)/

Step3: (**The case of** $H\in\mathcal{L}^2_{\text{a.s.}}$)

Lemma 7.3.10 *The linear map $H\mapsto H\cdot B$ from $\mathcal{L}^2_\infty\to\mathcal{M}^2_{c,\infty}$ defined in Lemma 7.3.9 is uniquely extended to a linear map from $\mathcal{L}^2_{\text{a.s.}}\to\mathcal{M}^2_{c,\text{loc}}$ for which the equality (7.27) holds for all $H\in\mathcal{L}^2_{\text{a.s.}}$ and all stopping time T.*

Proof: Let $H\in\mathcal{L}^2_{\text{a.s.}}$. To define a process $H\cdot B$, We introduce the stopping times

$$
S_n=S_n(H)=n\wedge\inf\left\{t>0\,;\;\int_0^t H_s^2ds\geq n\right\}. \tag{7.34}
$$

Then, $(S_n)_{n\geq 1}$ is a nondecreasing sequence of finite stopping times such that $S_n\nearrow\infty$ and

$$
\int_0^\infty (H\chi_{S_n})_s^2ds=\int_0^{S_n}H_s^2ds\leq n \text{ for all } n\geq 1.
$$

Hence $H\chi_{S_n}\in\mathcal{L}^2_\infty$. Consequently, $H\chi_{S_n}\cdot B\in\mathcal{M}^2_{c,\infty}$ by Step2. We define the process $H\cdot B$ by

1) $\quad (H\cdot B)_t=(H\chi_{S_n}\cdot B)_t$ for $t\leq S_n$.

The process is well defined, since if $m<n$ and $t\leq S_m$, then, for $s\leq t$, $\chi_{S_m}(s)=\chi_{S_n}(s)=1$ and hence $(H\chi_{S_m})_s=(H\chi_{S_n})_s$. Consequently, $(H\chi_{S_m}\cdot B)_t=(H\chi_{S_n}\cdot B)_t$.

We next prove that $H\cdot B\in\mathcal{M}^2_{c,\text{loc}}$. By Lemma 7.3.9, the equality (7.27) holds if H is replaced by $H\chi_{S_n}\in\mathcal{L}^2_\infty$. Thus, if a stopping time S satisfies $S\leq S_n$, then,

2) $\quad (H\cdot B)_{t\wedge S}\overset{1)}{=}(H\chi_{S_n}\cdot B)_{t\wedge S}=(H\chi_S\cdot B)_t$ for all $t\geq 0$.

In particular,

3) $\quad (H\cdot B)_{t\wedge S_n}=(H\chi_{S_n}\cdot B)_t$ for all $t\geq 0$.

Since $H\chi_{S_n}\cdot B\in\mathcal{M}^2_{c,\infty}$, the equality 3) implies that $H\cdot B\in\mathcal{M}^2_{c,\text{loc}}$.

We next prove the equality (7.27). We note that the equality 3) determines the values of of $H\cdot B$ on the set $\{S_n\nearrow\infty\}$. Therefore the property 3) characterizes the process $H\cdot B$ (up to the identification in the class $\mathcal{M}_{c,\text{loc}}$). Referring to (7.34), we set $U_n=S_n(H\chi_T)$. Then, the process $H\chi_T\cdot B$ is characterized by the equality

4) $\quad (H\chi_T\cdot B)_{t\wedge U_n}=(H\chi_{T\wedge U_n}\cdot B)_t$.

Therefore, to prove (7.27), it is enough to verify that

$\quad (H\cdot B)_{t\wedge T\wedge U_n}=(H\chi_{T\wedge U_n}\cdot B)_t$.

Since $S_n\to\infty$ a.s., the above equality follows from the equality with t replaced by $t\wedge S_n$, namely (Note that $S_n\leq U_n$.),

5) $\quad (H\cdot B)_{t\wedge S_n\wedge T}=(H\chi_{T\wedge U_n}\cdot B)_{t\wedge S_n}$.

It follows from 2) that

the LHS of 5) $=(H\chi_{S_n\wedge T}\cdot B)_t$.

On the other hand, noting that $H\chi_{T\wedge U_n}\in\mathcal{L}^2_\infty$ and applying Lemma 7.3.9,

the RHS of 5) $=(H\chi_{S_n\wedge T}\cdot B)_t$.

Thus, we have proved (7.27).

Finally prove the linearity (7.24). Let $H, K \in \mathcal{L}^2_{\text{a.s.}}$, $\alpha, \beta \in L^\infty(\Omega, \mathcal{F}_0, P)$ and $L = \alpha H + \beta K$. Referring to (7.34), we set $T_n = S_n(H) \wedge S_n(K)$. Then, $H\chi_{T_n}, H\chi_{T_n} \in \mathcal{L}^2_\infty$, and hence by linearity of the stochastic integral for $\mathcal{L}^2_\infty$ (Lemma 7.3.9), we have

$$(L\chi_{T_n} \cdot B)_t = \alpha(H\chi_{T_n} \cdot B)_t + \beta(K\chi_{T_n} \cdot B)_t.$$

Therefore,

$$\begin{aligned}(L \cdot B)_{t\wedge T_n} &\stackrel{(7.27)}{=} (L\chi_{T_n} \cdot B)_t = \alpha(H\chi_{T_n} \cdot B)_t + \beta(K\chi_{T_n} \cdot B)_t \\ &\stackrel{(7.27)}{=} \alpha(H \cdot B)_{t\wedge T_n} + \beta(K \cdot B)_{t\wedge T_n}.\end{aligned}$$

Letting $n \to \infty$, we obtain (7.24). \(^□^)/

Lemma 7.3.11 *For $H, K \in \mathcal{L}^2$, the processes (7.26) are martingales.*

Proof: It is enough to show that the processes $(H \cdot B)_{t\wedge t_0}$, $Q_{t\wedge t_0}(H, K)$, $t \geq 0$ are martingale for any fixed $t_0 > 0$. Moreover, if $H \in \mathcal{L}^2$, then $H\chi_{t_0} \in \mathcal{L}^2_\infty$ and

$$(H \cdot B)_{t\wedge t_0} = (H\chi_{t_0} \cdot B)_t \text{ and } Q_{t\wedge t_0}(H, K) = Q_t(H\chi_{t_0}, K\chi_{t_0}).$$

Therefore, it is enough to assume that $H, K \in \mathcal{L}^2_\infty$. The process $H \cdot B$ for $H \in \mathcal{L}^2_\infty$ is a continuous martingale by Lemma 7.3.9. It only remains to prove that $Q_t(H, K)$ is a martingale. Since $\mathcal{E}$ is dense in $\mathcal{L}^2_\infty$ (Lemma 7.3.7), there exists $H^{(n)}, K^{(n)} \in \mathcal{E}$ such that $H^{(n)} \stackrel{n\to\infty}{\longrightarrow} H$ and $K^{(n)} \stackrel{n\to\infty}{\longrightarrow} K$ in $\mathcal{L}^2_\infty$. Since the map (7.33) is continuous, we have that $(H^{(n)} \cdot B)_t \stackrel{n\to\infty}{\longrightarrow} (H \cdot B)_t$ and $(K^{(n)} \cdot B)_t \stackrel{n\to\infty}{\longrightarrow} (K \cdot B)_t$ in $L^2(P)$, which implies that

1) $(H^{(n)} \cdot B)_t(K^{(n)} \cdot B)_t \stackrel{n\to\infty}{\longrightarrow} (H \cdot B)_t(K \cdot B)_t$ in $L^1(P)$.

On the other hand, it is easy to see that

2) $\displaystyle\int_0^t H^{(n)}_s K^{(n)}_s ds \stackrel{n\to\infty}{\longrightarrow} \int_0^t H_s K_s ds$ in $L^1(P)$.

By 1) and 2), we have that

3) $Q_t(H^{(n)}, K^{(n)}) \stackrel{n\to\infty}{\longrightarrow} Q_t(H, K)$ in $L^1(P)$.

Since $Q_t(H^{(n)}, , K^{(n)})$ is a martingale by Lemma 7.3.6, we see from 3) that $Q_t(H, K)$ is also a martingale. \(^□^)/

Complement

Lemma 7.3.12 *If M is a continuous local martingale with $M_0 = 0$, which is of bounded variation on any finite interval. Then, $M_t = 0$ a.s. for all $t \geq 0$.*

Proof: We set

$$T_k = \inf\{t \geq 0\,; V_t > k\},$$

where V_t denotes the total variation of M on the interval $[0, t]$. T_k is a stopping time, since V_t is continuous in t. Thus, $M^{(k)}_t \stackrel{\text{def}}{=} M_{t\wedge T_k}$ is a local martingale. Note also that

1) $$|M_{t\wedge T_k}| \leq V_{t\wedge T_k} \leq k.$$

Therefore, $M^{(k)}$ is a bounded martingale (Exercise 7.3.1). On the other hand, we have $T_k \stackrel{k\to\infty}{\longrightarrow} \infty$, since $V_t < \infty$ for any $t > 0$. Therefore, it is enough to to prove that $M^{(k)}_t = 0$ a.s. for all fixed $k \geq 1$ and $t > 0$. Let $k \geq 1$ and $t > 0$ be fixed. Since $M^{(k)}$ is a bounded martingale, its differences

$$\Delta M^{(k,j)}_t \stackrel{\text{def}}{=} M^{(k)}(jt/n) - M^{(k)}((j-1)t/n), \;\; j = 1, \ldots, n$$

are orthognal. Thus,

2) $$E\left[|M_t^{(k)}|^2\right] = \sum_{j=1}^n E\left[|\Delta M_t^{(k,j)}|^2\right] = E\left[\sum_{j=1}^n |\Delta M_t^{(k,j)}|^2\right].$$

Moreover,

$$\sum_{j=1}^n |\Delta M_t^{(k,j)}|^2 \le V_{t\wedge T_k} \max_{1\le j\le n} |\Delta M_t^{(k,j)}|.$$

By 1), the RHS of the above display is bounded from above by the constant $2k^2$, while it converges to zero as $n \to \infty$, since M is uniformly continuous on the interval $[0, t]$. Therefore, by the bounded convergence theorem, the RHS of the display 2) converges to zero as $n \to \infty$, which shows that $M_t^{(k)} = 0$, a.s. \(^□^)/

Exercise 7.3.1 Prove the following. **i)** Suppose that M is a continuous local martingale and that $\sup_{t\le t_0} |M_t| \in L^1(P)$ for some $t_0 > 0$. Then, $(M_t)_{t\in[0,t_0]}$ is a martingale. **Hint**: Let $(T_n)_{n\ge1}$ be the stopping times in Definition 7.3.2. Then, $E[M(t \wedge T_n) : A] = E[M(s \wedge T_n) : A]$ for all $s < t \le t_0$ and $A \in \mathcal{F}_s$. **ii)** Suppose that $H \in \mathcal{L}_{\text{a.s.}}$ and that $\sup_{t\le t_0} |(H \cdot B)_t| \in L^1(P)$ for some $t_0 > 0$. Then, $((H \cdot B)_t)_{t\in[0,t_0]}$ is a martingale.

Exercise 7.3.2 Prove the following for $H \in \mathcal{L}^2_{\text{a.s.}}$. **i)** $\int_0^t H_s^2 ds = 0$ a.s. for all $t > 0 \iff \int_0^t H_s dB_s = 0$ a.s. for all $t > 0$. **Hint**: Referring to (7.26) and (7.34), apply the optional stopping theorem to the uniformly integrable martingale $Q_{t\wedge S_n}(H)$. **ii)** Suppose that S and T are stopping times such that $S \le T < \infty$ a.s. Then, $\int_S^T H_s^2 ds = 0$ a.s. $\iff \int_S^T H_s dB_s = 0$ a.s. **Hint**: Apply i) to $H_s \mathbf{1}_{(S,T]}(s)$

Exercise 7.3.3 Let $M_t^\alpha = \int_0^t H_s^\alpha dB_s^\alpha$, $(\alpha = 1, 2)$, where (B_t^1, B_t^2), $t \ge 0$ is a BM^2, and $H^1, H^2 \in \mathcal{L}^2$. Then, prove that the process $M_t^1 M_t^2$, $t \ge 0$ is a martingale. **Hint**: It is enough to assume that $H^1, H^2 \in \mathcal{L}^2_\infty$ (cf. proof of Lemma 7.3.11). Then, reduce the case of $H^1, H^2 \in \mathcal{L}^2_\infty$ to that of $H^1, H^2 \in \mathcal{E}$, by considering sequences $H^{\alpha,(n)} \in \mathcal{E}$ such that $H^{\alpha,(n)} \stackrel{n\to\infty}{\longrightarrow} H^\alpha$ in $\mathcal{L}^2_\infty$.

7.4 Itô's Formula I

In this subsection, we will explain Itô's formula for the Brownian motion and its applications. In what follows, $B_t = (B_t^\alpha)_{\alpha=1}^d$, $t \ge 0$ denotes a BM^d w.r.t. a filtration $(\mathcal{F}_t)_{t\ge0}$, cf. Definition 7.1.1. We first state the Itô's formula in its simplest form.

Theorem 7.4.1 (Itô's formula I) *Suppose that $f \in C^2(\mathbb{R}^d)$. Then, P-a.s., for all $t \ge 0$,*

$$f(B_t) - f(B_0) = \sum_{\alpha=1}^d \int_0^t \frac{\partial f}{\partial x^\alpha}(B_s) dB_s^\alpha + \tfrac{1}{2} \int_0^t \Delta f(B_s) ds, \tag{7.35}$$

where $\Delta f = \sum_{\alpha=1}^d \frac{\partial^2 f}{\partial x_\alpha^2}$

Example 7.4.2 (The Dirichlet problem) Let $D \subset \mathbb{R}^d$ be a domain, $f \in C(\partial D)$, and $g \in C(\overline{D})$ be given. A classical problem in the theory of partial differential equations is to show the existence and uniqueness of $u \in C(\overline{D}) \cap C^2(D)$ such that

a) $\frac{1}{2}\Delta u = -g$ in D,

b) $u|_{\partial D} = f$.

A special case where $g \equiv 0$ is especially celebrated as *Dirichlet problem.* Here, we suppose for simplicity that D is bounded. We will prove the uniqueness of the solution to a) and b) by running a Brownian motion. We adopt the notation introduced at the beginning of section 6.7. We will represent the solution as follows. Let B be a BM^d_x, $x \in D$ and

$$T = T_{D^{\mathrm{c}}} = \inf\{t > 0 \; ; \; B_t \in D^{\mathrm{c}}\}.$$

By Proposition 6.9.3, T is a stopping time. Moreover, by Example 7.1.4, there exists $\lambda > 0$ such that

$$\sup_{x \in D} E_x \exp(\lambda T) < \infty.$$

We will then prove that a solution u to a) and b) is represented as

1) $$u(x) = E_x f(B_T) + E_x \int_0^T g(B_s) ds,$$

hence is unique.
Proof: Suppose that $x \in D$ and $u \in C(\overline{D}) \cap C^2(D)$ satisfies a) and b). Let

$$\begin{aligned} D_n &= \{y \in D \; ; \; \mathrm{dist}(y, G^{\mathrm{c}}) > 1/n\}, \\ T_n &= \inf\{t > 0 \; ; \; B_t \in D_n^{\mathrm{c}}\}. \end{aligned}$$

Then, there exists $u_n \in C^2_{\mathrm{c}}(\mathbb{R}^d)$ such that $u_n = u$ on D_{n+1}. Take n large enough so that $x \in D_n$ and fix it. Then, for each $\alpha = 1, ..., d$, the process $(\partial_\alpha u_n(B_t))_{t \geq 0}$ is bounded, progressively measurable. Thus, by Theorem 7.3.3, the following process is a martingale:

$$M^{(n)}_t = \sum_{\alpha=1}^d \int_0^t \partial_\alpha u_n(B_s) dB^\alpha_s, \;\; t \geq 0.$$

Thus, $(M^{(n)}_{t \wedge T_n})_{t \geq 0}$ is also a martingale by Lemma 5.3.5. In particular,

2) $EM^{(n)}_{t \wedge T_n} = 0, \;\; \forall t \geq 0.$

On the other hand, we have by Itô's formula applied to u_n that,

$$\begin{aligned} u(B_{t \wedge T_n}) - u(x) - M^{(n)}_{t \wedge T_n} &= u_n(B_{t \wedge T_n}) - u_n(x) - M^{(n)}_{t \wedge T_n} \\ &= \frac{1}{2} \int_0^{t \wedge T_n} \Delta u_n(B_s) ds = \frac{1}{2} \int_0^{t \wedge T_n} \Delta u(B_s) ds \\ &\stackrel{\text{a)}}{=} - \int_0^{t \wedge T_n} g(B_s) ds. \end{aligned}$$

We then take expectation and use 2) to see that

$$E_x u(B_{t \wedge T_n}) - u(x) = -E_x \int_0^{t \wedge T_n} g(B_s) ds.$$

Since $T_n \stackrel{n \to \infty}{\longrightarrow} T$ by Exercise 4.2.4, we have by the bounded convergence theorem in the limit $n \to \infty$ that

$$E_x u(B_{t \wedge T}) - u(x) = -E_x \int_0^{t \wedge T} g(B_s) ds.$$

Finally by the assumption, we can use the bounded convergence theorem in the limit $t \to \infty$ to conclude 1) from the above displayed identity. \(^□^)/

It is also possible to show the existence of the solution u by Brownian moiton. In fact, it turns out that the function u defined by 1) gives a solution to a) and b). To do so, however, one has to assume the following regularity condition on ∂D to show the continuity of u at the boundary:

$$P_x(T = 0) = 1 \text{ for all } x \in \partial D.$$

See [Dur84, Sections 8.5,8.6], [KS91, Section 4.2] for the proofs and details.

Remark Of course, the existence and uniqueness of u discussed in Example 7.4.2 can be shown without using Brownian motion.

- Uniqueness is a consequence of the maximal principle for harmonic functions [Fol76,page 93].
- Existence can also be established via the existence of the Green function for the domain D assuming that D has a smooth boundary [Fol76,pages 112, 343].

Exercise 7.4.1 Let B be a BM_0^d and $h \in C^1([0,\infty) \to \mathbb{R}^d)$. Then, prove the following. **(i)**(**Integration by parts formula**)

$$\int_0^t h(s) \cdot dB_s = h(t) \cdot B_t - \int_0^t h'(s) \cdot B_s ds.$$

(ii) Use i) and Theorem 7.6.1 to show that

$$\mathcal{D}_t(h) \stackrel{\text{def}}{=} \exp\left(\int_0^t h(s) \cdot dB_s - \frac{1}{2}\int_0^t |h(s)|^2 ds\right) = \int_0^t \mathcal{D}_s(h) h(s) \cdot dB_s.$$

Then, use Exercise 7.3.1 that $\mathcal{D}_t(h)$ is a martingale. **(iii)** Suppose that $h(t) > 0$, $\forall t \geq 0$. Then, the process Y_t discussed in Exercise 6.1.9 (Y_t is the Ornstein-Uhlenbeck process if $h(t) = \exp(\lambda t)$ with $\lambda > 0$) can alternatively be written in terms of the stochastic integral as follows.

$$X_t = h(t)^{-1}\left(h(0)x + \int_0^t h(s) dB_s\right),$$

which, together with (7.28), implies that $EY_t = h(t)^{-1}h(0)x$ and that $E[Y_t^2] = h(t)^{-2}\int_0^t h(s)^2 ds$.

Exercise 7.4.2 Let $\varphi \in C^2(\mathbb{R}^d \to \mathbb{R})$ and suppose that there exists $C \in [0,\infty)$ such that

$$\varphi(x) \leq C(1+|x|), \quad \Delta\varphi(x) \geq -C(1+|x|).$$

for all $x \in \mathbb{R}^d$. For a BM_0^d denoted by B, we set $A_t = \frac{1}{2}\int_0^t \left(|\nabla\varphi(B_s)|^2 + \Delta\varphi(B_s)\right) ds.$

Use Theorem 7.4.1 and Theorem 7.6.1 to show that

$$\begin{aligned} \mathcal{D}_t(\varphi) &\stackrel{\text{def}}{=} \exp\left(\int_0^t \nabla\varphi(B_s) \cdot dB_s - \frac{1}{2}\int_0^t |\nabla\varphi(B_s)|^2 ds\right) \\ &= \exp\left(\varphi(B_t) - \varphi(B_0) - A_t\right) = \int_0^t \mathcal{D}_s(\varphi)\nabla\varphi(B_s) \cdot dB_s. \end{aligned}$$

Then, use Exercise 7.3.1 that $\mathcal{D}_t(\varphi)$ is a martingale.

Exercise 7.4.3 Let $Z_t = X_t + \mathbf{i}Y_t$, where (X_t, Y_t) is a BM^2. For $U, V \in \mathcal{L}^2_{\text{a.s.}}$. We define

$$\int_0^t (U_s + \mathbf{i}V_s) dZ_s = \int_0^t U_s dX_s - \int_0^t V_s dY_s + \mathbf{i}\int_0^t U_s dY_s + \mathbf{i}\int_0^t V_s dX_s.$$

Then, prove the following identity for a holomorphic function $f : \mathbb{C} \to \mathbb{C}$.

$$f(Z_t) - f(Z_0) = \int_0^t f'(Z_s)dZ_s.$$

Exercise 7.4.4 (**A uniformly integral local martingale which is not a martingale**) Let B be BM^d_c ($d \geq 3$, $c \neq 0$) and $\varphi(x) = |x|^{-(d-2)}$, $x \in \mathbb{R}^d$. Prove the following. **i)** $\varphi(B_t)$ is a local martingale. **Hint**: Let $T_n = \inf\{t \geq 0 \ ; \ |B_t| \leq 1/n\}$. Then, $\varphi(B(t \wedge T_n)) - \varphi(c) = -(d-2)\sum_{\alpha=1}^d \int_0^{t\wedge T_n} |B_s|^{-d} B_s^\alpha dB_s^\alpha$. **ii)** $\varphi(B_t)$ is not a martingale. **Hint**: By Exercise 1.2.10, $E[\varphi(B_t)]$ is strictly decreasing in t. **iii)** For $1 < p < d/(d-2)$ and $\varepsilon > 0$, $\sup_{t \geq \varepsilon} E[\varphi(B_t)^p] < \infty$. In particular, $\varphi(B_t)$ ($t \geq \varepsilon$) is uniformly integrable.

7.5 Semimartingales Generated by a Brownian Motion

Throughout this subsection, we let B denote a BM^d w.r.t. a filtration $(\mathcal{F}_t)_{t\geq 0}$. Recall the definition of the class of processes $\mathcal{L}^2_{\text{a.s.}}$ from Definition 7.3.1. We now define

$$(\mathcal{L}^2_{\text{a.s.}})^d = \{(H_t^1, \ldots, H_t^d)_{t\geq 0} \ ; \ (H_t^\alpha)_{t\geq 0} \in \mathcal{L}^2_{\text{a.s.}} \ \text{for all } \alpha = 1, \ldots, d\}.$$

For $H \in (\mathcal{L}^2_{\text{a.s.}})^d$, we write

$$\int_0^t H_s \cdot dB_s = \sum_{\alpha=1}^d \int_0^t H_s^\alpha dB_s^\alpha.$$

Definition 7.5.1 (**Local martingales generated by a Brownian motion**)
► A process M is called a **local martingale generated by** B, if there exists a process $\sigma = (\sigma_t)_{t\geq 0} \in (\mathcal{L}^2_{\text{a.s.}})^d$ such that

$$M_t = \int_0^t \sigma_s \cdot dB_s, \quad t \geq 0. \tag{7.36}$$

► Let M be a local martingale generated by B expressed in the form (7.36), and let H be a continuous, adapted process. We use the following notation.

$$\int_0^t H_s dM_t = \int_0^t H_s \sigma_s \cdot dB_t = \sum_{\alpha=1}^d \int_0^t H_s \sigma_s^\alpha dB_s^\alpha. \tag{7.37}$$

► Suppose that $M_t^\mu = \int_0^t \sigma_s^\mu \cdot dB_s$, $\mu = 1, 2$ are local martingales generated by B. Then, we define the process $\langle M^1, M^2 \rangle$ by

$$\langle M^1, M^2 \rangle_t = \int_0^t \sigma_s^1 \cdot \sigma_s^2 ds, \quad t \geq 0. \tag{7.38}$$

The above process is called the **quadratic variation** of M^μ ($\mu = 1, 2$). When $M^1 = M^2 = M$, we often write $\langle M \rangle$, instead of $\langle M, M \rangle$.

Lemma 7.5.2 *Suppose that M^1, M^2 are local martingales generated by B. Then, the quadratic variation $\langle M^1, M^2 \rangle$ is characterized as the unique process $Q = (Q)_{t\geq 0}$ with the following properties.*
Q1) *Q is locally of bounded variation.*
Q2) *$Q_0 = 0$ and $M_t^1 M_t^2 - Q_t$, $t \geq 0$ is a local martingale.*

Proof: Suppose that $Q = \langle M^1, M^2 \rangle$. We then verify properties Q1) and Q2). Q1) is obvious. To see Q2), we observe that

$$M_t^1 M_t^2 - Q_t = \sum_{\alpha,\beta=1}^{d} \left(\left(\int_0^t \sigma_s^{1,\alpha} dB_s^\alpha \right) \left(\int_0^t \sigma_s^{2,\beta} dB_s^\beta \right) - \delta_{\alpha,\beta} \int_0^t \sigma_s^{1,\alpha} \sigma_s^{2,\beta} ds \right).$$

By applying Theorem 7.3.3 b) and Exercise 7.3.3 respectively to the diagonal, and the off digonal terms of the summation on the RHS of the above display, we obtain the property Q2).

Suppose that a process Q satisfies the properties Q1) and Q2). Since

$$\langle M^1, M^2 \rangle_t - Q_t = (M_t^1 M_t^2 - Q_t) - (M_t^1 M_t^2 - \langle M^1, M^2 \rangle_t),$$

it follows that the process $\langle M^1, M^2 \rangle_t - Q_t$ is a local martingale and is at the same time of bounded variation. Therefore, by Lemma 7.3.12, $\langle M^\mu, M^\nu \rangle_t = Q_t$. \\(^□^)/

Definition 7.5.3 (**Semimartingales generated by a Brownian motion**)

▶ A process X is called a **semimartingale generated by** B, if

$$X_t = X_0 + M_t + A_t, \quad t \geq 0, \tag{7.39}$$

where M is a local martingale generated by B (Definition 7.5.1), and A is an adapted process with $A_0 = 0$ which is continuous and locally of bounded variation. The processes M and A are called respectively the **local martingale part** and **bounded variantion part** of X, cf. the remark after the definition.

▶ Let X be a semimartingale generated by B, decomposed in the form (7.39), and let H be a continuous, adapted process. Then, referring to (7.37), we use the following notation.

$$\int_0^t H_s dX_t = \int_0^t H_s dM_s + \int_0^t H_s dA_t. \tag{7.40}$$

▶ Let X^μ ($\mu = 1, 2$) be semimartingales generated by B and M^μ ($\mu = 1, 2$) be their respective martingale parts. Then, referring to (7.38), we define their **quadratic variation** $\langle X^1, X^2 \rangle$ by $\langle X^1, X^2 \rangle = \langle M^1, M^2 \rangle$.

Remark: Given a semimartingale generated by B, its local martingale part and bounded variantion part are uniquely determined (Lemma 7.3.12).

7.6 Itô's Formula II

Although Theorem 7.4.1 is already very useful, the scope of application can considerably be broadened by generalizing it in the following way.

Theorem 7.6.1 (**Itô's formula II**) *Let X^μ ($\mu = 1, \ldots, m$) be semimartingales generated by B (Definition 7.5.3) and $f \in C^2(\mathbb{R}^m)$. Then, P-a.s., for all $t \geq 0$,*

$$\begin{aligned} &f(X_t) - f(X_0) \\ &= \sum_{\mu=1}^{m} \int_0^t \frac{\partial f}{\partial x^\mu}(X_s) dX_s^\mu + +\tfrac{1}{2} \sum_{\mu,\nu=1}^{m} \int_0^t \frac{\partial^2 f}{\partial x^\mu \partial x^\nu}(X_s) d\langle X^\mu, X^\nu \rangle_s. \end{aligned} \tag{7.41}$$

We will prove Theorem 7.6.1 in section 7.7. For the rest of this section, we present applications of Theorem 7.4.1 and Theorem 7.6.1.

As an application of Theorem 7.6.1, the martingales in Proposition 7.1.3 are expressed in terms of the stochastic integral as follows.

Example 7.6.2 Let $g_\lambda : \mathbb{R}^2 \to \mathbb{R}$ ($\lambda \in \mathbb{R}$) and $H_n : \mathbb{R}^2 \to \mathbb{R}$ ($n \in \mathbb{N}$) be from Proposition 7.1.3. Then,

$$\begin{aligned} & H_n(B^\alpha_t - \lambda t, t) g_\lambda(B^\alpha_t, t) \\ = \; & H_n(B^\alpha_0, 0) g_\lambda(B^\alpha_0, 0) + \int_0^t (\lambda H_n + n H_{n-1})(B^\alpha_s - \lambda s, s) g_\lambda(B^\alpha_s, s) dB^\alpha_s. \end{aligned} \tag{7.42}$$

In particular,

$$\begin{aligned} g_\lambda(B^\alpha_t, t) &= g_\lambda(B^\alpha_0, 0) + \lambda \int_0^t g_\lambda(B^\alpha_s, s) dB^\alpha_s, \\ H_n(B^\alpha_t, t) &= H_n(B^\alpha_0, 0) + n \int_0^t H_{n-1}(B^\alpha_s, s) dB^\alpha_s. \end{aligned}$$

Proof: We have by Exercise 7.1.1 that

1) $\dfrac{\partial H_n}{\partial x}(x,t) = n H_{n-1}(x,t).$

On the other hand, it is easy to see that

2) $\dfrac{\partial g_\lambda}{\partial t}(x,t) + \frac{1}{2}\dfrac{\partial^2 g_\lambda}{\partial x^2}(x,t) = 0.$

Let $f(x,t) = H_n(x - \lambda t, t) g_\lambda(x,t) = \left(\frac{\partial}{\partial \lambda}\right)^n g_\lambda(x,t)$. Then,

3) $$\left\{ \begin{aligned} \frac{\partial f}{\partial x}(x,t) &= \left(\frac{\partial H_n}{\partial x} + \lambda H_n\right)(x - \lambda t, t) g_\lambda(x,t) \\ &\stackrel{1)}{=} (\lambda H_n + n H_{n-1})(x - \lambda t, t) g_\lambda(x,t), \end{aligned} \right.$$

4) $\left(\dfrac{\partial f}{\partial t} + \frac{1}{2}\dfrac{\partial^2 f}{\partial x^2}\right)(x,t) = \left(\dfrac{\partial}{\partial \lambda}\right)^n \left(\dfrac{\partial g_\lambda}{\partial t} + \frac{1}{2}\dfrac{\partial^2 g_\lambda}{\partial x^2}\right)(x,t) \stackrel{2)}{=} 0.$

Hence (7.42) follows from (7.41) as follows:

$$\begin{aligned} f(B^\alpha_t, t) - f(B^\alpha_0, 0) &= \int_0^t \frac{\partial f}{\partial x}(B^\alpha_s, s) dB^\alpha_s + \int_0^t \left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\right)(B^\alpha_s, s) ds \\ &\stackrel{3),4)}{=} \int_0^t (\lambda H_n + n H_{n-1})(B^\alpha_s - \lambda s, s) g_\lambda(B^\alpha_s, s) dB^\alpha_s. \end{aligned}$$

\(^□^)/

Lemma 7.6.3 *Let $\varphi \in C([0,\infty) \to \mathbb{R})$ with $\psi(t) = \sup_{s \le t} \varphi(s)$. Then,*

$$\int_G d\psi = 0, \ \textit{where } G = \{t \in (0,\infty) \; ; \; \varphi(t) < \psi(t)\}.$$

Proof: Let $S \subset [0,\infty)$ be the support of the measure $\mu(A) \stackrel{\text{def}}{=} \int_A d\psi$, $A \in \mathcal{B}([0,\infty))$ ($[0,\infty)\backslash S$ is the union of all open subsets of $[0,\infty)$ on which μ vanishes.) Then, it is enough to prove that $G \cap S = \emptyset$. To do so, we take an arbitrary $t \in G$. Since $\varphi(t) < \psi(t)$, there exists $t_* \in (0,t)$ such that $\psi(t) = \varphi(t_*)$. Then, by the continuity of φ, there exists $\varepsilon > 0$ such that $t_* < t - \varepsilon$ and that $\varphi(s) < \psi(t)$ for all $s \in [t-\varepsilon, t+\varepsilon]$. This implies that $\psi(t \pm \varepsilon) = \psi(t)$, and hence that $\int_{(t-\varepsilon,t+\varepsilon]} d\psi = 0$. Therefore, $t \notin S$. \(^□^)/

Example 7.6.4 (Position of the first decrease by length ℓ) Let B be BM_0^1, $S_t = \sup_{u \le t} B_u$, and

$$T = \inf\{t \geq 0 \; ; \; B_t = S_t - \ell\}, \quad \ell > 0.$$

Then, the r.v. $S_T (= B_T + \ell)$ is exponentially distributed with parameter $1/\ell$.

Proof: We start with a general consideration. Let $f \in C^2(\mathbb{R})$ and $F(x) = \int_0^x f$. Then,

1) $F(S_t) - (S_t - B_t) f(S_t) = \displaystyle\int_0^t f(S_u) dB_u.$

To see this, note first that

2) $F(S_t) = \displaystyle\int_0^t f(S_u) dS_u,$

which follows from Theorem 7.6.1 without Brownian motion. On the other hand, let $g(x,y) = (y-x)f(y)$. Then,

$$g_x(x,y) = -f(y), \quad g_{x,x}(x,y) = 0, \quad g_y(x,y) = f(y) + (y-x)f'(y).$$

Thus, by Theorem 7.6.1,

3) $(S_t - B_t) f(S_t) = -\displaystyle\int_0^t f(S_u) dB_u + \int_0^t f(S_u) dS_u + \int_0^t (S_u - B_u) f'(S_u) dS_u.$

By Lemma 7.6.3, the third term on the right-hand side of 3) vanishes. Therefore, 1) follows from 2) and 3). By applying 1) for $f(x) = -\alpha \exp(-\alpha x)$ with $\alpha > 0$, we see that the following process is a bounded martingale:

$$X_t \stackrel{\text{def}}{=} (1 + \alpha(S_t - B_t)) \exp(-\alpha S_t).$$

Hence by the optional stopping theorem,

$$EX_T = 1, \text{ i.e., } E \exp(-\alpha S_T) = \frac{1}{1 + \alpha \ell}.$$

Then, the result follows from the uniqueness of the Laplace transform (Example 1.8.3). \(^□^)/

Remark: See Exercise 4.5.3 for an analogy in the case of the random walk.

Example 7.6.5 (The heat equation in a domain) Let $D \subset \mathbb{R}^d$ be a domain. Following the convention in physics, we denote a point in $D \times [0,\infty)$ by (x,t) ($x \in D$, $t \geq 0$). Accordingly, for

$u : D\times(0,\infty) \to \mathbb{R}$, we write $\partial_t u = \partial_{d+1} u$. Suppose that $u \in C_{\rm b}(\overline{D}\times[0,\infty))\cap C^{2,1}(D\times(0,\infty))$ is such that

$$
\begin{array}{ll}
\text{a)} & \partial_t u = \frac{1}{2}\Delta u \ \text{ on } D\times(0,\infty),\\
\text{b)} & u = 0 \ \text{ on } \partial D\times[0,\infty).
\end{array}
$$

We adopt the notation introduced at the beginning of section 6.7. We will represent the solution of a) and b) as follows. Let B be a BM^d_x, $x \in D$ and

$$T = T_{D^{\rm c}} = \inf\{t > 0\ ;\ B_t \in D^{\rm c}\}.$$

By Proposition 6.9.3, T is a stopping time. We will then prove that a solution u to a) and b) is represented as

1) $\quad u(x,t) = E_x[u(B_t,0) : t < T].$

Proof: Let

$$
\begin{array}{rcl}
D_n &=& \{y \in D\ ;\ |y-x| < n,\ \mathrm{dist}(y, D^{\rm c}) > 1/n\},\\
T_n &=& n \wedge \inf\{t > 0\ ;\ B_t \in D_n^{\rm c}\}.
\end{array}
$$

Let $t > 0$ be fixed. Then, for $\varepsilon \in (0,t)$, there exists $u_n \in C^{2,1}_{\rm c}(\mathbb{R}^{d+1})$ such that $u_n = u$ on $D_{n+1}\times[\varepsilon, n+1]$. Take n large enough so that $x \in D_n$ and fix it. Then, for each $\alpha = 1, ..., d$, the process $(\partial_\alpha u_n(B_s, t-s))_{s\ge 0}$ is bounded, progressively measurable. Thus, by Theorem 7.3.3, the following process is a martingale:

$$M^{(t,n)}_s = \sum_{\alpha=1}^d \int_0^s \partial_\alpha u_n(B_r, t-r)dB^\alpha_r, \ \ s \ge 0.$$

Thus, $(M^{(t,n)}_{s\wedge T_n})_{s\ge 0}$ is also a martingale by Lemma 5.3.5. In particular,

2) $EM^{(t,n)}_{s\wedge T_n} = 0, \ \ \forall s \ge 0.$

On the other hand, we have by Itô's formula applied to the function

$$\mathbb{R}^{d+1} \ni (x,s) \mapsto u_n(x, t-s)$$

for $0 \le s \le (t-\varepsilon)\wedge T_n$ that,

$$
\begin{array}{rl}
& u(B_{(t-\varepsilon)\wedge T_n}, t-(t-\varepsilon)\wedge T_n) - u(x,t) - M^{(t,n)}_{(t-\varepsilon)\wedge T_n}\\
= & u_n(B_{(t-\varepsilon)\wedge T_n}, t-(t-\varepsilon)\wedge T_n) - u_n(x,t) - M^{(t,n)}_{(t-\varepsilon)\wedge T_n}\\
= & \displaystyle\int_0^{(t-\varepsilon)\wedge T_n} \left(\tfrac{1}{2}\Delta u_n(B_s, t-s) - \partial_t u_n(B_s, t-s)\right) ds\\
= & \displaystyle\int_0^{(t-\varepsilon)\wedge T_n} \left(\tfrac{1}{2}\Delta u(B_s, t-s) - \partial_t u(B_s, t-s)\right) ds \stackrel{\text{a)}}{=} 0.
\end{array}
$$

We then take expectation and use 2) to see that

$$u(x,t) = E_x u(B_{(t-\varepsilon)\wedge T_n}, t-(t-\varepsilon)\wedge T_n).$$

Since $T_n \stackrel{n\to\infty}{\longrightarrow} T$ by Exercise 4.2.4, we have by the bounded convergence theorem in the limit $n \to \infty$ that

$$u(x,t) = E_x u(B_{(t-\varepsilon)\wedge T}, t-(t-\varepsilon)\wedge T) \stackrel{\text{b)}}{=} E_x[u(B_{t-\varepsilon}, \varepsilon) : t-\varepsilon < T].$$

Finally, by taking the limit $\varepsilon \to 0$, we conclude 1) from the above displayed identity. \(^□^)/

Example 7.6.6 (The heat equation in a finite interval) Let $h_t(x) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2}\right)$ ($x \in \mathbb{R}$, $t > 0$), $\ell = b-a$, and

$$h_t^{a,b}(x,y) = \sum_{n\in\mathbb{Z}} (h_t(x-y-2\ell n) - h_t(x+y-2a-2\ell n)), \quad x,y \in \mathbb{R}.$$

Then, for $f \in C([a,b])$ with $f(a) = f(b) = 0$,

1) $E_x[f(B_t) : t < T_a \wedge T_b] = \int_a^b h_t^{a,b}(x,y) f(y) dy, \quad x \in [a,b], \ t > 0.$

Proof: We denote the RHS of 1) by $u(x,t)$. We will verify that

$$\begin{array}{ll} \text{a)} & \partial_t u = \frac{1}{2}\partial_x^2 u \ \text{ on } (a,b)\times(0,\infty), \\ \text{b)} & u(a,t) = u(b,t) = 0 \ \text{ for } t > 0, \\ \text{c)} & u(x,t) \stackrel{t\to 0}{\longrightarrow} f(x) \ \text{ for } x \in [a,b]. \end{array}$$

Then, by the result of Example 7.6.5, the function $u(x,t)$ is identified with the expectation on the LHS of 1). It is easy to see that $h_t^{a,b}(x,y) = 0$ if $x = a, b$, which implies b). Now, we define a continuous extention $\widetilde{f} : \mathbb{R} \to \mathbb{R}$ of f by

$$\widetilde{f}(x+2\ell n) = \left\{ \begin{array}{ll} f(x) & \text{if } x \in [a,b] \text{ and } n \in \mathbb{Z}, \\ -f(2a-x) & \text{if } x \in [2a-b,a] \text{ and } n \in \mathbb{Z}. \end{array} \right.$$

Since $f \in C([a,b])$ and $f(a) = f(b) = 0$, $\widetilde{f}$ is indeed a continuous extention of f. Note also that $\widetilde{f}$ has the period 2ℓ. We will show that

2) $u(x,t) = \int_{-\infty}^{\infty} h_t(x-y)\widetilde{f}(y)dy,$

which implies a) and c). 2) can be seen as follows.

$$\begin{aligned} \int_a^b h_t^{a,b}(x,y) f(y) dy &= \sum_{n\in\mathbb{Z}} \int_a^b (h_t(x-y-2\ell n) - h_t(x+y-2a-2\ell n)) f(y) dy \\ &= \sum_{n\in\mathbb{Z}} \left(\int_a^b h_t(x-y-2\ell n) f(y) dy + \int_{2a-b}^a h_t(x-y-2\ell n) f(2a-y) dy \right) \\ &= \sum_{n\in\mathbb{Z}} \left(\int_a^b h_t(x-y-2\ell n) \widetilde{f}(y) dy + \int_{2a-b}^a h_t(x-y-2\ell n) \widetilde{f}(y) dy \right) \\ &= \sum_{n\in\mathbb{Z}} \left(\int_{[a,b]+2\ell n} h_t(x-y) \widetilde{f}(y) dy + \int_{[2a-b,a]+2\ell n} h_t(x-y) \widetilde{f}(y) dy \right) \\ &= \int_{-\infty}^{\infty} h_t(x-y) \widetilde{f}(y) dy. \end{aligned}$$

\(^□^)/

Exercise 7.6.1 Suppose that $f \in C^2([0,\infty) \to (0,\infty))$ is nondecreaing, convex, $f(0) = 1$, and $f'(0) = 0$. We set

$$g_1 = f'/f,\ g_2 = f''/f \text{ and } h(t,x) = f(t)^{-d/2} \exp\left(-\tfrac{1}{2}g_1(t)|x|^2\right),\ (t,x) \in [0,\infty) \times \mathbb{R}^d.$$

Prove the following. **i)** $\dfrac{\partial}{\partial t}h(t,x) = \frac{1}{2}\Delta_x h(t,x) + \frac{1}{2}|x|^2 g_2(t)h(t,x)$. **ii)** With $t > 0$ and $\theta \in \mathbb{R}$ fixed, we define the process $(H_s)_{0 \le s \le t}$ by $H_s = h(t-s, B_s)$, where B is a BM_0^d. Then,

$$H_s = f(t)^{-d/2} - \int_0^s H_u g_1(t-u) B_u \cdot dB_u + \frac{1}{2}\int_0^s H_u g_2(t-u)|B_u|^2 du, \quad 0 \le s \le t.$$

iii) Let $Y_s = \exp\left(-\frac{1}{2}\int_0^s g_2(t-u)|B_u|^2 du\right)$. Then, the process $(H_s Y_s)_{0 \le s \le t}$ is a martingale, which implies that $EY_t = f(t)^{-d/2}$. In particular, taking $f(t) = \cosh(\theta t)$ $(\theta \in \mathbb{R})$,

$$E\exp\left(-\frac{\theta^2}{2}\int_0^t |B_s|^2 ds\right) = \cosh(\theta t)^{-d/2} \ (\textbf{Cameron-Martin formula I}).$$

vi) Let $Z_s = \exp\left(\mathbf{i}\int_0^s g_2(t-u)\sigma_u \cdot dB_u\right)$, where σ is a continuous process with values in $\mathbb{R}^d$ such that $|\sigma_s| = |B_s|$ and $\sigma_s \cdot B_s = 0$ a.s. for all $s \in [0,t]$. Then, the process $(H_s Z_s)_{0 \le s \le t}$ is a martingale, which implies that $EZ_t = f(t)^{-d/2}$. In particular, if $d = 2$ and $\mathcal{A}_t = \int_0^t B_s^1 dB_s^2 - \int_0^t B_s^2 dB_s^1$, then, taking $f(t) = \cosh(\theta t)$ $(\theta \in \mathbb{R})$,

$$E\exp(\mathbf{i}\theta\mathcal{A}_t) = \cosh(\theta t)^{-1} \ (\textbf{Lévy's area formula I}).$$

Remark For $d = 2$, it follows from Cameron-Martin formula I and Exercise 7.2.2 that the r.v. $\int_0^a |B_s|^2 ds$ $(a > 0)$ has the same law as the exit time from the interval $(-a,a)$ for a BM_0^1. On the other hand, By it follows from Lévy's area formula I and Exercise 2.2.7 that $\mathcal{A}_t \approx \frac{dx}{t\cosh\left(\frac{\pi x}{2t}\right)}$.

7.7 $(\star)$ Proof of Theorem 7.6.1

We start by stating a proposition, which the proof of Theorem 7.6.1 is based on. Let $t > 0$ be fixed. We divide the interval $(0,t]$ into $I_{n,j} = (t_{n,j-1}, t_{n,j}]$ $(j = 1,\ldots,n)$ in such a way that

$$0 = t_{n,0} < t_{n,1} < \ldots < t_{n,n} = t, \quad m_n \stackrel{\text{def}}{=} \max_{1 \le j \le n} |I_{n,j}| \stackrel{n\to\infty}{\longrightarrow} 0, \tag{7.43}$$

where $|I_{n,j}| = t_{n,j} - t_{n,j-1}$ Let $H = (H_t)_{t \ge 0}$ be a continuous process adapted to $(\mathcal{F}_t)_{t \ge 0}$ such that,

$$\sup_{(s,\omega) \in [0,t] \times \Omega} |H_t(\omega)| \le C < \infty. \tag{7.44}$$

To simplify the notation, we abbreviate

$$X^\mu(t_{n,j}),\ H(t_{n,j}),\ldots \text{ etc. as } X^\mu_{n,j},\ H_{n,j},\ldots \text{ etc.} \tag{7.45}$$

We also abbreviate

$$X^\mu_{n,j} - X^\mu_{n,j-1},\ A^\mu_{n,j} - A^\mu_{n,j-1},\ldots \text{ etc. as } \Delta X^\mu_{n,j},\ \Delta A^\mu_{n,j},\ldots \text{ etc.} \tag{7.46}$$

Then,

Proposition 7.7.1 *Referring to (7.45) and (7.46), The following convergences take place in probability.*

$$\sum_{j=1}^{n} H_{n,j-1}\Delta X^{\mu}_{n,j} \stackrel{n\to\infty}{\longrightarrow} \int_0^t H_s dX^{\mu}_s, \tag{7.47}$$

$$\sum_{j=1}^{n} H_{n,j-1}\Delta X^{\mu}_{n,j}\Delta X^{\nu}_{n,j} \stackrel{n\to\infty}{\longrightarrow} \int_0^t H_s d\langle\, X^{\mu}, X^{\nu}\,\rangle_s, \tag{7.48}$$

Proof of (7.47): Since

$$\sum_{j=1}^{n} H_{n,j-1}\Delta X^{\mu}_{n,j} = \sum_{\alpha=1}^{d}\sum_{j=1}^{n} H_{n,j-1}\int_{I_{n,j}} \sigma^{\mu,\alpha}_s dB^{\alpha}_s + \sum_{j=1}^{n} H_{n,j-1}\Delta A^{\mu}_{n,j},$$

It is enough to prove that for each fixed $\mu = 1,\ldots,m$ and $\alpha = 1,\ldots,d$ that

1) $\displaystyle\sum_{j=1}^{n} H_{n,j-1}\int_{I_{n,j}} \sigma^{\mu,\alpha}_s dB^{\alpha}_s \stackrel{n\to\infty}{\longrightarrow} \int_0^t H_s\sigma^{\mu,\alpha}_s dB^{\alpha}_s$ in probability.

and

2) $\displaystyle\sum_{j=1}^{n} H_{n,j-1}\Delta A^{\mu}_{n,j} \stackrel{n\to\infty}{\longrightarrow} \int_0^t H_s dA^{\mu}_s$ a.s.

We write $\sigma_s = \sigma^{\mu,\alpha}_s$, $B_s = B^{\alpha}_s$ and $A_s = A^{\mu}_s$ in what follows. we define $H^{(n)} \in \mathcal{L}^2$ by

$$H^{(n)}_s = \sum_{j=1}^{n} H_{n,j-1}\mathbf{1}_{I_{n,j}}(s), \quad s \in [0,t].$$

Then,

$$\text{the LHS of 1)} = \int_0^t H^{(n)}_s\sigma_s dB_s, \quad \text{the LHS of 2)} = \int_0^t H^{(n)}_s dA_s.$$

Therefore, the convergence 2) is a simple consequence of the uniform continuity of H_s on the interval $[0,t]$. To see the convergence 1), we introduce, for $m \geq 1$, the stopping time

$$T_m = \inf\left\{ s \geq 0 \quad \int_0^s \sigma_u^2 du \geq m \right\}.$$

Then, for m fixed, the process $\sigma^{(m)}_s = \sigma_s\mathbf{1}_{\{s<T_m\}}$ $(s>0)$ belongs to $\mathcal{L}^2$, hence, by Itô's isometry (7.28) and the dominated convergence theorem,

3) $$E\left[\left|\int_0^t (H^{(n)}_s - H_s)\sigma^{(m)}_s dB_s\right|^2\right] = E\left[\int_0^t (H^{(n)}_s - H_s)^2(\sigma^{(m)}_s)^2 ds\right] \stackrel{n\to\infty}{\longrightarrow} 0.$$

Note on the other hand that, on the event $\{t \leq T_m\}$, $\sigma^{(m)}_s = \sigma_s$ for $s \in [0,t]$. Thus, for arbitrary $\varepsilon > 0$,

$$\begin{aligned} & P\left(\left|\int_0^t (H^{(n)}_s - H_s)\sigma_s dB_s\right| > \varepsilon\right) \\ \leq\ & P\left(\left|\int_0^t (H^{(n)}_s - H_s)\sigma^{(m)}_s dB_s\right| > \varepsilon\right) + P(T_m > t). \end{aligned}$$

By 3), the first probability on the RHS of the above display converges to zero as $n \to \infty$, while the second probability converges to zero as $m \to \infty$. This proves 1). \\(^□^)/

To prove (7.48), we prepare the following lemma.

Lemma 7.7.2 *Suppose that* $(S_j)_{j=0}^n$ $(S_0 = 0)$ *is a martingale w.r.t. a filtration* $(\mathcal{G}_j)_{j=0}^n$ *and* $X_j = S_j - S_{j-1}$.

a) *If* $\{S_j\}_{j=0}^n \subset L^2(P)$*, then* $E[(S_n - S_{j-1})^2|\mathcal{G}_{j-1}] = \sum_{k=j}^n E[X_k^2|\mathcal{G}_{j-1}]$ *for* $j = 1, \ldots, n$.

b) *If* $\{S_j\}_{j=0}^n \subset L^\infty(P)$*, then* $E\left[\left(\sum_{j=1}^n X_j^2\right)^2\right] \leq 12 \max_{0 \leq j \leq n} \|S_j\|_\infty^4$.

Proof: a) We observe that if $1 \leq j \leq k < \ell \leq n$, then,

$$E[X_k X_\ell|\mathcal{G}_{j-1}] = E[X_k E[X_\ell|\mathcal{G}_k]|\mathcal{G}_{j-1}] = 0.$$

Therefore,

$$E[(S_n - S_{j-1})^2|\mathcal{G}_{j-1}] = \sum_{k,\ell=j}^n E[X_k X_\ell|\mathcal{G}_{j-1}] = \sum_{k=j}^n E[X_k^2|\mathcal{G}_{j-1}].$$

b)

1) $$E\left[\left(\sum_{j=1}^n X_j^2\right)^2\right] = \sum_{j,k=1}^n E\left[X_j^2 X_k^2\right] = \sum_{j=1}^n E\left[X_j^4\right] + 2 \sum_{1 \leq j < k \leq n} E\left[X_j^2 X_k^2\right]$$

Let $C_n = \max_{0 \leq j \leq n} \|S_j\|_\infty$. Then, the first summation on the RHS of 1) is bounded from above as follows.

$$\begin{aligned}\sum_{j=1}^n E\left[X_j^4\right] &= \sum_{j=1}^n E\left[(S_j - S_{j-1})^2 X_j^2\right] \\ &\leq 4C_n^2 \sum_{j=1}^n E\left[X_j^2\right] \stackrel{\text{a)}}{=} 4C_n^2 E\left[S_n^2\right] \leq C_n^4.\end{aligned}$$

As for the second summation on the RHS of 1),

$$\begin{aligned}\sum_{1 \leq j < k \leq n} E\left[X_j^2 X_k^2\right] &= \sum_{j=1}^{n-1} E\left[X_j^2 \sum_{k=j+1}^n X_k^2\right] = \sum_{j=1}^{n-1} E\left[X_j^2 E\left[\sum_{k=j+1}^n X_k^2 \middle| \mathcal{G}_j\right]\right] \\ &\stackrel{\text{a)}}{=} \sum_{j=1}^{n-1} E\left[X_j^2 E[(S_n - S_j)^2|\mathcal{G}_j]\right] \leq 4C_n^2 \sum_{j=1}^{n-1} E\left[X_j^2\right] \\ &\stackrel{\text{a)}}{=} 4C_n^2 E\left[S_{n-1}^2\right] \leq 4C_n^4.\end{aligned}$$

\(^□^)/

Proof of (7.48): For notational simplicity, we assume $d = m = 1$ and write accordingly

$$X_t = M_t + A_t \text{ with } M_t = \int_0^t \sigma_s dB_s.$$

We will then prove that

1) $$I_n \stackrel{\text{def}}{=} \sum_{j=1}^n H_{n,j-1}(\Delta X_{n,j})^2 \stackrel{n\to\infty}{\longrightarrow} \int_0^t H_s \sigma_s^2 ds \text{ in probability.}$$

After 1) is established, it is routine to obtain (7.48) in the case where $d \geq 2$ and/or $m \geq 2$.

Case1: We first consider the case of $A_t \equiv 0$. It is clear from the dininition of $\langle M \rangle$ that

$$J_n \stackrel{\text{def}}{=} \sum_{j=1}^n H_{n,j-1}\Delta\langle M \rangle_{n,j} \stackrel{n\to\infty}{\longrightarrow} \int_0^t H_s \sigma_s^2 ds \text{ a.s..}$$

Therefore, it is enough to prove that

2) $$I_n - J_n \stackrel{n\to\infty}{\longrightarrow} 0 \text{ in probability.}$$
To do so, we introduce the stopping times
$$T_\ell = \inf\{t \ge 0\ ;\ |M_t| + \int_0^t \sigma_s^2 ds \ge \ell\}, \quad \ell \ge 1.$$
Then,
$$\begin{aligned} M_t^{(\ell)} &\stackrel{\text{def}}{=} M(t \wedge T_\ell) = \int_0^t \mathbf{1}_{\{s \le T_\ell\}} \sigma_s dB_s, \\ \langle M^{(\ell)} \rangle_t &= \int_0^{t \wedge T_\ell} \sigma_s^2 ds \le \ell. \end{aligned}$$
Since $T_\ell \stackrel{\ell\to\infty}{\longrightarrow} \infty$ a.s., it is enough to prove 2), with M replaced by $M^{(\ell)}$ with large enough ℓ. For this reason, we may and will assume that both $\sup_{s\le t}|M_s|$ and $\langle M \rangle_t$ are bounded by a constant ℓ. Then, by applying Lemma 7.7.2 b) to the martingale $(M_{n,k})_{k=0}^n$, we have

3) $$E\left[\left(\sum_{j=1}^n (\Delta M_{n,j})^2\right)^2\right] \le C,$$
where $C \in (0,\infty)$ is a constant independent of n. On the other hand, let
$$X_{n,j} = (\Delta M_{n,j})^2 - \Delta\langle M \rangle_{n,j} \ (j = 1, \ldots, n).$$
We then define
$$S_{n,0} = 0, \quad S_{n,k} = \sum_{j=1}^k H_{n,j-1} X_{n,j}, \quad k = 1, \ldots, n.$$
It is easy to verify that $(S_{n,k})_{k=0}^n$ is a martingale w.r.t. the filtration $(\mathcal{F}_{n,k})_{k=0}^n$, and hence it follows from Lemma 7.7.2 a) that

4) $$\begin{aligned} E[|I_n - J_n|^2] &= E[S_{n,n}^2] = \sum_{j=1}^n E\left[(H_{n,j-1}X_{n,j})^2\right] \le C \sum_{j=1}^n E\left[X_{n,j}^2\right], \\ &\le 2CE\left[\sum_{j=1}^n (\Delta M_{n,j})^4\right] + 2CE\left[\sum_{j=1}^n (\Delta\langle M \rangle_{n,j})^2\right]. \end{aligned}$$
Therefore, it is enough to show that two expectations on the RHS of 4) converge to zero as $n \to \infty$. The first one is bounded from above as follows.
$$\begin{aligned} E\left[\sum_{j=1}^n (\Delta M_{n,j})^4\right] &\le E[\max_{1\le j\le n} (\Delta M_{n,j})^2 \sum_{j=1}^n (\Delta M_{n,j})^2] \\ &\le E\left[\max_{1\le j\le n} (\Delta M_{n,j})^4\right]^{1/2} E\left[\left(\sum_{j=1}^n (\Delta M_{n,j})^2\right)^2\right]^{1/2} \\ &\stackrel{3)}{\le} C^{1/2} E\left[\max_{1\le j\le n} (\Delta M_{n,j})^4\right]^{1/2}. \end{aligned}$$
By the continuity of M and the bounded convergence theorem, the expectation on the RHS of the above display vanishes as $n \to \infty$. As for the second expectation on the RHS of 4),
$$E\left[\sum_{j=1}^n (\Delta\langle M \rangle_{n,j})^2\right] \le E\left[\max_{1\le j\le n} \Delta\langle M \rangle_{n,j} \langle M \rangle_t\right].$$
By the continuity of $\langle M \rangle$ and the bounded convergence theorem, the expectation on the RHS of the above display vanishes as $n \to \infty$. This finishes the proof of Case1.
Case2: We treat the case of $A_t \not\equiv 0$. We decompose

$$\textbf{5)}\quad \begin{aligned}\sum_{j=1}^n H_{n,j-1}(\Delta X_{n,j})^2 &= \sum_{j=1}^n H_{n,j-1}(\Delta M_{n,j})^2 \\ &\quad +2\sum_{j=1}^n H_{n,j-1}(\Delta M_{n,j})(\Delta A_{n,j}) + \sum_{j=1}^n H_{n,j-1}(\Delta A_{n,j})^2.\end{aligned}$$

By Case1, the first term on the RHS of the above display converges in probability to $\int_0^t H_s d\langle X\rangle_s$. Therefore, it is enough to show that the other terms converge in probability to zero. Since the process H is bounded on $[0,t]$ and the process A is of bounded variation on $[0,t]$, the third term on the RHS of 5) converges a.s. to zero. The second term on the RHS of 5) is bounded by a constant multiple of

$$\sum_{j=1}^n |(\Delta M_{n,j})(\Delta A_{n,j})| \le \left(\sum_{j=1}^n |\Delta M_{n,j}|^2\right)^{1/2} \left(\sum_{j=1}^n |\Delta A_{n,j}|^2\right)^{1/2}.$$

The first summation on the RHS of the above display converges in probability to $\langle X \rangle_t$, while the second summation converges a.s. to zero. Getting things together, we obtain 1). \(^□^)/

Proof of Theorem 7.6.1: Since all the terms in (7.41) is a.s. continuous in t, it is enough to prove the formula a.s. for any fixed t. For $x, x_0 \in \mathbb{R}^m$, let

$$F(x,x_0) = F_1(x,x_0) + F_2(x,x_0),$$

where

$$F_1(x,x_0) = \sum_{\mu=1}^m \frac{\partial f}{\partial x^\mu}(x_0)(x^\mu - x_0^\mu), \quad F_2(x,x_0) = \tfrac{1}{2}\sum_{\mu,\nu=1}^d \frac{\partial^2 f}{\partial x^\mu \partial x^\nu}(x_0)(x^\mu - x_0^\mu)(x^\nu - x_0^\nu).$$

For $\delta, M > 0$, let

$$\rho_2(\delta,M) = \tfrac{1}{2}\sum_{\mu,\nu=1}^d \sup\left\{\left|\frac{\partial^2 f}{\partial x^\mu \partial x^\nu}(x) - \frac{\partial^2 f}{\partial x^\mu \partial x^\nu}(x_0)\right| \ ; \ |x-x_0|\le\delta,\ |x|\vee|x_0|\le M\right\}.$$

By Taylor's theorem, there exists a point x_1 on the line segment between x and x_0 such that,

$$\begin{aligned} f(x)-f(x_0) &= F_1(x,x_0) + \tfrac{1}{2}\sum_{\mu,\nu=1}^\ell \frac{\partial^2 f}{\partial x^\mu \partial x^\nu}(x_1)(x^\mu - x_0^\mu)(x^\nu - x_0^\nu) \\ &= F(x,x_0) \\ &\quad + \tfrac{1}{2}\sum_{\mu,\nu=1}^\ell \left(\frac{\partial^2 f}{\partial x^\mu \partial x^\nu}(x_1) - \frac{\partial^2 f}{\partial x^\mu \partial x^\nu}(x_0)\right)(x^\mu - x_0^\mu)(x^\nu - x_0^\nu).\end{aligned}$$

Therefore, if $|x-x_0|\le\delta$, and $|x|\vee|x_0|\le M$, then,

1) $|f(x)-f(x_0)-F(x,x_0)| \le \rho_2(\delta,M)|x-x_0|^2$

For $t>0$ fixed, let

$$0 = t_{n,0} < t_{n,1} < \ldots < t_{n,n} = t$$

be such that $\max_{1\le j\le n}(t_{n,j}-t_{n,j-1}) \stackrel{n\to\infty}{\longrightarrow} 0$. We write

$$\delta_{n,X} = \max_{1\le j\le n}|\Delta X_{n,j}|, \quad M_X = \sup_{0\le s\le t}|X_s|.$$

Then, it follows from 1) that

2) $$\begin{cases} |f(X_t) - f(X_0) - \sum_{j=1}^n F(X_{n,j-1}, X_{n,j})| \\ \quad \le \sum_{j=1}^n |f(X_{n,j}) - f(X_{n,j-1}) - F(X_{n,j-1}, X_{n,j})| \le \rho_2(\delta_{n,X}, M_X) \sum_{j=1}^n |\Delta X_{n,j}|^2. \end{cases}$$

Since $s \mapsto X_s$ is uniformly continuous on $[0,t]$, we have $\delta_{n,X} \stackrel{n\to\infty}{\longrightarrow} 0$, a.s. Then, since the derivatives of f, which appear in the definition of $\rho_2(\delta, M)$ are uniformly continuous inside the closed ball with radius M_X, we have

3) $\qquad \rho_2(\delta_{n,X}, M_X) \stackrel{n\to\infty}{\longrightarrow} 0$ a.s.

Let us assume for a moment that

4) all the first and second derivatives of f and are bounded.

Then, it follows from (7.47) that

5) $$\begin{cases} \sum_{j=1}^n F_1(X_{n,j-1}, X_{n,j}) &= \sum_{\mu=1}^d \sum_{j=1}^n \frac{\partial f}{\partial x^\mu}(X_{n,j-1})(\Delta X^\mu_{n,j}) \\ &\stackrel{n\to\infty}{\longrightarrow} \sum_{\mu=1}^d \int_0^t \frac{\partial f}{\partial x^\mu}(X_s) dX^\mu_s \ \text{ in probability.} \end{cases}$$

On the other hand, we see from (7.48) that

6) $$\begin{cases} \sum_{j=1}^n F_2(X_{n,j-1}, X_{n,j}) &= \frac{1}{2} \sum_{\mu,\nu=1}^\ell \sum_{j=1}^n \frac{\partial^2 f}{\partial x^\mu \partial x^\nu}(X_{n,j-1})(\Delta X^\mu_{n,j})(\Delta X^\nu_{n,j}) \\ &\stackrel{n\to\infty}{\longrightarrow} \frac{1}{2} \sum_{\mu,\nu=1}^\ell \int_0^t \frac{\partial^2 f}{\partial x^\mu \partial x^\nu}(X_s) d\langle X^\mu_s, X^\nu_s \rangle_s \ \text{ in probability,} \end{cases}$$

and that

7) $$\sum_{j=1}^n |\Delta X_{n,j}|^2 = \sum_{\mu=1}^\ell \sum_{j=1}^n (\Delta X^\mu_{n,j})^2 \stackrel{n\to\infty}{\longrightarrow} \sum_{\mu=1}^\ell \langle X^\mu, X^\mu \rangle_t \ \text{ in probability.}$$

We can take a subsequence, along which the convergences 5),6) and 7) take place a.s. Thus, by letting $n \to \infty$ in 2) along the subsequence, we have (7.41) a.s.

We now get rid of the assumption 4). Let $f_n \in C^2_c(\mathbb{R}^m)$ be such that $f_n(x) = f(x)$ if $|x| \le n+1$. Then, for $t > 0$ fixed,

$$I_n \stackrel{\text{def}}{=} \int_0^t \frac{\partial f_n}{\partial x^\mu}(X_s) dX^\mu_s - \int_0^t \frac{\partial f}{\partial x^\mu}(X_s) dX^\mu_s \stackrel{n\to\infty}{\longrightarrow} 0 \ \text{ in probability.}$$

In fact, since $\{\sup_{s\le t} |X_s| \le n\} \subset \{\frac{\partial f_n}{\partial x^\mu}(X_s) = \frac{\partial f}{\partial x^\mu}(X_s) \text{ for all } s \le t\}$,

$$P(I_n \neq 0) \le P(\sup_{s\le t} |X_s| \ge n) \stackrel{n\to\infty}{\longrightarrow} 0.$$

Similarly,

$$J_n \stackrel{\text{def}}{=} \int_0^t \frac{\partial^2 f_n}{\partial x^\mu \partial x^\nu}(X_s) d\langle X^\mu, X^\nu \rangle_s - \int_0^t \frac{\partial^2 f}{\partial x^\mu \partial x^\nu}(X_s) d\langle X^\mu, X^\nu \rangle_s \stackrel{n\to\infty}{\longrightarrow} 0 \ \text{ in probability.}$$

We can take a subsequence, along which I_n and J_n converge to zero a.s. Thus, by applying (7.41) for f_n, and then by letting $n \to \infty$, we have (7.41) a.s. \(^□^)/

7.8 Girsanov's Theorem and its Applications

Let B be a BM_0^d w.r.t. a filtration $(\mathcal{F}_t)_{t\geq 0}$, (cf. Definition 7.1.1). Recall that the definition of the class of processes $\mathcal{L}^2$ and $\mathcal{L}^2_{\text{a.s.}}$ from Definition 7.3.1. We now define

$$(\mathcal{L}^2_{\text{a.s.}})^d = \{(H^1_t,\ldots,H^d_t)_{t\geq 0}\ ;\ (H^\alpha_t)_{t\geq 0} \in \mathcal{L}^2_{\text{a.s.}} \text{ for all } \alpha = 1,\ldots,d\}.$$

For $H \in (\mathcal{L}^2_{\text{a.s.}})^d$, we define $(\mathcal{D}_t(H))_{t\geq 0}$ by

$$\mathcal{D}_t(H) = \exp\left(\int_0^t H_s \cdot dB_s - \frac{1}{2}\int_0^t |H_s|^2 ds\right), \tag{7.49}$$

where

$$\int_0^t H_s \cdot dB_s = \sum_{\alpha=1}^d \int_0^t H^\alpha_s dB^\alpha_s.$$

Theorem 7.8.1 (Girsanov's theorem) *Suppose that* $H \in (\mathcal{L}^2_{\text{a.s.}})^d$. *Then,*
a) *The following two conditions are equivalent.*
a1) $\mathcal{D}.(H)$ *is a martingale.*
a2) *There exists a measure* $Q \in \mathcal{P}(\Omega,\mathcal{F}_\infty)$ *such that*

$$Q(G) = E[\mathcal{D}_t(H) : G] \ \text{ for all } t > 0 \text{ and } G \in \mathcal{F}_t. \tag{7.50}$$

b) *Assuming a2) above, the following two conditions are equivalent.*
b1) $\mathcal{D}.(H+\theta)$ *is a martingale for each constant vector* $\theta \in \mathbb{R}^d$.
b2) *Under the measure* Q, *the process* B *satisfies the following integral equation,*

$$B_t = W_t + \int_0^t H_s ds \text{ for all } t \geq 0,$$

where W *is a* BM_0^d *and* $\int_0^t H_s ds = \left(\int_0^t H^\alpha_s ds\right)_{\alpha=1}^d$.

Proof: a) a1) $\Rightarrow$ a2): Let $I \subset [0,\infty)$ be a nonempty finite set, and let $\mathcal{F}_I = \{(B_I)^{-1}(H)\ ;\ H \in \mathcal{B}(\mathbb{R}^I)\}$, where the map $B_I : \Omega \to \mathbb{R}^I$ is defined by $B_I = (B_t)_{t\in I}$. Let also Q_I be the measure on $(\Omega,\mathcal{F}_I)$ defined by $Q_I(G) = E\left[\mathcal{D}_t(H) : G\right], \quad G \in \mathcal{F}_I$, where $t \geq \max I$. Since $\mathcal{D}.(H)$ is a martingale, the measure Q_I is independent of the choice of t and it is indeed a probability measure. Moreover, by the construction, the family $\{Q_I\}$ of all such measures are consistent in the following sense. If I and J are nonempty finite sets of $[0,\infty)$ and $I \subset J$, then for all $H \in \mathcal{B}(\mathbb{R}^I)$,

$$Q_J((B_J)^{-1}(H \times \mathbb{R}^{J\setminus I})) = Q_I((B_I)^{-1}(H)).$$

Then, by the Kolmogorov's extension theorem, there exists a uniquue measure $Q \in \mathcal{P}(\Omega,\mathcal{F}_\infty)$ such that for all nonempty finite set $I \subset [0,\infty)$, $Q(G) = Q_I(G), \quad G \in \mathcal{F}_I$. The measure Q satisfies (7.50), since $\mathcal{F}_t \subset \mathcal{F}^0_{t+1}$ for any $t > 0$, and $\mathcal{F}^0_{t+1}$ is generated by $\mathcal{F}_I$'s with $I \subset [0,t+1]$.
a2) $\Rightarrow$ a1): This follows from Example 4.3.2.
b) Let

$$W_t \stackrel{\text{def}}{=} B_t - \int_0^t H_s ds, \quad g_\theta(x,t) = \exp(\theta\cdot x - t|\theta|^2/2) \ (\theta, x \in \mathbb{R}^d,\ t > 0).$$

Then,

$$
\begin{aligned}
\mathcal{D}_t(H+\theta) &= \exp\left(\int_0^t H_s\cdot dB_s + \theta\cdot B_t - \frac{1}{2}\int_0^t |H_s|^2 ds - \int_0^t \theta\cdot H_s ds - \frac{t|\theta|^2}{2}\right)\\
&= \exp\left(\int_0^t H_s\cdot dB_s - \frac{1}{2}\int_0^t |H_s|^2 ds + \theta\cdot W_t - \frac{t|\theta|^2}{2}\right)\\
&= \mathcal{D}_t(H)g_\theta(W_t,t).
\end{aligned}
$$

Thus b1) is equivalent to

1) $E[\mathcal{D}_t(H)g_\theta(W_t,t):G] = E[\mathcal{D}_s(H)g_\theta(W_s,s):G]$ for all $0\le s<t$ and $G\in\mathcal{F}_s$.

By (7.50), 1) is equivalent to

2) $E^Q[g_\theta(W_t,t):G] = E^Q[g_\theta(W_s,s):G]$ for all $0\le s<t$ and $G\in\mathcal{F}_s$.

By Proposition 7.1.2, 2) is equivalent to that W_t is a BM_0^d under the measure Q, and this is equivalent to b2). \(^□^)/

As a special case of Theorem 7.8.1, where the process H is nonrandom, we obtain the following

Corollary 7.8.2 *Let $h\in C^1([0,\infty)\to\mathbb{R}^d)$. For a BM_0^d denoted by B, we set*

$$\mathcal{D}_t(h) = \exp\left(\int_0^t h(s)\cdot dB_s - \frac{1}{2}\int_0^t |h(s)|^2 ds\right).$$

Then,

a) *There exists a measure $Q\in\mathcal{P}(\Omega,\mathcal{F}_\infty)$ such that*

$$Q(G) = E\left[\mathcal{D}_t(h):G\right] \quad \text{for all } t>0 \text{ and } G\in\mathcal{F}_t. \tag{7.51}$$

b) *Under the measure Q,*

$$B_t = W_t + \int_0^t h(s)ds \quad \text{for all } t\ge 0, \tag{7.52}$$

where W is a BM_0^d and $\int_0^t h(s)ds = \left(\int_0^t h^\alpha(s)ds\right)_{\alpha=1}^d$.

Proof: Let $\theta\in\mathbb{R}^d$ be arbitrary constant vector. Then, by applying Exercise 7.4.1 to $h+\theta$, we see that $\mathcal{D}_\cdot(h+\theta)$ is a martingale. Then, this corollary follows from Theorem 7.8.1. \(^□^)/

Corollary 7.8.3 *Let* $\varphi \in C^2(\mathbb{R}^d \to \mathbb{R})$ *and suppose that there exists* $C \in [0,\infty)$ *such that*

$$\varphi(x) \le C(1+|x|), \quad \Delta\varphi(x) \ge -C(1+|x|).$$

for all $x \in \mathbb{R}^d$. *For a* BM_0^d *denoted by* B, *we set*

$$\mathcal{D}_t(\varphi) = \exp\left(\int_0^t \nabla\varphi(B_s)\cdot dB_s - \frac{1}{2}\int_0^t |\nabla\varphi(B_s)|^2 ds\right).$$

Then,

a) *There exist a measure* $Q \in \mathcal{P}(\Omega, \mathcal{F}_\infty)$ *such that*

$$Q(G) = E\left[\mathcal{D}_t(\varphi) : G\right] \text{ for all } t>0 \text{ and } G \in \mathcal{F}_t.$$

b) *Under the measure* Q, *the process* B *satisfies the following integral equation.*

$$B_t = W_t + \int_0^t \nabla\varphi(B_s)ds \text{ for all } t \ge 0,$$

where W *is a* BM_0^d.

c) *For* $t>0$, *set* $A_t = \frac{1}{2}\int_0^t \left(|\nabla\varphi(B_s)|^2 + \Delta\varphi(B_s)\right) ds$. *Then, for all measurable* F : $(\mathbb{R}^d)^{[0,t]} \to [0,\infty)$,

$$E\left[\exp\left(\varphi(B_t)-\varphi(0)-A_t\right)F(B)\right] = E^Q\left[F(B)\right], \tag{7.53}$$

$$E\left[\exp\left(-A_t\right)F(B)\right] = E^Q\left[\exp\left(\varphi(0)-\varphi(B_t)\right)F(B)\right]. \tag{7.54}$$

Proof: a),b): By applying Exercise 7.4.2 to the function $\varphi(x)+\theta\cdot x$, we see that the process $H_t = \nabla\varphi(B_t)$, $t \ge 0$ satisfies the condition b1) of Theorem 7.8.1. Thus the assertions a) and b) of this corollary follows from Theorem 7.8.1.

c) It follows from (7.50) that

$$E\left[\exp\left(\varphi(B_t)-\varphi(0)-A_t\right)F(B)\right] = E\left[\mathcal{D}_t(\varphi)F(B)\right] = E^Q\left[F(B)\right].$$

Replacing $F(B)$ by $\exp\left(\varphi(0)-\varphi(B_t)\right)F(B)$, we obtain (7.54). \(^□^)/

Example 7.8.4 Let B be a BM_0^d. Then, for any $t>0$, $tht \in \mathbb{R}$ and measurable function $f : \mathbb{R}^d \to [0,\infty)$,

$$E\left[\exp\left(-\frac{\theta^2}{2}\int_0^t |B_s|^2 ds\right) f(B_t)\right] = \cosh(\theta t)^{-d/2} E\left[f(\tau(t)^{1/2}X)\right]$$

(**Cameron-Martin formula II**),

where X is a r.v. with d-dimensional standard normal distribution and $\tau(t) = \tanh(\theta t)/\theta$.

Proof: Since the both-hand sides of the equality to be shown are even in θ, it is enough to prove it when $\theta > 0$. Let $\varphi(x) = -\frac{\theta}{2}|x|^2 \le 0$ ($x \in \mathbb{R}^d$). Then, $\nabla\varphi(x) = -\theta x$, $\Delta\varphi(x) = -\theta d$. Thus, by applying (7.54),

$$E\left[\exp\left(-\frac{\theta^2}{2}\int_0^t |B_s|^2 ds + \frac{d\theta t}{2}\right) f(B_t)\right] = E^Q\left[\exp\left(\frac{\theta}{2}|B_t|^2\right) f(B_t)\right],$$

and hence

1) $$E\left[\exp\left(-\frac{\theta^2}{2}\int_0^t |B_s|^2 ds\right) f(B_t)\right] = \exp\left(-\frac{d\theta t}{2}\right) E^Q\left[\exp\left(\frac{\theta}{2}|B_t|^2\right) f(B_t)\right].$$

The process B under the measure Q satisfies the following integral equation.

$$B_t = W_t + \theta\int_0^t B_s ds,$$

where W is a BM_0^1. This integral equation can be solved w.r.t.X, which gives

$$B_t = W_t - \theta\exp(-\theta t)\int_0^t \exp(\theta s) W_s ds.$$

Then, it follows from the above expression and Exercise 6.1.5 that B_t is a mean-zero Gaussian r.v. such that

$$\mathrm{cov}^Q(B_t^\alpha, B_t^\beta) = \sigma(t)\delta_{\alpha,\beta} \text{ with } \sigma(t) = \frac{1-\exp(-2\theta t)}{2\theta}.$$

Note that

$$1/(\sigma(t)^{-1}-\theta) = \tanh(\theta t)/\theta = \tau(t) \text{ and } \sigma(t)/\tau(t) = \exp(-\theta t)\cosh(\theta t).$$

Therefore,

$$\begin{aligned}
E^Q\left[\exp\left(\frac{\theta}{2}|B_t|^2\right) f(B_t)\right] &= (2\pi\sigma(t))^{-d/2}\int_{\mathbb{R}^d} \exp\left(-\frac{1}{2}\left(\sigma(t)^{-1}-\theta\right)|x|^2\right) f(x)dx \\
&= (\sigma(t)/\tau(t))^{-d/2}(2\pi\tau(t))^{-d/2}\int_{\mathbb{R}^d} \exp\left(-\frac{|x|^2}{2\tau(t)}\right) f(x)dx \\
&= \exp\left(\frac{d\theta t}{2}\right)\cosh(\theta t)^{-d/2} E\left[f(\tau(t)^{1/2}X)\right].
\end{aligned}$$

Plugging this into 1), we obtain the desired equality. \(^□^)/

7.9 The DDS Representation Theorem

In what follows, we let B denotes a BM_0^d w.r.t. a filtration $(\mathcal{F}_t)_{t\geq 0}$, cf. Definition 7.1.1.

Proposition 7.9.1 *Let $M_t = \int_0^t \sigma_s \cdot dB$ be a local martingale generated by B (Definition 7.5.1), where $\sigma \in (\mathcal{L}^2_{\text{a.s.}})^d$. Then,*

a)

$$\begin{aligned}
\mathcal{D}_t &\stackrel{\text{def}}{=} \exp\left(M_t - \tfrac{1}{2}\langle M\rangle_t\right) = 1 + \int_0^t \mathcal{D}_s\sigma_s \cdot dB_s, \\
\mathcal{E}_t &\stackrel{\text{def}}{=} \exp\left(\mathbf{i}M_t + \tfrac{1}{2}\langle M\rangle_t\right) = 1 + \mathbf{i}\int_0^t \mathcal{E}_s\sigma_s \cdot dB_s.
\end{aligned}$$

In particular, $\mathcal{D}_t$ and $\mathcal{E}_t$ are local martingales.

b) *Suppose that there exists $t_0 > 0$ such that*

$$E\exp\left(\sup_{0\leq s\leq t_0}|M_s|\right) < \infty.$$

Then, $(\mathcal{D}_t)_{0\leq t\leq t_0}$ is a martingale.

c) *Suppose that there exists $t_0 > 0$ such that*

$$E\exp\left(\frac{1}{2}\langle M\rangle_{t_0}\right) < \infty \quad \textbf{(Novikov's condition)}.$$

Then, $(\mathcal{E}_t)_{0\leq t\leq t_0}$ is a martingale.

Proof: a) To prove the first equality, we apply Itô's formula II to a function $f(x,y) = \exp(x - \frac{1}{2}y)$ of $(x,y) \in \mathbb{R}^2$, and the process $(M_t, \langle M\rangle_t)$. Then,

$$\begin{aligned}
\mathcal{D}_t &= 1 + \int_0^t \frac{\partial f}{\partial x}(M_s, \langle M\rangle_s)dM_s + \int_0^t \frac{\partial f}{\partial y}(M_s, \langle M\rangle_s)d\langle M\rangle_s \\
&+ \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(M_s, \langle M\rangle_s)d\langle M\rangle_s \\
&= 1 + \int_0^t \mathcal{D}_s dM_s - \frac{1}{2}\int_0^t \mathcal{D}_s d\langle M\rangle_s + \frac{1}{2}\int_0^t \mathcal{D}_s d\langle M\rangle_s \\
&= 1 + \int_0^t \mathcal{D}_s\sigma_s \cdot dB_s.
\end{aligned}$$

The proof of the second equality is similar.

b) Since $\mathcal{D}_t$ is a local martingale, it is enough to verify the condition of Exercise 7.3.1. Note that $\langle M\rangle_t \geq 0$, and hence

$$\mathcal{D}_t \le \exp\left(\sup_{0\le s\le t_0} |M_s|\right) \in L^1(P).$$

Therefore, by Exercise 7.3.1, $(\mathcal{D}_t(M))_{0\le t\le t_0}$ is a martingale.
c) Since $\mathcal{E}_t$ is a local martingale, it is enough to verify the condition of Exercise 7.3.1. For $t \le t_0$,

$$|\mathcal{E}_t| = \exp\left(\frac{1}{2}\langle M \rangle_t\right) \le \exp\left(\frac{1}{2}\langle M \rangle_{t_0}\right) \in L^1(P).$$

Therefore, by Exercise 7.3.1, $(\mathcal{E}_t)_{0\le t\le t_0}$ is a martingale. \(^□^)/

Remark: It is known that Novikov's condition also implies that $(\mathcal{E}_t(M))_{0\le t\le t_0}$ is a martingale.

Setting 7.9.2 Let $M_t^\mu = \int_0^t \sigma_s^\mu \cdot dB$ ($\mu = 1,\ldots,m$) be local martingales generated by B (Definition 7.5.1), where $\sigma^\mu \in (\mathcal{L}^2_{\text{a.s.}})^d$ ($\mu = 1,\ldots,m$). We consider the process $M_t = (M_t^\mu)_{\mu=1}^d$, $t \ge 0$ with values in $\mathbb{R}^m$.

Let M be defined in Setting 7.9.2. Then, for $\theta \in \mathbb{R}^m$, the inner product $\theta \cdot M_t$ is again a local martingales generated by B. Applying Proposition 7.9.1 to $\theta \cdot M_t$, we obtain the following

Corollary 7.9.3 *Let M be defined in Setting 7.9.2 and $\theta \in \mathbb{R}^m$. Then, the following processes are local martingales generated by B.*

$$\exp\left(\theta \cdot M_t - \tfrac{1}{2}\langle \theta \cdot M \rangle_t\right), \quad \exp\left(\mathbf{i}\theta \cdot M_t + \tfrac{1}{2}\langle \theta \cdot M \rangle_t\right).$$

By combining Proposition 7.1.2, Proposition 7.1.3, and Corollary 7.9.3, we obtain

Corollary 7.9.4 (Lévy's chracterization of the Brownian motion) *Let M be defined in Setting 7.9.2. Then, the following conditions are equivalent.*
a) *M is a BM_0^m;*
b) *$(M_t^\mu M_t^\nu - \delta_{\mu,\nu} t)_{t\ge 0}$ is a local martingale for all $\mu,\nu = 1,\ldots,m$;*
c) *$\{\sigma_t^\mu\}_{\mu=1}^m$ are a.s. orthonormal ($\sigma_t^\mu \cdot \sigma_t^\nu = \delta_{\mu,\nu}$, $\mu,\nu = 1,\ldots,m$) for all $t > 0$.*

Proof: **Hint**: a) $\Rightarrow$ b):This follows from Proposition 7.1.3.
b) $\Rightarrow$ c):Suppose b). Then, it follows from Lemma 7.5.2 that $\langle M^\mu, M^\nu \rangle_t = \delta_{\mu,\nu} t$, for all $\mu,\nu = 1,\ldots,m$, which implies c).
c) $\Rightarrow$ a): It follows from the condition c) that $\langle M^\mu, M^\nu \rangle_t = \delta_{\mu\,\nu} t$, and hence by Corollary 7.9.3, $\exp\left(\mathbf{i}\theta \cdot M_t + \frac{t|\theta|^2}{2}\right)$ is a martingale. Thus, a) follows from Proposition 7.1.2. \(^□^)/

Example 7.9.5 (Bessel process) For a BM_0^d, denoted by B, the following process is a BM_0^1.

$$B_t^+ = \int_0^t |B_s|^{-1} B_s \cdot dB_s,\ t \ge 0.$$

Moreover, for $d \ge 2$, $p > 0$, and $t \ge 0$,

$$|B_t|^p = p\int_0^t |B_s|^{p-2} B_s \cdot dB_s + \frac{p(d+p-2)}{2}\int_0^t |B_s|^{p-2} ds, \tag{7.55}$$

$$\sigma(|B_s|\ ;\ s \le t) = \sigma(B_s^+\ ;\ s \le t). \tag{7.56}$$

Proof: Since the process $(|B_t|^{-1}B_t)_{t\ge 0} \in (\mathcal{L}^2_{\text{a.s.}})^d$ cosists of unit vectors, it follows from Corollary 7.9.4 that B^+ is a BM_0^1. We next turn to (7.55). We first verify that two integrals on the RHS are well-defined. Indeed, it follows from Exercise 6.1.4 that $|B_\cdot|^{p-2} B_\cdot^\alpha \in L^2([0,t] \times \Omega)$

($\alpha = 1,2$) and $|B_\cdot|^{p-2} \in L^1([0,t]\times\Omega)$. Therefore, the stochastic integral $\int_0^t |B_s|^{p-2}B_s \cdot dB_s$ and the integral $\int_0^t |B_s|^{p-2}ds$ exists. We would like to apply Itô's formula to conclude (7.55). However, for $p<2$, the function $|x|^p$ fails to be twice differentiable at $x=0$. To circumvent this obstacle, we fix $0<b<a<\infty$ and define

$$S_a = \inf\{t>0\ ;\ |B_t|\ge a\},\quad T_{a,b} = \inf\{t>0\ ;\ |B(t+S_a)|\le b\}.$$

Then, $(\mathcal{F}_{S_a+\cdot})$ is a filtration w.r.t which $B(S_a+\cdot)$ is a BM^d, and $T_{a,b}$ is a stopping time. Note that, outside the closed ball $|x|\le b$, the function $|x|^p$ is smooth. We apply Itô's formula (Theorem 7.4.1) to this function and the stopped Brownian motion $(B(S_a+t\wedge T_{a,b}))_{t\ge0}$ to obtain

1)
$$\begin{aligned}&|B(S_a+t\wedge T_{a,b})|^p\\ &= |a|^p + p\int_{S_a}^{S_a+t\wedge T_{a,b}} |B_s|^{p-2}B_s\cdot dB_s + \frac{p(d+p-2)}{2}\int_{S_a}^{S_a+t\wedge T_{a,b}} |B_u|^{p-2}du,\quad t\ge0.\end{aligned}$$

Then, we see from 1) with $p=1$ and Lemma 7.9.8 that

2) $\sigma(|B(S_a+s\wedge T_{a,b})|\ ;\ s\le t) = \sigma\left(B^+(S_a+s\wedge T_{a,b}) - B^+(S_a)\ ;\ s\le t\right)$.

We now let b tend to zero. then, $T_{a,b}\stackrel{b\to0}{\longrightarrow}\infty$ a.s. Consequently, it follows from 1) and 2) that

$$|B(S_a+t)|^p = |a|^p + p\int_{S_a}^{S_a+t} |B_s|^{p-2}B_s\cdot dB_s + \frac{p(d+p-2)}{2}\int_{S_a}^{S_a+t}|B_u|^{p-2}du,\quad t\ge0,$$
$$\sigma(|B(S_a+s)|\ ;\ s\le t) = \sigma\left(B^+(S_a+s)-B^+(S_a)\ ;\ s\le t\right).$$

Then, by letting a tend to zero, and noting that $S_a\stackrel{a\to0}{\longrightarrow}0$ a.s., we obtain (7.55) and (7.56). \(^□^)/

Finally, we present the following representation theorem due to Dambis, Dubins, Schwartz (for $m=1$) and Knight ($m\ge2$).

Proposition 7.9.6 (The DDS Representation Theorem) *Referring to Setting 7.9.2, suppose that for all $\mu,\nu=1,\ldots,m$,*

$$\sigma^\mu_t\cdot\sigma^\nu_t = 0 \ \text{ a.s. for } t>0, \tag{7.57}$$
$$\int_0^\infty |\sigma^\mu_s|^2ds = \infty \text{ a.s.} \tag{7.58}$$

Then, there exist m independdent BM^1_0's denoted by W^μ, ($\mu=1,\ldots,m$) such that for all $\mu=1,\ldots,m$ and $t\ge0$,

$$M^\mu_t = W^\mu(\langle M^\mu\rangle_t),\quad \text{where } \langle M^\mu\rangle_t = \int_0^t|\sigma^\mu_s|^2ds. \tag{7.59}$$

More precisely, W^μ, ($\mu=1,\ldots,m$) are defined as follows.

$$W^\mu_t = M^\mu_{T^\mu(t)},\quad \text{where } T^\mu(t) = \inf\{s\ge0\ ;\ \langle M^\mu\rangle_s>t\}. \tag{7.60}$$

Proof: Step1: We first prove that the process W^μ defined by (7.60) is continuous and satisfies (7.59). The coordinate μ is fixed throughtout Step1 and hence is dropped from the notation. We write $A_t = \langle M\rangle_t$ for simplify the notation. Define

$$S(t) = \inf\{s\ge0\ ;\ \langle M\rangle_s\ge t\}\ \text{ and }\ T(t) = \inf\{s\ge0\ ;\ \langle M\rangle_s>t\}.$$

Then, they have the following properties.

1) $S(\cdot)$ (resp. $T(\cdot)$) is left-continuous (resp. right-continuous).

2) For all $t \ge 0$, $S(t) \le T(t)$, $\langle M \rangle_{S(t)} = \langle M \rangle_{T(t)} = t$, $S(t) = T(t-)$. Moreover, $S(\langle M \rangle_t) = T(\langle M \rangle_t) = t$, since $\langle M \rangle_\cdot$ is continuous.

3) For all $t \ge 0$, $S(t)$ and $T(t)$ are stopping times.

We have $W_{\langle M \rangle_t} = M_{T(\langle M \rangle_t)} = M_t$ by 2). Thus, it only remains to prove that W is continuous. W is right-continuous, because of the right-continuity of T. Its left-continuity can be seen as follows. Let $t > 0$. Then, $\langle M \rangle_{S^\mu(t)} = \langle M \rangle_{T(t)} = t$ by 2). This implies, via Exercise 7.3.2 that $M_{S^\mu(t)} = M_{T(t)}$, and hence

$$W_{t-} = M_{T(t-)} = M_{S(t)} = M_{T(t)} = W_t.$$

Step2: We next prove that W^μ ($\mu = 1, \ldots, m$) are independent BM^1's. By Step1, the process W^μ is continuous each $\mu = 1, \ldots, m$. Therefore, it is enough to show that for each fixed $1 \le \mu < \nu \le m$, the process (W^μ, W^ν) is a BM^2_0. Thus we assume henceforce that $m = 2$ and set $A_t = \langle M^1 \rangle_t \vee \langle M^2 \rangle_t$. Then, it is not difficult to see that

4) $$T(t) \stackrel{\text{def}}{=} T^1(t) \vee T^2(t) = \inf\{s \ge 0 \; ; \; A_s > t\}.$$

Then, $W = (W^1, W^2)$ is continuous and adapted to the filtration $(\mathcal{F}_{T(\cdot)})$. Therefore, by Proposition 7.1.2, it is enough to prove that, for all $\theta \in \mathbb{R}^2$,

5) $\mathcal{E}_t(\theta) \stackrel{\text{def}}{=} \exp\left(\theta \cdot W_t + \frac{1}{2}|\theta|^2 t\right)$, $t \ge 0$ is an $(\mathcal{F}_{T(\cdot)})$-martingale.

For $s \ge 0$, A_s is an $(\mathcal{F}_{T(\cdot)})$-stopping time. Moreover, it follows from 4) that $T(t \wedge A(s)) = T(t) \wedge s$. Therefore, for $\mu = 1, 2$,

$$W^\mu(t \wedge A(s)) = \sum_{\alpha=1}^d \int_0^{T(t)\wedge s} \sigma_u^{\mu,\alpha} dB_u^\alpha = \sum_{\alpha=1}^d \int_0^s \mathbf{1}_{\{u \le T(t)\}} \sigma_u^{\mu,\alpha} dB_u^\alpha.$$

The above display shows that, with $t \ge 0$ fixed, the process $N_s^{\mu,t} \stackrel{\text{def}}{=} W^\mu(t \wedge A(s))$, $s \ge 0$ is an $(\mathcal{F}_\cdot)$-local martingale generated by B with the quadratic variation

$$\begin{aligned}\langle N^{\mu,t}, N^{\nu,t} \rangle_s &= \int_0^s \mathbf{1}_{\{u \le T(t)\}} \sigma_u^\mu \cdot \sigma_u^\nu du = \delta_{\mu,\nu} \int_0^s \mathbf{1}_{\{u \le T(t)\}} |\sigma_u^\mu|^2 du \\ &= \delta_{\mu,\nu} A(T(t) \wedge s) = \delta_{\mu,\nu}(t \wedge A(s)).\end{aligned}$$

Hence,

$$\mathbf{i}\theta \cdot W(t \wedge A(s)) + \frac{1}{2}|\theta|^2 (t \wedge A(s)) = \mathbf{i} \sum_{\mu=1}^2 \theta_\mu N_s^{\mu,t} + \frac{1}{2} \sum_{\mu=1}^2 \theta_\mu \theta_\nu \langle N^{\mu,t}, N^{\nu,t} \rangle_s$$

It follows from the above display and Corollary 7.9.3 that

6) $$\mathcal{E}_{t \wedge A(s)}(\theta), \; s \ge 0$$

is an $(\mathcal{F}_\cdot)$-local martingale. Moreover, since $|\mathcal{E}_{t \wedge A(s)}(\theta)| \le \exp(|\theta|^2 t/2)$, the process 5) is a bounded $(\mathcal{F}_\cdot)$-martingale. Therefore, by applying the optional stopping theorem to the martingale 6) and the pair $T(s) \le T(t)$ of stopping times, we obtain

$$E[\mathcal{E}_t(\theta) | \mathcal{F}_{T(s)}] = E[\mathcal{E}_{t \wedge A(T(t))}(\theta) | \mathcal{F}_{T(s)}] = \mathcal{E}_{t \wedge A(T(s))}(\theta) = \mathcal{E}_s(\theta),$$

which proves 5). \(^□^)/

Example 7.9.7 (Stochastic area, revisited) Let B be a BM^2_0 and

$$\mathcal{A}_t^{(p)} = \int_0^t |B|^{p-2} (B_s^2 dB_s^1 - B_s^1 dB_s^2), \quad p > 0.$$

In particular, $\mathcal{A}^{(2)}$ is the stochastic area (Exercise 7.6.1). Then, there exists a BM^1_0, denoted by X such that

$$\mathcal{A}_t^{(p)} = X\left(\int_0^t |B_s|^{2p-2} ds\right), \; t \ge 0.$$

Moreover, for $\theta \in \mathbb{R}$,

$$E\left[\exp\left(\mathbf{i}\theta\mathcal{A}_t^{(p)}\right)|\mathcal{F}_t^{|B|}\right] = \exp\left(-\frac{\theta^2}{2}\int_0^t |B_s|^{2p-2}ds\right), \tag{7.61}$$

$$E\left[\exp\left(\mathbf{i}\theta\mathcal{A}_t^{(2)}\right)\right] = \cosh(\theta t)^{-1}, \tag{7.62}$$

where $\mathcal{F}_t^{|B|} = \sigma(|B_s|\ ;\ s \le t)$.

Proof: Let $B_t^+ = \int_0^t |B_s|^{-1}(B_s^1 dB_s^1 + B_s^2 dB_s^2)$, $t \ge 0$. Then,

$$\langle\, \mathcal{A}^{(p)}\, \rangle_t = \int_0^t |B_s|^{2p-2}ds,\ \langle\, B^+\, \rangle_t = t, \text{ and } \langle\, \mathcal{A}^{(p)}, B^+\, \rangle_t = 0.$$

Therefore, by Proposition 7.9.6, there exist independent BM_0^1, denoted by X and Y such that

$$\mathcal{A}_t^{(p)} = X\left(\int_0^t |B_s|^{2p-2}ds\right) \text{ and } B_t^+ = Y_t,\ t \ge 0.$$

In particular, X is independent of B^+. On the other hand, we know from Example 7.9.5 that

$$\sigma(B_s^+\ ;\ s \le t) = \mathcal{F}_t^{|B|}.$$

Therefore X is independent of $|B|$. As a consequence,

$$\begin{aligned} E\left[\exp\left(\mathbf{i}\theta\mathcal{A}_t^{(p)}\right)|\mathcal{F}_t^{|B|}\right] &= E\left[\exp\left(\mathbf{i}\theta X\left(\int_0^t |B_s|^{2p-2}ds\right)\right)|\mathcal{F}_t^{|B|}\right] \\ &= \exp\left(-\frac{\theta^2}{2}\int_0^t |B_s|^{2p-2}ds\right). \end{aligned}$$

This proves (7.61). For $p = 2$, by taking the expectation of both-hand sides of the above display and recalling Example 7.8.4, we obtain (7.62). \(^□^)/

Remark See Exercise 7.9.1 for a generalization of (7.62).

Complement

Lemma 7.9.8 *Suppose that X and Y are continuous process with values in $\mathbb{R}^d$, that $b : \mathbb{R}^d \to \mathbb{R}^d$ is a bounded Lipchitz continuous function such that*

$$(*) \qquad Y_t = X_t + \int_0^t b(Y_s)ds \text{ for all } t > 0.$$

Then, $\mathcal{F}_t^X = \sigma(X_s\ ;\ s \le t)$ and $\mathcal{F}_t^Y = \sigma(Y_s\ ;\ s \le t)$ are the same for all $t > 0$.

Proof: Since $X_t = Y_t + \int_0^t b(Y_s)ds$, it is obvious that $\mathcal{F}_t^X \subset \mathcal{F}_t^Y$. The opposite inclusion can be shown by express the process Y as a limit of Picard approximation as follows. Let $Y_t^{(0)} = X_t$, $t \ge 0$, and for $n \ge 1$,

$$Y_t^{(n)} = X_t + \int_0^t b(Y_s^{(n-1)})ds, \quad t \ge 0.$$

Then, by induction, it is easy to see that there exists a constant C such that

$$\sup_{s \le t} |Y_s^{(n)} - Y_s^{(n-1)}| \le \frac{(Ct)^n}{n!},$$

which implies that the processes $Y^{(n)}$ converge locally uniformly, and hence that the limit, say $\widetilde{Y}$, solves the equation $(*)$. Then, $Y = \widetilde{Y}$, since the soltion to the equation $(*)$ is unique, as can easily be seen from the Gronwall inequality. Since $\mathcal{F}_t^{\widetilde{Y}} \subset \mathcal{F}_t^X$ by the way $\widetilde{Y}$ is obtained, it follows that $\mathcal{F}_t^Y = \mathcal{F}_t^{\widetilde{Y}} \subset \mathcal{F}_t^X$. \(^□^)/

Exercise 7.9.1 Let $f : \mathbb{R}^2 \to \mathbb{R}$ be bounded and Borel measurable. Prove the following. **i)** Suppose that a function $F : (\mathbb{R}^2)^{[0,\infty)} \to \mathbb{R}$ satisfies the following properties. $F(B) \in L^1(P)$ for each BM_0^2 denoted by B, and that $F(R(\alpha)B) = F(B)$ a.s. for all $\alpha \in \mathbb{R}$, where $R(\alpha) = \left(\begin{smallmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{smallmatrix}\right)$. Then,

$$E\left[F(B)f(B_t)\right] = E\left[E[F(B)|\mathcal{F}_t^{|B|}]\widetilde{f}(B_t)\right],$$

where $\mathcal{F}_t^{|B|} = \sigma(|B_s| \ ; \ s \le t)$ and $\widetilde{f}(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(R(\alpha)x)d\alpha$. **Hint**: For all $\alpha \in \mathbb{R}$,

$$E\left[F(B)f(B_t)\right] = E\left[F(R(-\alpha)B)f(B_t)\right] = E\left[F(B)f(R(\alpha)B_t)\right].$$

Hence, $E\left[F(B)f(B_t)\right] = E\left[F(B)\widetilde{f}(B_t)\right]$. Moreover, $\widetilde{f} \circ R(\alpha) = \widetilde{f}$. **ii)** The formula (7.62) can be generalized as follows.

$$E\left[\exp\left(\mathbf{i}\theta\mathcal{A}_t^{(2)}\right) f(B_t)\right] = \cosh(\theta t)^{-1} E\left[f(\tau(t)^{1/2}X)\right],$$

where X is a r.v. with 2-dimensional standard normal distribution and $\tau(t) = \tanh(\theta t)/\theta$.

8 Appendix to Section 1

8.1 Some Fundamental Inequalities

Proposition 8.1.1 (**Hölder's inequality**) *Suppose that* $(S, \mathcal{A}, \mu)$ *is a measure space, and that* $p, q \in (1, \infty)$, $\frac{1}{p} + \frac{1}{q} = 1$. *Then, for* $f \in L^p(\mu)$ *and* $g \in L^q(\mu)$,

$$\int_S |fg| d\mu \le \left(\int_S |f|^p d\mu \right)^{1/p} \left(\int_S |g|^q d\mu \right)^{1/q}. \tag{8.1}$$

Proof: We recall that for $s, t \ge 0$,

1) $$st \le \frac{s^p}{p} + \frac{t^q}{q}.$$

Thus, for $\varepsilon > 0$,

2) $$\frac{|fg|}{(\|f\|_p + \varepsilon)(\|g\|_q + \varepsilon)} \overset{1)}{\le} \frac{|f|^p}{p(\|f\|_p + \varepsilon)^p} + \frac{|g|^q}{q(\|g\|_q + \varepsilon)^q}.$$

Therefore,

$$\frac{1}{(\|f\|_p + \varepsilon)(\|g\|_q + \varepsilon)} \int_S |fg| d\mu \overset{2)}{\le} \frac{\|f\|_p^p}{p(\|f\|_p + \varepsilon)^p} + \frac{\|g\|_q^q}{q(\|g\|_q + \varepsilon)^q} \le 1.$$

Multiplying the both hands sides of the above inequality by $(\|f\|_p + \varepsilon)(\|g\|_q + \varepsilon)$, and letting $\varepsilon \to 0$, we get (8.1). \(^□^)/

Proposition 8.1.2 (**Jensen's inequality**) *Let* $I \subset \mathbb{R}$ *be an open interval and* $\varphi : I \to \mathbb{R}$ *be convex. Suppose that* X *be a r.v. with values in* I *such that* $X, \varphi(X) \in L^1(P)$. *Then,*

$$\varphi(EX) \le E[\varphi(X)]. \tag{8.2}$$

Proof: Let $m = EX$. As is well known, for $y \in I$, the limit

$$\varphi'_+(y) \overset{\text{def}}{=} \lim_{\substack{h \to 0 \\ h > 0}} \frac{\varphi(y+h) - \varphi(y)}{h}$$

exists and is non decreasing in y. Moreover,

$$\varphi(x) \ge \varphi(y) + \varphi'_+(y)(x - y), \quad \text{for all } x, y \in I.$$

Thus,

$$\varphi(X) \ge \varphi(m) + \varphi'_+(m)(X - m), \quad \text{a.s.}$$

By taking the expectation, we have that

$$E[\varphi(X)] \ge \varphi(m) + \varphi'_+(m)(EX - m) = \varphi(m).$$

\(^□^)/

8.2 Polar Decomposition of a Matrix

Notation:

- $\mathcal{S}_d^+$ denotes the totality of symmetric, non-negative definite $d \times d$ real matrices.
- $\mathcal{O}_d$ denotes the totality of $d \times d$ real orthogonal matrices.
- For a real matrix A, A^* denotes its transposition.

We recall that for a symmetric, $d \times d$ real matrix S, there exists a $U \in \mathcal{O}_d$ such that

$$SU = UD(s_1, \ldots, s_d), \tag{8.3}$$

where $s_1 \geq ... \geq s_d$ are eigenvalues of S and $D(s_1, \ldots, s_d) = (s_\alpha \delta_{\alpha,\beta})_{\alpha,\beta=1}^d$. Let $u_1, \ldots, u_d$ be column vectors of U, so that $U = (u_1, \ldots, u_d)$. Then, (8.3) reads

$$Su_\alpha = s_\alpha u_\alpha, \quad \alpha = 1, \ldots, d. \tag{8.4}$$

Lemma 8.2.1 *For $S \in \mathcal{S}_d^+$, there exists a unique $R \in \mathcal{S}_d^+$ such that $S = R^2$. The matrix R is called the* **square root** *of S and is denoted by $\sqrt{S}$.*

Proof: We take $U \in \mathcal{O}_d$ so that (8.3), or equivalently (8.4) holds. Note that $s_\alpha \geq 0$ ($\alpha = 1, \ldots, d$).
Existence of R: $R \stackrel{\text{def}}{=} UD(\sqrt{s_1}, \ldots, \sqrt{s_d})U^*$ satisfies the desired property.
Uniqueness of R: Let $R \in \mathcal{S}_d^+$ be such that $S = R^2$. We will show that

1) $Ru_\alpha = \sqrt{s_\alpha} u_\alpha$, ($\alpha = 1, \ldots, d$),

which implies that $R = UD(\sqrt{s_1}, \ldots, \sqrt{s_d})U^*$. If $s_\alpha = 0$, then $Ru_\alpha = 0$, since

$$|Ru_\alpha|^2 = Ru_\alpha \cdot Ru_\alpha = Su_\alpha \cdot u_\alpha = s_\alpha |u_\alpha|^2 = 0.$$

Suppose on the other hand that $s_\alpha > 0$. Then,

$$(R + \sqrt{s_\alpha} I)(R - \sqrt{s_\alpha} I) = R^2 - s_\alpha I = S - s_\alpha I,$$

and hence

2) $(R + \sqrt{s_\alpha} I)(R - \sqrt{s_\alpha} I)u_\alpha = 0$.

$R + \sqrt{s_\alpha} I$ is strictly positive definite and hence invertible. Thus, 2) implies 1). \(^□^)/

For $d, k \in \mathbb{N} \backslash \{0\}$, we define a subset $\mathcal{O}_{d,k}$ of $d \times k$ real matrices as follows.

$$V \in \mathcal{O}_{d,k} \iff \begin{cases} \text{The colomn vectors of } V \text{ are orthonormal,} & \text{if } d \geq k, \\ \text{The raw vectors of } V \text{ are orthonormal,} & \text{if } d \leq k. \end{cases}$$

Lemma 8.2.2 *Let $V \in \mathcal{O}_{d,k}$.*

$$d \geq k \implies V^*V = I_k, \ \ (VV^* - I_d)|_{\text{Ran}V} = 0, \tag{8.5}$$
$$d \leq k \implies \text{Ran}(V^*V - I_k) \subset \text{Ker}V, \ \ VV^* = I_d, \tag{8.6}$$
$$U \in \mathcal{O}_k \implies VU \in \mathcal{O}_{d,k}. \tag{8.7}$$

Proof: (8.5): The first identity is equivalent to the definition of $\mathcal{O}_{d,k}$ for $d \geq k$. Using the first identity, we have

$$(VV^*)V = V(V^*V) = VI_k = I_dV,$$

which implies the second identity.
(8.6): The second identity is equivalent to the definition of $\mathcal{O}_{d,k}$ for $d \leq k$. Using the second identity, we have

$$V(V^*V) = (VV^*)V = I_dV = VI_k,$$

which implies the first identity.
(8.7): Let $u_1, ..., u_k \in \mathbb{R}^k$ be the column vectors of U and $v_1, ..., v_d \in \mathbb{R}^k$ be the raw vectors of V^*. Then,

$$VU = V(u_1, ..., u_k) = (Vu_1, ..., Vu_k), \quad (VU)^* = U^*V^* = U^*(v_1^*, ..., v_k^*) = (U^*v_1^*, ..., U^*v_k^*).$$

For $d \geq k$, we have $V^*V = I_k$ and hence for $\alpha, \beta = 1, ..., k$,

$$Vu_\alpha \cdot Vu_\beta = V^*Vu_\alpha \cdot u_\beta = u_\alpha \cdot u_\beta = \delta_{\alpha,\beta}.$$

Thus, the column vectors of VU are orthonormal. For $d \leq k$, we have $v_\alpha^* \cdot v_\beta^* = \delta_{\alpha,\beta}$ for $\alpha, \beta = 1, ..., d$,

$$U^*v_\alpha^* \cdot U^*v_\beta^* = UU^*v_\alpha^* \cdot v_\beta^* = v_\alpha^* \cdot v_\beta^* = \delta_{\alpha,\beta}.$$

Thus, the column vectors of $(VU)^*$ are orthonormal, i.e., the raw vectors of VU are orthonormal. \(^□^)/

Lemma 8.2.3 *Let A be a $d \times d$ real matrix, $s_1 \geq ... \geq s_k$ be the eigenvalues of A^*A, and $D = D(\sqrt{s_1}, \ldots, \sqrt{s_k})$. Then, there exist $U \in \mathcal{O}_k$ and $V \in \mathcal{O}_{d,k}$ such that*

$$AU = VD, \tag{8.8}$$

$$V^*V = \begin{pmatrix} I_d & 0 \\ 0 & 0 \end{pmatrix} \text{ if } d < k. \tag{8.9}$$

Proof: Let $S = A^*A$, and we take $U \in \mathcal{O}_d$ so that (8.3), or equivalently (8.4) holds. We then note that for $\alpha, \beta = 1, \ldots, d$,

1) $Au_\alpha \cdot Au_\beta = Su_\alpha \cdot u_\beta = s_\alpha \delta_{\alpha,\beta}$.

Let $m \stackrel{\text{def}}{=} \max\{\alpha\ ;\ s_\alpha > 0\} = \text{rank}\, S \leq d \wedge k$. Then, we see from 1) that

$$v_\alpha \stackrel{\text{def}}{=} Au_\alpha/\sqrt{s_\alpha} \in \mathbb{R}^d, \quad \alpha = 1, \ldots, m$$

are orthonormal and that $Au_\alpha = 0$ for $\alpha > m$. If $m = d \wedge k$, then, $v_1, \ldots, v_{d\wedge k}$ are orthonormal. If $m < d \wedge k$, then, we add orthonormal vectors $v_{m+1}, \ldots, v_{d\wedge k} \in \mathbb{R}^d$ so that $v_1, \ldots, v_{d\wedge k}$ are orthonormal. In particular, if $d < k$, we define $v_{d+1} = ... = v_k = 0$. Finally, we set $V = (v_1, ..., v_k)$. Then, $V \in \mathcal{O}_{d,k}$ and $Au_\alpha = \sqrt{s_\alpha} v_\alpha$ for all $\alpha = 1, \ldots, k$. Therefore,

$$AU = (Au_1, \ldots, Au_k) = (\sqrt{s_1}v_1, \ldots, \sqrt{s_k}v_k) = VD.$$

\(^□^)/

Remark: Referring to Lemma 8.2.3 and its proof, we see that rank$A^*A = m$ and Ran$A = \bigoplus_{\alpha=1}^m \mathbb{R}v_\alpha$, and hence rank$A^*A =$ rank$A$. By interchanging the role of A and A^*, and recalling that rank$A^* =$ rank$A$, we have rank$AA^* =$ rank$A^* =$ rank$A =$ rankA^*A. Combinning this with obvious inclusions Ran$AA^* \subset$ Ran$A$, Ran$A^*A \subset$ RanA^*, we obtain also Ran$A =$ RanAA^*, Ran$A^* =$ RanA^*A.

Proposition 8.2.4 *Let A be a $d \times k$ real matrix, $Q \in \mathcal{S}_d^+$, and $\sqrt{Q}$ be the square root of Q (Lemma 8.2.1). If $d \le k$, then*

$$Q = AA^* \iff \text{There exists } T \in \mathcal{O}_{d,k} \text{ such that } A = \sqrt{Q}T.$$

If $d > k$, then

$$Q = AA^* \iff \text{There exists } T \in \mathcal{O}_{d,k} \text{ such that } A = \sqrt{Q}T \text{ and } \mathrm{Ran}Q \subset \mathrm{Ran}T.$$

Proof: We treat the two cases ($d \le k$ and $d > k$) at the same time.
($\Rightarrow$) For A, we take $U \in \mathcal{O}_k$, $V \in \mathcal{O}_{d,k}$ and D as in Lemma 8.2.3. Then,

1) $A = VDU^*$ and $A^* = UDV^*$,

and hence

2) $Q = AA^* = VD^2V^*$.

We verify that

3) $D = DV^*V$.

This is obvious if $d \ge k$, since $V^*V \stackrel{(8.5)}{=} I_k$. If $d < k$, then, as is mentioned in the proof of Lemma 8.2.3, $s_\alpha = 0$ for $\alpha > d$, and hence by denoting $D_0 = (\sqrt{s_\alpha}\delta_{\alpha,\beta})_{\alpha,\beta=1}^d$,

$$DV^*V \stackrel{(8.9)}{=} \begin{pmatrix} D_0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} I_d & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} D_0 & 0 \\ 0 & 0 \end{pmatrix} = D.$$

We use 3) to prove that

4) $\sqrt{Q} = V^*DV$.

Note that $VDV^* \in \mathcal{S}_d^+$. Thus, by the uniqueness of the square root (Lemma 8.2.1), it is enough to show that $Q = (VDV^*)^2$.

$$(VDV^*)^2 = VDV^*VDV^* \stackrel{3)}{=} VD^2V^* \stackrel{2)}{=} Q.$$

Finally, with $T \stackrel{\text{def}}{=} VU^* \stackrel{(8.7)}{\in} \mathcal{O}_{d,k}$,

$$A \stackrel{1)}{=} VDU^* \stackrel{3)}{=} VDV^*VU^* \stackrel{4)}{=} \sqrt{Q}T.$$

Moreover, if $d > k$, then Ran$Q \stackrel{Q=AA^*}{=}$ Ran$A \stackrel{1)}{\subset}$ Ran$V =$ Ran$T$.
($\Leftarrow$) We verify that

5) $\sqrt{Q} = TT^*\sqrt{Q}$.

This is obvious if $d \le k$, since $TT^* \stackrel{(8.6)}{=} I_d$. Suppose that $d > k$. Then, $\mathrm{Ran}Q \subset \mathrm{Ran}T$ by the assumption. Moreover, $\mathrm{Ran}\sqrt{Q} = \mathrm{Ran}Q$, as can be seen from the proof of Lemma 8.2.1. Thus, $\mathrm{Ran}\sqrt{Q} \subset \mathrm{Ran}T$. Since $(TT^* - I_d)|_{\mathrm{Ran}T} \stackrel{(8.5)}{=} 0$, we have $(TT^* - I_d)|_{\mathrm{Ran}\sqrt{Q}} = 0$, which implies 5). Using 5) we conclude that

$$AA^* = \sqrt{Q}TT^*\sqrt{Q} \stackrel{5)}{=} (\sqrt{Q})^2 = Q.$$

\(^□^)/

8.3 Uniform Distribution and an Existence Theorem for Independent Random Variables

To define a random walk (cf. Definition 3.1.1 below), we will need countably many independent r.v's. A question[24] then arises: "Do such independent r.v's exist?" This subsection is devoted to answer this question. Throughout this subsection, we fix a probability sapce $(\Omega, \mathcal{F}, P)$ and r.v. U with the uniform distribution on $[0, 1)$, i. e., $P\{U \in B\} = \int_B dt$ for all $B \in \mathcal{B}([0, 1))$. The simplest example is provided by $\Omega = [0, 1)$, $\mathcal{F} = \mathcal{B}([0, 1))$ and $U(\omega) = \omega$. We will prove the following existence theorem for independent r.v.'s;

Proposition 8.3.1 *Consider a sequence of probability spaces $\{(S_n, \mathcal{B}_n, \mu_n)\}_{n\ge1}$ where for each n, S_n is a complete separable metric space and $\mathcal{B}_n$ is the Borel σ-algebra. Then, there is a sequence of independent r.v.'s $\{X_n : \Omega \to S_n\}_{n\ge1}$ such that $\mu_n(B) = P(X_n \in B)$ for all $n \ge 1$ and $B \in \mathcal{B}_n$.*

Remark: Proposition 8.3.1 can be considered as a special case of Kolmogorov's extension theorem (See e.g., [Dur95, page 26 (4.9)] for the case $S_n = \mathbb{R}^d$). Kolmogorov's extension theorem is so powerful that it allows us to construct not only independent r.v.'s but also *any* r.v.'s which exsit at all. However, the proof usually requires another extention theorem in measure theory (e.g., Carathéodory's extention theorem). Here, to make the exposition more self-contained, we restrict our attention only to independent cases and give an elementary proof of Proposition 8.3.1 without relying on any big theorem from measure theory.

We begin with examples:

Example 8.3.2 Let us now construct an i.i.d. sequence $\{U_n\}_{n\ge1}$ of $[0, 1)$-valued r.v.'s with the uniform distribution. By Example 1.9.4, there is an i.i.d. sequence $\{X_{n,k}\}_{n,k\ge1}$ of $\{0, 1\}$-valued r.v.'s with $P\{X_{n,k} = 1\} = 1/2$. We define $\{U_n\}_{n\ge1}$ by

$$U_n = \sum_{k\ge1} 2^{-k} X_{n,k}.$$

Then, each U_n is uniformly distributed by Lemma 8.5.1. Moreover, $\{U_n\}_{n\ge1}$ are independent by Exercise 1.6.9.

[24]This may be a question which a physicist would not care about. Those who do not worry about this question can skip this subsection.

To prove Proposition 8.3.1, we will use Example 1.9.4, Example 8.3.2 and the following lemma.

Lemma 8.3.3 *Suppose that $(S, \mathcal{B}, \mu)$ is a probability space where S is a complete separable metric space and $\mathcal{B}$ is the Borel σ-algebra. Then, there is a measurable map $\varphi : [0,1) \to S$ such that*

$$P\{\varphi(U) \in B\} = \mu(B), \quad \text{for all } B \in \mathcal{B}, \tag{8.10}$$

where $U : \Omega \to [0,1)$ is a uniformly distributed r.v.

Lemma 8.3.3 is quite surprising in the sense that it claims *any* r.v. with values in a complete separable metric space can be constructed just by using a single uniformly distributed r.v. The proof of Lemma 8.3.3 will be presented in subsection 8.4.

We now prove Proposition 8.3.1.

Proof of Proposition 8.3.1: Let $\{U_n\}_{n \geq 1}$ be $[0,1)$-valued r.v.'s with the uniform distribution constructed in Example 8.3.2. For each $\mu_n \in \mathcal{P}(S_n, \mathcal{B}_n)$, we can find a measurable map $\varphi_n : [0,1) \to S_n$ such that $P\{\varphi_n(U_n) \in \cdot\} = \mu_n$ by Lemma 8.3.3. We also see that $\{\varphi_n(U_n)\}_{n \geq 1}$ are independent since $\{U_n\}_{n \geq 1}$ are. Therefore the r.v.'s $X_n = \varphi_n(U_n)$ $(n \geq 1)$ have desired properties claimed in Proposition 8.3.1. \(^□^)/

Exercise 8.3.1 For $\mu \in \mathcal{P}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, define

$$\begin{aligned} f(s) &= \mu(-\infty, s], \quad s \in \mathbb{R}, \\ f_-^{-1}(t) &= \inf\{s \in \mathbb{R} \mid t \leq f(s)\ \} \\ &= \sup\{s \in \mathbb{R} \mid f(s) < t\ \}, \quad t \in \mathbb{R}. \end{aligned}$$

Prove the following; (i) $f(s)$ is right-continuous at any $s \in \mathbb{R}$. (ii) $f_-^{-1}(t)$ is left-continuous at all $t \in (0,1)$. (iii) For $s \in \mathbb{R}$ and $t \in (0,1)$, $f_-^{-1}(t) \leq s \iff t \leq f(s)$

Exercise 8.3.2 Let $\mu_n \in \mathcal{P}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ $(n = 1, \ldots)$ be a sequence of probability measures. Use Example 8.3.2 and Exercise 8.3.1 to construct a sequence of independent r.v.'s $X_n : \Omega \to \mathbb{R}$ such that $P(X_n \in \cdot) = \mu_n$ for all $n \geq 1$. Hint: Define $f_n(s) = \mu_n(-\infty, s]$ and $\varphi_n(\theta) = (f_n)_-^{-1}(\theta)$. Then, for all $s \in \mathbb{R}$,

$$P\{\varphi_n(U_n) \leq s\} = P\{U_n \leq f_n(s)\} = f_n(s).$$

Then, recall Exercise 1.3.2.

8.4 Proof of Lemma 8.3.3

The proof of Lemma 8.3.3 is not very difficult and the argument involved there is a rather standard way to take advantage of the completeness and the separability of the metric space S. However, the proof may look a little complicated at first sight. We therefore present also a proof for the case of $S = \mathbb{R}^d$, which is less abstract and which is the only case we need in this course. The proof for this special case might be useful to understand the idea behind the proof of general case.

Those who are interested only in the case $S = \mathbb{R}^d$ can skip the proof for the general case. On the other hand, it is also possible to skip the proof for the case $S = \mathbb{R}^d$ to proceed directly to that in general case.

Proof of Lemma 8.3.3 in the case $S = \mathbb{R}^d$:

Step 1: We begin by constructing a sequence of intervals (in $\mathbb{R}^d$)

$$Q_{s_1} \supset Q_{s_1 s_2} \supset \ldots \supset Q_{s_1 \cdots s_n} \supset \ldots,$$

inductively, where the running indices $s_1, s_2, \ldots$ are diadic rational points. As the first step of the induction, we find a subset $C \subset 2^{-1}\mathbb{Z}^d$ and disjoint intervals $\{Q_{s_1}\}_{s_1 \in C}$ such that

$$\begin{array}{rcl} Q_{s_1} & \ni & s_1 \quad \text{for all } s_1 \in C, \\ \mu(N) & = & 0, \quad \text{where } N \stackrel{\text{def.}}{=} S \setminus \cup_{s_1 \in C} Q_{s_1}, \\ \mu(Q_{s_1}) & > & 0, \quad \text{for all } s_1 \in C. \end{array} \tag{8.11}$$

In fact, this can be done just by setting

$$\begin{array}{rcl} Q_{s_1} & = & \prod_{j=1}^d [s_1^j, s_1^j + 2^{-1}), \quad \text{for } s_1 = (s_1^j)_{j=1}^d \in 2^{-1}\mathbb{Z}^d, \\ C & = & \{s_1 \in 2^{-1}\mathbb{Z}^d \ ; \ \mu(Q_{s_1}) > 0\}. \end{array} \tag{8.12}$$

The second step of the induction is as follows. For each $s_1 \in C$, we repeat the same argument as in the first step of the induction to find a subset $C(s_1) \subset Q_{s_1} \cap 2^{-2}\mathbb{Z}^d$ and disjoint intervals $\{Q_{s_1,s_2}\}_{s_2 \in C(s_1)}$ with the side-length 2^{-2} such that

$$\begin{array}{rcl} Q_{s_1} & \supset & Q_{s_1 s_2} \ni s_2 \quad \text{for all } s_2 \in C(s_1), \\ \mu(N_{s_1}) & = & 0, \quad \text{where } N_{s_1} \stackrel{\text{def.}}{=} Q_{s_1} \setminus \cup_{s_2 \in C(s_1)} Q_{s_1,s_2}, \\ \mu(Q_{s_1 s_2}) & > & 0 \quad \text{for all } s_2 \in C(s_1). \end{array}$$

Suppose as the n^{th} step of the induction that we have an interval $Q_{s_1 \cdots s_n}$ with non-zero μ-measure and the side-length 2^{-n} for $s_1 \in C, \ldots, s_n \in C(s_1 \cdots s_{n-1})$. Then, we can find $C(s_1 \cdots s_n) \subset Q_{s_1 \cdots s_n} \cap 2^{-(n+1)}\mathbb{Z}^d$ and intervals $Q_{s_1 \cdots s_{n+1}}$ for $s_{n+1} \in C(s_1 \cdots s_n)$ such that

$$Q_{s_1 \cdots s_n} \supset Q_{s_1 \cdots s_{n+1}} \ni s_{n+1} \quad \text{for all } s_{n+1} \in C(s_1, \ldots, s_n). \tag{8.13}$$

$$\mu(N_{s_1 \cdots s_n}) = 0, \quad \text{where } N_{s_1 \cdots s_n} \stackrel{\text{def.}}{=} Q_{s_1 \cdots s_n} \setminus \cup_{s_{n+1} \in C(s_1, \ldots, s_n)} Q_{s_1 \cdots s_{n+1}}, \tag{8.14}$$

$$\mu(Q_{s_1 \cdots s_{n+1}}) > 0 \quad \text{for all } s_{n+1} \in C(s_1, \ldots, s_n).$$

Step 2: We next construct a sequence

$$I_{s_1} \supset I_{s_1 s_2} \supset \ldots \supset I_{s_1 \cdots s_n} \supset \ldots,$$

of sub-intervals of $[0,1)$ with positive lengths, where $I_{s_1 \cdots s_n}$ corresponds to $Q_{s_1 \cdots s_n}$ in a way as is explained below. We first split $[0,1)$ into disjoint intervals $\{I_{s_1}\}_{s_1 \in C}$ with length $|I_{s_1}| = \mu(Q_{s_1})$ for each $s_1 \in C$. Then, for each $s_1 \in C$, we split I_{s_1} into disjoint intervals $\{I_{s_1,s_2}\}_{s_2 \in C(s_1)}$ with length $|I_{s_1,s_2}| = \mu(Q_{s_1,s_2})$ for each $s_2 \in C(s_1)$. We then inductively iterate this procedure to get $\{I_{s_1 \cdots s_n}\}$ such that

$$[0,1) = \cup_{s_1 \in C} I_{s_1}, \tag{8.15}$$

$$I_{s_1 \cdots s_{n-1}} = \cup_{s_n \in C(s_1, \ldots, s_{n-1})} I_{s_1 \cdots s_n}, \tag{8.16}$$

$$|I_{s_1 \cdots s_n}| = \mu(Q_{s_1 \cdots s_n}). \tag{8.17}$$

Step 3: We now define $\varphi_n : [0,1) \to S$ by

$$\varphi_n(\theta) = s_n \quad \text{if } \theta \in I_{s_1 \cdots s_n}.$$

Let us check the following;

$$\varphi_n : [0,1) \to S \text{ is well defined and measurable for all } n \geq 1. \tag{8.18}$$

$$(\varphi_n(\theta))_{n\geq 1} \text{ is a Cauchy sequence for for all } \theta \in [0,1). \tag{8.19}$$

By (8.15) and (8.16), any $\theta \in [0,1)$ belongs to a unique interval $I_{s_1 \cdots s_n}$. Therefore, φ_n is well defined. The measurability is obvious, since φ_n is a constant s_n on each measurable set $I_{s_1 \cdots s_n}$. To see (8.19), just observe that

$$\varphi_{m+n}(\theta) \in Q_{\varphi_1(\theta),\ldots,\varphi_{m+n}(\theta)} \subset Q_{\varphi_1(\theta),\ldots,\varphi_n(\theta)},$$

and hence that

$$|\varphi_{m+n}(\theta) - \varphi_n(\theta)| \leq 2^{-n}\sqrt{d}.$$

Step 4: By, (8.19) and (8.19), we can define a measurable map $\varphi : [0,1) \to \mathbb{R}^d$ by $\varphi(\theta) = \lim_{n\to\infty} \varphi_n(\theta)$ for all $\theta \in [0,1)$. Let us see that φ satisfies (8.10). To do so, define a set

$$N_0 = \cup_{n\geq 1} \cup_{s_1 \in C} \cup_{s_2 \in C(s_1)} \ldots \cup_{s_n \in C(s_1,\ldots,s_{n-1})} N \cup N_{s_1} \cup N_{s_1 s_2} \cup \ldots \cup N_{s_1 \ldots s_n}$$

which is μ-measure zero by (8.11) and (8.14). Moreover, for each $x \in \mathbb{R}^d \backslash N_0$ and $n \geq 1$, there is a unique $Q_{s_1,\ldots,s_n}$ such that $x \in Q_{s_1,\ldots,s_n}$. Therefore, for any $f \in C_{\mathrm{b}}(\mathbb{R}^d)$ we can define function $f_n : \mathbb{R}^d \backslash N_0 \to \mathbb{R}$ by

$$f_n(x) = \sum_{s_1 \in C} \sum_{s_2 \in C(s_1)} \cdots \sum_{s_n \in C(s_1,\ldots,s_{n-1})} f(s_n) \mathbf{1}\{x \in Q_{s_1 \ldots s_n}\}.$$

We see that

$$\lim_{n\to\infty} f_n(x) = f(x) \quad \text{for all } x \in S \backslash N_0, \tag{8.20}$$

since $|x - s_n| \leq 2^{-n}\sqrt{d}$ if $x \in Q_{s_1 \ldots s_n}$. Therefore,

$$\begin{aligned}
Ef(\varphi(U)) &= \lim_{n\to\infty} Ef(\varphi_n(U)) \quad \text{by definition of } \varphi, \\
&= \lim_{n\to\infty} \sum_{s_1 \in C} \sum_{s_2 \in C(s_1)} \cdots \sum_{s_n \in C(s_1,\ldots,s_{n-1})} f(s_n) |I_{s_1,\ldots,s_n}| \quad \text{by definition of } \varphi_n, \\
&= \lim_{n\to\infty} \sum_{s_1 \in C} \sum_{s_2 \in C(s_1)} \cdots \sum_{s_n \in C(s_1,\ldots,s_{n-1})} f(s_n) \mu(Q_{s_1,\ldots,s_n}) \quad \text{by (8.27)}, \\
&= \lim_{n\to\infty} \int f_n d\mu \quad \text{by definition of } f_n, \\
&= \int f d\mu \quad \text{by (8.20)}.
\end{aligned}$$

This proves (8.10) (cf. Lemma 1.3.2). \(^□^)/

Proof of Lemma 8.3.3 in general case: Most of the arguments presented below are repetitions of the ones in the case of $S = \mathbb{R}^d$. However, we do repeat the every detail, so that this proof for the general case can be read independently.

Step 1: We begin by constructing a sequence of measurable subsets

$$Q_{s_1} \supset Q_{s_1 s_2} \supset \ldots \supset Q_{s_1 \cdots s_n} \supset \ldots,$$

inductively, where the running indices $s_1, s_2, \ldots$ are elements in S. The first step of the induction is as follows. Since S is separable, we can find a countable subset $C \subset S$ and disjoint measurable subsets $\{Q_{s_1}\}_{s_1 \in C}$ such that

$$\begin{aligned} Q_{s_1} &\ni s_1 \quad \text{for all } s_1 \in C, \\ \mu(N) &= 0, \quad \text{where } N \stackrel{\text{def.}}{=} S \setminus \cup_{s_1 \in C} Q_{s_1}, \\ \text{diam}(Q_{s_1}) &\leq 2^{-1}, \\ \mu(Q_{s_1}) &> 0, \end{aligned} \tag{8.21}$$

In fact, let $\{B_n\}_{n \geq 1}$ be a covering of S by balls (open or closed) with the diameter 2^{-1} and define $\{\underline{B}_n\}_{n \geq 1}$ by $\underline{B}_1 = B_1$ and

$$\underline{B}_n = B_n \setminus \cup_{j=1}^{n-1} B_j \quad \text{n=1,2,} \ldots.$$

Then, $\{\underline{B}_n\}_{n \geq 1}$ are covering of S by disjoint measurable sets and $\text{diam}(\underline{B}_n) \leq 2^{-1}$. Now let $\{Q_n\}_{n \geq 1}$ be a subsequence of $\{\underline{B}_n\}_{n \geq 1}$ which is obtained by throwing away all $\underline{B}_n$'s which have μ-measure zero. Finally, we take $s_n \in Q_n$ for each $n \geq 1$ and define $Q_{s_n} = Q_n$ and $C = \{s_n\}_{n \geq 1}$.

The second step of the induction is as follows. Since any subset in S is separable, we can find a countable subset $C(s_1) \subset Q_{s_1}$ for each $s_1 \in C$, and disjoint measurable subsets $\{Q_{s_1, s_2}\}_{s_2 \in C(s_1)}$ such that

$$\begin{aligned} Q_{s_1} &\supset Q_{s_1 s_2} \ni s_2 \quad \text{for all } s_2 \in C(s_1). \\ \mu(N_{s_1}) &= 0, \quad \text{where } N_{s_1} \stackrel{\text{def.}}{=} Q_{s_1} \setminus \cup_{s_2 \in C(s_1)} Q_{s_1, s_2}, \\ \text{diam}(Q_{s_1 s_2}) &\leq 2^{-2}, \\ \mu(Q_{s_1 s_2}) &> 0. \end{aligned}$$

Suppose as the n^{th}-step of the induction that we have a measurable set $Q_{s_1 \cdots s_n}$ with non-zero μ-measure and the diameter $\leq 2^{-n}$ for $s_1 \in C, \ldots, s_n \in C(s_1 \cdots s_{n-1})$. Then, we can find a countable subset $C(s_1 \cdots s_n) \subset Q_{s_1 \cdots s_n}$ and disjoint measurable sets $\{Q_{s_1 \cdots s_{n+1}}\}$ for $s_{n+1} \in C(s_1 \cdots s_n)$ such that

$$Q_{s_1 \cdots s_n} \supset Q_{s_1 \cdots s_{n+1}} \ni s_{n+1} \quad \text{for all } s_{n+1} \in C(s_1, \ldots, s_n). \tag{8.22}$$

$$\mu(N_{s_1 \cdots s_n}) = 0, \quad \text{where } N_{s_1 \cdots s_n} \stackrel{\text{def.}}{=} Q_{s_1 \cdots s_n} \setminus \cup_{s_{n+1} \in C(s_1, \ldots, s_n)} Q_{s_1 \cdots s_{n+1}}, \tag{8.23}$$

$$\text{diam}(Q_{s_1 \cdots s_n}) \leq 2^{-n}, \tag{8.24}$$

$$\mu(Q_{s_1 \cdots s_{n+1}}) > 0 \quad \text{for all } s_{n+1} \in C(s_1, \ldots, s_n).$$

Step 2: We next construct a sequence

$$I_{s_1} \supset I_{s_1 s_2} \supset \ldots \supset I_{s_1 \cdots s_n} \supset \ldots,$$

of sub-intervals of $[0, 1)$ with positive lengths, where $I_{s_1 \cdots s_n}$ corresponds to $Q_{s_1 \cdots s_n}$ in a way as is explained below. We first split $[0, 1)$ into disjoint intervals $\{I_{s_1}\}_{s_1 \in C}$ with length $|I_{s_1}| = \mu(Q_{s_1})$

for each $s_1 \in C$. Then, for each $s_1 \in C$, we split I_{s_1} into disjoint intervals $\{I_{s_1,s_2}\}_{s_2 \in C(s_1)}$ with length $|I_{s_1,s_2}| = \mu(Q_{s_1,s_2})$ for each $s_2 \in C(s_1)$. We then inductively iterate this procedure to get $\{I_{s_1 \cdots s_n}\}$ such that

$$[0,1) = \cup_{s_1 \in C} I_{s_1}, \tag{8.25}$$

$$I_{s_1 \cdots s_{n-1}} = \cup_{s_n \in C(s_1,\ldots,s_{n-1})} I_{s_1 \cdots s_n}, \tag{8.26}$$

$$|I_{s_1 \cdots s_n}| = \mu(Q_{s_1 \cdots s_n}). \tag{8.27}$$

Step 3: We now define $\varphi_n : [0,1) \to S$ by

$$\varphi_n(\theta) = s_n \quad \text{if } \theta \in I_{s_1 \cdots s_n}.$$

Let us check the following;

$$\varphi_n : [0,1) \to S \text{ is well defined and measurable for all } n \geq 1. \tag{8.28}$$

$$(\varphi_n(\theta))_{n \geq 1} \text{ is a Cauchy sequence for for all } \theta \in [0,1). \tag{8.29}$$

By (8.25) and (8.26), any $\theta \in [0,1)$ belongs to a unique interval $I_{s_1 \cdots s_n}$. Therefore, φ_n is well defined. The measurability is obvious, since φ_n is a constant s_n on each measurable set $I_{s_1 \cdots s_n}$. To see (8.29), just observe that

$$\varphi_{m+n}(\theta) \in Q_{\varphi_1(\theta),\ldots,\varphi_{m+n}(\theta)} \subset Q_{\varphi_1(\theta),\ldots,\varphi_n(\theta)},$$

and hence by (8.24) that

$$\text{dist.}(\varphi_{m+n}(\theta), \varphi_n(\theta)) \leq 2^{-n}.$$

Step 4: By, (8.29) and (8.29), we can define a measurable map $\varphi : [0,1) \to S$ by $\varphi(\theta) = \lim_{n\to\infty} \varphi_n(\theta)$ for all $\theta \in [0,1)$. Let us see that φ satisfies (8.10). To do so, take $f \in C_{\mathrm{b}}(S)$ and define a set

$$N_0 = \cup_{n \geq 1} \cup_{s_1 \in C} \cup_{s_2 \in C(s_1)} \ldots \cup_{s_n \in C(s_1,\ldots,s_{n-1})} N \cup N_{s_1} \cup N_{s_1 s_2} \cup \ldots \cup N_{s_1 \ldots s_n}$$

which is μ-measure zero, and function $f_n : S \backslash N_0 \to \mathbb{R}$ by

$$f_n(x) = \sum_{s_1 \in C} \sum_{s_2 \in C(s_1)} \cdots \sum_{s_n \in C(s_1,\ldots,s_{n-1})} f(s_n) 1\{x \in Q_{s_1,\ldots,s_n}\},$$

which is well defined, by (8.21) and (8.23). Moreover, we see from (8.24) that

$$\lim_{n\to\infty} f_n(x) = f(x) \quad \text{for all } x \in S \backslash N_0. \tag{8.30}$$

Therefore,

$$\begin{aligned}
Ef(\varphi(U)) &= \lim_{n\to\infty} Ef(\varphi_n(U)) \quad \text{by definition of } \varphi, \\
&= \lim_{n\to\infty} \sum_{s_1 \in C} \sum_{s_2 \in C(s_1)} \cdots \sum_{s_n \in C(s_1,\ldots,s_{n-1})} f(s_n) |I_{s_1,\ldots,s_n}| \quad \text{by definition of } \varphi_n, \\
&= \lim_{n\to\infty} \sum_{s_1 \in C} \sum_{s_2 \in C(s_1)} \cdots \sum_{s_n \in C(s_1,\ldots,s_{n-1})} f(s_n) \mu(Q_{s_1,\ldots,s_n}) \quad \text{by (8.27)}, \\
&= \lim_{n\to\infty} \int f_n d\mu \quad \text{by definition of } f_n, \\
&= \int f d\mu \quad \text{by (8.30)}.
\end{aligned}$$

This proves (8.10) (cf. Lemma 1.3.2). \(^□^)/

8.5 Complement to Section 1.9

Lemma 8.5.1 *Suppose that $q \geq 2$ is an integer and that $V = \sum_{k\geq 1} q^{-k} Y_k$, where $\{Y_k\}_{k\geq 1}$ are $\{0, 1, \ldots, q-1\}$-valued r.v. and V is a $[0,1)$-valued r.v. Then, the following conditions are related as " (a1) & (a2) $\iff$ (b)";*

a1) $\{Y_k\}_{k\geq 1}$ *are i.i.d.*

a2) Y_k *is uniformly distributed, i.e.,* $P\{Y_k = s\} = q^{-1}$ *for any* $s = 1, \ldots, q-1$.

b) V *is uniformly distributed on* $[0,1)$.

Proof: (a1) & (a2) $\Rightarrow$ (b) : Suppose that (a1) & (a2) holds. Then, $(X_n)_{n\geq 1}$ in Example 1.9.1 and $(Y_n)_{n\geq 1}$ have the same distribution. Therefore, U and V have the same distribution, which proves (b).

(b) $\Rightarrow$ (a1) & (a2) : Suppose that (b) holds. Then, outside an event

$$\cup_{n\geq 1} \cup_{0\leq s\leq q^n - 1} \{V = sq^{-n}\},$$

and therefore for P-almost all $\omega \in \Omega$, $Y_k(\omega)$ is uniquely determied as the $k^{\rm th}$ digit of the q-adic expansion of the number $V(\omega)$. We therefore see from (1.75) that $(X_n)_{n\geq 1}$ in Example 1.9.1 and $(Y_n)_{n\geq 1}$ have the same distribution, which proves (a1) & (a2). \(^□^)/

Exercise 8.5.1 Check an alternative proof of Lemma 8.5.1, (a1) & (a2) $\Rightarrow$ (b) presented below. It is enough to prove that for any $t \in [0,1)$

$$P\{V \leq t\} = t \tag{8.31}$$

(cf. Exercise 1.3.2). Let us expand $t \in [0,1)$ as $t = \sum_{k=1}^{\infty} q^{-k} s_k$ ($s_k \in \{0, \ldots, q-1\}$) and denote the left-hand side of (8.31) by $f(s_1, s_2, \ldots)$. We will prove that

$$f(s_1, s_2, \ldots) = q^{-1} s_1 + q^{-1} f(s_2, s_3, \ldots). \tag{8.32}$$

We have that

$$\begin{aligned}
\{U \leq t\} &= \{Y_1 < s_1\} \cup \left\{ Y_1 = s_1, \ \sum_{k=2}^{\infty} q^{-k} Y_k \leq \sum_{k=2}^{\infty} q^{-k} s_k \right\} \\
&= \{Y_1 < s_1\} \cup \left\{ Y_1 = s_1, \ \sum_{k=1}^{\infty} q^{-k} Y_{k+1} \leq \sum_{k=1}^{\infty} q^{-k} s_{k+1} \right\}. \qquad (8.33)
\end{aligned}$$

We are now going to use the two facts;

i) Y_1 and $(Y_{k+1})_{k=1}^{\infty}$ are independent,

ii) $(Y_{k+1})_{k=1}^{\infty}$ and $(Y_k)_{k=1}^{\infty}$ have the same distribution.

Facts (i),(ii) and (8.33) imply that

$$\begin{aligned}
P\{V \leq t\} &= P\{Y_1 < s_1\} + P\{Y_1 = s_1\} P\left\{ \sum_{k=1}^{\infty} q^{-k} Y_{k+1} \leq \sum_{k=1}^{\infty} q^{-k} s_{k+1} \right\} \quad \text{by (i)} \\
&= s_1 q^{-1} + q^{-1} P\left\{ \sum_{k=1}^{\infty} q^{-k} Y_k \leq \sum_{k=1}^{\infty} q^{-k} s_{k+1} \right\} \quad \text{by (ii)} \\
&= s_1 q^{-1} + q^{-1} f(s_2, s_3, \ldots), \qquad (8.34)
\end{aligned}$$

which proves (8.32).

With (8.32) in hand, proof of (8.31) is easy. In fact, we have for any $n = 1, 2, \ldots$

$$f(s_1, s_2, \ldots) = \sum_{k=1}^{n} q^{-k} s_k + q^{-n} f(s_{n+1}, s_{n+2}, \ldots) \tag{8.35}$$

by induction. Then (8.32) follows by letting $n \nearrow \infty$. \(^□^)/

8.6 Convolution

Definition 8.6.1 For Borel finite measures $\{\mu_j\}_{j=1}^n$ on $\mathbb{R}^d$, their *convolution* $\mu_1 * \cdots * \mu_n$ is a Borel finite measure defined by

$$(\mu_1 * \cdots * \mu_n)(B) = \left(\otimes_{j=1}^n \mu_j\right) \left\{(x_j)_{j=1}^n \in (\mathbb{R}^d)^n \; ; \; x_1 + \ldots + x_n \in B\right\}, \quad B \in \mathcal{B}(\mathbb{R}^d). \tag{8.36}$$

Suppose that $\mathbb{R}^d$-valued r.v.'s $\{X_j\}_{j=1}^n$ are independent and $P\{X_j \in \cdot\} = \mu_j$. We then have by Proposition 1.6.1 that

$$P(X_1 + \ldots + X_n \in \cdot) = \mu_1 * \cdots * \mu_n. \tag{8.37}$$

Lemma 8.6.2 **(a)** *For Borel finite measures μ_1, μ_2 on $\mathbb{R}^d$,*

$$(\mu_1 * \mu_2)^\wedge(\theta) = \widehat{\mu_1}(\theta)\widehat{\mu_2}(\theta) \quad \text{for all } \theta \in \mathbb{R}^d. \tag{8.38}$$

(b) *Suppose that μ_j $(j = 1, 2)$ are Borel finite measures on $\mathbb{R}^d$ with density f_j with respect to the Lebesgue measure $(j = 1, 2)$. Then $\mu_1 * \mu_2$ has a density*

$$(f_1 * f_2)(x) = \int f_1(x - y) f_2(y) dy \tag{8.39}$$

with respect to the Lebesgue measure.

(c) *Suppose that μ_j $(j = 1, 2)$ are Borel finite measures on $\mathbb{R}^d$ such that $\mu_j(B) = \sum_{x \in \mathbb{Z}^d \cap B} f_j(x)$ for some $f_j : \mathbb{Z}^d \to [0, \infty)$ for all $B \in \mathcal{B}(\mathbb{R}^d)$. Then, $\mu_1 * \mu_2(B) = \sum_{x \in \mathbb{Z}^d \cap B} (f_1 * f_2)(x)$ for all $B \in \mathcal{B}(\mathbb{R}^d)$, where*

$$(f_1 * f_2)(x) = \sum_{y \in \mathbb{Z}^d} f_1(x - y) f_2(y) dy. \tag{8.40}$$

Proof: It is easy to see (8.38). (8.39) can be seen as follows;

$$\begin{aligned} \mu_1 * \mu_2(B) &= \int \mu_1 \otimes \mu_2(dzdy) 1_B(z + y) \\ &= \int f_1(z) f_2(y) dzdy 1_B(z + y) \\ &= \int f_1(x - y) f_2(y) dxdy 1_B(x) \\ &= \int_B (f_1 * f_2)(x) dx. \end{aligned} \tag{8.41}$$

The proof of (8.40) is similar to that of (8.39). \\(^□^)/

Example 8.6.3 Let χ_1 and χ_2 be independent Gaussian r.v.'s such that $P(\chi_j \in \cdot) = \nu_{V_j}$ $(j = 1, 2)$. Then, by Exercise 2.2.4,

$$P(\chi_1 + \chi_2 \in \cdot) = \nu_{V_1} * \nu_{V_2} = \nu_{V_1+V_2}. \tag{8.42}$$

Example 8.6.4 Let X and Y be independent real r.v.'s such that $P((X,Y) \in \cdot) = \gamma_{r,a} \otimes \gamma_{r,b}$. Then, by Example 1.7.5,

$$P(X + Y \in \cdot) = \gamma_{r,a} * \gamma_{r,b} = \gamma_{r,a+b}. \tag{8.43}$$

Example 8.6.5 Then, by (1.65),

$$P(N_1 + N_2 \in \cdot) = \pi_{r_1} * \pi_{r_2} = \pi_{r_1+r_2}. \tag{8.44}$$

Exercise 8.6.1 Suppose that r.v.'s U_j $(j = 1, 2)$ are independent and have the uniform distribution on an interval $[a, b]$, i. e., $P\{U_j \in B\} = \int_B u(t)dt$ for all $B \in \mathcal{B}(\mathbb{R})$ $(j = 1, 2)$, where $u(t) = (b-a)^{-1}1_{[a,b]}(t)$. Prove then that the r.v.$U_1 + U_2$ has the *triangular distribution* on $[2a, 2b]$, i. e.,

$$P\{U_1 + U_2 \in B\} = \int_B v(t)dt, \tag{8.45}$$

where

$$v(t) = (u * u)(t) = \frac{t-2a}{(b-a)^2}\mathbf{1}_{[2a,a+b]}(t) + \frac{2b-t}{(b-a)^2}\mathbf{1}_{[a+b,2b]}(t).$$

Then, conclude from (2.7) and (8.45) that

$$\widehat{v}(\theta) = \widehat{u}(\theta)^2 = \left(\tfrac{\exp(\mathbf{i}\theta b)-\exp(\mathbf{i}\theta a)}{(b-a)\theta}\right)^2. \tag{8.46}$$

Exercise 8.6.2 Suppose that X_j $(j \geq 1)$ are r.v.'s with $P\{X_j \in \cdot\} = \mu_j \in \mathcal{P}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and that N is a r.v. with (r)-Poisson distribution (cf. (1.18)). Suppose also that $\{N, X_1, X_2, \ldots\}$ are independent. Prove then that

$$P\{X_1 + \ldots + X_N \in \cdot\} = \sum_{n \geq 1} e^{-r}r^n(\mu_1 * \cdots * \mu_n)/n! \tag{8.47}$$

The distribution on the right-hand side of (8.47) is called the *compound Poisson distribution.* Poisson distribution is a compound Poisson distribution with $X_j \equiv 1$.

8.7 Independent Families of Random Variables

Definition 8.7.1 a) Independent events: Suppose that $\mathcal{A} \subset \mathcal{F}$. Then, $\mathcal{A}$ said to be *independent*, if

$$P\left(\cap_{A \in \mathcal{A}_0} A\right) = \prod_{A \in \mathcal{A}_0} P(A) \quad \text{for any finite subset } \mathcal{A}_0 \text{ in } \mathcal{A}. \tag{8.48}$$

b) Independence for families of events: Suppose that $\mathcal{A}_\lambda \subset \mathcal{F}$ for each $\lambda \in \Lambda$. Then, the families $\{\mathcal{A}_\lambda\}_{\lambda \in \Lambda}$ are said to be *quasi-independent*, if

$$\{A_\lambda\}_{\lambda \in \Lambda} \subset \mathcal{F} \text{ is independent in the sense of (a) for any } A_\lambda \in \mathcal{A}_\lambda \ (\lambda \in \Lambda). \tag{8.49}$$

The families $\{\mathcal{A}_\lambda\}_{\lambda \in \Lambda}$ are said to be *independent* if the σ-algebras $\{\sigma[\mathcal{A}_\lambda]\}_{\lambda \in \Lambda}$ are quasi-independent.

Remark: 1) The condition (8.49) does not imply that $\{\sigma[\mathcal{A}_\lambda]\}_{\lambda\in\Lambda}$ are independent σ-algebras (cf. Exercise 8.7.2). This is the reason we do not define it as the "independence" for the families $\{\mathcal{A}_\lambda\}_{\lambda\in\Lambda}$. If $\{\mathcal{A}_\lambda\}_{\lambda\in\Lambda}$ are σ-algebras, then the notion of independence and quasi-indpendence coincide.

2) The terminology "quasi independence" does not appear in standard text books in probability theory. It is introduced by the author of this notes for the convenience.

Exercise 8.7.1 Prove the following: (i) $\sigma[\{A\}] = \{\emptyset, \Omega, A, A^c\}$ for a set A. (ii) For $\mathcal{A} \subset \mathcal{F}$, the following conditions (a)–(c) are equivalent. (a):$\mathcal{A}$ is independent. (b):$\{1_A\}_{A\in\mathcal{A}}$ are independent r.v.'s. (c):$\{\sigma[\{A\}]\}_{A\in\mathcal{A}}$ are independent σ-algebras.

Exercise 8.7.2 In the setting of Definition 8.7.1(a), events in $\mathcal{A} \subset \mathcal{F}$ are (or $\mathcal{A}$ is) said to be *pairwise independent*, if any two events in $\mathcal{A}$ are independent. Consider a probability space $(\Omega, \mathcal{F}, P)$ defined by $\Omega = \{0,1,2,3\}$, $\mathcal{F} = 2^S$ and $P(\{i\}) = 1/4$ for $i \in \Omega$. Check the following statements for events $A_1 = \{1,2\}$, $A_2 = \{2,3\}$ and $A_3 = \{3,1\}$.
i) $\{A_i\}_{i=1}^3$ are pairwise independent, but not independent in the sense of Definition 8.7.1 (a).
ii) $\mathcal{A}_1 = \{A_1\}$ and $\mathcal{A}_{23} = \{A_2, A_3\}$ are quasi-independent in the sense of Definition 8.7.1 (b).
iii) $\sigma(\mathcal{A}_1) = \{\emptyset, \Omega, A_1, A_1^c\}$ and $\sigma(\mathcal{A}_{23}) = \mathcal{F}$. In particular, $\sigma(\mathcal{A}_1)$ and $\sigma(\mathcal{A}_{23})$ are not independent while $\mathcal{A}_1$ and $\mathcal{A}_{23}$ are quasi-independent.

Remark: In Exercise 8.7.2, $P(B|A_1) = P(B)$ for all $B \in \mathcal{A}_{23}$, but not for all $B \in \sigma(\mathcal{A}_{23})$. In particlar, $\{B \in \mathcal{F} \ ; \ P(B|A_1) = P(B)\}$ is not a σ-algebra. cf. Lemma 1.3.1.

Throughout this subsection, we consider the following items;

- A probability space $(\Omega, \mathcal{F}, P)$,
- Measurable spaces $\{(S_\lambda, \mathcal{B}_\lambda)\}_{\lambda\in\Lambda}$ indexed by a set Λ,
- R.v. $X_\lambda : \Omega \to S_\lambda$ for each $\lambda \in \Lambda$.

Definition 8.7.2 A σ-algebra:

$$\sigma\left[X_\lambda^{-1}(B_\lambda) \ ; \ B_\lambda \in \mathcal{B}_\lambda, \ \lambda \in \Lambda\right] \tag{8.50}$$

is called the σ-algebra generated by maps $\{X_\lambda\}_{\lambda\in\Lambda}$ and is denoted by

$$\sigma\left[\{X_\lambda\}_{\lambda\in\Lambda}\right] \ \text{ or } \ \sigma\left[X_\lambda \ ; \ \lambda \in \Lambda\right].$$

The σ-algebra $\sigma\left[\{X_\lambda\}_{\lambda\in\Lambda}\right]$ (cf. (8.50)) is all the information needed to know how the values of $\{X_\lambda\}_{\lambda\in\Lambda}$ *for all* λ are distributed *at the same time*.

Proposition 8.7.3 *For a disjoint decomposition $\Lambda = \cup_{\gamma \in \Gamma} \Lambda(\gamma)$ of the index set Λ, the following conditions are equivalent:*

a) *The σ-algebras*

$$\sigma[X_\lambda \ ; \ \lambda \in \Lambda(\gamma)], \quad \gamma \in \Gamma$$

are independent (cf. Definition 8.7.1(b)).

b) *R.v.'s $\{\widetilde{X}\}_{\gamma \in \Gamma}$ defined by*

$$\widetilde{X}_\gamma : \omega \mapsto (X_\lambda(\omega))_{\lambda \in \Lambda(\gamma)} \in \prod_{\lambda \in \Lambda(\gamma)} S_\lambda, \quad \gamma \in \Gamma. \tag{8.51}$$

are independent.

Definition 8.7.4 Families of r.v.'s

$$\{X_\lambda \ ; \ \lambda \in \Lambda(\gamma)\}, \quad \gamma \in \Gamma \tag{8.52}$$

in Proposition 8.7.3 are said to be *independent* if they satisfy one of (therefore all of) conditions in the corollary.

Proof of Proposition 8.7.3: The equivalence is a consequence of Proposition 1.6.1 and an identity $\sigma[\widetilde{X}_\gamma] = \sigma[X_\lambda \ ; \ \lambda \in \Lambda(\gamma)]$, which can be seen from Lemma 1.5.2. \(^□^)/

Remarks:
1) The independence of the families of r.v.'s (Definition 8.7.4) can be considered as is a special case of the independence of r.v.'s (Proposition 1.6.1), if we consider r.v.'s $\{\widetilde{X}\}_{\gamma \in \Gamma}$ defined by (8.51).

2) In the setting of Proposition 8.7.3, let us consider the following condition:

$$\{X_{\lambda(\gamma)}\}_{\gamma \in \Gamma} \text{ are independent r.v.'s for any choice of } \lambda(\gamma) \in \Lambda(\gamma) \ (\gamma \in \Gamma). \tag{8.53}$$

This condition follows from the independence of the families (8.52). However, the converse is not true. A counterexample is again provided by Exercise 8.7.2. Consider $\{1_{A_1}\}$ and $\{1_{A_2}, 1_{A_3}\}$ there. Since, $\{A_i\}_{i=1}^3$ are pairwise independent, we have (8.53) by Exercise 8.7.1. However, $\{1_{A_1}\}$ and $\{1_{A_2}, 1_{A_3}\}$ are not independent, since $\sigma[\{A_1, A_2\}] = 2^\Omega$.

Exercise 8.7.3 Suppose that $(X_n)_{n \ge 1}$ are $\mathbb{R}^d$-valued independent r.v.'s and let $S_n = X_1 + \cdots + X_n$. Prove then that, for each fixed $m \ge 1$, two families of r.v.'s

$$\{S_n\}_{n=1}^m, \quad \{S_{n+m} - S_m\}_{n \ge 1}$$

are independent. Hint: Note that $\sigma(\{S_n\}_{n=1}^m) = \sigma(\{X_n\}_{n=1}^m)$ and that $\sigma(\{S_{n+m} - S_m\}_{n \ge 1}) = \sigma(\{X_{n+m}\}_{n \ge 1})$. Then, use Exercise 1.6.9.

8.8 (⋆) Order Statistics

Example 8.8.1 $X_1, \ldots, X_n$ be real i.i.d. such that $F(t) = P(X_i \le t)$ is continuous in $t \in \mathbb{R}$. Define $X_{n,k}$ to be the k-th smallest number in $\{X_1, \ldots, X_n\}$ ($k = 1, \ldots n$). Then the distribution of $X_{n,k}$ can be computed as:

$$P\{X_{n,k} \in A\} = n \binom{n-1}{k-1} E\left[F(X_1)^{k-1}(1-F(X_1))^{n-k} 1\{X_1 \in A\}\right] \quad A \in \mathcal{B}(\mathbb{R}).$$

Proof: An rough explanation can be given as follows. First of all, there are n ways to choose $X_{n,k}$ form $X_1, ..., X_n$ and the probability of all such selections are the same (This explains the first factor n). Now, suppose that $X_1 = X_{n,k}$. Then, there are $\binom{n-1}{k-1}$ ways to choose $k-1$ numbers from $X_2, ..., X_n$ which are smaller than X_1 and again by symmetry, these selections have equal probability (This explains the factor $\binom{n-1}{k-1}$). Finally, once such $k-1$ numbers are choosen, say, $X_2, ..., X_k$, then, the probability that

$$X_2, ..., X_k < X_1 < X_{k+1}, ..., X_n, \text{ and } X_1 \in A$$

is $E[F(X_1)^{k-1}(1-F(X_1))^{n-k} : X_1 \in A]$.

We now present a less intuitive, but mathematically clearer proof. Let $\mathcal{S}_n$ denote the set of all permutation of $\{1, 2, ..., n\}$. Then,

$$\begin{aligned}
&P\{X_{n,k} \in A\} \\
&= \sum_{\sigma \in \mathcal{S}_n} P\{X_{\sigma(1)} < X_{\sigma(2)} < ... < X_{\sigma(k)} < ... < X_{\sigma(n)},\ X_{\sigma(k)} \in A\} \\
&= \sum_{\sigma \in \mathcal{S}_n} \int_A P\{X_{\sigma(k)} \in dx\} P\{X_{\sigma(1)} < X_{\sigma(2)} < ... < X_{\sigma(k-1)} < x < X_{\sigma(k+1)} < ... < X_{\sigma(n)}\} \\
&= \sum_{\sigma \in \mathcal{S}_n} \int_A P\{X_{\sigma(k)} \in dx\} P\{X_{\sigma(1)} < X_{\sigma(2)} < ... < X_{\sigma(k-1)} < x\} P\{x < X_{\sigma(k+1)} < ... < X_{\sigma(n)}\} \\
&= \sum_{\sigma \in \mathcal{S}_n} \int_A P\{X_{\sigma(k)} \in dx\} \frac{F(x)^{k-1}}{(k-1)!} \frac{(1-F(x))^{n-k}}{(n-k)!} \\
&= n! \int_A P\{X_1 \in dx\} \frac{F(x)^{k-1}}{(k-1)!} \frac{(1-F(x))^{n-k}}{(n-k)!} \\
&= n \binom{n-1}{k-1} E\left[F(X_1)^{k-1}(1-F(X_1))^{n-k} 1\{X_1 \in A\}\right].
\end{aligned}$$

\(^□^)/

Exercise 8.8.1 Let $U_1, \ldots, U_n$ be i.i.d. with uniform distribution on $[0, 1]$ and $X_1, \ldots, X_{n+1}$ be i.i.d. with $P(X_i \in \cdot) = \gamma_{r,1}$, cf. (1.27). Define $U_{n,k}$ to be the k th smallest number in $\{U_1, \ldots, U_n\}$ ($k = 1, \ldots n$). Prove then that $(U_{n,k})_{k=1}^n$ and $(\sum_{j=1}^k X_j / \sum_{j=1}^{n+1} X_j)_{k=1}^n$ have the same distribution on $\mathbb{R}^n$. In particular, $P(U_{n,k} \in \cdot) = \beta_{k,n+1-k}$ by Example 1.7.5.

8.9 Proof of the Law of Large Numbers: L^1 Case

We may and will assume that $X_n \geq 0$. In fact, $X_n^+ = \max\{X_n, 0\}$ and $X_n^- = \max\{-X_n, 0\}$ satisfy the assumption of the theorem and $X_n = X_n^+ - X_n^-$. Therefore, it is enough to prove the theorem for $X_n^\pm$ separately. Define r.v.'s Y_n and T_n by :

$$Y_n = X_n 1\{X_n \leq n\}, \quad T_n = Y_1 + \ldots + Y_n.$$

We first observe that

1) $$\sum_{n\geq 1} 1\{X_n \neq Y_n\} < \infty \text{ a.s.}$$

This can be seen as follows;

$$\begin{aligned} E\sum_{n\geq 1} 1\{X_n \neq Y_n\} &\stackrel{\text{Fubini}}{=} \sum_{n\geq 1} P\{X_n \neq Y_n\} \\ &\leq \sum_{n\geq 1} P\{X_n > n\} = \sum_{n\geq 1} P\{X_1 > n\} \\ &\leq \sum_{n\geq 1} \int_{n-1}^{n} dt P\{X_1 > t\} = \int_0^\infty dt P\{X_1 > t\} \\ &\stackrel{(1.11)}{=} EX_1 < \infty, \end{aligned}$$

which in particular implies (1).

We see from (1) that Theorem 1.10.2 follows from:

2) $$\lim_{n\to\infty} \frac{T_n}{n} = E[X_1] \text{ a.s.}$$

We first prove (2) along the subsequence $l(n) = \lfloor q^n \rfloor$, where $q > 1$:

3) $$\lim_{n\to\infty} \frac{T_{l(n)}}{l(n)} = E[X_1] \text{ a.s.}$$

Since

$$EY_n = EX_n 1\{X_n \leq n\} = EX_1 1\{X_1 \leq n\} \to EX_1,$$

we have

$$\lim_{n\to\infty} \frac{E[T_n]}{n} = EX_1.$$

Thus, (3) follows from:

4) $$\lim_{n\to\infty} \frac{T_{l(n)} - E[T_{l(n)}]}{l(n)} = E[X_1] \text{ a.s.}$$

To show (4), we prepare the following estimate:

5) $$\text{var}\,(T_n) \leq nE[X_1^2 \mathbf{1}\{X_1 \leq n\}]$$

Indeed,

$$\begin{aligned}\text{var}(T_n) &\stackrel{(1.54)}{=} \sum_{j=1}^n \text{var}(Y_j) \le \sum_{j=1}^n E[Y_j^2] \\ &= \sum_{j=1}^n E[X_1^2 \mathbf{1}\{X_1 \le j\}] \le nE[X_1^2\mathbf{1}\{X_1 \le n\}].\end{aligned}$$

We next observe that

6) $$\sum_{n:l(n)\ge x} \frac{1}{l(n)} \le \frac{2q}{(q-1)x} \text{ for any } x > 0.$$

In fact, let M be the smallest $n \in \mathbb{N}$ such that $l(n) \ge x$. Then, $q^M \ge x$. Note also that $l(n) \ge q^n/2$ for all $n \in \mathbb{N}$. Thus,

$$\sum_{n:l(n)\ge x} \frac{1}{l(n)} \le 2\sum_{n\ge M} q^{-n} = 2q^{-M}\sum_{n\ge 0} q^{-n} \le \frac{2q}{(q-1)x}.$$

With (5) and (6), we proceed as follows:

$$\begin{aligned}E\sum_{n\ge 1}\left|\frac{T_{l(n)} - E[T_{l(n)}]}{l(n)}\right|^2 &= \sum_{n\ge 1} l(n)^{-2}\text{var}(S_{l(n)}) \stackrel{(5)}{\le} E\left[X_1^2 \sum_{n\ge 1} l(n)^{-1}\mathbf{1}\{X_1 \le n\}\right] \\ &\stackrel{(6)}{\le} \frac{2q}{q-1}E[X_1] < \infty.\end{aligned}$$

This implies that $\sum_{n\ge 1}\left|\frac{T_{l(n)}-E[T_{l(n)}]}{l(n)}\right|^2 < \infty$, P-a.s. and therefore (4).

Finally, we get rid of the subsequence in (3). For any n, there is a unique integer k such that

$$l(k) \le n < l(k+1).$$

We have by the positivity of $\{X_m\}$ that

$$l(k+1)^{-1}T_{l(k)} \le n^{-1}T_n \le l(k)^{-1}T_{l(k+1)}.$$

By letting $n \nearrow \infty$, we see from (3) that

$$q^{-1}EX_1 \le \varliminf_{n\nearrow\infty} n^{-1}T_n \le \varlimsup_{n\nearrow\infty} n^{-1}T_n, \le qEX_1,$$

which conclude the proof, since $q > 1$ is arbitrary. \(^□^)/

9 Appendix to Sections 2

We prepare

1) $h_t * f \longrightarrow f$ in $L^1(\mathbb{R}^d)$ as $t \to 0$, where $h_t(x) = (2\pi t)^{-d/2}\exp(-|x|^2/2t)$

We have that

$$|h_t * f - f|(x) \le \int_{\mathbb{R}^d} h_t(y)|f(x-y) - f(x)|dy = \int_{\mathbb{R}^d} h_1(y)|f(x-\sqrt{t}y) - f(x)|dy$$

and hence

2) $$\int_{\mathbb{R}^d} |h_t * f - f|(x)dx \le \int_{\mathbb{R}^d} h_1(y)g_t(y)dy \quad \text{where} \quad g_t(y) = \int_{\mathbb{R}^d} |f(x-\sqrt{t}y) - f(x)|dx.$$

We have for any $y \in \mathbb{R}^d$ that

$$\lim_{t\to 0} g_t(y) = 0 \quad \text{and} \quad 0 \le g_t(y) \le 2\int_{\mathbb{R}^d} |f(x)|dx.$$

Thus, by (2) and the dominated convergence theorem,

$$\lim_{t\to 0} \int_{\mathbb{R}^d} |h_t * f - f|(x)dx = 0.$$

We set $f^\vee(x) = (2\pi)^{-d}\widehat{f}(-x)$ $(x \in \mathbb{R}^d)$. We will next show that:

3) $f * h_t = (f^\wedge h_t^\wedge)^\vee$, where $h_t(x) = (2\pi t)^{-d/2}\exp(-|x|^2/2t)$ $(x \in \mathbb{R}^d,\ t > 0)$.

By (2.10),

4) $h_t^\wedge(\theta) = \exp(-t|\theta|^2/2)$.

Using (2.10) again, we see that $h_t = h_t^{\wedge\vee}$. Therefore,

$$\begin{aligned} f * h_t(x) &= f * h_t^{\wedge\vee}(x) \\ &= (2\pi)^{-d}\int f(x-y)dy\int \underbrace{\exp(-\mathbf{i}\theta\cdot y)}_{=\exp(-\mathbf{i}\theta\cdot x)\exp(\mathbf{i}(\theta\cdot(x-y)))} h_t^\wedge(\theta)d\theta \\ &\overset{\text{Fubini}}{=} (2\pi)^{-d}\int \exp(-\mathbf{i}\theta\cdot x)h_t^\wedge(\theta)d\theta \underbrace{\int f(x-y)\exp(\mathbf{i}(\theta\cdot(x-y)))dy}_{=f^\wedge(\theta)} \\ &= (f^\wedge h_t^\wedge)^\vee(x). \end{aligned}$$

We see from (4) and the dominated convergence theorem that

$$\lim_{t\to 0}(f^\wedge h_t^\wedge)^\vee(x) = f^{\wedge\vee}(x) \quad \text{for all } x \in \mathbb{R}^d.$$

Combining this, (1) and (3), we arrive at $f^{\wedge\vee} = f$,a.e., which is (2.37). \(^□^)/

9.1 Weak Convergence of Finite Measures on a Metric Space

Theorem 9.1.1 *Let S is a metric space with the metric ρ, and let μ_n ($n = 0, 1, ..$) be finite Borel measures on S. Then, the following conditions are equivalent.*

a1)

$$\int f d\mu_n \stackrel{n\to\infty}{\longrightarrow} \int f d\mu_0 \tag{9.1}$$

for any bounded Borel $f : S \to \mathbb{R}$ for which the set of discontinuities is a μ_0-null set.

a2) *(9.1) holds for all $f \in C_{\rm b}(S)$.*

a3) *(9.1) holds for all bounded, Lipschitz continuous $f : S \to \mathbb{R}$.*

b1)

$$\mu_0(B^\circ) \le \varliminf_{n\to\infty} \mu_n(B) \le \varlimsup_{n\to\infty} \mu_n(B) \le \mu_0(\overline{B}) \quad \text{for any Borel } B \subset S. \tag{9.2}$$

b2) $\mu_n(B) \stackrel{n\to\infty}{\longrightarrow} \mu_0(B)$ *for any Borel $B \subset S$ such that $\mu_0(\partial B) = 0$.*

Proof: a1) ⇒ a2) ⇒ a3), and b1) ⇒ b2) are obvious.
a3) ⇒ b1): We see from the proof of Lemma 1.3.2 that

1) for any closed $F \subset S$, there is a sequence of Lipschitz continuous $f_m : S \to [0, 1]$ such that $f_m \searrow \mathbf{1}_F$.

and hence that

2) for any open $G \subset S$, there is a sequence of Lipschitz continuous $g_m : S \to [0, 1]$ such that $g_m \nearrow \mathbf{1}_G$.

By taking $F = \overline{B}$ in 1), we have that

$$\mu_0(\overline{B}) \stackrel{1)}{=} \lim_{m\to\infty} \int f_m d\mu_0 \stackrel{a3)}{=} \lim_{m\to\infty} \lim_{n\to\infty} \int f_m d\mu_n \ge \varlimsup_{n\to\infty} \mu_n(\overline{B}) \ge \varlimsup_{n\to\infty} \mu_n(B).$$

Similarly, by taking $G = B^\circ$ in 2), we have that

$$\mu_0(B^\circ) \stackrel{2)}{=} \lim_{m\to\infty} \int g_m d\mu_0 \stackrel{a3)}{=} \lim_{m\to\infty} \lim_{n\to\infty} \int g_m d\mu_n \le \varliminf_{n\to\infty} \mu_n(B^\circ) \le \varliminf_{n\to\infty} \mu_n(B).$$

b2) ⇒ a1): Let $D_f = \{x \in S \; ; \; f \text{ is discontinuous at } x\}$, which is a μ_0-null set. We first verify that

3) $\partial f^{-1}(A) \subset D_f \cup f^{-1}(\partial A)$ for any $A \subset \mathbb{R}$.

Let us show 3) in the form $\partial f^{-1}(A) \backslash D_f \subset f^{-1}(\partial A)$. Indeed, if $x \in \partial f^{-1}(A) \backslash D_f$, there are sequences $x_n \to x$, $y_n \to x$ such that $f(x_n) \in A$ and $f(y_n) \not\in A$. Since f is continuous at x, we have

$$f(x) = \lim_{n\to\infty} f(x_n) \in \overline{A}, \quad f(x) = \lim_{n\to\infty} f(y_n) \not\in A^\circ,$$

hence $f(x) \in \partial A$.
We next note that the set

$$E_f = \{t \in \mathbb{R} \; ; \; \mu(f^{-1}(t)) > 0\}$$

is at most countable, since E_f is exactly the set of discontinuities of the bounded monotone function $t \mapsto \mu(f^{-1}([0,t]))$. We see from this observation that, for any $\varepsilon > 0$ there are $c_1, ..., c_k \in \mathbb{R} \backslash E_f$ such that

$$f(S) \subset [c_1, c_k), \quad 0 < c_{j+1} - c_j < \varepsilon, \;\; j = 1, ..., k-1.$$

Let $f_\varepsilon : S \to \mathbb{R}$ be defined by

$$f_\varepsilon = \sum_{j=1}^{k-1} c_j \mathbf{1}_{f^{-1}(I_j)}, \;\; \text{with } I_j = [c_j, c_{j+1}).$$

Then, $\sup_S |f - f_\varepsilon| \leq \varepsilon$. Note also that

$$\partial f^{-1}(I_j) \stackrel{3)}{\subset} D_f \cup f^{-1}(\{c_j, c_{j+1}\}),$$

and hence that $\mu_0(\partial f^{-1}(I_j)) = 0$. Therefore, as $n \to \infty$,

$$\Delta_{n,\varepsilon} \stackrel{\text{def}}{=} \left| \int f_\varepsilon d\mu_n - \int f_\varepsilon d\mu_0 \right| \leq \sum_{j=1}^{k-1} |c_j| \mu_n(f^{-1}(I_j)) - \mu_0(f^{-1}(I_j))| \stackrel{\text{b2)}}{\longrightarrow} 0.$$

Finally, we write

$$\left| \int f d\mu_n - \int f d\mu_0 \right| \leq \int |f - f_\varepsilon| d\mu_n + \Delta_{n,\varepsilon} + \int |f - f_\varepsilon| d\mu_0 \leq \Delta_{n,\varepsilon} + 2\varepsilon.$$

By letting $n \to \infty$ first, and then $\varepsilon \searrow 0$, we get (9.1). \(^□^)/

9.2 Some Results from Fourier Transform

Theorem 9.2.1 (Lévy's convergence theorem) *Let $\mu_n \in \mathcal{P}(\mathbb{R}^d)$ $(n \in \mathbb{N})$ and $f : \mathbb{R}^d \to \mathbb{C}$. Suppose that $\lim_{n\to\infty} \mu_n^\wedge(\theta) = f(\theta)$ for all $\theta \in \mathbb{R}^d$ and that the convergence is uniform in $|\theta| \leq \delta$ for some $\delta > 0$. Then, there exists a $\mu \in \mathcal{P}(\mathbb{R}^d)$ such that $f = \mu^\wedge$.*

Theorem 9.2.2 (Bochner's theorem) *Let $f \in C_{\mathrm{b}}(\mathbb{R}^d \to \mathbb{C})$. Then, the following are equivalent:*

a) *There exists a finite measure μ on $\mathbb{R}^d$ such that $f = \mu^\wedge$.*

b) *For any $N \in \mathbb{N}\backslash\{0\}$ and $x_1, ..., x_N \in \mathbb{R}^d$, the $N \times N$ matrix $(f(x_i - x_j))_{i,j=1}^N$ is non-negative definite.*

10 Appendix to Section 3

10.1 True d-dimensionality and Aperiodicity

Definition 10.1.1 A random walk in $\mathbb{R}^d$ is said to be *truly d-dimensional* if

$$\Theta_1 \stackrel{\text{def.}}{=} \{\theta \in \mathbb{R}^d \; ; \; \theta \cdot X_1 = 0, \; P\text{-a.s.}\} = \{0\}. \tag{10.1}$$

Condition (10.1) says that the random walk is not confined in a subspace with positive codimension.

Lemma 10.1.2 *Consider a random walk such that $E[|X_1|^2] < \infty$, and denote its mean vector by m and the covariance matrix by V.*

a)

$$\begin{aligned} \Theta_2 &\stackrel{\text{def}}{=} \{\theta \in \mathbb{R}^d \; ; \; \theta \cdot V\theta = 0\} \\ &= \{\theta \in \mathbb{R}^d \; ; \; \theta \cdot (X_1 - m) = 0, \; P\text{-a.s.}\} \\ &= \{\theta \in \mathbb{R}^d \; ; \; \theta \cdot (X_1 - X_2) = 0, \; P\text{-a.s.}\}. \end{aligned}$$

b) *If* $\det V > 0$, *then the random walk is truly d-dimensional.*

c) *If the random walk is truly d-dimensional and $m = 0$, then* $\det V > 0$.

Proof: **a)**: It is easy to see that for $\theta \in \mathbb{R}^d$,

$$\theta \cdot V\theta = E\left[|(X_1 - m) \cdot \theta|^2\right] = \tfrac{1}{2} E\left[|(X_1 - X_2) \cdot \theta|^2\right],$$

from which the equalities follow.
b): $\det V > 0$ is equivalent to that $\Theta_2 = \{0\}$. Hence, it is enough to prove that $\Theta_1 \subset \Theta_2$. But this is clear from a).
c): If $m = 0$, then a) shows that $\Theta_1 = \Theta_2$. \(^□^)/

Example 10.1.3 Suppose that $P(X_1 \in \{0, \pm e_1, ..., \pm e_d\}) = 1$ and set $p(x) = P(X_1 = x)$ $(x \in \mathbb{Z}^d)$. Then, the random walk is truely d-dimensional iff

$$p(e_\alpha) \vee p(-e_\alpha) > 0 \text{ for all } \alpha = 1, ..., d. \tag{10.2}$$

(See also Example 3.2.3.)

Proof: Suppose (10.2) and define, for $\alpha = 1, ..., d$,

$$\widetilde{e}_\alpha = \begin{cases} e_\alpha & \text{if } p(e_\alpha) > 0, \\ -e_\alpha & \text{if } p(e_\alpha) = 0 \text{ and } p(-e_\alpha) > 0. \end{cases}$$

Then, $\{\widetilde{e}_\alpha\}_{\alpha=1}^d$ is a basis of $\mathbb{R}^d$. Now, take any $\theta \in \Theta_1$. Then, $\theta \cdot \widetilde{e}_\alpha = 0$ for all $\alpha = 1, ..., d$, since $p(\widetilde{e}_\alpha) > 0$. Hence $\theta = 0$.
Suppose on the contrary that (10.2) fails. Then, there is an $\alpha = 1, ..., d$ such that $p(\pm e_\alpha) = 0$. Then, $e_\alpha \in \Theta_1$. \(^□^)/

Proposition 10.1.4 *Let $(S_n)_{n\geq 0}$ be a truly d-dimensional random walk with $\nu = P\{X_1 \in \cdot\}$. Then,*

a) *There exist $\delta_i > 0$, $i = 1, 2$ such that*

$$1 - \mathrm{Re}\widehat{\nu}(\theta) \geq \delta_1 |\theta|^2 \quad \text{if } |\theta| \leq \delta_2. \tag{10.3}$$

b) *The random walk is transient if $d \geq 3$.*

Proof: a) The proof is based on the observation that the expectation $E\left[|\sigma \cdot X_1|^2\right]$ (can be $+\infty$, but) can never be zero for $\sigma \neq 0$. Recall that

$$\begin{aligned} 1 - \cos t &= 2\sin^2(t/2) \leq t^2/2, \quad t \in \mathbb{R}, && (10.4) \\ |\sin t| &\geq \frac{2}{\pi}|t|, \quad |t| \leq \frac{\pi}{2}. && (10.5) \end{aligned}$$

We now use (10.4) and (10.5) as follows;

$$\begin{aligned} 1 - \mathrm{Re}\widehat{\nu}(\theta) &= E[1 - \cos(\theta \cdot X_1)] \\ &= 2E[\sin^2(\theta \cdot X_1/2)] \\ &\geq 2E\left[\frac{4}{\pi^2}\frac{|\theta \cdot X_1|^2}{4} \ : \ |\theta \cdot X_1| \leq \pi\right] \\ &= \frac{2|\theta|^2}{\pi^2} F(|\theta|, \theta/|\theta|), \end{aligned}$$

where on the last line, we have introduced

$$\begin{aligned} F(\delta, \sigma) &= E\left[|\sigma \cdot X_1|^2 \ : \ |\sigma \cdot X_1| \leq \pi/\delta\right], \\ &\quad \delta > 0, \ \sigma \in S^{d-1} = \{y \in \mathbb{R}^d \ ; \ |y| = 1\}. \end{aligned}$$

Hence it is enough to show that there exists $\delta_2 > 0$ such that

$$\inf\{F(\delta, \sigma) \ ; \ \delta < \delta_2, \sigma \in S^{d-1}\} > 0. \tag{10.6}$$

Since $F(\delta, \sigma)$ is decreasing in δ, (10.6) is equivalent to;

$$\inf\{F(\delta, \sigma) \ ; \ \sigma \in S^{d-1}\} > 0 \ \text{ for some } \delta > 0. \tag{10.7}$$

We prove (10.7) by contradiction. Suppose that (10.7) is false. Then, there is $\delta_n \searrow 0$ and $\{\sigma_n\}_{n\geq 1} \subset S^{d-1}$ such that $\lim\limits_{n\to\infty} F(\delta_n, \sigma_n) = 0$. By the compactness of S^{d-1} and by taking a subsequence, we may assume that $\lim\limits_{n\to\infty} \sigma_n = \sigma$ for some $\sigma \in S^{d-1}$. Then, by Fatou's lemma,

$$\lim_{n\to\infty} F(\delta_n, \sigma_n) \geq E\left[|\sigma \cdot X_1|^2\right] \neq 0,$$

which is a contradiction.
b) This follows from (10.3) and Proposition 3.4.1 with $\alpha = 2$. \\(^□^)/

Definition 10.1.5 • For $\mathbb{Z}^d$-valued random walk, we define

$$\mathcal{R}_n = \{z \in \mathbb{Z}^d \ ; \ P\{S_n = z\} > 0\}. \tag{10.8}$$

- A $\mathbb{Z}^d$-valued random walk is said to be *aperiodic* if

$$\{x-y\ ;\ x,y \in \cup_{n\geq 1}\mathcal{R}_n\} = \mathbb{Z}^d. \tag{10.9}$$

  If otherwise, the random walk is called *periodic.*

Remark 1) The left-hand side of (10.9) is nothing but the Abelian subgroup of $\mathbb{Z}^d$ generated by $\mathcal{R}_1$.
2) The definition of aperiodicity is the same as that in [Spi76, page 20]. However, the aperiodicity defined here is weaker notion than the "aperiodicity" as a Markov chain. The "aperiodicity" as a Markov chain is called "strong aperiodicity" in [Spi76, page 42].

Lemma 10.1.6 *Let* $(S_n)_{n\geq 0}$ *be a* $\mathbb{Z}^d$*-valued random walk.*

(a)

$$\mathcal{R}_n = \{x_1+...+x_n\ ;\ x_i \in \mathcal{R}_1\}. \tag{10.10}$$

(b) $(S_n)_{n\geq 0}$ *is truly d-dimensional if and only if* $\mathcal{R}_1$ *contains a linear basis of* $\mathbb{R}^d$.

(c) *Aperiodicity implies true d-dimensionality.*

Proof: **(a) & (b)**: Obvious from the definitions.
(c): This follows from (a),(b) and simple linear algebra. \(^□^)/

Example 10.1.7 If $\{e_1,...,e_d\} \subset \mathcal{R}_1$, where $e_i = (\delta_{ij})_{i=1}^d \in \mathbb{Z}^d$, we then see from (10.10) that the random walk is aperiodic. In particular, the simple random walk is aperiodic.

Proposition 10.1.8 *Let* $(S_n)_{n\geq 0}$ *be an aperiodic random walk with with* $\nu = P\{X_1 \in \cdot\}$. *Then,*

a)

$$\{\theta \in \mathbb{R}^d\ ;\ \widehat{\nu}(\theta) = 1\} = \{2\pi m\ ;\ m \in \mathbb{Z}^d\}. \tag{10.11}$$

b) *There exists* $\delta > 0$ *such that*

$$1 - \mathrm{Re}\widehat{\nu}(\theta) \geq \delta|\theta|^2 \quad \textit{if } \theta \in [-\pi,\pi]^d. \tag{10.12}$$

c) *The random walk is transient if* $d \geq 3$.

Proof: a) Let $(S'_n)_{n\geq 0}$ be an independent copy of $(S_n)_{n\geq 0}$. We first observe that

$$\cup_{n,n'\geq 0}\{x \in \mathbb{Z}^d\ ;\ P\{S_n - S'_{n'} = x\} > 0\} = \mathbb{Z}^d. \tag{10.13}$$

This can be seen as follows. For any $x \in \mathbb{Z}^d$, there are $n,n' \geq 0$ and $y \in \mathcal{R}_n$, $y' \in \mathcal{R}_{n'}$ such that $x = y - y'$. Then,

$$\begin{aligned} P\{S_n - S'_{n'} = x\} &\geq P\{S_n = y, S'_{n'} = y'\} \\ &= P\{S_n = y\}P\{S'_{n'} = y'\} > 0. \end{aligned}$$

We also observe that for $t \in \mathbb{R}$ and a real r.v. X,

$$E\exp(\mathbf{i}X) = \exp(\mathbf{i}t) \iff E\cos(X-t) = 1 \iff X \in \{t + 2\pi m\}_{m\in\mathbb{Z}},\ P\text{-a.s.} \qquad (10.14)$$

Let $S'_n = X'_1 + ... + X'_n$. We then have that

$$\begin{aligned}
\widehat{\nu}(\theta) = 1 &\iff E\exp(\mathbf{i}\theta\cdot X_1) = E\exp(\mathbf{i}\theta\cdot X'_1) = 1\\
&\iff E\exp(\mathbf{i}\theta\cdot S_n) = E\exp(\mathbf{i}\theta\cdot S'_{n'}) = 1, \quad \text{for all } n, n' \geq 1,\\
&\Rightarrow E\exp(\mathbf{i}\theta\cdot(S_n - S'_{n'})) = 1, \quad \text{for all } n, n' \geq 1,\\
&\iff \theta\cdot(S_n - S'_{n'}) \in \{2\pi m\}_{m\in\mathbb{Z}}, \quad P\text{-a.s. for all } n, n' \geq 1, \quad \text{by (10.14)}\\
&\iff \theta\cdot x \in \{2\pi m\}_{m\in\mathbb{Z}}, \quad \text{for all } x \in \mathbb{Z}^d, \quad \text{by (10.13)}\\
&\iff \theta \in \{2\pi m\}_{m\in\mathbb{Z}^d}
\end{aligned}$$

b) We see from (10.3) that (10.12) is valid for $|\theta| \leq \delta_2$. We next prove (10.3) for the case $|\theta| \geq \delta_2$. By (10.11), $\{\theta \in \pi I\ ;\ \widehat{\nu}(\theta) = 1\} = \{0\}$. Therefore, if we set $K = \{\theta \in \pi I\ ;\ |\theta| \geq \delta_2\}$, then $\theta \in K \mapsto 1 - \mathrm{Re}\widehat{\nu}(\theta)$ attains a positive minimum $=: \delta_3 > 0$. Hence for $|\theta| \geq \delta_2$,

$$1 - \mathrm{Re}\widehat{\nu}(\theta) \geq \delta_3 \geq \delta_3\delta_2^{-1}|\theta|^2.$$

c) This follows from (10.12) and Proposition 3.4.1 with $\alpha = 2$.

\(^□^)/

10.2 Strong Markov Property for IID Sequence

Lemma 10.2.1 (Strong Markov Property) *Let $(S, \mathcal{B})$ be a measurable space and $X_n : \Omega \to S$, $n \in \mathbb{N}\backslash\{0\}$ be i.i.d. Suppose that T is a stopping time such that $P(T < \infty) > 0$. Then, under the measure $P(\cdot|T < \infty)$,*

a) *$\mathcal{F}_T$ and $(X_{T+n})_{n\geq 1}$ are independent,*

b) *$(X_{T+n})_{n\geq 1}$ is an i.i.d. $\approx X_1$.*

Proof: It is enough to prove that

1) $P(A \cap \{(X_{T+k})_{k=1}^n \in B\} \mid T < \infty) = P(A \mid T < \infty)P((X_k)_{k=1}^n \in B)$

for all $A \in \mathcal{F}_T$, $n \geq 1$ and $B \in \mathcal{B}(S^n)$. This can be seen as follows,

$$\begin{aligned}
&P(\{T < \infty\} \cap A \cap \{(X_{T+k})_{k=1}^n \in B\})\\
&= \sum_{m\geq 1} P(\{T = m\} \cap A \cap \{(X_{m+k})_{k=1}^n \in B\})\\
&= \sum_{m\geq 1} P(\{T = m\} \cap A)P((X_{m+k})_{k=1}^n \in B)\\
&= P(\{T < \infty\} \cap A)P((X_k)_{k=1}^n \in B).
\end{aligned}$$

which is equivalent to 1). \(^□^)/

Exercise 10.2.1 The purpose of this exercise is to illustrate that property **(a)** in Lemma 10.2.1 is not true in general if we assume $\{X_n\}_{n\geq 1}$ just to be independent (not necessarily identically distributed). Consider $S_n = X_1 + \ldots + X_n$ where $\{X_n\}_{n\geq 1}$ are $\{1,2\}$-valued independent r.v.'s such that $P(X_j = 1) = 1/2$, $(j \leq 2)$ $P(X_k = 1) = p$ $(k \geq 3)$. We set $t = \inf\{n \geq 1 \,|S_n \geq 2\,\}$. Prove then that two events $\{T = 1\}$ and $\{X_{T+1} = 1\}$ are independent if and only if $p = 1/2$.

10.3 Green Function and Hitting Times

Exercise 10.3.1 Prove that for any $x, y \in \mathbb{R}^d$,

$$1 - h(x+y) \geq \max\{P\{T_x < T_{x+y}\}(1 - h(y)),\ P\{T_y < T_{x+y}\}(1 - h(x))\}. \tag{10.15}$$

Hint: Let us prove that $1 - h(x+y) \geq P\{T_x < T_{x+y}\}(1-h(y))$. To do so, we may assume that $h(x) > 0$ ($P\{T_x < T_{x+y}\} = 0$ if otherwise). Since $h(x) = P\{T_x < \infty\}$, we have

$$\begin{aligned} 1 - h(x+y) &= P\{T_{x+y} = \infty\} \\ &\geq P\{T_x < T_{x+y},\ \widetilde{T}_y = \infty\}, \end{aligned}$$

where

$$\widetilde{T}_y = \inf\{n \geq 1\ ;\ X_{T_x+1} + \ldots + X_{T_x+n} = y\}.$$

Therefore, by Lemma 10.2.1,

$$\begin{aligned} P\{T_x < \infty,\ \widetilde{T}_y = \infty\} &= P\{T_x < T_{x+y}\}P\{\widetilde{T}_y = \infty \mid T_x < \infty\} \\ &= P\{T_x < T_{x+y}\}P\{T_y = \infty\} \\ &= P\{T_x < T_{x+y}\}(1 - h(y)). \end{aligned}$$

By exchanging the role of x and y, we also see that $1 - h(x+y) \geq P\{T_y < T_{x+y}\}(1 - h(x))$.

Exercise 10.3.2 Use a similar argument in the proof (10.15) to show that

$$h(x+y) \geq h(x)h(y) \quad \text{for any } x, y \in \mathbb{R}^d. \tag{10.16}$$

Exercise 10.3.3 Generalize (3.13) by showing

$$h_s(z) = s(1 - h_s(0))P\{X_1 = z\} + sPh_s(z - X_1), \quad z \in \mathbb{R}^d,\ 0 \leq s < 1. \tag{10.17}$$

Exercise 10.3.4 Consider a symmetric, $\mathbb{Z}^d$-valued, aperiodic random walk such that $E[|X_1|^2] < \infty$.

i) Use (3.29) to prove that

$$P\{S_n = x\} = (2\pi)^{-d} \int_{\pi I} d\theta \cos(\theta \cdot x)\widehat{\nu}(\theta)^n \tag{10.18}$$

Hint: $P\{S_n = x\} = \frac{1}{2}P\{S_n = x\} + \frac{1}{2}P\{S_n = -x\}$ by symmetry.

ii) Use (10.18) to show that the following for any $d \geq 1$;

$$a(x) \stackrel{\text{def.}}{=} \lim_{n\to\infty} \sum_{k=0}^{n} \{P(S_k = 0) - P(S_k = x)\} \tag{10.19}$$

$$= (2\pi)^{-d} \int_{\pi I} d\theta \frac{1-\cos(\theta\cdot x)}{1-\widehat{\nu}(\theta)}, \tag{10.20}$$

$$= \lim_{s \nearrow 1} (g_s(0) - g_s(x)). \tag{10.21}$$

The function $a(x)$ is called the *potential kernel* of the random walk. Hint: Use (10.12) and an inequality $1 - \cos(\theta \cdot x) \leq (\theta \cdot x)^2/2$ to prove

$$\int_{\pi I} d\theta \sup_{0\leq s\leq 1} \left| \frac{1-\cos(\theta\cdot x)}{1-s\widehat{\nu}(\theta)} \right| < \infty. \tag{10.22}$$

Then, use (10.18), (10.22) and the dominated convergence theorem to prove (10.20) and (10.21).

Remark 10.3.1 **i)** We will see in (10.24) that $a(z)$ has the following probabilistic meaning;

$$a(z) = E\left[\sum_{n=0}^{T_z-1} 1\{S_n = 0\}\right] / (1 + h(z)).$$

ii) The symmetry we have assumed to prove the existence of the limit (10.19) is not essential, but to simplify the discussion for $d = 1$. In fact, for $d \geq 2$, we can prove the existence of the limit (10.19) and (10.21) without symmetry by (3.29), since $|1 - \exp(\mathbf{i}\theta \cdot x)| \leq |\theta \cdot x|$. Even for $d = 1$, it is known that the limit (10.19) exists without symmetry [Spi76, page 352].

Exercise 10.3.5 Consider a $\mathbb{Z}$-valued random walk such that $P\{X_1 = 0\} = r$ and $P\{X_1 = \pm 1\} = \frac{1-r}{2}$. Use Exercise 3.4.3 and (10.21) to compute $a(x)$ in Exercise 10.3.4 explicitly;

$$a(x) = |x|/(1-r).$$

Exercise 10.3.6 Consider a symmetric, $\mathbb{Z}^d$-valued, aperiodic random walk such that $E[|X_1|^2] < \infty$. Use (10.21) and (3.43) to prove that

$$g_1^{\mathbb{Z}^d\setminus\{z\}}(x,y) = a(z-x) + h(z-x)a(y-z) - a(y-x). \tag{10.23}$$

and in particular ($x = y = 0 \neq z$) that

$$a(z) = g_1^{\mathbb{Z}^d\setminus\{z\}}(0,0)/(1+h(z)). \tag{10.24}$$

Exercise 10.3.7 Consider a symmetric, $\mathbb{Z}^d$-valued, aperiodic random walk such that $E[|X_1|^2] < \infty$. Use (10.21) and (3.43) to prove that, if $A \subset \mathbb{Z}^d$ is finite, then

$$a(y-x) = -g_1^A(x,y) + \sum_{z\in\mathbb{Z}^d\setminus A} H_1^A(x,z)a(y-z), \quad x,y \in A. \tag{10.25}$$

cf. [Law91, Proposition 1.6.3] for the simple random walk case.

References

[Ahl79] Ahlfors, L. V.: Complex Analysis, 3rd. ed. McGraw-Hill.

[Bil95] Billingsley, P: "Probability and Measure", 3rd Ed. Wiley Series in Probability and Statistics, 1995.

[Chu74] Chung, K.-L. :"A Course in Probability Theory", Harcourt Brace & World

[DG78] Dubins, Lester E.; Gilat, David On the distribution of maxima of martingales.Proc. Amer. Math. Soc.68(1978), no.3, 337–338.

[Dud89] Dudley, R.: "Real Analysis and Probability", Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, Calfornia.

[Dur84] Durrett, R. :"Brownian Motion and Martingales in Analysis", Wadsworth & Brooks/Cole Advanced Books & Software, Belmont, Calfornia.

[Dur95] Durrett, R. :"Probability–Theory and Examples", 2nd Ed., Duxbury Press, 1995.

[Fol76] Folland, G. B.: Introduction to partial differential equations, Princeton University Press, 1976.

[IkWa89] Ikeda, N. and Watanabe, S. : Stochastic Differential Equations and Diffusion Processes (2nd ed.), North-Holland, Amsterdam / Kodansha,Tokyo (1989).

[KS91] Karatzas, I. and Shreve, S. E.: Brownian Motion and Stochastic Calculus, Second Edition. Springer Verlag (1991).

[Kre89] Kreyszig, E. : Introductory Functional Analysis with Applications (Wiley Classics Library)

[Law91] Lawler, G. F.: Intersections of Random Walks: Birkhäuser.

[Leb72] Lebedev, N. N.: Special Functions & Their Applicatinns, Dover, 1972.

[LeG16] Le Gall, J.-F.: Brownian Motion, Martingales, and Stochastic Calculus, Springer Verlag (2016).

[MP10] Möters, P., Peres, Y. "Brownian Motion" Cambridge University Press (2010).

[RS80] Reed, M. and Simon, B. "Method of Modern Mathematical Physics II" Academic Press 1980.

[Rud87] Rudin, W.: "Real and Complex Analysis—3rd ed." McGraw-Hill Book Company, 1987.

[Spi76] Spitzer, F.: "Principles of Random Walks", Springer Verlag, New York, Heiderberg, Berlin (1976).

[Wal82] Walters, P.: "An Introduction to Ergodic Theory", Springer Verlag, New York, Heiderberg, Berlin (1976).

Index