A short course in probability[1]

Nobuo Yoshida[2]

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

[1]Notes for a course for undergraduates "Kakuritsuron" given at Kyoto University. printed out November 5, 2009.

[2][Mailing address] Division of Mathematics, Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan, [e-mail] nobuo@math.kyoto-u.ac.jp, [URL] http://www.math.kyoto-u.ac.jp/~nobuo

0.1 Notations

For a set S,

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

0.2 References

References

[Dud] Dudley, R.: "Real Analysis and Probability", Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, Calfornia.

[Dur95] Durrett, R. :"Probability–Theory and Examples", 2nd Ed., Duxbury Press, 1995.

[Law91] Lawler, G. F.: Intersections of Random Walks: Birkhäuser.

[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).

1 Random Variables

1.1 Measurability and probability

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 $\mathcal{B}$ is a *σ-field* on a set S, 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 $(\Omega, \mathcal{F})$, $(S, \mathcal{B})$ be measurable spaces.

• A map $X : \Omega \to S$ is called *measurable* if

$$\{X^{-1}(B) \ ; \ B \in \mathcal{B}\} \subset \mathcal{F}. \tag{1.1}$$

Example 1.1.2 (*The Borel σ-field*) When S is a topological space, we let $\mathcal{B}(S)$ denote the Borel σ-field of S, i.e., the σ-field generated by 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) $\mu(S) = 1$.

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

• A measure space $(S, \mathcal{B}, \mu)$ is called a *a probability space* if μ is a probability measure on (S, μ).

1.2 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;

$$X = \begin{cases} +1 & \text{if the coin falls head,} \\ -1 & \text{if the coin falls tail.} \end{cases} \tag{1.2}$$

The value X is not always the same (may be -1 today and $+1$ tomorrow) 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" $P\{X = +1\}$ of the "event" $\{\omega \in \Omega; X = +1\}$. We now set up a mathematical ground on which this situation can be described.

Definition 1.2.1 (Events and random variables) 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.

• An element of the σ-field $\mathcal{F}$ is called an *event*.

• $X : \Omega \to S$ 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.

• We introduce the following notation

$$\mathcal{P}(S, \mathcal{B}) = \{\mu \ ; \ \mu \text{ is a probability measure on } (S, \mathcal{B})\}. \tag{1.3}$$

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

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

$$\mu(B) = P\{\omega \in \Omega \ ; \ X(\omega) \in B\}, \quad B \in \mathcal{B}. \tag{1.4}$$

• For a r.v. $X : \Omega \to S$, *the σ-field generated by* X is defined by

$$\sigma[X] = \{X^{-1}(B) \ ; \ B \in \mathcal{B}\}. \tag{1.5}$$

The σ-field $\sigma[X]$ (cf. (1.5)) is, roughly speaking, all the information needed to know how the values of X are distributed.

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

• The set $\{\omega \in \Omega \ ; \ X(\omega) \in B\}$ will often be abbreviated by $\{X \in B\}$, so that the right-hand-side of (1.4) becomes $P\{X \in B\}$.

• The distribution of a r.v. X, i.e., the measure defined by the right-hand-side of (1.4) will often be denoted by $P\{X \in \cdot\}$.

Exercise 1.2.1 Prove that $\{X^{-1}(B) \ ; \ B \in \mathcal{B}\}$ in (1.5) is a σ-field.

• Let $(\Omega, \mathcal{F}, P)$, $(S, \mathcal{B})$ and $X : \Omega \to S$ be as in Definition 1.2.1 for the rest of this subsection.

Example 1.2.2 (*Identity map on the state space*) Let a measurable space $(S, \mathcal{B})$ and a measure $\mu \in \mathcal{P}(S, \mathcal{B})$ be given. If we take $(\Omega, \mathcal{F}, P) = (S, \mathcal{B}, \mu)$ and $X(\omega) = \omega$, then X is a r.v. on $(\Omega, \mathcal{F}, P)$ with the distribution μ. In fact, X is clearly measurable and

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

Moreover, $\sigma[X] = \mathcal{F}$.

Example 1.2.3 (*Unit interval as a probability space*) Let:

• $\Omega = [0,1)$, $\mathcal{F} = \mathcal{B}([0,1))$, P=the Lebesgue measure on $[0,1)$,

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

Let us find a r.v. $X : \Omega \to S$ with the distribution μ. To do so, we split $[0,1)$ into disjoint intervals $\{I_s\}_{s\in S}$ with length $|I_s| = \mu(s)$ for each $s \in S$ and define

$$X(\omega) = s \ \text{ if } \omega \in I_s.$$

Then, X is measurable. Moreover, we have that

$$P\{X = s\} = |I_s| = \mu(s), \ \text{ for any } s \in S.$$

and hence for any $B \in \mathcal{B}$,

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

Moreover, $\sigma[X] = \left\{\bigcup_{s\in T} I_s \ ; \ T \subset S\right\}$.

Definition 1.2.4 (Expectation and (co)variance)

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

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

for the traditional[3] reason.

• For $X, Y \in L^2(P)$, we set

$$\begin{aligned} \mathrm{cov}(X,Y) &= E\left((X - EX)(Y - EY)\right) \\ &= E(XY) - E(X)E(Y), \qquad (1.7) \\ \mathrm{var}(X) &= \mathrm{cov}(X,X). \qquad (1.8) \end{aligned}$$

$\mathrm{cov}(X,Y)$ is called the *covariance* or *correlation* of X and Y. $\mathrm{var}(X)$ is called the *variance* of X.

Exercise 1.2.2 Let $X : \Omega \to S$ be a r.v. with the distribution μ. For a measurable function $f : S \to [0,\infty]$, prove that

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

Use this to conclude that $f(X) \in L^1(P) \iff f \in L^1(\mu)$ and that (1.9) holds true for $f \in L^1(\mu)$.

Exercise 1.2.3 (**Chebyshev's inequality**[4]) Prove that

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

[3]Though this may not be a good tradition, we follow it to be consistent with the other standard texts in probability theory. Notations (1.6) are also used to denote the expectations for complex or vector valued r.v.

[4]Pafnuty L. Chebyshev, 1821–1894

Exercise 1.2.4 Let $-\infty < a < b < \infty$ and suppose that $X \in L^1(P)$ satisfies $X \le b$ a.s. Prove then that

$$P\{X \le a\} \le \frac{b - EX}{b-a}.$$

Exercise 1.2.5 Suppose that $f \in C^1([0,\infty) \to \mathbb{R})$ is non-decreasing. Use $f(x) - f(0) = \int_0^x dt f'(t)$ and Fubini's theorem to prove;

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

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

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

Exercise 1.2.6 Suppose that $f : \mathbb{N} \to \mathbb{R}$ is non-decreasing. Use $f(n) - f(0) = \sum_{j=1}^x (f(j) - f(j-1))$ and Fubini's theorem to prove that

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

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

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

Definition 1.2.5 (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

$$P(A|B) = P(A \cap B)/P(B), \quad A \in \mathcal{F}. \tag{1.13}$$

Exercise 1.2.7 Suppose that $B = \sum_{i=1}^n B_i$, where $B_i \in \mathcal{F}$ and $P(B_i) > 0$. Prove then that $P(A|B) = \sum_{i=1}^n P(A|B_i)P(B|B_i)$ for any $A \in \mathcal{F}$.

1.3 Examples

Example 1.3.1 (*Uniform distribution*) Let $I = (a,b) \subset \mathbb{R}$. A real r.v. U is said to be *uniformly distributed* on I if

$$P\{U \in B\} = \tfrac{1}{b-a} \int_{B \cap I} dt \text{ for all } B \in \mathcal{B}(\mathbb{R}). \tag{1.14}$$

A Borel probability measure on $\mathbb{R}$ defined by the right hand side of (1.14) is called the *uniform distribution* on I. Uniform distribution is concentrated on I and therefore, U is in fact an I-valued r.v.

Exercise 1.3.1 For U in Example 1.3.1, show that $E[U] = (a+b)/2$ and $\mathrm{var}(U) = (b-a)^2/12$.

Example 1.3.2 (*Gaussian distribution*) Let V be a symmetric, strictly positive definite $d \times d$-matrix. An $\mathbb{R}^d$-valued r.v. X is called a *Gaussian* r.v. with the covariance matrix V if

$$P\{X \in B\} = \nu_V(B) \stackrel{\text{def.}}{=} (\det(2\pi V))^{-1/2} \int_B dx \exp\left(-\tfrac{1}{2} x \cdot V^{-1} x\right) \quad \text{for } B \in \mathcal{B}(\mathbb{R}^d). \tag{1.15}$$

The measure ν_V is called *normal* (or *Gaussian*) distribution with the covariance matrix V. When V is the identity matrix I, ν_I is called the *standard normal* (or *standard Gaussian*) distribution. If $d = 1$, then the matrix V is just a positive number $V = v > 0$. In this case, ν_v is called normal (or Gaussian) distribution with the variance v.

Exercise 1.3.2 For $X = (X_i)_{i=1}^d$ in Example 1.3.2, prove that $V = (\mathrm{cov}(X_i, X_j))_{1 \le i,j \le d}$.

Example 1.3.3 (*Gamma, Beta, exponential, and χ^2 distributions*) We define the Gamma function and the Beta function as usual;

$$\Gamma(s) = \int_{(0,\infty)} x^{s-1} e^{-x} dx, \quad s \in \mathbb{C},\ \mathrm{Re}(s) > 0 \tag{1.16}$$

$$B(a,b) = \int_{(0,1)} x^{a-1}(1-x)^{b-1} dx, \quad a > 0, b > 0. \tag{1.17}$$

Let $a > 0$, $b > 0$ and $r > 0$.

• We define (r,a)-*gamma distribution* $\gamma_{r,a} \in \mathcal{P}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and (a,b)-*beta distribution* $\beta_{a,b} \in \mathcal{P}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ respectively by

$$\gamma_{r,a}(B) = \Gamma(a)^{-1} \int_{B \cap (0,\infty)} (rx)^{a-1} e^{-rx} r dx, \quad \text{for } B \in \mathcal{B}(\mathbb{R}), \tag{1.18}$$

$$\beta_{a,b}(B) = B(a,b)^{-1} \int_{B \cap (0,1)} x^{a-1}(1-x)^{b-1} dx \quad \text{for } B \in \mathcal{B}(\mathbb{R}), \tag{1.19}$$

There are two important special cases of $\gamma_{r,a}$:

• $\gamma_{r,1}$ is called the (r)-*exponential distribution.*

• $\gamma_{1/2,n/2}$ ($n \in \mathbb{N}^*$) is called the χ_n^2-*distribution.*

In particular, χ_2^2 is $(1/2)$-exponential distribution. Note also that $\beta_{1,1}$ is the uniform distribution on $(0,1)$.

Exercise 1.3.3 Let $P(X \in \cdot) = \gamma_{r,a}$ and $P(Y \in \cdot) = \beta_{a,b}$ (cf. Example 1.3.3). Verify then that $EX = a/r$, $\mathrm{var}(X) = a/r^2$, $EY = a/(a+b)$, $\mathrm{var}(Y) = \frac{ab}{(a+b)^2(a+b+1)}$.

Exercise 1.3.4 Let $P(X \in \cdot) = \gamma_{r,a}$ (cf. Example 1.3.3) and show that $E[X^p] = \frac{\Gamma(p+a)}{\Gamma(a)} r^{-p}$ for $p \ge 0$.

Exercise 1.3.5 Let $X = (X_i)_{i=1}^d$ be a r.v. with standard normal distribution on $\mathbb{R}^d$ (Example 1.3.2). Prove then the following.
(i) $|X|^2 = X_1^2 + ... + X_d^2$ has χ_d^2-distribution. (Hint: Polar coodinates).
(ii) $E[|X|^p] = \frac{\Gamma(\frac{p+1}{2})}{\Gamma(\frac{d}{2})} 2^{\frac{p}{2}}$ for $p \ge 0$ (Hint: Exercise 1.3.4).

Exercise 1.3.6 Let X be a positive r.v. Prove then that the following are equivalent: (i) $P(X \in \cdot) = \gamma_{r,1}$ for some $r \in (0, \infty)$. (ii) $P(X > t+s|X > s) = P(X > t) > 0$ for any $t, s \geq 0$. (The property (ii) is referred to as the "memoryless property".)

Exercise 1.3.7 Let $P(X \in \cdot) = \gamma_{r,a}$. Prove then that
(i) $P(X/c \in \cdot) = \gamma_{rc,a}$ for $c > 0$,
(ii) $P(X^{1/2} \in B) = (2r^a/\Gamma(a)) \int_{B\cap(0,\infty)} x^{2a-1}e^{-rx^2}dx$ for $B \in \mathcal{B}(\mathbb{R})$,
(iii) $P(1/X \in B) = (r^a/\Gamma(a)) \int_{B\cap(0,\infty)} x^{-(a+1)}e^{-r/x}dx$ for $B \in \mathcal{B}(\mathbb{R})$.

Exercise 1.3.8 (i) Let U be a r.v. with uniform distrubution on $(0, 1)$. Prove then that $X = (1/r)\log(1/U)$ has (r)-exponential distribution. (ii) Let X has (r)-exponential distribution. Prove then that $U = \exp(-rX)$ has uniform distrubution on $(0, 1)$.

Exercise 1.3.9 A beta $(1/2, 1/2)$ distribution is more commonly called an *arcsin law* for the following reason; for $0 < a < b < 1$,

$$\beta_{1/2,1/2}([a, b]) = \frac{2}{\pi}\int_{\sqrt{a}}^{\sqrt{b}} (1 - x^2)^{-1/2}dx = \frac{2}{\pi}\left(\arcsin\sqrt{b} - \arcsin\sqrt{a}\right).$$

Prove this.

Exercise 1.3.10 ($\star$) Let $p \in [0, 1]$ and $1 \leq k \leq n$ be integers. Show that

$$\beta_{k,n-k+1}((0, p]) = \sum_{r=k}^{n}\binom{n}{r}p^r(1 - p)^{n-r}.$$

Exercise 1.3.11 ($\star$) Let X be a r.v. with $(r, 1/2)$-Gamma distribution. Prove the following;

$$E\exp(-\theta/X) = \exp(-2\sqrt{r\theta}), \quad \text{for } \theta > 0. \tag{1.20}$$

Hint: Define $f_a : (0, \infty) \to (0, \infty)$ by $f_a(x) = \int_0^\infty t^{a-1}\exp\left(-t - \frac{x^2}{t}\right)dt$ and prove the following; (i) $\lim_{x\to0} f_a(x) = \Gamma(a)$ if $a > 0$, (ii) $f_a(x) = x^{2a}f_{-a}(x)$, (iii)$f_a'(x) = -2xf_{a-1}(x)$, (iv) $f_{1/2}(x) = \exp(-2x)\sqrt{\pi}$ (v) $E\exp(-\theta/X) = \sqrt{1/\pi}f_{1/2}(\sqrt{r\theta})$,

Example 1.3.4 (*Cauchy distribution*) An $\mathbb{R}^d$-valued r.v. Y is said to have (c)-*Cauchy distribution* $(c > 0)$ on $\mathbb{R}^d$ if

$$P(Y \in B) = \frac{c}{\omega_d}\int_B \frac{dx}{(c^2+|x|^2)^{\frac{d+1}{2}}} \quad \text{for } B \in \mathcal{B}(\mathbb{R}^d), \tag{1.21}$$

where $\omega_d = \pi^{\frac{d+1}{2}}/\Gamma(\frac{d+1}{2})$ is the area of the unit sphere in $\mathbb{R}^{d+1}$. For $d = 1$ and $B = [a, b]$, one can compute the integral as follows;

$$P(Y \in [a, b]) = \frac{c}{\pi}\int_a^b \frac{dx}{c^2 + x^2} = \frac{1}{\pi}\left(\arctan(b/c) - \arctan(a/c)\right).$$

The Cauchy distribution has the following interpretation in electromagnetism. Imagine that a unit electric charge is put at $z_0 = (0,\dots,0,c) \in H \stackrel{\text{def.}}{=} \mathbb{R}^d \times (0,\infty) \subset \mathbb{R}^{d+1}$, which generates the electric field $E(z) = \frac{z-z_0}{\omega_d |z-z_0|^{d+1}}$, $z \in \mathbb{R}^{d+1}$. The density of the Cauchy distribution is the inner product of $E(x_1,\dots,x_d,0)$ and $(0,\dots,0,-1)$. In this interpretation, the fact $\frac{c}{\omega_d}\int_{\mathbb{R}^d} \frac{dx}{(c^2+|x|^2)^{\frac{d+1}{2}}} = 1$ is exactly *Gauss' law*;

$$\int_{\partial H} E(x_1,\dots,x_d,0)\cdot(0,\dots,0,-1)dx_1\cdots dx_d = \text{total charge in } H = 1.$$

Exercise 1.3.12 Let U be a r.v. with uniform distrubution on $(-\frac{\pi}{2},\frac{\pi}{2})$. Prove then that $X = c\tan U$ has (c)-Cauchy distribution on $\mathbb{R}$.

Example 1.3.5 (*Poisson distribution*) A real r.v. N is called a *Poisson r.v.* with the parameter $r > 0$ if

$$P\{N \in B\} = \pi_r(B) \stackrel{\text{def.}}{=} \sum_{n\in\mathbb{N}\cap B} e^{-r}r^n/n!, \quad B \in \mathcal{B}(\mathbb{R}). \tag{1.22}$$

A probability measure π_r defined above is called *Poisson distribution* with parameter $r > 0$ or (r)-*Poisson distribution*. Poisson distribution is concentrated on $\mathbb{N}$ and therefore, N is in fact an $\mathbb{N}$-valued r.v.

Exercise 1.3.13 For N in Example 1.3.5, show that $E[N] = \text{var}(N) = r$.

Exercise 1.3.14 ($\star$) Let N be a Poisson r.v. with the parameter $r > 0$. Prove then that $P(N \geq a) = \gamma_{1,a}((0,r])$, where $\gamma_{1,a}$ denotes the $(1,a)$-gamma distribution.

Exercise 1.3.15 ($\star$) Show the following: (i) For any $n \in \mathbb{N}^*$, there exist coefficinets $c_{n,k} \in \mathbb{N}$ ($k \in \mathbb{Z}$, $1 \leq |k| < n$) such that

$$Q_n(z) \stackrel{\text{def.}}{=} \frac{1}{n}\frac{2-z^n-z^{-n}}{2-z-z^{-1}} = 1 + \sum_{1\leq|k|<n} c_{n,k}z^k, \quad \text{for } z \in \mathbb{C}\backslash\{0\}, \tag{1.23}$$

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) Show that

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

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*.

2 Independent Random Variables I: Definition

2.1 When do two measures coincide?

In this subsection, we take up a question as follows; Let μ and ν be probability measures on a measurable space $(S, \mathcal{B})$. Suppose that $\mathcal{A} \subset \mathcal{B}$ and that $\mu(A) = \nu(A)$ for all $A \in \mathcal{A}$. Then, is it the case that $\mu(A) = \nu(A)$ for all $A \in \sigma[\mathcal{A}]$?, where

$$\sigma[\mathcal{A}] = \text{the smallest } \sigma\text{-field that contains } \mathcal{A}. \tag{2.1}$$

We will see that the answer is yes, if $\mathcal{A}$ is closed under intersection:

Lemma 2.1.1 (Dynkin's lemma[5]) *Let μ and ν be measures on a measurable space $(S, \mathcal{B})$ and that $\mu(S) = \nu(S) < \infty$. Suppose that $\mathcal{A} \subset \mathcal{B}$ is closed under intersection (i.e., if $A_1 \cap A_2 \in \mathcal{A}$ for any $A_1, A_2 \in \mathcal{A}$) and that $\mu(A) = \nu(A)$ for all $A \in \mathcal{A}$. Then, $\mu(A) = \nu(A)$ for all $A \in \sigma[\mathcal{A}]$.*

The proof of this lemma is presened in Section 9.1. It is more important to know how to apply Lemma 2.1.1 than to know how to prove it. An example of such application is the following

Lemma 2.1.2 *Suppose that $\{\mu, \nu\} \subset \mathcal{P}(S, \mathcal{B})$, where S is a metric space and $\mathcal{B}$ is the Borel σ-field. Then, the following are equivalent:*

a) $\mu = \nu$

b) $\mu(G) = \nu(G)$ *for any open subset* $G \subset S$.

c) $\int f d\mu = \int f d\nu$ *for any bounded, uniformly continuous function* f.

Proof: (a) $\Rightarrow$ (c): Obvious.
(c) $\Rightarrow$ (b): For an indicator function 1_G of an open set G, there is a sequence $\{f_n\}_{n\geq 1}$ of bounded, uniformly continuous functions such that $0 \leq f_n \leq 1$ and $\lim_{n\nearrow\infty} f_n = 1_G$. If $G = S$, we just take $f_n \equiv 1$. If $G \neq S$, then we can find such f_n by;

$$\begin{aligned} F_n &= \{x \in G \ ; \ \text{dist.}(x, G^{\mathsf{c}}) \geq 1/n\}, \\ f_n(x) &= \frac{\text{dist.}(x, G^{\mathsf{c}})}{\text{dist.}(x, G^{\mathsf{c}}) + \text{dist.}(x, F_n)}. \end{aligned}$$

We now have by the bounded convergence theorem that

$$\mu(G) = \lim_{n\nearrow\infty} \int f_n d\mu = \lim_{n\nearrow\infty} \int f_n d\nu = \nu(G).$$

(b) $\Rightarrow$ (a): This can be seen from Lemma 2.1.1. □

Exercise 2.1.1 Suppose that $\{\mu, \nu\} \subset \mathcal{P}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Use Lemma 2.1.1 to prove that $\mu = \nu$ if and only if

$$\mu\left(\prod_{j=1}^d (-\infty, b_j]\right) = \nu\left(\prod_{j=1}^d (-\infty, b_j]\right) \quad \text{for any } (b_j)_{j=1}^d \in \mathbb{R}^d. \tag{2.2}$$

[5] E. B. Dynkin, 1924–

2.2 Product measures

Definition 2.2.1 (Cylinder sets, the product σ-field,...etc.) Suppose that $\{(S_\lambda, \mathcal{B}_\lambda)\}_{\lambda \in \Lambda}$ are measurable spaces indexed by a set Λ. Let $S = \prod_{\lambda \in \Lambda} S_\lambda$ be the direct product and $\pi_\lambda : S \to S_\lambda$ be the canonical projection.

• A subset C of S is called a *cylinder set* if

$$C = \cap_{\lambda \in \Lambda_0} \pi_\lambda^{-1}(B_\lambda), \tag{2.3}$$

where Λ_0 is a finite subset of Λ and $B_\lambda \in \mathcal{B}_\lambda$.

• We define

$$\mathcal{C} \;=\; \text{the set of all cylinder sets in } S, \tag{2.4}$$

$$\mathcal{B} \;=\; \otimes_{\lambda \in \Lambda} \mathcal{B}_\lambda \stackrel{\text{def.}}{=} \sigma[\mathcal{C}], \quad \text{cf. (2.1)}. \tag{2.5}$$

The σ-field $\otimes_{\lambda \in \Lambda} \mathcal{B}_\lambda$ is called the *cylindrical σ-field* or the *product σ-field* on S.

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

Let $\lambda \in \Lambda$ and $\mu \in \mathcal{P}(S, \mathcal{B})$.

• The measure $\mu \circ \pi_\lambda^{-1} \in \mathcal{P}(S_\lambda, \mathcal{B}_\lambda)$ defined as follows is called a *marginal* of μ (with respect to the λ-coordinate);

$$\mu \circ \pi_\lambda^{-1}(B) = \mu(\pi_\lambda^{-1}(B)), \quad B \in \mathcal{B}_\lambda. \tag{2.6}$$

Exercise 2.2.1 (i) Let everything be as in Definition 2.2.1. Prove then that $\mathcal{C}$ is closed under intersection. (ii) 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 2.2.2 Let everything be as in Definition 2.2.1 and $X(\omega) = (X_\lambda(\omega))_{\lambda \in \Lambda}$ be a map from Ω to S. Prove then the following:
i) $\sigma[X] = \sigma\left[X_\lambda^{-1}(B_\lambda)\ ;\ B_\lambda \in \mathcal{B}_\lambda,\ \lambda \in \Lambda\right]$. Hint: The "$\supset$" part is obvious, since $X_\lambda^{-1} = X^{-1} \circ \pi_\lambda^{-1}$. To prove "$\subset$" part, we have to prove that

$(*)$ $X^{-1}(A) \in \sigma\left[X_\lambda^{-1}(B_\lambda)\ ;\ B_\lambda \in \mathcal{B}_\lambda,\ \lambda \in \Lambda\right]$

for all $A \in \mathcal{B}$. We define $\mathcal{A}$ as the set of $A \in 2^S$ for which $(*)$ holds. It is then easy to see that $\mathcal{A}$ is a σ-field on S and contains $\mathcal{C}$. Therefore $\mathcal{B} \subset \mathcal{A}$, by the definition of $\mathcal{B}$.
ii) $X : (\Omega, \mathcal{F}) \to (S, \mathcal{B})$ is measurable if and only if $X_\lambda : (\Omega, \mathcal{F}) \to (S_\lambda, \mathcal{B}_\lambda)$ is measurable for all $\lambda \in \Lambda$.

Lemma 2.2.2 *(***Cylinder sets determin the distribution***) Let everything be as in Definition 2.2.1 and suppose that $\{\mu, \nu\} \subset \mathcal{P}(S, \mathcal{B})$. Then, the following are equivalent;*

a) $\mu = \nu$

b) $\mu(C) = \nu(C)$ *for all* $C \in \mathcal{C}$.

Proof: (a) $\Rightarrow$ (b): Obvious.
(b) $\Rightarrow$ (a): This follows from Lemma 2.1.1, since $\mathcal{C}$ is closed under intersection (Exercise 2.2.1) and $\mathcal{B} = \sigma[\mathcal{C}]$. $\square$

Example 2.2.3 (*Marginals do not determin the distribution*) Concerning Lemma 2.2.2, note that

$$\mu = \nu \quad \begin{array}{c}\Longrightarrow \\ \not\Longleftarrow\end{array} \quad \mu \circ \pi_\lambda^{-1} = \nu \circ \pi_\lambda^{-1} \text{ for all } \lambda \in \Lambda.$$

Take for example $S_1 = S_2 = \{0, 1\}$ and $\mu_i \in \mathcal{P}(S_i)$, $i = 1, 2$. We define $\nu_\theta \in \mathcal{P}(S_1 \times S_2)$ by

$$\nu_\theta(0,0) = \theta, \;\; \nu_\theta(0,1) = \mu_1(0) - \theta, \;\; \nu_\theta(1,0) = \mu_2(0) - \theta,$$

where, for ν_θ to be a probability measure, we suppose that

$$0 \vee (\mu_1(0) + \mu_2(0) - 1) \leq \theta \leq \mu_1(0) \wedge \mu_2(0).$$

Then, $\nu_\theta \circ \pi_i^{-1} = \mu_i$, $i = 1, 2$ for all θ with the above constraint. As is shown already by this simple example, many different measures on a product space can share the same marginals.

Theorem 2.2.4 (Product measures) *Let everything be as in Definition 2.2.1. Suppose that* $\mu_\lambda \in \mathcal{P}(S_\lambda, \mathcal{B}_\lambda)$ *for each* $\lambda \in \Lambda$. *Then, there exists a unique* $\mu \in \mathcal{P}(S, \mathcal{B})$ *such that*

$$\mu\left(\bigcap_{\lambda \in \Lambda_0} \pi_\lambda^{-1}(B_\lambda)\right) = \prod_{\lambda \in \Lambda_0} \mu_\lambda\left(B_\lambda\right) \quad \text{for any finite subset } \Lambda_0 \text{ of } \Lambda \text{ and } B_\lambda \in \mathcal{B}_\lambda. \tag{2.7}$$

• *The measure* μ *defined by (2.7) 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 2.2.2. For the existence[6], we refer the reader to [Dud, page 201, Theorem 8.2.2]. A self-contained exposition is given by Proposition 9.2.1 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-field. □

Remark: Concerning Theorem 2.2.4, note that:

$$\mu = \otimes_{\lambda \in \Lambda} \mu_\lambda \quad \begin{array}{c}\Longrightarrow \\ \not\Longleftarrow\end{array} \quad \mu \circ \pi_\lambda^{-1} = \mu_\lambda, \quad \text{for all } \lambda \in \Lambda.$$

This can be seen form Example 2.2.3, where $\nu_\theta = \mu_1 \otimes \mu_2$ only when $\theta = \mu_1(0)\mu_2(0)$.

Exercise 2.2.3 Suppose in Theorem 2.2.4 that each S_λ contains at most countable elements. Prove then that (2.7) is equivalent to that

$$\mu\left(\cap_{\lambda \in \Lambda_0} \pi_\lambda^{-1}(x_\lambda)\right) = \prod_{\lambda \in \Lambda_0} \mu_\lambda\left(x_\lambda\right) \quad \text{for any } (x_\lambda)_{\lambda \in \Lambda_0} \in \prod_{\lambda \in \Lambda_0} S_\lambda.$$

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

2.3 Independent random variables

Let us now come back to our informal description (1.2) 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 2.3.1 (Independent r.v.'s) *Suppose that* $\{(S_\lambda, \mathcal{B}_\lambda)\}_{\lambda \in \Lambda}$ *are measurable spaces indexed by a set* Λ *and that* $X_\lambda : \Omega \to S_\lambda$ *is a r.v. with the distribution* $\mu_\lambda \in \mathcal{P}(S_\lambda, \mathcal{B}_\lambda)$ *for each* $\lambda \in \Lambda$. *Then the following conditions (a)–(c) are equivalent:*

a) *For any finite subset* Λ_0 *in* Λ *and for any* $B_\lambda \in \mathcal{B}_\lambda$ $(\lambda \in \Lambda_0)$,

$$P\left(\bigcap_{\lambda \in \Lambda_0} \{X_\lambda \in B_\lambda\}\right) = \prod_{\lambda \in \Lambda_0} P(X_\lambda \in B_\lambda). \tag{2.8}$$

b) *The r.v.* $(X_\lambda)_{\lambda \in \Lambda}$ *has* $\otimes_{\lambda \in \Lambda} \mu_\lambda$ *as its distribution, i.e.,*

$$P\left((X_\lambda)_{\lambda \in \Lambda} \in \cdot\right) = \otimes_{\lambda \in \Lambda} \mu_\lambda. \tag{2.9}$$

c) *For any finite subset* Λ_0 *in* Λ *and for any* $f_\lambda \in L^1(\mu_\lambda)$ $(\lambda \in \Lambda_0)$

$$E\left[\prod_{\lambda \in \Lambda_0} f_\lambda(X_\lambda)\right] = \prod_{\lambda \in \Lambda_0} E\left[f_\lambda(X_\lambda)\right]. \tag{2.10}$$

• *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* **i.i.d.** *(independent identically distributed) r.v.'s.*

Proof: (a) $\iff$ (b): Let μ be the distribution of $(X_\lambda)_{\lambda \in \Lambda}$, i.e., $\mu = P\left((X_\lambda)_{\lambda \in \Lambda} \in \cdot\right)$. Then, for any finite subset Λ_0 in Λ and for any $B_\lambda \in \mathcal{B}_\lambda$ $(\lambda \in \Lambda_0)$,

$$\begin{aligned} P\left(\cap_{\lambda \in \Lambda_0} \{X_\lambda \in B_\lambda\}\right) &= P\left((X_\lambda)_{\lambda \in \Lambda} \in \cap_{\lambda \in \Lambda_0} \pi_\lambda^{-1}(B_\lambda)\right) = \mu\left(\cap_{\lambda \in \Lambda_0} \pi_\lambda^{-1}(B_\lambda)\right), \\ \textstyle\prod_{\lambda \in \Lambda_0} P(X_\lambda \in B_\lambda) &= \textstyle\prod_{\lambda \in \Lambda_0} \mu_\lambda(B_\lambda) \end{aligned}$$

Therefore, condition (2.8) is equivalent to $\mu = \otimes_{\lambda \in \Lambda} \mu_\lambda$ (cf. (2.7)).
(a) $\Rightarrow$ (c): We see from the proof of "(a) $\iff$ (b) " that "(a) $\iff$ (b') ", where

b') For any finite subset Λ_0 in Λ, the r.v. $(X_\lambda)_{\lambda\in\Lambda_0}$ has $\otimes_{\lambda\in\Lambda_0}\mu_\lambda$ as its distribution, i.e.,

$$P\left((X_\lambda)_{\lambda\in\Lambda_0} \in \cdot\right) = \otimes_{\lambda\in\Lambda_0}\mu_\lambda.$$

Then, (b') implies (c) by Fubini's theorem.
(c) $\Rightarrow$ (a): This can be seen by considering $f_\lambda = 1_{B_\lambda}$. □

Remarks:
1) The condition (a) in Proposition 2.3.1 amounts to saying that the σ-fields $\{\sigma(X_\lambda)\}_{\lambda\in\Lambda}$ (cf. (1.5)) are independent in the sense of Definition 9.6.1 (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

$$P\{X_\lambda \in \cdot\} = \mu_\lambda \quad \text{for all } \lambda \in \Lambda,$$

which are *not* independent. For example, consider the measure ν_θ in Example 2.2.3 and $\{0,1\}$-valued r.v.'s $\{X_i\}_{i=1,2}$ such that $P\{(X_1,X_2) \in \cdot\} = \nu_\theta$ with $\theta \neq \mu_1(0)\mu_2(0)$.

Exercise 2.3.1 Suppose that $\zeta_n = (\xi_n, \eta_n)$ $(n \geq 1)$ are $\mathbb{R}^2$-valued r.v.'s and that $\{\xi_n, \eta_n\} \subset L^2(P)$ for all $n \geq 1$. Prove then that conditions (a)–(c) listed below are related as (a1) $\Rightarrow$ (a2) $\Rightarrow$ (b) $\Rightarrow$ (c);

a1) $\{\zeta_n\}_{n\geq 1}$ are independent.

a2) $\{\zeta_n\}_{n\geq 1}$ are pairwise independent.

b) ξ_i and η_j for $i \neq j$ are *uncorrelated*, i.e., $E(\xi_i;\eta_j) = 0$ if $i \neq j$.

c)

$$\mathrm{cov}\Big(\sum_{i=1}^m \xi_i; \sum_{j=1}^n \eta_j\Big) = \sum_{i=1}^m \mathrm{cov}(\xi_i;\eta_i) \quad \text{if } m \leq n. \tag{2.11}$$

Exercise 2.3.2 Let a r.v. U be uniformly distributed on $(0, 2\pi)$. Prove than that $X = \cos U$ and $Y = \sin U$ are not independent and that $\mathrm{cov}(X,Y) = 0$.

Exercise 2.3.3 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 2.3.4 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 2.3.5 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 distribution for any $n \in \mathbb{N}^*$.

Exercise 2.3.6 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 $\mathbb{N}^*$ 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 distribution.

Exercise 2.3.7 ($\star$) 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 2.3.8 ($\star$) 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 distribution 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 distribution. This generalizes Exercise 2.3.5.

Exercise 2.3.9 ($\star$) Let everything be as in Proposition 2.3.1. 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 $\{\widetilde{X}\}_{\gamma \in \Gamma}$ are independent if $\{X_\lambda\}_{\lambda \in \Lambda}$ are. Hint: Condition (b) of Proposition 2.3.1.

3 Independent Random Variables II: Examples

3.1 Some functions of independent r.v.'s.

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

Example 3.1.1 (*Relation between gamma and beta distributions*) Let X and Y be real r.v.'s with $P((X,Y) \in \cdot) = \gamma_{r,a} \otimes \gamma_{r,b}$ (cf.Example 1.3.3). Then,

$$P\left((X+Y, \tfrac{X}{X+Y}) \in \cdot\right) = \gamma_{r,a+b} \otimes \beta_{a,b}. \tag{3.1}$$

In particular,

$$P\left(X+Y \in \cdot\right) = \gamma_{r,a+b} \quad \text{and} \quad P\left(\tfrac{X}{X+Y} \in \cdot\right) = \beta_{a,b}.$$

Let us prove (3.1).The well-known formula

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

will also be reproduced in the course of the proof. We first note the following simple equality:

(2) $$\int_0^{tz} x^{a-1}(z-x)^{b-1}dx = z^{a+b-1}B(a,b)\beta_{a,b}((0,t]).$$

By Exercise 2.1.1, (3.1) is equivalent to that

(3) $$P\left((X+Y, \tfrac{X}{X+Y}) \in (0,s] \times (0,t]\right) = \gamma_{r,a+b}((0,s])\beta_{a,b}((0,t]) \text{ for all } s,t>0.$$

We first show that

(4) $$\text{LHS of (3)} = \frac{B(a,b)\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\gamma_{r,a+b}((0,s])\beta_{a,b}((0,t]).$$

Let us write $D = \{(x,y) \in (0,\infty)^2 \; ; \; (x+y, \frac{x}{x+y}) \in (0,s] \times (0,t]\}$. Then,

$$\begin{aligned}
\text{LHS of (3)} &= \gamma_{r,a} \otimes \gamma_{r,b}(D) \\
&= \frac{1}{\Gamma(a)\Gamma(b)} \int_D (rx)^{a-1}(ry)^{b-1}e^{-(x+y)r}r^2 dxdy \\
&\stackrel{z=x+y}{=} \frac{1}{\Gamma(a)\Gamma(b)} \int_0^s r^{a+b}e^{-zr}dz \int_0^{tz} x^{a-1}(z-x)^{b-1}dx \\
&\stackrel{(2)}{=} \frac{B(a,b)}{\Gamma(a)\Gamma(b)} \int_0^s (rz)^{a+b-1}e^{-zr}rdz\beta_{a,b}((0,t]) \\
&= \frac{B(a,b)\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\gamma_{r,a+b}((0,s])\beta_{a,b}((0,t]) = \text{RHS (4)}.
\end{aligned}$$

Letting $s \nearrow \infty$ and $t \nearrow \infty$ in (4), we get

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

Finally, plugging this back in (4), we arrive at (3).

Exercise 3.1.1 Let X,Y and Z be r.v.'s with $P((X,Y) \in \cdot) = \gamma_{r,a} \otimes \gamma_{s,b}$. and let $P(Z \in \cdot) = \beta_{a,b}$. Prove then that

$$P(X/Y \in A) = P\left(\frac{s}{r}\frac{Z}{1-Z} \in A\right) = \frac{(r/s)^{a-1}}{B(a,b)} \int_A \frac{x^{a-1}dx}{(1+rx/s)^{a+b}}, \quad A \in \mathcal{B}((0,\infty)).$$

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 $P((X_1, Y_1) \in \cdot) = \gamma_{1,a} \otimes \gamma_{1,b}$. Then, $P((X,Y) \in \cdot) = P((X_1/r, Y_1/s) \in \cdot)$ and $\frac{X_1}{Y_1} = \frac{\frac{X_1}{X_1+Y_1}}{1-\frac{X_1}{X_1+Y_1}}$. Then use (3.1).

Exercise 3.1.2 Prove the following extension of Exercise 3.1.2. Let $X_j > 0$, $j = 1,..,n+1$ be independent r.v.'s with $P(X_j \in \cdot) = \gamma_{r,a_j}$ 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 $P(S \in \cdot) = \gamma_{r,a_1+..+a_{n+1}}$ and

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

for any Borel set $B \subset \left\{x \in (0,1)^n \ ; \ \sum_{j=1}^n x_j < 1\right\}$.

Exercise 3.1.3 Let $X_1,..,X_n$ be real i.i.d. with $P(X_j \in \cdot) = \nu_v$ ($v > 0$, $j = 1,..,n$, cf. (1.15)). Prove then that $P(\sum_{j=1}^n X_j^2 \in \cdot) = \gamma_{1/(2v),n/2}$ When $v = 1$, the distribution $\gamma_{1/2,m/2}$ of $\sum_{j=1}^n X_j^2$ is called the χ_n^2 *distribution* and is used in statistics.

Hint: Exercise 1.3.7 and Example 3.1.1.

Exercise 3.1.4 Let e and U are independent r.v. such that $P(e \in \cdot) = \gamma_{1,1}$ and U is uniformly distributed on $(0, 2\pi)$. Prove then that $\sqrt{2e}(\cos U, \sin U)$ have the standard normal distribution on $\mathbb{R}^2$.

Exercise 3.1.5 ($\star$) Let $S_n = X_1^2 + .. + X_n^2$, where $(X_j)_{j\geq 1}$ are real i.i.d. with $P(X_j \in \cdot) = \nu_v$ ($v > 0$, cf. (1.15)) Prove then that for $m, n = 1, 2, ..$,

$$\begin{aligned} P\left(\left(S_{m+n}, \frac{S_m}{S_{m+n}}\right) \in \cdot\right) &= \gamma_{1/(2v),(m+n)/2} \otimes \beta_{m/2,n/2}. \\ P\left(\left(S_n, \frac{(S_{m+n} - S_n)/m}{S_n/n}\right) \in \cdot\right) &= \gamma_{1/(2v),n/2} \otimes F_n^m. \quad \text{(cf. Exercise 3.1.1).} \end{aligned}$$

Hint: Exercise 1.3.7, Example 3.1.1 and Exercise 3.1.1.

Exercise 3.1.6 ($\star$) 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. (1.18). 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 (3.1).

Example 3.1.2 (*Poisson process*) For $t \geq 0$, we define

$$N_t = \sup\{n \in \mathbb{N}\ ;\ \tau_1 + ... + \tau_n \leq t\},$$

where $\tau_1, \tau_2, \ldots$ be i.i.d. with $P(\tau_j \in \cdot) = \gamma_{r,1}$ (cf. (1.18)). Then, $P(N_t \in \cdot) = \pi_{rt}$ (cf. (1.22)). This can be seen as follows: We set $T_0 \equiv 0$ and $T_n = \sum_{j=1}^n \tau_j$ ($n \geq 1$). Then,

$$\begin{aligned} P(N_t = n) &= P(T_n \leq t < T_n + \tau_{n+1}) \\ &= \int_0^t P(T_n \in ds) \int P(\tau_{n+1} \in du) 1_{\{t \leq s+u\}}, \end{aligned}$$

where we have used the independence of T_n and τ_{n+1} on the second line. Since $P(\tau_{n+1} \in du) = \gamma_{r,1}(du) = e^{-ru} r du$,

$$\int P(\tau_{n+1} \in du) 1_{\{t \leq s+u\}} = \int_{t-s}^\infty e^{-ru} r du = e^{-(t-s)r}.$$

On the other hand, we see from (3.1) that

$$P(T_n \in ds) = \gamma_{r,n}(ds) = \frac{(rs)^{n-1} e^{-rs} r ds}{(n-1)!}$$

Putting these together, we conclude that

$$P(N_t = n) = \int_0^t r ds \frac{(rs)^{n-1} e^{-rs}}{(n-1)!} e^{-(t-s)r} = \frac{r^n e^{-rt}}{(n-1)!} \int_0^t s^{s-1} ds = \frac{e^{-rt} (rt)^n}{n!}.$$

$(N_t)_{t \geq 0}$ is called the *Poisson process* with the parameter r. N_t has, for example, the following interpretation; T_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.

Exercise 3.1.7 Let $\{X_i\}_{i=1}^n$ be r.v.'s with $P((X_i)_{i=1}^n \in \cdot) = \otimes_{i=1}^n \gamma_{r_i,1}$ (cf. (1.18)) and $M_n = \min\limits_{i=1,\ldots,n} X_i$. Prove then that for any $j = 1, \ldots, n$ and $x \geq 0$,

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

In particular, $P(M_n \in \cdot) = \gamma_{r_1+\ldots+r_n,1}$

Exercise 3.1.8 (*Thinning of a Poisson r.v.*) Let N be a r.v. with $P(N \in \cdot) = \pi_r$ 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 $P(N_s \in \cdot) = \pi_{p(s)r}$, where $p(s) = P(X_0 = s)$.

Exercise 3.1.9 (*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 $P(G = n) = p(1-p)^{n-1}$, $E[G] = 1/p$, and $\text{var}(G) = (1-p)/p$. 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 3.1.10 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 $P(\tau_j \in \cdot) = \gamma_{r,1}$ (cf. (1.18)). Prove then that $P(\tau_1 + ... + \tau_G \in \cdot) = \gamma_{pr,1}$.

Exercise 3.1.11 (*binomial distribution*) Let $S_n = X_1 + \ldots + X_n$, where $(X_n)_{n\geq 1}$ are $\{0,1\}$-valued i.i.d. with $P(X_n = 1) = p$. Prove then that

$$\begin{aligned} E[S_n] &= np, \quad \text{var}\,(S_n) = np(1-p) \leq n/4, \\ P(S_n = r) &= \binom{n}{r} p^r (1-p)^{n-r} \quad \text{for } r = 0, 1, \ldots, n. \end{aligned}$$

The distribution of S_n is called the *p-binomial distribution*. Hint: $\text{var}(S_n) = \text{var}(X_1) + ... + \text{var}(X_n)$ (cf. (2.11)).

Exercise 3.1.12 ($\star$) (*Relation between geometric and binomial distributions*) Let $G_1, G_2, ...$ be i.i.d. such that $P(G_1 = n) = p(1-p)^{n-1}$ (p-geometric distribution) and let

$$S_n = \sup\{r \geq 0\,; G_1 + ... + G_r \leq n\} \text{ for } n \in \mathbb{N}.$$

Prove then that $X_n = S_n - S_{n-1}$, $n \in \mathbb{N}^*$ are $\{0,1\}$-valued i.i.d. such that $P(X_n = 1) = p$. In particular, S_n has p-binomial distribution (Exercise 3.1.11). The r.v.'s $(S_n)_{n\geq 1}$ above can be thought of as a discrete-time analogue of Poisson process (Example 3.1.2).

Example 3.1.3 ($\star$) Let $X_1, X_2, ..., X_d, Y$ be independent r.v.'s with $P(Y \in \cdot) = \gamma_{r,a}$ (cf. (1.18)), $P(X_j \in \cdot) = \nu_1$, $1 \leq j \leq d$ (cf. (1.15)). Let us write $X = (X_j)_{j=1}^d$ for simplicity. We can then, compute the distribution of the $\mathbb{R}^d$-valued r.v.$Y^{-1/2}X$ as follows;

$$P(Y^{-1/2}X \in B) = \frac{(2r)^a \Gamma(\frac{d}{2} + a)}{\pi^{d/2}\Gamma(a)} \int_B \frac{dx}{(2r+|x|^2)^{\frac{d}{2}+a}} \quad \text{for } B \in \mathcal{B}(\mathbb{R}^d). \tag{3.2}$$

There are two important special cases:

• $(r, a) = (c^2/2, 1/2)$: In this case, $P(Y^{1/2} \in \cdot) = P(|X_0| \in \cdot)$, where X_0 is a r.v. with $P(X_0 \in \cdot) = \nu_{c^{-2}}$ (Exercise 1.3.7) and the right-hand-side of (3.2) is the (c)-Cauchy distribution. Therfore, (3.2) says that a r.v. with the (c)-Cauchy distribution can be obtained by dividing a Gaussian r.v. by the square root of an independent gamma r.v.(or equivalently, the absolute value of another independent Gaussian r.v.).

• $d = 1$, $r = a = n/2$ with $n \in \mathbb{N}$. In this case, the distribution given by (3.2) is called the T_n-*distribution* used in statistics.

The proof of (3.2) goes as follows. We set $r(x) = r + \frac{|x|^2}{2}$ and $a_d = a + \frac{d}{2}$. Then,

$$\begin{aligned} P(Y^{-1/2}X \in B) &= \int_0^\infty P(Y \in dy) \int_{\mathbb{R}^d} P(X \in dx) 1_B(y^{-1/2}x) \\ &= \frac{r^a}{(2\pi)^{d/2}\Gamma(a)} \int_0^\infty y^{a-1}e^{-ry}dy \int_{\mathbb{R}^d} dx 1_B(y^{1/2}x)e^{-|x|^2/2} \\ &= \frac{r^a}{(2\pi)^{d/2}\Gamma(a)} \int_0^\infty y^{a_d-1}dy \int_B dx e^{-r(x)y} \\ &= \frac{r^a}{(2\pi)^{d/2}\Gamma(a)} \int_B dx \int_0^\infty y^{a_d-1}e^{-r(x)y}dy \end{aligned}$$

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

$$\int_0^\infty y^{a_d-1}e^{-r(x)y}dy = \Gamma\left(a_d\right)r(x)^{-a_d}.$$

Thus, we conclude that

$$P(Y^{-1/2}X \in B) = \frac{r^a\Gamma(a_d)}{(2\pi)^{d/2}\Gamma(a)}\int_B r(x)^{-a_d}dx = \text{the right-hand-side of (3.2).}$$

Example 3.1.4 ($\star$) $X_1, \ldots, X_n$ be real i.i.d. such that $F(t) = P(X_i \leq t)$ 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:

$$P\{X_{n,k} \in A\} = n\begin{pmatrix} n-1 \\ k-1 \end{pmatrix} E\left[F(X_1)^{k-1}\left(1-F(X_1)\right)^{n-k}1\{X_1 \in A\}\right] \quad A \in \mathcal{B}(\mathbb{R}).$$

This can roughly be explained as follows. First of all, there are n ways to choose $X_{n,k}$ form $X_1, ..., X_n$ and the probabiliy of all such selections are the same (This explains the first factor n). Now, suppose that $X_1 = X_{n,k}$. Then, there are $\begin{pmatrix} n-1 \\ k-1 \end{pmatrix}$ 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 $\begin{pmatrix} n-1 \\ k-1 \end{pmatrix}$). 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,

$$\begin{aligned}
&P\{X_{n,k} \in A\} \\
&= \sum_{\sigma\in\mathcal{S}_n} P\{X_{\sigma(1)} < X_{\sigma(2)} < ... < X_{\sigma(k)} < ... < X_{\sigma(n)},\ X_{\sigma(k)} \in A\} \\
&= \sum_{\sigma\in\mathcal{S}_n}\int_A P\{X_{\sigma(k)} \in dx\}P\{X_{\sigma(1)} < X_{\sigma(2)} < ... < X_{\sigma(k-1)} < x < X_{\sigma(k+1)} < ... < X_{\sigma(n)}\} \\
&= \sum_{\sigma\in\mathcal{S}_n}\int_A P\{X_{\sigma(k)} \in dx\}P\{X_{\sigma(1)} < X_{\sigma(2)} < ... < X_{\sigma(k-1)} < x\}P\{x < X_{\sigma(k+1)} < ... < X_{\sigma(n)}\} \\
&= \sum_{\sigma\in\mathcal{S}_n}\int_A P\{X_{\sigma(k)} \in dx\}\frac{F(x)^{k-1}}{(k-1)!}\frac{(1-F(x))^{n-k}}{(n-k)!} \\
&= n!\int_A P\{X_1 \in dx\}\frac{F(x)^{k-1}}{(k-1)!}\frac{(1-F(x))^{n-k}}{(n-k)!} \\
&= n\begin{pmatrix} n-1 \\ k-1 \end{pmatrix} E\left[F(X_1)^{k-1}\left(1-F(X_1)\right)^{n-k}1\{X_1 \in A\}\right].
\end{aligned}$$

3.2 A proof of Weierstrass' approximation theorem

Example 3.2.1 (*Weierstrass' approximation theorem*) Let $I = [0,1]$ and $f \in C(I \to \mathbb{R})$. Then, there exist polynomials $f_n : \mathbb{R} \to \mathbb{R}$ $(n \geq 1)$ such that

1) $$\lim_{n \nearrow \infty} \max_{\theta \in I} |f_n(\theta) - f(\theta)| = 0.$$

To prove this, we fix $\theta \in I$ and $n \in \mathbb{N}^*$ for a moment and let S_n be a r.v. such that

$$P(S_n = r) = \binom{n}{r} \theta^r (1-\theta)^{n-r} \quad \text{for } r = 0, ..., n.$$

Then,

$$f_n(\theta) \stackrel{\text{def.}}{=} Ef(n^{-1}S_n) = \sum_{r=0}^{n} f(n^{-1}r) P(S_n = r)$$

is a polynomial in θ. On the other hand, we see from Exercise 3.1.11 that

2) $$\text{var}(S_n) \leq n/4.$$

The key to prove (1) is [7]:

3) $$P\left(|n^{-1}S_n - \theta| \geq \varepsilon\right) \leq \frac{1}{4\varepsilon^2 n} \text{ for any } \varepsilon > 0.$$

In fact, using Chebyshev's inequality (Exercise 1.2.3) and (2),

$$\begin{aligned} P\left(|n^{-1}S_n - \theta| \geq \varepsilon\right) &\stackrel{\text{Chebyshev}}{\leq} \varepsilon^{-2} E\left[|n^{-1}S_n - \theta|^2\right] \\ &= \varepsilon^{-2} n^{-2} \text{var}(S_n) \stackrel{(2)}{\leq} \frac{1}{4\varepsilon^2 n}. \end{aligned}$$

We now conclude (1) from (3) as follows:

$$\begin{aligned} |f_n(\theta) - f(\theta)| &\leq E\left|f(n^{-1}S_n) - f(\theta)\right| \\ &= E\left[\left|f(n^{-1}S_n) - f(\theta)\right| 1\{|n^{-1}S_n - \theta| \geq n^{-1/3}\}\right] \\ &\quad + E\left[\left|f(n^{-1}S_n) - f(\theta)\right| 1\{|n^{-1}S_n - \theta| < n^{-1/3}\}\right] \\ &\stackrel{(3)}{\leq} 2\sup_{\theta \in I} |f(\theta)| \cdot \frac{d}{4n^{1/3}} + \sup_{\substack{\theta, \theta' \in I \\ |\theta - \theta'| < n^{-1/3}}} |f(\theta') - f(\theta)| \\ &\longrightarrow 0, \quad \text{as } n \nearrow \infty \text{ uniformly in } \theta, \end{aligned}$$

where in the last line, we have used the uniform continuity of f. □

Exercise 3.2.1 (*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 $\lim_{n \nearrow \infty} \max_{\theta \in I} |f_n(\theta) - f(\theta)| = 0$. Hint: Fix $\theta = (\theta^\nu)_{\nu=1}^d \in I$ and $n \in \mathbb{N}^*$ 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 $P(S_n^\nu = r) = \binom{n}{r} (\theta^\nu)^r (1-\theta^\nu)^{n-r}$ $(0 \leq r \leq n,\ 1 \leq \nu \leq d)$. Then, $P(S_n = x) = \prod_{\nu=1}^d \binom{n}{x^\nu} (\theta^\nu)^{x^\nu} (1-\theta^\nu)^{n-x^\nu}$.

[7] This is a special case of the weak law of large numbers.

Exercise 3.2.2 (i) Let $f \in C_{\rm b}([0,\infty))$ and

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

Prove then that $\lim_{n\nearrow\infty} f_n(x) = f(x)$ for all $x \geq 0$. Hint: We may assume $x > 0$, since $f_n(0) = f(0)$. Let S_n be r.v. with $P(S_n \in \cdot) = \pi_{nx}$ (cf. (1.22)). Then, $f_n(x) = E[f(\frac{S_n}{n})]$.
(ii) (*Injectivity of the Laplace transform*) Let $\mu_1, \mu_2 \in \mathcal{P}([0,\infty))$ be such that

$$\int_{[0,\infty)} e^{-sx} d\mu_1(x) = \int_{[0,\infty)} e^{-sx} d\mu_2(x) \quad \text{for all } s \geq 0.$$

Use (i) to show that $\mu_1 = \mu_2$. Hint: Show that $\int_{[0,\infty)} f_n d\mu_1 = \int_{[0,\infty)} f_n d\mu_2$ for any $f \in C_{\rm b}([0,\infty))$.

Exercise 3.2.3 ($\star$) (*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 1.3.15). Prove then that f_n is a trigonometric polynomial and that

$$\lim_{n\nearrow\infty} \sup_{0\leq\theta\leq 1} |f_n(\theta) - f(\theta)| = 0.$$

Hint: $f_n(\theta) = \int_0^1 f(\varphi) F_n(\theta - \varphi) d\varphi$ by the periodicity. Then, use (1.23) to see that f_n is a trigonometric polynomial.

3.3 Decimal fractions as i.i.d.

In this subsection, we consider a probability sapce $(\Omega, \mathcal{F}, P)$ and a r.v. U with the uniform distribution on $(0,1)$, i. e., $P\{U \in B\} = \int_B dt$ for all $B \in \mathcal{B}((0,1))$.

Example 3.3.1 (*Decimal fractions are i.i.d.*) Suppose that $q \geq 2$ is an integer. For $n \geq 1$ and $s_1, \ldots, s_n \in \{0, \ldots, q-1\}$, we define $I_{s_1 \ldots, s_n} \subset [0,1)$ and $d_n : \Omega \to \{0, \ldots, q-1\}$ by

$$\begin{aligned} I_{s_1 \ldots s_n} &= \left\{ \sum_{1\leq k\leq n} q^{-k} s_k + x \ ; \ x \in \big[0, \ q^{-n}\big) \right\}, \\ d_n(\omega) &= s \ \ \text{if } U(\omega) \in \textstyle\bigcup_{s_1,\ldots,s_{n-1}} I_{s_1 \cdots s_{n-1} s}. \end{aligned}$$

Note that $\{I_{s_1 \ldots, s_{n-1} s}\}_{s=0}^{q-1}$ are obtained by dividing $I_{s_1 \ldots, s_{n-1}}$ into q smaller intervals with equal length ($=q^{-n}$) and that the interval $I_{s_1 \ldots, s_{n-1} s}$ is the $(s+1)$-th one from the left. This means that $d_n(\omega)$ is nothing but the k-th digit of the q-adic expansion of the number $U(\omega) \in [0,1)$ and therefore that

$$U(\omega) = \sum_{k\geq 1} q^{-k} d_k(\omega) \quad \text{for all } \omega \in \Omega, \tag{3.3}$$

We will prove that

$$(d_n)_{n\geq 1} \text{ are i.i.d. with } P(d_n = s) = q^{-1},\ s = 0, \ldots, q-1. \tag{3.4}$$

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

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

and hence that

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

Moreover, this implies

2) $$P(d_n = s_n) = q^{-1} \text{ for all } n \geq 1 \text{ and } s_n \in \{0, ..., q-1\},$$

since

$$\begin{aligned} P(d_n = s_n) &= \sum_{s_1,\ldots,s_{n-1}} P\left(\bigcap_{j=1}^{n}\{d_j = s_j\}\right) \\ &\stackrel{(1)}{=} \sum_{s_1,\ldots,s_{n-1}} q^{-n} = q^{-1}. \end{aligned}$$

We now conclude (3.4) from (1) and (2) (cf. Exercise 2.2.3). □

Example 3.3.2 *Construction of a sequence of independent random variables with discrete state spaces:* Let $\mu_n \in \mathcal{P}(S_n, \mathcal{B}_n)$ $(n = 1, \ldots)$ be a sequence of probability measures, where for each $n \geq 1$, 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 $P(X_n = \cdot) = \mu_n$ for all $n \geq 1$. The construction is just a slight extension of Example 3.3.1. 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

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

We see from the definition that

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

and hence that

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

Moreover, this implies:

2) $\quad P(X_n = s_n) = \mu_n(s_n)$ for all $n \geq 1$,

since

$$\begin{aligned} P(X_n = s_n) &= \sum_{s_1,\dots,s_{n-1}} P\left(\cap_{j=1}^n \{X_j = s_j\}\right) \\ &\overset{(1)}{=} \sum_{s_1,\dots,s_{n-1}} \mu_1(s_1)\cdots\mu_{n-1}(s_{n-1})\mu_n(s_n) = \mu_n(s_n). \end{aligned}$$

We conclude from (1) and (2) that $(X_n)_{n\geq 1}$ are independent and that $P(X_n \in \cdot) = \mu_n$ (cf. Exercise 2.2.3). □

4 In which direction does a random walk travel?

4.1 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

$$\frac{S_n}{n} \longrightarrow \frac{1}{2} (= E[X_1]), \quad \text{as } n \nearrow \infty. \tag{4.1}$$

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 (4.1) 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.

Definition 4.1.1 Let $A = \{\omega \in \Omega \; ; \; .....\}$ be an event in a certain probability space $(\Omega, \mathcal{F}, P)$.
• We say "..... *almost surely*" ("..... *a.s.*" for short) if $P(A) = 1$.
Therefore, "almost surely" ("a.s.") just synonymizes "almost everywhere" ("a.e.") in measure theory.

Theorem 4.1.2 (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,*

$$\lim_{n\nearrow\infty} \frac{S_n}{n} = E[X_1], \quad P\text{-a.s.} \tag{4.2}$$

Here, we give a proof of Theorem 4.1.2 only 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 10.2. See also Exercise 4.1.4 and Exercise 4.1.6 below to see what happens if we do not assume $E[|X_n|] < \infty$.

Proof of Theorem 4.1.2 in a special case $X_i \in L^4(P)$: Observe first that (4.2) is equivalent to that

1) $$\lim_{n\nearrow\infty} n^{-1} \sum_{k=1}^{n} (X_k - EX_k) = 0, \quad P\text{-a.s.}$$

We will in fact show (1) for independent r.v.'s $\{X_n\}_{n\geq 1}$ such that $\sup_{n\geq 1} E[X_n^4] < \infty$ (without assuming that $\{X_n\}_{n\geq 1}$ are identically distributed).

Define r.v.'s $\widetilde{X}_n$ and $\widetilde{S}_n$ respectively by

$$\widetilde{X}_n = X_n - EX_n \;\text{ and }\; \widetilde{S}_n = \widetilde{X_1} + \ldots + \widetilde{X_n}.$$

We prove (1) by showing a stronger statement that

2) $$\sum_{n\geq 1} |n^{-1}\widetilde{S}_n|^4 < \infty, \quad P\text{-a.s.}$$

We have

$$\begin{aligned} E[|\widetilde{S}_n|^4] &= \sum_{i,j,k,\ell=1}^{n} E[\widetilde{X}_i\widetilde{X}_j\widetilde{X}_k\widetilde{X}_\ell] \\ &= \sum_{i=1}^{n} E[\widetilde{X}_i^4] + 6 \sum_{1\leq r<s\leq n} E[\widetilde{X}_r^2]E[\widetilde{X}_s^2]. \end{aligned}$$

Here is an explanation for the second equality. The only terms in $\sum_{i,j,k,\ell=1}^n$ that do not vanish are those of the form either $E[\widetilde{X}_i^4]$ or $E[\widetilde{X}_r^2\widetilde{X}_s^2] = E[\widetilde{X}_r^2]E[\widetilde{X}_s^2]$ $(1 \leq r < s \leq n)$. In the second case, there are $\binom{4}{2} = 6$ ways to choose r and s from i, j, k, ℓ. Note then that there is a constant C such that

$$E[\widetilde{X}_m^2]^2 \leq E[\widetilde{X}_m^4] \leq C, \quad m = 1, 2, ...$$

We therefore see that

$$E[|\widetilde{S}_n|^4] \leq Cn + 3Cn(n-1),$$

and hence that

$$\begin{aligned} E\left[\sum_{n\geq 1} |n^{-1}\widetilde{S}_n|^4\right] &\overset{\text{Fubini}}{=} \sum_{n\geq 1} E\left[|n^{-1}\widetilde{S}_n|^4\right] \\ &\leq \sum_{n\geq 1} \frac{Cn + 3Cn(n-1)}{n^4} < \infty, \end{aligned}$$

which in particular implies (2). □

Example 4.1.3 (*Almost all numbers are normal.*) Let $q \geq 2$ be an integer and $d_k(\omega)$ is the k-th digit of the q-adic expansion of the number $\omega \in [0, 1)$. A number ω is said to be *normal* in base q if

1) $$\lim_{n\nearrow\infty} \frac{1}{n}(\text{the number of } 1 \leq k \leq n \text{ such that } d_k(\omega) = s) = q^{-1}.$$

for any $s = 0, \ldots, q-1$. Let N_q be the set of all normal numbers in base q. Then, *Borel's theorem* asserts that

$$P\left(\bigcap_{q\geq 2} N_q\right) = 1,$$

where P is the Lebesgue measure on $[0, 1)$. This can be seen as follows. It is enough to show that $P(N_q) = 1$ for any fixed $q \geq 2$. By Example 3.3.1 (or Lemma 9.4.1), we know that the digits $\{d_k\}_{k\geq 1}$ in base q are i.i.d. with $P\{d_k(\omega) = s\} = 1/q$, $s = 0, \ldots, q-1$. We now fix any s and set $X_k = 1\{d_k(\omega) = s\}$. Then, $(X_k)_{k\geq 1}$ are i.i.d. with $E[X_k] = 1/q$. Now, (1) turns out to be a special case of (4.2).

Exercise 4.1.1 For $p \in (0, 1)$, we let $\{X_n^{(p)}\}_{n\geq 1}$ be $\{0, 1\}$-valued i.i.d. with $P(X_n^{(p)} = 1) = p$ and define $\mu_p = P((X_n^{(p)})_{n\geq 1} \in \cdot)$. Prove then that μ_p and μ_q are singular if $p \neq q$.

Exercise 4.1.2 (*Shannon's theorem*) For a finite set A be and $\mu \in \mathcal{P}(A)$ such that $0 < \mu(\alpha) < 1$ for all $\alpha \in A$, we define the *entropy* $H(\mu)$ of μ by

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

Let $\{\alpha_n\}_{n\geq 1}$ be A-valued i.i.d. with $P(\alpha_n \in \cdot) = \mu$. Prove that

$$\lim_{n \nearrow \infty} \left(\prod_{j=1}^{n} \mu(\alpha_j) \right)^{1/n} = e^{-H(\mu)}, \quad P\text{-a.s.}$$

Let us interpret A as the set of letters. Then, the above result says that the probability $\prod_{j=1}^{n} \mu(\alpha_j)$ of almost all randomly generated sentence $\alpha_1\alpha_2....\alpha_n$ decays like $e^{-nH(\mu)}$ as $n \nearrow \infty$.

Exercise 4.1.3 (*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 $E[T_n] < \infty$ for all n (cf. Example 3.1.2 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 $\lim_{t \nearrow \infty} N_t/t = 1/E[T_1]$, P-a.s.
Hint: $T_{N_t} \leq t < T_{N_t+1}$ and $\lim_{t \nearrow \infty} T_{N_t}/N_t = E[T_1]$ by Theorem 4.1.2.

Exercise 4.1.4 (*Infinite mean*) Let $S_n = X_1 + ... + X_n$, where $\{X_n\}_{n\geq 1}$ are i.i.d. such that $E[X_n] = \infty$, i.e., $E[X_n^+] = \infty$ and $E[X_n^-] < \infty$. Prove then that $\lim_{n \nearrow \infty} \frac{S_n}{n} = \infty$ a.s. Hint: $X_n \wedge m \in L^1(P)$ for any fixed $m \in (0, \infty)$.

Exercise 4.1.5 (*The second Borel-Cantelli lemma*) Let $(X_n)_{n\geq 1}$ be independent r.v.'s with values in $[0, 1]$. Prove then that the following (a)–(c) are equivalent: (a): $\sum_{n\geq 1} E[X_n] = \infty$. (b): $\prod_{n=1}^{\infty}(1 - X_n) = 0$, a.s. (c): $\sum_{n\geq 1} X_n = \infty$, a.s.

Exercise 4.1.6 (*Indefinite mean*) Let $S_n = X_1 + ... + X_n$, where $\{X_n\}_{n\geq 1}$ are i.i.d. such that $E[X_n^\pm] = \infty$. Prove then that $P(S_n/n \text{ converges}) = 0$. Hint: Use Exercise 4.1.5 to show that $\sum_{n\geq 1} \mathbf{1}\{X_n > n\} = \infty$, a.s. Then, note that $\frac{S_{n+1}}{n+1} - \frac{S_n}{n} = \frac{X_{n+1}}{n+1} - \frac{S_n}{n(n+1)}$.

4.2 What is a random walk?

Definition 4.2.1 Suppose that $(X_n)_{n\geq 1}$ are $\mathbb{R}^d$-valued i.i.d. with the distribution $\nu \in \mathcal{P}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ on a probability space $(\Omega, \mathcal{F}, P)$;

$$P(X_n \in B) = \nu(B) \text{ for all } B \in \mathcal{B}(\mathbb{R}^d) \text{ and } n \geq 1.$$

A *random walk* is a sequence $(S_n)_{n\geq 0}$ of $\mathbb{R}^d$-valued r.v.'s defined[8] by

$$S_n = \begin{cases} 0 & \text{if } n = 0, \\ X_1 + \cdots + X_n & \text{if } n \geq 1. \end{cases}$$

[8] 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. The the $\mathbb{Z}^d$-valued random walk.

Note that independent r.v.'s $(X_n)_{n\geq 1}$ referred to above certainly exist by Proposition 9.2.1 and so does the random walk $(S_n)_{n\geq 0}$. The measure ν will be called the *single step distribution* of the random walk $(S_n)_{n\geq 0}$. If $|X_1| \in L^p(P)$, or equivalently, $|x| \in L^p(\nu)$, we say that the random walk is an *L^p-random walk.*

Example 4.2.2 Examples we introduce here are familiar and important ones which we should always keep in mind.

• 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 a *nearest neighbor* random walk if

$$P(X_1 \in \{0, e_1, \dots, e_d, -e_1, \dots, -e_d\}) = 1. \tag{4.3}$$

where $e_i = (\delta_{ij})_{i=1}^d$.

• A $\mathbb{Z}^d$-valued random walk is said to be a *simple* random walk if

$$P(X_1 = e_i) = P(X_1 = -e_i) = (2d)^{-1} \quad \text{for all } i = 1, \dots, d. \tag{4.4}$$

Exercise 4.2.1 Consider a $\mathbb{Z}$-valued random walk such that

$$P\{X_1 = 1\} = p > 0,\ P\{X_1 = -1\} = q > 0,\ P\{X_1 = 0\} = r = 1 - p - q.$$

Let $\{y_1, \dots, y_n\} \subset \mathbb{Z}$ be such that $x_p \stackrel{\text{def.}}{=} y_p - y_{p-1} \in \{0, \pm 1\}$ for $1 \leq p \leq n$. We set $N^+ = \sum_{p=1}^n 1\{x_p = +1\}$, $N^- = \sum_{p=1}^n 1\{x_p = -1\}$ and $N^0 = \sum_{p=1}^n 1\{x_p = 0\}$. Show that

$$P\{S_1 = y_1, \dots S_n = y_n\} = p^{N^+} q^{N^-} r^{N^0} = p^{\frac{n-N^0+y_n}{2}} q^{\frac{n-N^0-y_n}{2}} r^{N^0}.$$

In particular, if we fix n, N^0 and y_n, then all the events of the form $\{S_1 = y_1, \dots S_n = y_n\}$ have the same probability. Next, use the observation made above to conclude

$$P(S_n = y) = \sum_{\substack{|y| \leq m \leq n \\ m \pm y \text{ are even}}} \binom{m}{\frac{m+y}{2}} p^{\frac{m+y}{2}} q^{\frac{m-y}{2}} r^{n-m}.$$

4.3 The law of large numbers for the random walk

Theorem 4.1.2 implies;

Theorem 4.3.1 *For an L^1-random walk (cf. Definition 4.2.1), define its mean velocity by*

$$v = (v^i)_{i=1}^d = \left(E[X_1^i]\right)_{i=1}^d. \tag{4.5}$$

Then,

$$\lim_{n\nearrow\infty} n^{-1} S_n = v, \quad P\text{-a.s.} \tag{4.6}$$

Remark: If we write S_n in a silly expression:

$$S_n = nv + (S_n - nv),$$

then (4.6) says that $\{S_n\}_{n\geq 1}$ almost surely follows a deterministic constant velocity motion $\{nv\}_{n\geq 1}$ by the error term $S_n - nv$ which is of order $o(n)$. In this sense, one can conclude that the random walk travels in the direction of v.

Exercise 4.3.1 Suppose that the random walk is of nearest neighbor (cf. (4.3)). Prove then that

$$v^i = P\{X_1 = e_i\} - P\{X_1 = -e_i\}.$$

Exercise 4.3.2 An $\mathbb{R}^d$-valued r.v. X is said to be *symmetric* if $P\{-X \in \cdot\} = P\{X \in \cdot\}$. A random walk is said to be *symmetric* if X_1 is symmetric. Check that a symmetric random walks with $P|X_1| < \infty$ has the mean velocity $v = 0$.

Exercise 4.3.3 Consider an L^1-random walk. Use Theorem 4.3.1 to prove that, if $v^i > 0$ (resp. $v^i < 0$), for some $i = 1, \ldots, d$, then

$$P\{\lim_{n\nearrow\infty} S^i_n = +\infty\} = 1, \quad (\text{resp. } P\{\lim_{n\nearrow\infty} S^i_n = -\infty\} = 1.)$$

5 Characteristic functions

5.1 Fourier transform

Definition 5.1.1 (**Fourier transform**) For $t \in \mathbb{R}$, we set

$$\mathbf{e}(t) = \exp(\mathbf{i}t), \quad \text{where } \mathbf{i} = \sqrt{-1} \tag{5.1}$$

• For $f \in L^1(\mathbb{R}^d)$, the *Fourier transform of* f is a function $\widehat{f} : \mathbb{R}^d \to \mathbb{C}$ defined by

$$\widehat{f}(\theta) = \int_{\mathbb{R}^d} \mathbf{e}(\theta \cdot x) f(x) dx. \tag{5.2}$$

• 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

$$\widehat{\mu}(\theta) = \int \mathbf{e}(\theta \cdot x) d\mu(x). \tag{5.3}$$

Remark: (5.2) can be thought of as a special case of (5.3). In fact, if we identify $f \in L^1(\mathbb{R}^d)$ with a signed Borel measure μ on $\mathbb{R}^d$ defined by $\mu(B) = \int_B f(x)dx$, $B \in \mathcal{B}(\mathbb{R}^d)$, we then have $\widehat{\mu} = \widehat{f}$.

The following fact has an important application to probability theory.

Lemma 5.1.2 (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) $\lim_{n \nearrow \infty} \widehat{\mu_n}(\theta) = \widehat{\mu_0}(\theta)$ *for all* $\theta \in \mathbb{R}^d$ *(cf. (5.3)).*

b) *For all* $f \in C_{\mathrm{b}}(\mathbb{R}^d)$,

$$\lim_{n \nearrow \infty} \int f d\mu_n = \int f d\mu_0. \tag{5.4}$$

• *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*

$$\lim_{n \nearrow \infty} \mu_n = \mu_0 \quad \textit{weakly}. \tag{5.5}$$

Before we prove Lemma 5.1.2, let us look at a simple example to get familiar with the notion of weak convergnce.

Example 5.1.3 Let μ be the uniform distribution on $(0,1)$ and $\mu_n = \frac{1}{n}\sum_{k=1}^n \delta_{k/n} \in \mathcal{P}(\mathbb{R})$ for $n \in \mathbb{N}^*$, where δ_x is a point mass at $x \in \mathbb{R}$. Then $\mu_n \to \mu$ weakly, since for any $f \in C_{\mathrm{b}}(\mathbb{R})$,

$$\lim_{n \nearrow \infty} \int f d\mu_n = \lim_{n \nearrow \infty} \frac{1}{n} \sum_{k=1}^n f\left(\frac{k}{n}\right) = \int_0^1 f(x)dx = \int f d\mu.$$

To prove Lemma 5.1.2, we will use:

Lemma 5.1.4 *Suppose* $f, \widehat{f} \in L^1(\mathbb{R}^d)$.

a) (Inversion formula) *For a.e.* $x \in \mathbb{R}^d$,

$$f(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} \mathbf{e}(-\theta \cdot x) \widehat{f}(\theta) d\theta. \tag{5.6}$$

In particular, if f is continuous, then, (5.6) holds for all $x \in \mathbb{R}^d$.

b) (Plancherel's formula) *Suppose in addition that f is bounded. Then, for a Borel finite measure μ on $\mathbb{R}^d$,*

$$\int f d\mu = (2\pi)^{-d} \int_{\mathbb{R}^d} d\theta \widehat{f}(\theta) \widehat{\mu}(-\theta). \tag{5.7}$$

Proof: See section 11.1.

Proof of Lemma 5.1.2: "(**b**) $\Rightarrow$ (**a**)" is obvious. We decompose "(**a**) $\Rightarrow$ (**b**)" into:

$$(\mathbf{a}) \;\Rightarrow\; (\mathbf{d})\&(\mathbf{e}) \;\Rightarrow\; (\mathbf{c})\&(\mathbf{e}) \;\Rightarrow\; (\mathbf{b})$$

where conditions **(c)–(e)** have been introduced as:

c) (5.4) holds for all $f \in C_{\mathrm{c}}(\mathbb{R}^d)$.

d) (5.4) holds for all $f \in C_{\mathrm{c}}^{\infty}(\mathbb{R}^d)$.

e) (5.4) holds for $f \equiv 1$.

(**a**) $\Rightarrow$ (**e**): Take $\theta = 0$.
(**a**) $\Rightarrow$ (**d**): We have for $f \in C_{\mathrm{c}}^{\infty}(\mathbb{R}^d)$ that $\widehat{f} \in L^1(\mathbb{R}^d)$, which is a well-known properties of the Fourier transform for the Schwartz space of rapidly decreasing (cf. [RS80, page 3, Theorem IX.1]). Thus, (d) follows from (5.7) and the dominated convergence theorem.
(**d**) $\Rightarrow$ (**c**): Take $f \in C_{\mathrm{c}}(\mathbb{R}^d)$ and note that

1) there is a sequence $f_n \in C_{\mathrm{c}}^{\infty}(\mathbb{R}^d)$ such that $\lim_{n \nearrow \infty} \sup_x |f_n(x) - f(x)| = 0$.

cf. Exercise 5.1.1. The desired convergence **(c)** then follows from a simple approximation argument (Exercise 5.1.2)..
(**c**)&(**e**) $\Rightarrow$ (**b**): Take $f \in C_{\mathrm{b}}(\mathbb{R}^d)$ and a sequence $g_\ell \in C_{\mathrm{c}}(\mathbb{R}^d \to [0,1])$ such that

$$g_1(x) \le g_2(x) \le ... \le g_\ell(x) \stackrel{\ell \to \infty}{\longrightarrow} 1 \quad \text{for all } x \in \mathbb{R}^d.$$

Suppose now that $\sup_x |f(x)| \le M$. We then have that (with "MC" referring to the monotone convergence theorem),

$$\begin{aligned}
\int f d\mu_0 + M\mu_0(\mathbb{R}^d) &= \int (f+M) d\mu \stackrel{\text{MC}}{=} \sup_\ell \int (f+M) g_\ell d\mu_0 \\
&= \sup_\ell \lim_{n \nearrow \infty} \int (f+M) g_\ell d\mu_n \quad \text{since } (f+M)g_\ell \in C_{\mathrm{c}}(\mathbb{R}^d) \\
&\le \underline{\lim}_{n \nearrow \infty} \sup_\ell \int (f+M) g_\ell d\mu_n \stackrel{\text{MC}}{=} \underline{\lim}_{n \nearrow \infty} \int (f+M) d\mu_n \\
&\stackrel{\text{(e)}}{=} \underline{\lim}_{n \nearrow \infty} \int f d\mu_n + M\mu_0(\mathbb{R}^d),
\end{aligned}$$

and hence that

2) $$\int f d\mu_0 \leq \underline{\lim}_{n \nearrow \infty} \int f d\mu_n.$$

By replacing f by $-f$ in (4) we have

$$\int f d\mu_0 \geq \overline{\lim}_{n \nearrow \infty} \int f d\mu_n,$$

which, together with (2), proves the desired convergence. □

Exercise 5.1.1 Check (1) in the proof of Lemma 5.1.2.

Exercise 5.1.2 Implement the "approximation" in the proof of Lemma 5.1.2, **(d)** $\Rightarrow$ **(c)**.

Exercise 5.1.3 Referring to Lemma 5.1.2 and its proof, is it true that **(c)** $\Rightarrow$ **(b)**?
Hint $\mu_n = \delta_{x_n}$, where $|x_n| \to \infty$.

Exercise 5.1.4 $(\star)$ Referring to Lemma 5.1.2, prove that (b) implies[9] the following (b1)–(b3):
b1) $\mu_0(G) \leq \underline{\lim}_{n \to \infty} \mu_n(G)$ for any open subset $G \subset \mathbb{R}^d$.
b2) $\overline{\lim}_{n \to \infty} \mu_n(F) \leq \mu_0(F)$ for any closed subset $F \subset \mathbb{R}^d$.
b3) $\lim_{n \to \infty} \mu_n(B) = \mu_0(B)$ for any Borel subset $B \subset \mathbb{R}^d$ such that $\partial B = 0$.
Hint: Proof of Lemma 2.1.2 (c) $\Rightarrow$ (b).

Exercise 5.1.5 $(\star)$ Suppose that $X, X_1, X_2...$ are real r.v's and that $X_n \to X$ weakly. 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 $P(Y > m) > 0$.

Corollary 5.1.5 (Injectivity of the Fourier transform) *Suppose that μ and ν are Borel finite measures on $\mathbb{R}^d$. Then, the following are equivalent.*

a) $\widehat{\mu} = \widehat{\nu}$.

b) $\mu = \nu$.

Proof: (b) $\Rightarrow$ (a): Obvious.
(a) $\Rightarrow$ (b): We consider a "sequence" $\mu_0 = \mu$ and $\mu_n = \nu$, $n \geq 1$. We have for all $\theta \in \mathbb{R}^d$ that

1) $$\lim_{n \nearrow \infty} \widehat{\mu_n}(\theta) \stackrel{(\mu_n = \nu)}{=} \widehat{\nu}(\theta) \stackrel{\text{(a)}}{=} \widehat{\mu}(\theta).$$

Thus, we have by Lemma 5.1.2 that for all $f \in C_b(\mathbb{R}^d)$,

$$\int f d\mu \stackrel{(1),\text{Lemma 5.1.2}}{=} \lim_{n \nearrow \infty} \int f d\mu_n \stackrel{(\mu_n = \nu)}{=} \int f d\nu,$$

which implies (b) by Lemma 2.1.2. □

[9] In fact, each of (b1)–(b3) is equivalent to (b).

5.2 Characteristic functions

Definition 5.2.1 (**Characteristic function**) • For an $\mathbb{R}^d$-valued r.v. X, a function:

$$\mathbb{R}^d \ni \theta \mapsto E[\mathbf{e}(\theta \cdot X)] \in \mathbb{C}$$

is called the *characteristic function* (*ch.f.* for short) of X.

Remark: If $P(X \in \cdot) = \mu$, then

$$E[\mathbf{e}(\theta \cdot X)] = \int \mathbf{e}(\theta \cdot x) d\mu(x) = \widehat{\mu}(\theta).$$

Therefore,

the ch.f. of a r.v. = the Fourier transform of its distribution.

Lemma 5.2.2 (Criterion of the independence) *Let X_j be r.v. with values in $\mathbb{R}^{d_j}$ ($j = 1, ..., n$). Then, the following are equivalent:*

a) $E\left[\prod_{j=1}^n \mathbf{e}(\theta_j \cdot X_j)\right] = \prod_{j=1}^n E[\mathbf{e}(\theta_j \cdot X_j)]$ *for all* $\theta_j \in \mathbb{R}^{d_j}$ ($j = 1, ..., n$).

b) $\{X_j\}_{j=1}^n$ *are independent.*

Proof: Let

$$\begin{aligned} \mu &= P((X_j)_{j=1}^n \in \cdot) \in \mathcal{P}(\mathbb{R}^d), \quad \text{where } d = d_1 + ... + d_n, \\ \nu &= \otimes_{j=1}^n \mu_j \in \mathcal{P}(\mathbb{R}^d), \quad \text{where } \mu_j = P(X_j \in \cdot) \in \mathcal{P}(\mathbb{R}^{d_j}) \ (1 \le j \le n). \end{aligned}$$

Then,

$$E\left[\prod_{j=1}^n \mathbf{e}(\theta_j \cdot X_j)\right] = E\left[\mathbf{e}\left(\sum_{j=1}^n \theta_j \cdot X_j\right)\right] = \widehat{\mu}\left((\theta_j)_{j=1}^n\right)$$

and

$$\prod_{j=1}^n E\left[\mathbf{e}(\theta_j \cdot X_j)\right] = \prod_{j=1}^n \widehat{\nu_j}(\theta_j) \stackrel{\text{Fubini}}{=} \widehat{\nu}\left((\theta_j)_{j=1}^n\right).$$

Therefore,

$$(\text{a}) \stackrel{\text{Corollary 5.1.5}}{\Longleftrightarrow} \mu = \nu \stackrel{\text{Proposition 2.3.1}}{\Longleftrightarrow} (\text{b}).$$

□

Definition 5.2.3 (**Weak convergence of r.v.'s**) Let $X, X_1, X_2, ...$ be $\mathbb{R}^d$-valued r.v.'s.

• The sequence $(X_n)_{n \ge 1}$ is said to *converge weakly* (or *converge in law*) to X if

$$\lim_{n \nearrow \infty} P(X_n \in \cdot) = P(X \in \cdot) \text{ weakly},$$

or equivalently (cf. Lemma 5.1.2),

$$\lim_{n \nearrow \infty} E[f(X_n)] = E[f(X)] \quad \text{for all } f \in C_{\mathrm{b}}(\mathbb{R}^d).$$

We will henceforth denote this convergence by

$$X_n \longrightarrow X \text{ weakly} \quad \text{or} \quad X_n \longrightarrow X \text{ in law}$$

Lemma 5.2.4 (Criterion of the weak convergence) *Let X_n be r.v. with values in $\mathbb{R}^d$ $(n = 0, 1, ..)$. Then, the following are equivalent:*

a) $\lim_{n\nearrow\infty} E[\mathbf{e}(\theta \cdot X_n)] = E[\mathbf{e}(\theta \cdot X_0)]$ *for all* $\theta \in \mathbb{R}^d$.

b) $\lim_{n\nearrow\infty} X_n = X_0$ *weakly.*

Proof: Let $\mu_n = P(X_n \in \cdot)$. Then,

$$E[\mathbf{e}(\theta \cdot X_n)] = \widehat{\mu_n}(\theta), \quad n = 0, 1, ...$$

Thus, by Lemma 5.1.2, (a) is equivalent to $\mu_n \longrightarrow \mu_0$ weakly. □

Exercise 5.2.1 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).
(a) $\lim_{n\nearrow\infty} X_n = X$, P-a.s. (b) $\lim_{n\nearrow\infty} X_n = X$ in $L^p(P)$ for some $p \geq 1$.
(c) $\lim_{n\nearrow\infty} X_n = X$ in probability, i.e., $\lim_{n\nearrow\infty} P\{|X_n - X| > \varepsilon\}$ =0 for any $\varepsilon > 0$.
(d) $X_n \longrightarrow X$ in law.

Exercise 5.2.2 Show by an example that (d) $\not\Rightarrow$ (c) in Exercise 5.2.1.
Hint: $X_n = (-1)^n X$ for a symmetric r.v. X.

Exercise 5.2.3 Let $X, Y, X_1, X_2, \ldots$ be $\mathbb{R}^d$-valued r.v.'s such that $X_n \to X$ weakly. Is it true in general that $X_n + Y \to X + Y$ weakly?

Exercise 5.2.4 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 $X_n \to c$ in law.
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 5.2.5 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 $X_n \to X$ and $Y_n \to Y$ in law. Prove then that $(X_n, Y_n) \to (X, Y)$ in law, and hence that $F(X_n, Y_n) \to F(X, Y)$ in law for any $F \in C(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2})$.

Exercise 5.2.6 Let (X_n, Y_n) be r.v.'s with values in $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. Suppose that $X_n \to X$ and $Y_n \to c$ in law (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 $(X_n, Y_n) \to (X, c)$ in law, and hence that $F(X_n, Y_n) \to F(X, c)$ in law for any $F \in C(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2})$. Hint: It is enough to show that

$$\lim_{n\nearrow\infty} E[\mathbf{e}(\theta_1 \cdot X_n + \theta_2 \cdot Y_n)] = E[\mathbf{e}(\theta_1 \cdot X + \theta_2 \cdot c)] \quad \text{for } (\theta_1, \theta_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}.$$

In doing so, uniform continuity of the map $(x, y) \mapsto \mathbf{e}(\theta_1 \cdot x + \theta_2 \cdot y)$ would help.

Exercise 5.2.7 Let $X, Y_1, Y_2, ...$ be r.v.'s with $P\{X \in \cdot\} = \gamma_{r,a}$ and $P\{Y_n \in \cdot\} = \beta_{a,n}$ ($n = 1, 2, ..$ cf. (1.18), (1.19)). Prove then that $nY_n \to X$ in law.
Hint: Let $S_n = X_1 + ... + X_n$ where $X_1, X_2, ...$ be i.i.d. such that $P\{X_n \in \cdot\} = \gamma_{r,1}$. Then, $P\{nY_n \in \cdot\} = P\{\frac{nX}{X+S_n} \in \cdot\}$ by (3.1). Moreover, $\frac{nX}{X+S_n} \to X$, P-a.s. by Theorem 4.1.2.

Exercise 5.2.8 ($\star$) Let X be a real r.v.and $\varphi(\theta) = E[\mathbf{e}(\theta X)]$. Then, $\varphi \in C^2 \iff X \in L^2(P)$. Prove this by assuming that X is symmetric (cf. Exercise 5.2.9 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 5.2.9 ($\star$) 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 5.2.8

Exercise 5.2.10 ($\star$) (*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 $\lfloor y \rfloor = \max\{n \in \mathbb{Z}\ ;\ n \le y\}$ for $y \in \mathbb{R}$. Prove then that the measures $\mu_n = \frac{1}{n}\sum_{k=1}^n \delta_{\alpha_n}$ converges weakly to the uniform distribution on $(0,1)$. Hint: We want to prove that $\lim_{n\nearrow\infty} \frac{1}{n}\sum_{k=1}^n f(\alpha_n) = \int_0^1 f(\theta)d\theta$ for all $f \in C_{\mathrm{b}}(\mathbb{R})$. By, Exercise 3.2.3, it is enough to prove this for all trigonometric polynomial f.

5.3 Examples

Example 5.3.1 (*ch.f. of a Uniform r.v.*) Suppose that a r.v. U is uniformly distributed on an interval $[a,b]$ (cf. (1.14)). We then see from a direct computation that

$$E\mathbf{e}(\theta \cdot U) = \widehat{u}(\theta) = \frac{\mathbf{e}(\theta b) - \mathbf{e}(\theta a)}{\mathbf{i}(b-a)\theta}, \tag{5.8}$$

where $u(t) = (b-a)^{-1}1_{[a,b]}(t)$.

Exercise 5.3.1 Let U_1, U_2 be i.i.d. with uniform distribution on $(-1,1)$. (i) Show that $P(\frac{U_1+U_2}{2} \in \cdot)$ has the density $f(x) \stackrel{\text{def.}}{=} (1-|x|)^+$ and that $f^\wedge(\theta) = \frac{\sin^2(\theta/2)}{(\theta/2)^2}$. (ii) Show that $\int_{-\infty}^{\infty} f^\wedge(\theta)d\theta = 2\pi$. $\frac{1}{2\pi}f^\wedge$ is the density of *Polya's distribution.* Hint: (5.6).

Exercise 5.3.2 Let $2 \le q \in \mathbb{N}$. Then, show that the following infinite product converges for all $\theta \in \mathbb{R}$ and that $\mathbf{e}(\theta) = 1 + \theta P(\theta)$:

$$P(\theta) = \prod_{n=1}^{\infty} \frac{1 + \mathbf{e}(\theta q^{-n}) + \mathbf{e}(2\theta q^{-n}) + \ldots + \mathbf{e}((q-1)\theta q^{-n})}{q}.$$

Example 5.3.2 (*ch.f. of a Gaussian r.v.*) Let X be the Gaussian r.v. with the strictly positive definite covariance matrix V, cf. Example 1.3.2. Let us prove that

1) $\quad E\exp(z\theta \cdot X) = \exp\left(\frac{1}{2}z^2\theta \cdot V\theta\right)$ for all $\theta \in \mathbb{R}^d$ and $z \in \mathbb{C}$

and hence that

$$\widehat{\nu_V}(\theta) = \exp\left(-\tfrac{1}{2}\theta \cdot V\theta\right). \tag{5.9}$$

Note first that both hand sides of (1) are holomorphic in z. Therefore, it is enough to prove it for all $z = t \in \mathbb{R}$. Since

2) $\quad t\theta \cdot x - \frac{1}{2}x \cdot V^{-1}x = \frac{1}{2}t^2\theta \cdot V\theta - \frac{1}{2}(x - tV\theta) \cdot V^{-1}(x - tV\theta),$

we have

$$
\begin{aligned}
E\exp(t\theta\cdot X) &= (\det(2\pi V))^{-1/2}\int dx \exp\left(t\theta\cdot x-\tfrac{1}{2}x\cdot V^{-1}x\right)\\
&\stackrel{(2)}{=} \exp\left(\tfrac{1}{2}t^2\theta\cdot V\theta\right)(\det(2\pi V))^{-1/2}\int dx\exp\left(-\tfrac{1}{2}(x-tV\theta)\cdot V^{-1}(x-tV\theta)\right)\\
&= \exp\left(\tfrac{1}{2}t^2\theta\cdot V\theta\right).
\end{aligned}
$$

We take this opportunity to generalize the definition of Gaussian measure. Let V be a symmetric, non-negative definite (not necessarily positive definite) $d\times d$-matrix.

- We define $\nu_V \in \mathcal{P}(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d))$ as a unique measure such that (5.9) holds.

The existence of such measure can be seen as follows. By diagonalizing V, we find a symmetric, non-negative definite matrix W such that $W^2 = V$. Let X be r.v. with the standard normal distribution on $\mathbb{R}^d$ and set $\nu_V = P(WX\in\cdot)$. Then, it is easy to see (5.9). The uniqueness follows from Corollary 5.1.5.

Exercise 5.3.3 (*Stability of Gaussian distribution*) Let X_1, X_2 be $\mathbb{R}^d$-valued r.v.'s such that $P((X_1,X_2)\in\cdot) = \nu_{V_1}\otimes\nu_{V_2}$, cf. (1.15) and A_1, A_2 be $d\times d$ matrices. Prove then that

$$P(A_1X_1 + A_2X_2\in\cdot) = \nu_V,\quad \text{where } V = A_1V_1A_1^* + A_2V_2A_2^*$$

Hint: Prove that $E\mathbf{e}(\theta\cdot(A_1X_1+A_2X_2)) = \widehat{\nu}_V(\theta)$ and use Corollary 5.1.5.

Exercise 5.3.4 Let $X = (X^i)_{i=1}^d$ be a mean-zero $\mathbb{R}^d$-valued r.v. Prove then that X is a Gaussian r.v. if and only if $\sum_{i=1}^d \theta^i X^i$ is a Gaussian r.v. for any $\theta = (\theta^i)_{i=1}^d\in\mathbb{R}^d$. Hint: (5.9), Corollary 5.1.5.

Exercise 5.3.5 Suppose that $X = (X^i)_{i=1}^d$ is a mean-zero $\mathbb{R}^d$-valued Gaussian r.v. Prove then that coordinates $\{X^i\}_{i=1}^d$ are independent if and only if $E[X^iX^j] = 0$ for $i\neq j$. This shows in particular that the independence for r.v.'s $\{X^i\}_{i=1}^d$ above follows from the pairwise independence. Hint: (5.9), Lemma 5.2.2.

Exercise 5.3.6 Let $X, X_1, X_2, ...$ $\mathbb{R}^d$-valued r.v.'s. Suppose that X_n $(n=1,2,\ldots)$ are mean-zero Gaussian r.v.'s and that they converge in law to X. Prove then that X is a mean-zero Gaussian r.v. and that the covariance matrix $\Gamma = (\gamma_{ij})_{i,j=1}^n$ is given by $\gamma_{ij} = \lim\limits_{n\nearrow\infty} E[X_n^iX_n^j]$.
Hint: Consider characteristic functions to see that limits γ_{ij} $(1\le i,j\le n)$ exist. Prove then that $E[\mathbf{e}(\theta\cdot X)] = \exp(-\theta\cdot\Gamma\theta/2)$.

Example 5.3.3 (*ch.f. of a Gamma r.v.*) Let X be a real r.v. such that $P(X\in\cdot) = \gamma_{r,a}$. We then see from a direct computation that

$$E\exp(-zX) = \left(1+\frac{z}{r}\right)^{-a}\quad \text{for any } z\in\mathbb{C} \text{ with } \mathrm{Re}(z)\ge 0,$$

In particular,

$$\widehat{\gamma_{r,a}}(\theta) = \left(1-\frac{\mathbf{i}\theta}{r}\right)^{-a}. \tag{5.10}$$

Exercise 5.3.7 Let (X, Y) be $\mathbb{R}^2$-valued r.v. such that $P((X, Y) \in \cdot) = \gamma_{r,a} \otimes \gamma_{r,b}$, where $r, a, b > 0$. (i) Prove that the distribution of $X - Y$ has the density:

$$f_{r,a,b}(z) = \frac{r^{a+b}e^{rz}}{\Gamma(a)\Gamma(b)} \int_{z\vee 0}^{\infty} x^{a-1}(x-z)^{b-1}e^{-2rx}dx$$

and that $f_{r,a,b}^{\wedge}(\theta) = \left(1 - \frac{\mathbf{i}\theta}{r}\right)^{-a}\left(1 + \frac{\mathbf{i}\theta}{r}\right)^{-b}$. In particular, $f_{r,a,a}^{\wedge}(\theta) = \left(1 + \frac{\theta^2}{r^2}\right)^{-a}$.
(ii) Suppose that $a + b > 1$, so that $f_{r,a,b}^{\wedge} \in L^1(\mathbb{R}^d)$. Then, use the inversion formula (5.6) to show that

$$f_{r,a,b}(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathbf{e}(-\theta z) f_{r,a,b}^{\wedge}(\theta)d\theta.$$

(iii)($\star$) The value of the integral $\int_{-\infty}^{\infty} f_{1,\frac{a+1}{2},\frac{a+1}{2}}^{\wedge}(\theta)d\theta$ can be evaluated in two different way: one is to use (i) and (ii) above, and another is to use (3.2). Compare these two to deduce:

$$\Gamma(a) = \frac{2^{a-1}}{\sqrt{\pi}}\Gamma\left(\frac{a}{2}\right)\Gamma\left(\frac{a+1}{2}\right) \quad a > 0.$$

Exercise 5.3.8 ($\star$) (i) Let $f : [0, \infty) \to \mathbb{R}$ be a Borel function such that $\int_0^{\infty} r^2|f(r)|dr < \infty$. Prove then that $F \in L^1(\mathbb{R}^3)$ and that $\widehat{F}(\theta) = 4\pi \int_0^{\infty} r^2 f(r) \frac{\sin(r|\theta|)}{r|\theta|}dr$, where $F(x) = f(|x|)$ and $\frac{\sin 0}{0} = 1$. (ii) Use (i) and Example 5.3.3 to show that

$$\widehat{F_{c,a}}(\theta) = \frac{c^{a+1}}{a|\theta|}\mathrm{Im}\left(\frac{c + \mathbf{i}|\theta|}{c^2 + |\theta|^2}\right)^a \quad \text{for} \quad F_{c,a}(x) = \frac{c^{a+1}}{4\pi\Gamma(a+1)}|x|^{a-2}e^{-c|x|}, \quad a, c > 0.$$

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 5.3.6.

Example 5.3.4 (*ch.f. of a Poisson r.v.*) Let N be a r.v. with $P(N \in \cdot) = \pi_r$, ($r > 0$, cf. (1.22)). We then see from a direct computation that $E[z^N] = \exp((z-1)r)$ for any $z \in \mathbb{C}$, which shows in particular that

$$\widehat{\pi_r}(\theta) = \exp((\mathbf{e}(\theta) - 1)r). \tag{5.11}$$

Exercise 5.3.9 (*Stability of Poisson distribution*) Let N_1, N_2 be r.v.'s such that $P((N_j)_{j=1}^2 \in \cdot) = \otimes_{j=1}^2 \pi_{r_j}$, cf. (1.22). Prove then that

$$P(N_1 + N_2 \in \cdot) = \pi_{r_1+r_2}.$$

Hint: Prove that $E\mathbf{e}(\theta(N_1 + N_2)) = \widehat{\pi}_{r_1+r_2}(\theta)$ and use Corollary 5.1.5.

Example 5.3.5 (*ch.f. of a Cauchy r.v.: dimension one*) Suppose that an $\mathbb{R}$-valued r.v. Y has (c)-Cauchy distribution (cf. (1.21)). Then,

$$E\mathbf{e}(\theta Y) = \exp(-c|\theta|), \quad \theta \in \mathbb{R}. \tag{5.12}$$

Let $f(x) = \frac{e^{-c|x|}}{2c}$ (the density of the *double exponential distribution*). Then,

$$\begin{aligned}\widehat{f}(\theta) &= \int_0^\infty \frac{e^{-(c-\mathbf{i}\theta)x}}{2c}dx + \int_0^\infty \frac{e^{-(c+\mathbf{i}\theta)x}}{2c}dx \\ &= \frac{1}{2c}\left(\frac{1}{c-\mathbf{i}\theta} + \frac{1}{c+\mathbf{i}\theta}\right) = \frac{1}{c^2+\theta^2}.\end{aligned}$$

Thus,

$$f(x) \stackrel{(5.6)}{=} \frac{1}{2\pi}\int_{-\infty}^\infty \mathbf{e}(-\theta x)\widehat{f}(\theta)d\theta = \frac{1}{2\pi}\int_{-\infty}^\infty \frac{\mathbf{e}(-\theta x)}{c^2+\theta^2}d\theta,$$

from which (5.12) follows.

Exercise 5.3.10 Apply the residue theorem to a melomorphic function $\frac{\exp(i\theta z)}{c^2+z^2}$ to give an alternative proof of (5.12).

Exercise 5.3.11 (*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 and Y_1 have the same distribution for all $n \geq 1$. This shows that S_n/n does not converge to a constant, even weakly (cf. Theorem 4.1.2).

Exercise 5.3.12 Let $S_n = X_1 + ... + X_n$, where $(X_n)_{n\geq 1}$ are i.i.d. with Polya distributions (Exercise 5.3.1). Prove then that $P(S_n/n \in \cdot)$ converges weakly to (1)-Cauchy distribution as $n \to \infty$.

Example 5.3.6 ($\star$) (*ch.f. of a Cauchy r.v.: higher dimensions*) Suppose that an $\mathbb{R}^d$-valued r.v. Y has (c)-Cauchy distribution (cf. (1.21)). Then,

$$E\mathbf{e}(\theta \cdot Y) = \exp(-c|\theta|), \quad \theta \in \mathbb{R}^d. \tag{5.13}$$

We will use (3.2) to prove this. Let $\xi, X_1, X_2, ..., X_d$ be independent r.v.'s with $P(\xi \in \cdot) = \gamma_{c^2/2,1/2}$ (cf. (1.18)) and $P(X_j \in \cdot) = \nu_1$, $1 \leq j \leq d$ (cf. (1.15)). Let us write $X = (X_j)_{j=1}^d$ for simplicity. Then,

$$\begin{aligned}E\mathbf{e}(\theta\cdot Y) &\stackrel{(3.2)}{=} E\mathbf{e}(\theta\cdot\xi^{-1/2}X) = \int \gamma_{c^2/2,1/2}(dx)E\mathbf{e}(\theta\cdot x^{-1/2}X) \\ &\stackrel{(5.9)}{=} \int \gamma_{c^2/2,1/2}(dx)\exp\left(-\frac{|\theta|^2}{2x}\right) \stackrel{(1.20)}{=} \exp(-c|\theta|).\end{aligned}$$

Remark: An alternative proof of (5.13) for $d = 3$ can be given by applying the inversion formula (5.6) to $\widehat{F_{c,2}}(\theta) = \frac{c^4}{(c^2+|\theta|^2)^2}$ (Exercise 5.3.8) as in Example 5.3.5.

5.4 Convergence of moments

Let $(Y_n)_{n\geq 0}$ be $\mathbb{R}^d$-valued r.v.'s such that $Y_n \to Y_0$ weakly, and let $f \in C(\mathbb{R}^d)$. If f is bounded, we have

$(*)$ $\lim_{n\nearrow\infty} E[f(Y_n)] = E[f(Y_0)]$.

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 5.4.1 (**uniform integrability**) Real r.v.'s $(X_n)_{n\geq 1}$ are said to be *uniformly integrable* (*u.i.* in short) if

$$\sup_{n\geq 1} E[|X_n|\mathbf{1}\{|X_n| > m\}] \longrightarrow 0 \text{ as } m \to \infty.$$

The next lemma shows that the uniform integrability is close to, but slightly more than that

$$\sup_{n\geq 1} E[|X_n|] < \infty. \tag{5.14}$$

Lemma 5.4.2 *Let $(X_n)_{n\geq 1}$ be real r.v.'s.*

a) *If $(X_n)_{n\geq 1}$ are u.i., then (5.14) holds.*

b) *Suppose that there exists a non-decreasing $\varphi : [0, \infty) \to [0, \infty)$ such that*

$$\lim_{|x|\to\infty} \varphi(x) = \infty, \quad \sup_{n\geq 1} E[|X_n|\varphi(|X_n|)] < \infty.$$

Then, $(X_n)_{n\geq 1}$ are u.i.

Proof: Let $\varepsilon_m = \sup_{n\geq 1} E[|X_n|\mathbf{1}\{|X_n| > m\}]$.
a):$\varepsilon_m \leq 1$ for large enough m, and for such m and for all $n \geq 1$,

$$E[|X_n|] \leq E[|X_n|\mathbf{1}\{|X_n| \leq m\}] + E[|X_n|\mathbf{1}\{|X_n| > m\}] \leq m + \varepsilon_m < m + 1.$$

b): By the monotonicity of φ and (a variant of) Chebychev's inequality (Exercise 1.2.3),

$$E[|X_n|\mathbf{1}\{|X_n| > m\}] \leq E[|X_n|\mathbf{1}\{\varphi(|X_n|) > \varphi(m)\}] \leq \varphi(m)^{-1}E[|X_n|\varphi(|X_n|)].$$

Thus, $\varepsilon_m \leq \varphi(m)^{-1}C \to 0$, as $m \to \infty$, where $C = \sup_{n\geq 1} E[|X_n|\varphi(X_n)] < \infty$. □

Exercise 5.4.1 Prove that real r.v.'s $(X_n)_{n\geq 1}$ are u.i. if $E[\sup_{n\geq 1} |X_n|] < \infty$.

Example 5.4.3 Let $S_n = X_1 + ... + X_n$, where $(X_n)_{n\geq 1}$ are real r.v.'s such that

$$\sup_{n\geq 1} \text{var}(X_n) \leq M < \infty, \quad \text{cov}(X_m, X_n) = 0 \text{ if } m \neq n.$$

Then, $Y_n = (S_n - E[S_n])/\sqrt{n}$ are u.i. In fact,

$$E[|Y_n|^2] = \frac{1}{n}\text{var}(S_n) = \frac{1}{n}\sum_{k=1}^{n} \text{var}(X_k) \leq M.$$

Thus, Lemma 5.4.2(b) applies.

Proposition 5.4.4 *Suppose that $(X_n)_{n\geq 0}$ be real r.v.'s such that $X_n \to X_0$ weakly.*

a) (Fatou's lemma for weak convergence)

$$E[|X_0|] \le \varliminf_{n \nearrow \infty} E[|X_n|]. \tag{5.15}$$

b) (Convergence of moments) *Supppose in addition that $(X_n)_{n\ge 1}$ are uniformly integrable (Definition 5.4.1). Then,*

$$\sup_{n\ge 0} E[|X_n|] \;<\; \infty, \tag{5.16}$$

$$\lim_{n\nearrow\infty} E[X_n] \;=\; E[X_0]. \tag{5.17}$$

Proof: a): Since $\mathbb{R} \ni x \mapsto |x| \wedge m$ is in $C_{\rm b}(\mathbb{R})$ for any $m > 0$, we have that

$$E[|X_0|] = \sup_{m\ge 0} E[|X_0| \wedge m] = \sup_{m\ge 0} \lim_{n\nearrow\infty} E[|X_n| \wedge m] \le \varliminf_{n\nearrow\infty} E[|X_n|].$$

b): (5.16) follows from Lemma 5.4.2(a) and (5.15). To prove (5.17), we first consider:
Case 1: $X_n \ge 0$ for all $n \in \mathbb{N}$. By (5.15), it is enough to show that

1) $$\varlimsup_{n\nearrow\infty} E[X_n] \le E[X_0].$$

Note that

2) $$\varlimsup_{n\nearrow\infty} E[X_n \mathbf{1}\{X_n \le m\}] \le E[X_0] \quad \text{for any } m > 0.$$

In fact,

$$\begin{aligned}\varlimsup_{n\nearrow\infty} E[X_n \mathbf{1}\{X_n \le m\}] &\le \varlimsup_{n\nearrow\infty} E[X_n \wedge m] \\ &= E[X_0 \wedge m] \le E[X_0].\end{aligned}$$

Then,

$$\begin{aligned}\varlimsup_{n\nearrow\infty} E[X_n] &= \varlimsup_{n\nearrow\infty} \left(E[X_n \mathbf{1}\{X_n \le m\}] + E[X_n \mathbf{1}\{X_n > m\}]\right) \\ &\overset{(2)}{\le} E[X_0] + \varepsilon_m,\end{aligned}$$

where $\varepsilon_m = \sup_{n\ge 1} E[X_n \mathbf{1}\{X_n > m\}]$. Since m is arbitrary, we get (1).
Case 2: general case. This can be reduced to Case 1. In fact,

$$X_n \longrightarrow X_0 \text{ weakly} \implies X_n^{\pm} \longrightarrow X_0^{\pm} \text{ weakly}$$

and

$$(X_n)_{n\ge 1} \text{ are uniformly integrable.} \implies \text{so are } (X_n^{\pm})_{n\ge 1}.$$

Thus, by the result for Case 1,

$$\sup_{n\ge 0} E[|X_n|] \le \sup_{n\ge 0} E[X_n^+] + \sup_{n\ge 0} E[X_n^-] < \infty$$

and

$$E[X_0] = E[X_0^+] - E[X_0^-] = \lim_{n\nearrow\infty} (E[X_n^+] - E[X_n^-]) = \lim_{n\nearrow\infty} E[X_n].$$

□

Remark: In Proposition 5.4.4(b), it is not true in general that

$$\lim_{n \nearrow \infty} E[|X_n - X_0|] = 0. \tag{5.18}$$

For example, if $0 \neq X_0 \in L^1(P)$, X_0 is symmetric, and $X_n = (-1)^n X_0$. Then, all the assumptions for Proposition 5.4.4(b) are satisfied. But $|X_n - X_0| = 2|X_0|$ for odd n's.

Exercise 5.4.2 Suppose that $(X_n)_{n \geq 0}$ be uniformly integrable real r.v.'s such that $X_n \to X_0$ in probability. Prove then that (5.18) is true.

6 How much does a random walk fluctuate?

6.1 The central limit theorem

Definition 6.1.1 (covariance matrix) • For an $\mathbb{R}^d$-valued L^2-r.v. X, we define its *covariance matrix* $\Gamma = (\gamma^{ij})_{i,j=1}^d$ by

$$\gamma^{ij} = \mathrm{cov}(X^i, X^j) = E[X^i X^j] - E[X^i]E[X^j]. \tag{6.1}$$

• For an L^2-random walk $S_n = X_1 + ... + X_n$ (cf Definition 4.2.1), we define its *covariance matrix* $\Gamma = (\gamma^{ij})_{i,j=1}^d$ as that of X_1.

Exercise 6.1.1 Suppose that the random walk is of nearest neighbor (cf. (4.3)). Prove the following. (i)$v^i = p(e_i) - p(-e_i)$ and $\gamma^{ij} = \delta_{ij}(p(e_i) + p(-e_i)) - v^i v^j$, where $p(x) = P(X_1 = x)$. ii)Two different coordinates X_n^i, X_n^j $(i \neq j)$ of X_n are not independent of each other, even though they are uncorrelated if $v = 0$.

Theorem 6.1.2 (The Central Limit Theorem[10]) *Let $S_n = X_1 + ... + X_n$ be an L^2-random walk (Definition 4.2.1) with $v = (E[X_1^i])_{i=1}^d \in \mathbb{R}^d$ and $\Gamma = (\mathrm{cov}(X_1^i, X_1^j))_{i,j=1}^n$ ((4.5), (6.1)). Then,*

$$P\left(\frac{S_n - nv}{\sqrt{n}} \in \cdot\right) \stackrel{n\to\infty}{\longrightarrow} \nu_\Gamma \quad \textit{weakly} \tag{6.2}$$

where ν_Γ is the Gaussian measure with the covariance matrix Γ (cf. Example 5.3.2).

Remark : Theorem 6.1.2 tells us the following information on the distribution of S_n for large n. Let χ be r.v. such that $P(\chi \in \cdot) = \nu_\Gamma$. Roughly speaking, Theorem 6.1.2 says that for large n,

$$P(n^{-1/2}(S_n - nv) \in \cdot) \approx P(\chi \in \cdot)$$

or

$$P(S_n \in \cdot) \approx P(nv + n^{1/2}\chi \in \cdot).$$

Although it requires some work to prove CLT in the generality of Theorem 6.1.2, the proof is remarkably easy in some examples:

Example 6.1.3 Note that

$$\textbf{1)} \qquad \mathbf{e}(\theta) = 1 + \mathbf{i}\theta - \frac{\theta^2}{2} + O(\theta^3) \ \text{ as } \theta \to 0.$$

Thus, if S_n is an (nr)-Poisson r.v. (Exercise 5.3.9),

$$\begin{aligned} E\left[\mathbf{e}\left(\theta \frac{S_n - rn}{\sqrt{n}}\right)\right] &\stackrel{(5.11)}{=} \exp\left(nr\left(\mathbf{e}\left(\frac{\theta}{\sqrt{n}}\right) - 1\right) - \mathbf{i}\theta r\sqrt{n}\right) \\ &\stackrel{(1)}{=} \exp\left(-\frac{r\theta^2}{2} + nrO\left(\frac{\theta^3}{n^{3/2}}\right)\right) \\ &\stackrel{n\to\infty}{\longrightarrow} \exp\left(-\frac{r\theta^2}{2}\right) \stackrel{(5.9)}{=} \widehat{\nu_r}(\theta), \end{aligned}$$

[10] Often abbreviated as CLT.

We see from this and Lemma 5.2.4 that

$$P\left(\frac{S_n - nr}{\sqrt{n}} \in \cdot\right) \stackrel{n\to\infty}{\longrightarrow} \nu_r \text{ weakly}$$

When $(X_n)_{n\geq 1}$ in Theorem 6.1.2 are (r)-Poisson r.v., then, $S_n \stackrel{\text{def}}{=} X_1+...+X_n$ is an (nr)-Poisson r.v. (Exercise 5.3.9). Thus we have proved Theorem 6.1.2 in this special case.

Exercise 6.1.2 Let $(N_r)_{r>0}$ be r.v.'s such that $P(N_r = k) = e^{-r}r^k/k!$ for all $k \in \mathbb{N}$ and $r > 0$. Generalize the computation in Example 6.1.3 to prove that $P(\frac{N_r - r}{\sqrt{r}} \in \cdot)$ converges weakly to the standard normal distribution as $r \to \infty$.

Exercise 6.1.3 Suppose that $P(X_n \in \cdot) = \nu_\Gamma$ in Theorem 6.1.2. Prove then that $P\left(\frac{S_n - nv}{\sqrt{n}} \in \cdot\right) = \nu_\Gamma$ for any $n \geq 1$. Thus the theorem in this special case is trivial.

Exercise 6.1.4 Let $f_{r,a}(x) = \frac{r^a x^{a-1}}{\Gamma(a)}e^{-rx}\mathbf{1}\{x > 0\}$ (the density of (r,a)-Gamma distribution, $r, a > 0$). (i) Show that $f^{\wedge}_{\sqrt{a},a}(\theta)e^{-\mathbf{i}\theta\sqrt{a}} \stackrel{a\to\infty}{\longrightarrow} \exp(-\theta^2/2)$ for all $\theta \in \mathbb{R}$ and conclude from this that

$$P\left(X_a - \sqrt{a} \in \cdot\right) = P\left(\frac{Y_a - a}{\sqrt{a}} \in \cdot\right) \stackrel{a\to\infty}{\longrightarrow} \nu_1, \quad \text{weakly,}$$

where $P(X_a \in \cdot) = \gamma_{\sqrt{a},a}$ and $P(Y_a \in \cdot) = \gamma_{1,a}$. Hint: (5.10) and $-\text{Log}(1-z) = \sum_{n\geq 1}\frac{z^n}{n} = z + \frac{z^2}{2} + O(|z|^3)$ as $z \to 0$. (ii) Show that $|f^{\wedge}_{\sqrt{a},a}(\theta)| = \left(1 + \frac{\theta^2}{a}\right)^{-a/2}$.

Example 6.1.4 (*Stirling's formula*) Let us prove as an application of CLT for Poisson r.v.'s (Example 6.1.3) that

1) $\qquad n! \sim \sqrt{2\pi n}(n/e)^n \quad \text{as } n \to \infty.$

Let N be a r.v. with $P(N = n) = \frac{r^n e^{-r}}{n!}$ ((r)-Poisson r.v.), Then,

2) $\qquad E[(N-r)^-] = r\,\dfrac{r^{\lfloor r\rfloor}e^{-r}}{\lfloor r\rfloor!}.$

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,

$$\begin{aligned} E\left[\left(\frac{S_n - n}{\sqrt{n}}\right)^-\right] &= n^{-1/2}E\left[(S_n - n)^-\right] \\ &\stackrel{(2)}{=} n^{-1/2}\cdot n\cdot\frac{n^n e^{-n}}{n!} = \frac{n^{n+\frac{1}{2}}e^{-n}}{n!}. \end{aligned}$$

Since $(S_n - n)/\sqrt{n}$ $(n \geq 1)$ are uniformly integrable by Example 5.4.3, so are their negative parts. Thus, by CLT (Example 6.1.3) and Proposition 5.4.4,

$$\begin{aligned} \lim_{n \nearrow \infty} \frac{n^{n+\frac{1}{2}} e^{-n}}{n!} &= \lim_{n \nearrow \infty} E\left[\left(\frac{S_n - n}{\sqrt{n}}\right)^-\right] = \sqrt{\frac{1}{2\pi}} \int_{-\infty}^0 x^- e^{-x^2/2} dx \\ &= \sqrt{\frac{1}{2\pi}} \int_0^\infty x e^{-x^2/2} dx = \sqrt{\frac{1}{2\pi}}, \end{aligned}$$

which proves (1). □

Exercise 6.1.5 (*Stirlings' formula revisited*) For $f_{r,a}$ in Exercise 6.1.4, show that $\frac{1}{\sqrt{a}} f_{a,a}(1) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f^{\wedge}_{\sqrt{a},a}(\theta) e^{-\mathbf{i}\theta\sqrt{a}} d\theta$ for $a > 1$. Then, use this and the results from Exercise 6.1.4 to prove the Stirling's formula in the form: $\Gamma(a+1) \stackrel{a\to\infty}{\sim} \sqrt{2\pi a} \left(\frac{a}{e}\right)^a$.

Exercise 6.1.6 (*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 (Example 6.1.4). (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 6.1.4.

Exercise 6.1.7 Let $(S_n)_{n \geq 0}$ and χ be as in Theorem 6.1.2. Suppose that $f : \mathbb{R}^d \to \mathbb{R}^m$ be measurable, differentiable at v, and that

$$|f(v+x) - f(v) - f'(v)x| \leq C|x|^2 \quad \text{for all } x \in \mathbb{R}^d$$

where C is a constant. Use (6.2) to show that

$$\sqrt{n}\,(f(S_n/n) - f(v)) \longrightarrow f'(v)\chi \quad \text{in law as } n \nearrow \infty,$$

This result includes (6.2) as a special case that $f(x) = x$.
Hint: Set $\chi_n = (S_n - vn)/\sqrt{n}$ and $g(x) = f(v+x) - f(v) - f'(v)x$ to write

$$\sqrt{n}\,(f(S_n/n) - f(v)) = f'(v)\chi_n + \sqrt{n} g(\chi_n/\sqrt{n}).$$

Then, apply Exercise 5.2.6 to $F(x,y) = x + y$.

6.2 Proof of the central limit theorem

We will prove that

$$\lim_{n \nearrow \infty} E\left[\mathbf{e}\left(\theta \cdot \frac{S_n - nv}{\sqrt{n}}\right)\right] = \exp\left(-\tfrac{1}{2}\theta \cdot \Gamma\theta\right) \quad \text{for all } \theta \in \mathbb{R}^d \tag{6.3}$$

By Lemma 5.1.2, (6.3) finishes the proof of Theorem 6.1.2. We now present a lemma which plays an important role not only to prove (6.3) but also to prove Theorem 12.3.1.

Lemma 6.2.1 *Let $Y = X_1 - v$. Then,*

$$\begin{aligned} \left|E\left[\mathbf{e}(\theta \cdot Y)\right] - (1 - \tfrac{1}{2}\theta \cdot \Gamma\theta)\right| &\leq 4|\theta|^2 E\left(|Y|^2 \min\{|Y||\theta|,\ 1\}\right) && (6.4) \\ &= o(|\theta|^2) \quad \text{as } |\theta| \searrow 0. && (6.5) \end{aligned}$$

Proof: Let us prove that

$$\text{1)} \qquad \left| \mathbf{e}(t) - \sum_{m=0}^{n} \frac{(\mathbf{i}t)^m}{m!} \right| \le \frac{|t|^{n+1}}{(n+1)!} \wedge \frac{2|t|^n}{n!} \quad \text{for all } t \in \mathbb{R} \text{ and } n = 1, 2, \ldots,$$

We have

$$\begin{aligned} \mathbf{e}(t) - \sum_{m=0}^{n} \frac{(\mathbf{i}t)^m}{m!} &= \mathbf{i}^{-n} \int_0^t dt_1 \int_0^{t_1} dt_2 \ldots \int_0^{t_n} \mathbf{e}(t_{n+1}) dt_{n+1} \\ &= \mathbf{i}^{-n} \int_0^t dt_1 \int_0^{t_1} dt_2 \ldots \int_0^{t_{n-1}} \frac{\mathbf{e}(t_n) - 1}{\mathbf{i}} dt_n. \end{aligned}$$

We get the bound $|t|^{n+1}/(n+1)!$ from the integral on the first line, while another bound $2|t|^n/n!$ is obtained from the integral on the second line.

It is now easy to prove (6.4). We have that

$$E[Y \cdot \theta] = 0, \quad E[(Y \cdot \theta)^2] = \sum_{i,j=1}^{d} \theta_i \theta_j E[Y^i Y^j] = \theta \cdot \Gamma \theta.$$

Therefore, we see (6.4) as follows:

$$\begin{aligned} & \left| E\left[\mathbf{e}(Y \cdot \theta)\right] - 1 + \tfrac{1}{2} \theta \cdot \Gamma \theta \right| \\ = \; & \left| E\left[\mathbf{e}(Y \cdot \theta) - 1 - \mathbf{i} Y \cdot \theta - \tfrac{1}{2} (\mathbf{i} Y \cdot \theta)^2\right] \right| \\ \le \; & E\left[|Y \cdot \theta|^2 (|Y \cdot \theta| \wedge 1)\right] \quad \text{by (1) with } n = 2, \\ \le \; & |\theta|^2 E\left[|Y|^2 (|Y||\theta| \wedge 1)\right] \end{aligned}$$

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 (6.4), proves (6.5). □

Exercise 6.2.1 (*More than* L^2) Use the argument in the proof of Lemma 6.2.1 to prove the following:
(i) If $X_1 \in L^{2+q}(P)$ for some $q \in [0, 1]$, then,

$$\left| E\left[\mathbf{e}(Y \cdot \theta)\right] - 1 + \tfrac{1}{2} \theta \cdot \Gamma \theta \right| \le |\theta|^{2+q} P[|Y|^{2+q}] = O(|\theta|^{2+q}) \quad \text{as } |\theta| \searrow 0.$$

Hint: $\min\{|Y||\theta|, \ 1\} \le |Y|^q |\theta|^q$.
(ii) If Y is symmetric and $X_1 \in L^{3+q}(P)$ for some $q \in [0, 1]$, then,

$$\left| E\left[\mathbf{e}(Y \cdot \theta)\right] - 1 + \tfrac{1}{2} \theta \cdot \Gamma \theta \right| \le |\theta|^{3+q} P(|Y|^{3+q}) = O(|\theta|^{3+q}) \quad \text{as } |\theta| \searrow 0.$$

Exercise 6.2.2 ($\star$)(*Less than* L^2) Let X real r.v. with the density $\frac{\alpha}{2}|x|^{-(\alpha+1)} \mathbf{1}\{|x| \ge 1\}$, where $0 < \alpha \le 2$. Show that

$$\begin{aligned} \varphi(\theta) \stackrel{\text{def}}{=} E\left[\mathbf{e}(\theta X)\right] &= \cos \theta - |\theta|^\alpha \int_{|\theta|}^{\infty} \frac{\sin y}{y^\alpha} dy \\ &= \begin{cases} 1 - \theta^2 \ln(1/|\theta|) + O(\theta^2) & \text{if } \alpha = 2 \\ 1 - c(\alpha)|\theta|^\alpha + o(|\theta|^\alpha) & \text{if } 0 < \alpha < 2 \end{cases} \quad \text{as } \theta \to 0, \end{aligned}$$

where $c(\alpha) = \int_0^\infty \frac{\sin y}{y^\alpha} dy > 0$.

We next present a general Lemma:

Lemma 6.2.2 *Let $\alpha > 0$, $\varphi : \mathbb{R}^d \to \mathbb{C}$ and $h : \mathbb{R}^d \to \mathbb{C}$ be such that*

a) $M_0(R) \stackrel{\text{def}}{=} \sup_{|\theta| \le R} |h(\theta)| < \infty$ *for any* $R > 0$,

b) $h(r\theta) = r^\alpha h(\theta)$ *for all* $\theta \in \mathbb{R}^d$ *and* $r \in (0, 1]$,

c) $\varphi(\theta) = 1 - h(\theta) + o(|\theta|^\alpha)$ *as* $\theta \to 0$.

Then,

$$\lim_{n \nearrow \infty} \varphi \left(\frac{\theta}{n^{1/\alpha}} \right)^n = \exp(-h(\theta)) \ \ \textit{for all } \theta \in \mathbb{R}^d.$$

Moreover, the convergence is uniform in $|\theta| \le R$ for any $R > 0$.

Remarks: 1) The lemma can be understood as an extension of the easy example: $\varphi(\theta) = 1 - |\theta|$ (thus $\alpha = 1$), in which the lemma is nothing but $\lim_{n \nearrow \infty} \left(1 - \frac{|\theta|}{n}\right)^n = e^{-|\theta|}$.
2) Suppose that $\varphi = \mu^\wedge$ for some $\mu \in \mathcal{P}(\mathbb{R}^d)$. Then $\varphi(n^{-1/\alpha}\theta)^n$ is the ch.f.of

$$Y_n = (X_1 + ... + X_n)/n^{-1/\alpha},$$

where $X_1, X_2, ...$ are i.i.d. with the common distribution μ. Thus, Lemma 6.2.2, together with Lévy's convergence theorem (Theorem 11.2.1) shows that e^{-h} is a ch.f. of a random variable Y and $Y_n \to Y$ weakly.

Proof of Lemma 6.2.2: We first prepare an elementary estimate:

1) $$\left| \left(1 + \frac{z}{n}\right)^n - e^z \right| \le \frac{|z|^2 e^{|z|}}{2n} \quad \text{for any } z \in \mathbb{C} \text{ and } n \in \mathbb{N}^*.$$

We define

$$\ell_n(s) = (1 - s) \exp(z/n) + s \left(1 + \tfrac{z}{n}\right), \quad s \in [0, 1].$$

We then have

2) $\left(1 + \frac{z}{n}\right)^n - e^z = \ell_n(1)^n - \ell_n(0)^n$

3) $|\ell_n(s)| \le |\exp(z/n)| \vee |1 + \frac{z}{n}| \le \exp(|z|/n)$,

4) $|\ell_n'(s)| \le \left|\exp(z/n) - \left(1 + \frac{z}{n}\right)\right| \le \frac{1}{2}|z/n|^2 \exp(|z/n|)$.

Combining these, we get (1):

$$\begin{aligned} \left|\left(1 + \tfrac{z}{n}\right)^n - e^z\right| &\stackrel{(2)}{=} \left| \int_0^1 ds\, n \ell_n(s)^{n-1} \ell_n'(s) \right| \\ &\stackrel{(3)-(4)}{\le} n \cdot \exp((n-1)|z|/n) \cdot \tfrac{1}{2}|z/n|^2 \exp(|z/n|) \\ &= (2n)^{-1}|z|^2 \exp(|z|). \end{aligned}$$

Next, difine a function h_n by $\varphi\left(\frac{\theta}{n^{1/\alpha}}\right) = 1 - \frac{h_n(\theta)}{n}$. We take arbitrary $R > 0$ and fix it. We then have for $|\theta| \le R$ that:

$$\begin{aligned} h_n(\theta) &= n\left(1-\varphi\left(\frac{\theta}{n^{1/\alpha}}\right)\right) \overset{\text{(c)}}{=} n\, h\left(\frac{\theta}{n^{1/\alpha}}\right) + n\, o\left(\frac{|\theta|^\alpha}{n}\right) \\ &\overset{\text{(b)}}{=} h(\theta) + n\, o\left(\frac{R^\alpha}{n}\right). \end{aligned}$$

This implies that

5) $h_n(\theta) \overset{n\to\infty}{\longrightarrow} h(\theta)$ uniformly in $|\theta| \le R$,

and thus,

6) $M_1(R) \overset{\text{def.}}{=} \sup_{n\ge 0} \sup_{|\theta|\le R} |h_n(\theta)| < \infty$.

We now write

$$e^{-h(\theta)} - \varphi\left(\frac{\theta}{n^{1/\alpha}}\right)^n = \underbrace{e^{-h(\theta)} - e^{-h_n(\theta)}}_{\delta_n(\theta)} + \underbrace{e^{-h_n(\theta)} - \left(1 - \frac{h_n(\theta)}{n}\right)^n}_{\varepsilon_n(\theta)}$$

Using (5)–(6) and

$$|e^z - e^w| \le |z-w| \exp\left((\operatorname{Re} z)^+ + (\operatorname{Re} w)^+\right) \quad z, w \in \mathbb{C},$$

we see that

$$|\delta_n(\theta)| \le |h(\theta) - h_n(\theta)| e^{M_0(R)+M_1(R)} \overset{n\to\infty}{\longrightarrow} 0 \quad \text{uniformly in } |\theta| \le R.$$

On the other hand, we see from (1) and (6) that

$$|\varepsilon_n(\theta)| \le \frac{M_1(R)^2 e^{M_1(R)}}{2n} \overset{n\to\infty}{\longrightarrow} 0 \quad \text{uniformly in } |\theta| \le R.$$

This completes the proof of the lemma. □

Let us finish the proof of (6.3). Note first that

$$\begin{aligned} E\left[\mathbf{e}\left(\theta \cdot \frac{S_n - nv}{\sqrt{n}}\right)\right] &= E\left[\prod_{j=1}^n \mathbf{e}\left(\theta \cdot \frac{X_j - v}{\sqrt{n}}\right)\right] \\ &= E\left[\mathbf{e}\left(\frac{\theta}{\sqrt{n}} \cdot Y\right)\right]^n \quad \text{since } \{X_j\}_{j\ge 1} \text{ are i.i.d.} \end{aligned}$$

By Lemma 6.2.1, all the assumptions in Lemma 6.2.2 are satisfied by

$$\varphi(\theta) = E\left[\mathbf{e}\left(\theta \cdot Y\right)\right], \quad h(\theta) = \tfrac{1}{2}\theta \cdot \Gamma\theta$$

with $\alpha = 2$. Therefore (6.3) follows from Lemma 6.2.2. □

Exercise 6.2.3 (*CLT for continuous-time RW*) Let $S_n = X_1 + ... + X_n$ be as in Theorem 6.1.2 and $(N_t)_{t \ge 0}$ be Poisson process with parameter $r > 0$ (Example 3.1.2). We suppose that $(X_n)_{n \ge 1}$ and $(N_t)_{t \ge 0}$ are independent and define $\widetilde{S}_t = S_{N_t}$. Then, show the following: (i) $E[\mathbf{e}(\theta \cdot \widetilde{S}_t)] = \exp\left((E[\mathbf{e}(\theta \cdot X_1)] - 1)\, rt\right)$. (ii) $P(\frac{\widetilde{S}_t - vrt}{\sqrt{t}} \in \cdot) \stackrel{t\to\infty}{\longrightarrow} \nu_{r\widetilde{\Gamma}}$ weakly, where $\nu_{\widetilde{\Gamma}}$ is the matrix $\widetilde{\Gamma}$ is given by $\widetilde{\Gamma}_{ij} = E[X_1^i X_1^j]$ $(i, j = 1, ..., d)$.

Exercise 6.2.4 $(\star)$(*Logarithmic correction to Lemma 6.2.2*) Replace the condition (c) in Lemma 6.2.2 by:

$$\varphi(\theta) = 1 - \alpha\, h(\theta) \ln(1/|\theta|) + O(|\theta|^\alpha), \quad \theta \to 0,$$

while keeping all the other assumptions. Prove then that

$$\lim_{n \nearrow \infty} \varphi\left(\frac{\theta}{(n \ln n)^{1/\alpha}}\right)^n = \exp(-h(\theta)) \text{ for all } \theta \in \mathbb{R}^d.$$

Exercise 6.2.5 $(\star)$(α-*stable law*) Let $S_n = X_1 + ... + X_n$, where $(X_n)_{n \ge 1}$ are real i.i.d. with the density $\frac{\alpha}{2}|x|^{-(\alpha+1)}\mathbf{1}\{|x| \ge 1\}$, where $0 < \alpha \le 2$. (i) For $\alpha = 2$, use Exercise 6.2.2 and Exercise 6.2.4 to prove that $\frac{S_n}{\sqrt{n \ln n}} \stackrel{n\to\infty}{\longrightarrow} \chi$ weakly, where χ is a r.v. with $P(\chi \in \cdot) = \nu_1$. (ii) $(\star)$ For $0 < \alpha < 2$ and $c > 0$, use Exercise 6.2.2 and Lemma 6.2.2 to show that

$$\lim_{n \nearrow \infty} \varphi(n^{-1/\alpha}\theta)^n = \exp(-c(\alpha)|\theta|^\alpha) \text{ uniformly in } |\theta| < R \text{ for any } R > 0,$$

or equivalently, for any $c > 0$,

$$\lim_{n \nearrow \infty} \varphi(n^{-1/\alpha} r\theta)^n = \exp(-c|\theta|^\alpha) \text{ uniformly in } |\theta| < R \text{ for any } R > 0,$$

where $r = (c/c(\alpha))^{1/\alpha}$. This shows that there exists $\mu_{c,\alpha} \in \mathcal{P}(\mathbb{R})$ such that $\mu_{c,\alpha}^\wedge(\theta) = \exp(-c|\theta|^\alpha)$ $(\theta \in \mathbb{R})$ and that $\frac{rS_n}{n^{1/\alpha}} \stackrel{n\to\infty}{\longrightarrow} \chi$ weakly, where χ is a r.v. with $P(\chi \in \cdot) = \mu_{c,\alpha}$ (cf. the remark after Lemma 6.2.2). $\mu_{c,\alpha}$ is called the *symmetric* α*-stable law* (For $\alpha = 2$, it is ν_{2c}, and for $\alpha = 1$, it is the (c)-Cauchy distribution).

7 Does a random walk come back?

7.1 Transience and recurrence

In this section, we will take up a question whether a random walk $(S_n)_{n\ge 0}$ comes back to its starting point with probability one. In particular, we will prove the following

Theorem 7.1.1 *For the simple random walk $(S_n)_{n\in\mathbb{N}}$ in $\mathbb{Z}^d$ (cf. (4.4)),*

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

This theorem is often explained with a joke:

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

We also make observation for more general RW in the course of the proof of Theorem 7.1.1.

- Throughout this section, we will restrict ourselves to $\mathbb{Z}^d$-valued random walks.

This is to avoid being bothered by inessential complication. It is convenient to introduce the following notations. For $x \in \mathbb{Z}^d$, we set

$$\begin{aligned} V(x) &= \sum_{n\ge 1} \mathbf{1}\{S_n = x\} && (7.1) \\ &= \text{"the number of visits to } x\text{"}. \\ h^{(m)}(x) &= P(V(x) \ge m), \quad m = 1, \ldots, \infty && (7.2) \\ &= \text{"probability that } x \text{ is visited at least } m \text{ times"}. \\ h(x) &= h^{(1)}(x) && (7.3) \\ &= \text{"probability that } x \text{ is visited at least once"}. \\ g_s(x) &= \sum_{n\ge 0} s^n P(S_n = x), \quad 0 \le s \le 1, && (7.4) \end{aligned}$$

The function $g_s(x)$ above is called the *Green function* of the random walk.

Exercise 7.1.1 ($\star$)(*Green function for continuous-time RW*) Let $S_n = X_1 + ... + X_n$ be a $\mathbb{Z}^d$-valued random walk and $(N_t)_{t\ge 0}$ be Poisson process with parameter $r > 0$ (Example 3.1.2). We suppose that $(X_n)_{n\ge 1}$ and $(N_t)_{t\ge 0}$ are independent. Then, show that $\int_0^\infty P(S_{N_t} = x)dt = \frac{1}{r} g_1(x), \quad x \in \mathbb{Z}^d$, where g_1 is the Green function for $(S_n)_{n\in\mathbb{N}}$.

Proposition 7.1.2 (**Transience/Recurrence**) *Let $(S_n)_{n\ge 0}$ be a $\mathbb{Z}^d$-valued random walk. Then, the following conditions (T1)–(T5) are equivalent:*

T1) $h(0) < 1$.

T2) $h^{(\infty)}(0) = 0$.

T3) $g_1(0) < \infty$.

T4) $h^{(\infty)}(x) = 0$ *for all* $x \in \mathbb{Z}^d$.

T5) $g_1(x) < \infty$ *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) $h^{(\infty)}(0) = 1$.

R3) $g_1(0) = \infty$.

R4) $h^{(\infty)}(x) = 1$ *if* $h(x) > 0$.

R5) $g_1(x) = \infty$ *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.*

Remark: Note that

$$h(0) = P\left(0 \in S_n \text{ for some } n \geq 1\right).$$

Thus, Theorem 7.1.1 says that the simple random walk is recurrent for $d \leq 2$ and is transient for $d \geq 3$.

Example 7.1.3 Suppose that you and your friend 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 $d \geq 3$. This can be seen as follows. Let $(S'_n)_{n\geq 1}$ and $(S''_n)_{n\geq 1}$ be independent random walks. Then, $S_n = S'_n - S''_n$ $n \geq 0$ is again a random walk and

1) $\qquad P(S_n = 0) = P(S'_n - S''_n = 0) = P(S'_{2n} = 0)$

Let g_1 and g'_1 be the Green functions of $S_\cdot$ and $S'_\cdot$ respectively. Then,

$$g_1(0) = \sum_{n\geq 0} P(S_n = 0) \stackrel{(1)}{=} \sum_{n\geq 0} P(S'_{2n} = 0) = g'_1(0),$$

where the reason for the last identity is that $P(S'_{2n+1} = 0) = 0$. Thus, we see the claim from Theorem 7.1.1 and Proposition 7.1.2. □

Exercise 7.1.2 Prove that

$$P(\lim_{n\nearrow\infty} |S_n| = +\infty) = \begin{cases} 0 & \text{for a recurrent RW,} \\ 1 & \text{for a transient RW.} \end{cases}$$

Exercise 7.1.3 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 12.1.3 below.

Exercise 7.1.4 Prove that for all $0 < s \leq 1$ and $z \in \mathbb{Z}^d$,

$$\begin{aligned} g_s(z) &= \delta_{0,z} + sEg_s(z - X_1), && (7.5) \\ h(z) &= (1 - h(0))P\{X_1 = z\} + Eh(z - X_1), && (7.6) \\ h^{(\infty)}(z) &= Eh^{(\infty)}(z - X_1). && (7.7) \end{aligned}$$

7.2 Proof of Proposition 7.1.2

We begin by proving the following

Lemma 7.2.1 *For $x \in \mathbb{Z}^d$,*

$$h^{(m)}(x) = h(x)h(0)^{m-1}, \tag{7.8}$$

As consequences,

$$h^{(\infty)}(x) = \begin{cases} 0 & \text{if } h(0) < 1, \\ h(x) & \text{if } h(0) = 1. \end{cases} \tag{7.9}$$

and

$$g_1(x) = \delta_{0,x} + \frac{h(x)}{1 - h(0)}. \tag{7.10}$$

Remark Intuition behind (7.8) can be explained as follows; A trajectry 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 returns 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

$$\eta_x^{(m)} = \inf \left\{ n \geq 1 \mid \sum_{k=1}^{n} 1\{S_k = x\} = m \right\}.$$

Then,

$$h^{(m)}(x) = P(\eta_x^{(m)} < \infty) = \sum_{\ell \geq 1} P \left(\underbrace{\eta_x^{(m-1)} = \ell}_{=:E_\ell}, \underbrace{\cup_{n \geq 1}\{S_{n+\ell} - S_\ell = 0\}}_{=:F_\ell} \right)$$

We observe that

1) E_ℓ and F_ℓ are independent,

since

$$E_\ell \in \sigma[X_j \ ; \ j \leq \ell], \quad F_\ell \in \sigma[X_j \ ; \ j > \ell].$$

We have also that

2) $\sum_{\ell \geq 1} P(E_\ell) \stackrel{(7.2)}{=} h^{(m-1)}(x), \quad P(F_\ell) = P(F_0) \stackrel{(7.3)}{=} h(0).$

(Note that $S_{n+\ell} - S_\ell = X_{\ell+n} + ... + X_{\ell+1}$ to see $P(F_\ell) = P(F_0)$ above.) Therefore,

$$h^{(m)}(x) \stackrel{(1)}{=} \sum_{\ell \geq 1} P(E_\ell)P(F_\ell) \stackrel{(2)}{=} \sum_{\ell \geq 1} P(E_\ell)P(F_0) \stackrel{(2)}{=} h^{(m-1)}(x)h(0).$$

We then get (7.8) by induction.

To see (7.9), we use the monotone convergence theorem (MCT) and (7.8):

$$h^{(\infty)}(x) \stackrel{\text{MCT}}{=} \lim_{m \nearrow \infty} h^{(m)}(x) \stackrel{(7.8)}{=} \begin{cases} 0 & \text{if } h(0) < 1, \\ h(x) & \text{if } h(0) = 1. \end{cases}$$

Equality (7.10) can be seen as follows;

$$\begin{aligned} g_1(x) &\stackrel{(7.4)}{=} \delta_{0,x} + \sum_{n\ge 1} P(S_n = x) \stackrel{\text{Fubini}}{=} \delta_{0,x} + E[V(x)] \\ &\stackrel{(1.12)}{=} \delta_{0,x} + \sum_{m\ge 1} \underbrace{P(V(x)\ge m)}_{=h^{(m)}(x)} \stackrel{(7.8)}{=} \delta_{0,x} + \frac{h(x)}{1-h(0)}. \end{aligned}$$

□

Exercise 7.2.1 Conclude from (7.8) that $V(0)$ for a transient RW is a r.v. with geometric distribution with the parameter $1-h(0)$ (cf. Exercise 3.1.9).

Exercise 7.2.2 (i) Show that $h(x+y)\ge h(x)h(y)$ 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 to prove (7.8) above. (ii) Use (i) and (7.10) to show that $g_1(x+y)g_1(0)\ge g_1(x)g_1(y)$ for all $x,y\in\mathbb{Z}^d$.

Proof of Proposition 7.1.2: **(T1)** $\Rightarrow$ **(T2)-(T5)**: This follows from (7.9)–(7.10).
(**T2**) $\Rightarrow$ (**T1**): This follows from (7.9) for $x=0$.
(**T3**) $\Rightarrow$ (**T1**): This follows from (7.10) for $x=0$.
Therefore, **(T1)**, **(T2)** and **(T3)** are equivalent. Moreover,
(**T4**) $\Rightarrow$ (**T2**): Obvious.
(**T5**) $\Rightarrow$ (**T3**): Obvious.
(R1) $\Rightarrow$ **(R1)**–**(R3)**: This follows from (7.9)–(7.10).
(**R2**) $\Rightarrow$ (**R1**): Obvious.
(**R3**) $\Rightarrow$ (**R1**): This follows from (7.10) for $x=0$.
Therefore, **(R1)**–**(R3)** are equivalent. Moreover,
(**R1**) $\Rightarrow$ (**R4**): We will show that

1) $\qquad h(x)>0,\ h(0)=1 \implies h^{(\infty)}(x)=1.$

To do so, we first prove that

2) $\qquad h(x)>0,\ h^{(\infty)}(y)=1 \implies h^{(\infty)}(y-x)=1.$

This can be seen as follows. Since $h(x)>0$, there is a $k\ge 1$ such that $P(S_k=x)>0$. We have for any m that

$$\begin{aligned} P\left(\cap_{n\ge m+k}\{S_n\ne y\}\right) &\ge P\left(\{S_k=x\}\cap\left(\cap_{n\ge m+k}\{S_n-S_k\ne y-x\}\right)\right) \\ &= P(S_k=x)P\left(\cap_{n\ge m}\{S_n\ne y-x\}\right). \end{aligned}$$

Note that $h^{(\infty)}(y)=1$ is equivalent to that $P(\cap_{n\ge m}\{S_n\ne y\})=0$ for any m. Therefore (2) can be seen from the above relation. Now it is easy to conclude (1) from (2). Suppose $h(x)>0$ and $h(0)=1$. We then see $h^{(\infty)}(0)=1$ (**(R1)** $\iff$ **(R3)**). Now $h(x)>0$ and $h^{(\infty)}(0)=1$ imply $h^{(\infty)}(-x)=1$ by (2). Finally, $h^{(\infty)}(0)=1$ and $h^{(\infty)}(-x)=1$ imply $h^{(\infty)}(x)=1$ again by (2).
(R4) $\Rightarrow$**(R5)**: Suppose that $h(x)>0$. Then, $1=h^{(\infty)}(x)=P(V(x)=\infty)$ by (R4). This implies that $g_1(x)=\delta_{0,x}+E[V(x)]=\infty$.
(R5) $\Rightarrow$**(R1)**: This follows from (7.10), since $h(x)>0$ for some $x\in\mathbb{Z}^d$. □

Exercise 7.2.3 Conclude from (2) in the proof of Proposition 7.1.2 that the set $\{x\in\mathbb{Z}^d\ ;\ h^{(\infty)}(x)=1\}$ is either empty or a subgroup of $\mathbb{Z}^d$.

7.3 Proof of Theorem 7.1.1

We first consider a general $\mathbb{Z}^d$-valued random walk with $\nu = P(X_1 \in \cdot)$. We will use the following notation;

$$\begin{aligned} \delta I &= [-\delta, \delta]^d \subset \mathbb{R}^d \quad \text{for } \delta > 0. & (7.11) \\ \widehat{\nu}(\theta) &= E\mathbf{e}(\theta \cdot X_1) = \sum_{x \in \mathbb{Z}^d} \mathbf{e}(\theta \cdot x)\nu(x), \quad \theta \in \pi I. & (7.12) \end{aligned}$$

where $\mathbf{e}(t) = \exp(\mathbf{i}t)$, $t \in \mathbb{R}$.

Lemma 7.3.1

$$P(S_n = x) = \frac{1}{(2\pi)^d} \int_{\pi I} d\theta \mathbf{e}(-\theta \cdot x)\widehat{\nu}(\theta)^n, \quad \textit{for } x \in \mathbb{Z}^d \textit{ and } n \in \mathbb{N}. \tag{7.13}$$

Proof: Note that

1) $\quad E\mathbf{e}(\theta \cdot S_n) = \widehat{\nu}(\theta)^n$ by Lemma 5.2.2.

and that

2) $\quad \delta_{x,y} = \frac{1}{(2\pi)^d} \int_{\pi I} d\theta \mathbf{e}(-\theta \cdot x)\mathbf{e}(\theta \cdot y), \quad x, y \in \mathbb{Z}^d.$

By taking the expectation of (2) with $y = S_n$, we have

$$P(S_n = x) = \frac{1}{(2\pi)^d} E\left[\int_{\pi I} d\theta \mathbf{e}(-\theta \cdot x)\mathbf{e}(\theta \cdot S_n)\right] \stackrel{\text{Fubini}}{=} \frac{1}{(2\pi)^d} \int_{\pi I} d\theta \mathbf{e}(-\theta \cdot x) \underbrace{E\left[\mathbf{e}(\theta \cdot S_n)\right]}_{=\widehat{\nu}(\theta)^n \text{ by } (1)},$$

which proves (7.13). □

Lemma 7.3.2 *For the simple random walk on $\mathbb{Z}^d$, there exists constants $b_i \in (0, \infty)$ ($i = 1, 2$) such that for all $n \geq 1$,*

$$\begin{aligned} \sup_{x \in \mathbb{Z}^d} P(S_n = x) &\leq \frac{b_1}{n^{d/2}} & (7.14) \\ P(S_{2n} = 0) &\geq \frac{b_2}{n^{d/2}} & (7.15) \end{aligned}$$

Proof of Theorem 7.1.1 assuming Lemma 7.3.2: We see from Lemma 7.3.2 that

$$g_1(0) = \sum_{n \geq 0} P(S_{2n} = 0) \begin{cases} \stackrel{(7.15)}{=} \infty & \text{if } d = 1, 2 \\ \stackrel{(7.14)}{<} \infty & \text{if } d \geq 3 \end{cases}$$

Thus, the theorem follows from Proposition 7.1.2. □

Proof of Lemma 7.3.2: We have

$$\widehat{\nu}(\theta) = \frac{1}{d} \sum_{j=1}^{d} \cos \theta_j \in \mathbb{R}.$$

for all $\theta = (\theta_j)_{j=1}^d \in \mathbb{R}^d$. Recall that

$$\begin{aligned} 1 - \cos t &= 2\sin^2(t/2) \le t^2/2, \quad t \in \mathbb{R}, && (7.16) \\ |\sin t| &\ge \frac{2}{\pi}|t|, \quad |t| \le \frac{\pi}{2}. && (7.17) \end{aligned}$$

We then see that

$$1 - \widehat{\nu}(\theta) = \frac{1}{d}\sum_{j=1}^d (1 - \cos\theta_j) \stackrel{(7.16)}{=} \frac{2}{d}\sum_{j=1}^d \sin^2(\theta_j/2) \begin{cases} \stackrel{(7.16)}{\le} c_1|\theta|^2 \\ \stackrel{(7.17)}{\ge} c_2|\theta|^2 \end{cases},$$

where $c_1 = \frac{1}{d}$ and $c_2 = \frac{2}{d\pi^2}$. In particular,

1) $\quad 0 \le 1 - c_1|\theta|^2 \le \widehat{\nu}(\theta) \le 1 - c_2|\theta|^2 \quad \text{if } |\theta| \le \varepsilon,$

where we choose $\varepsilon > 0$ small enough. We start with (7.15).

$$\begin{aligned} P(S_{2n} = 0) &\stackrel{(7.13)}{=} \frac{1}{(2\pi)^d}\int_{\pi I} \widehat{\nu}(\theta)^{2n} d\theta \ge \frac{1}{(2\pi)^d}\int_{|\theta|\le\varepsilon} \widehat{\nu}(\theta)^{2n} d\theta \\ &\stackrel{(1)}{\ge} \frac{1}{(2\pi)^d}\int_{|\theta|\le\varepsilon} (1 - c_1|\theta|^2)^{2n} d\theta = c_3\int_0^\varepsilon (1 - c_1 r^2)^{2n} r^{d-1} dr \\ &\stackrel{s=r\sqrt{n}}{=} \frac{c_3}{n^{d/2}}\int_0^{\varepsilon\sqrt{n}} \left(1 - \frac{c_1 s^2}{n}\right)^{2n} s^{d-1} ds. \end{aligned}$$

from which (7.15) follows, since for any $c > 0$,

2) $$\lim_{n\nearrow\infty}\int_0^{c\sqrt{n}} \left(1 - \frac{cs^2}{n}\right)^{2n} s^{d-1} ds \stackrel{\text{MCT}}{=} \int_0^\infty e^{-2cs^2} s^{d-1} ds \in (0, \infty)$$

To show (7.14), we introduce some notation. Let $v_0 = 0$ and $\{v_k\}_{k=1}^{2^d}$ are vertices of the cube πI. We also set

$$W = \{\theta \in \pi I \; ; \; |\theta - v_j| \ge \varepsilon \text{ for all } j = 0, .., 2^d\}.$$

We write

$$P(S_n = x) \stackrel{(7.13)}{\le} \frac{1}{(2\pi)^d}\int_{\pi I} |\widehat{\nu}(\theta)|^n d\theta = \sum_{k=0}^{2^d} J_k(n) + J_W(n),$$

where

$$J_k(n) = \frac{1}{(2\pi)^d}\int_{\substack{\theta\in\pi I, \\ |\theta - v_k|\le\varepsilon}} |\widehat{\nu}(\theta)|^n d\theta, \quad J_W(n) = \frac{1}{(2\pi)^d}\int_W |\widehat{\nu}(\theta)|^n d\theta.$$

We see that

$$\{\theta \in \pi I \; ; \; |\widehat{\nu}(\theta)| = 1\} = \{v_k\}_{k=0}^{2^d},$$

and thus, there exists $\delta \in (0, 1)$ such that

$$\max_W |\widehat{\nu}| \le \delta.$$

This implies that

3) $|J_W(n)| \leq \delta^n.$

For $J_0(n)$, we have

$$J_0(n) = \frac{1}{(2\pi)^d} \int_{|\theta| \leq \varepsilon} |\widehat{\nu}(\theta)|^n d\theta \overset{(1)}{\leq} \frac{1}{(2\pi)^d} \int_{|\theta| \leq \varepsilon} (1 - c_2|\theta|^2)^n d\theta$$

Thus, we see as in the proof of (7.15) that there exists $c_4 \in (0, \infty)$ such that

4) $J_0(n) \leq \frac{c_4}{n^{d/2}}.$

Since $\widehat{\nu}(\theta + v_k) = -\widehat{\nu}(\theta)$, we have

5) $\sum_{k=1}^{2^d} J_k(n) = J_0(n).$

Finally, we have the desired estimate:

$$P(S_n = x) \leq \sum_{k=0}^{2^d} |J_k(n)| + |J_W(n)| \overset{(3)-(5)}{\leq} \frac{2c_4}{n^{d/2}} + \delta^n \leq \frac{c_5}{n^{d/2}},$$

for some c_5. □

8 Brownian Motion

8.1 What is a Brownian motion?

We fix a probability space $(\Omega, \mathcal{F}, P)$ in this subsection.

Definition 8.1.1 (**Brownian motion**) A family $\{B_t : \Omega \to \mathbb{R}^d\}_{t \geq 0}$ of r.v.'s is called a *d-dimensional Brownian Motion* (BM^d for short) if the following properties are satisfied;

B0) $B_0 = 0$, a.s.

B1) There is an $\Omega_0 \in \mathcal{F}$ such that $P(\Omega_0) = 1$ and $t \mapsto B_t(\omega)$ is continuous for all $\omega \in \Omega_0$.

B2) $\{B_{t_j} - B_{t_{j-1}}\}_{j=1}^n$ are independent if $n \geq 2$ and $0 = t_0 < t_1 < \ldots < t_n$.

B3) For any $0 \leq s < t$ and $A \in \mathcal{B}(\mathbb{R}^d)$,

$$P(B_t - B_s \in A) = \int_A p_{t-s}(x)dx, \quad \text{where } p_t(x) = (2\pi t)^{-d/2} \exp\left(-\frac{|x|^2}{2t}\right). \tag{8.1}$$

Brownian motion came into the history in 1827, when R. Brown, a British botanist, observed that pollen grains suspended in water preform a contunual swarming motion. In 1905, A. Einstein derived (8.1) from the moleculer physics point of view. A mathematically rigorous construction with a proof of the continuity ((B1) above) was given by N. Wiener (1923).

Remark: It is useful to note that following consequence of (8.1): for $0 \leq s < t$,

$$P(B_t - B_s \in \cdot) = P(B_{t-s} \in \cdot) = P(\sqrt{t-s}B_1 \in \cdot). \tag{8.2}$$

Exercise 8.1.1 Verify (8.2).

We will prove in subsection 8.2 that a BM^d exists on a suitable probability space $(\Omega, \mathcal{F}, P)$. Once we are given a BM^d, then, we can construct many other BM^d's (Exercise 8.1.4 below). However, "the law of BM^d is unique" in the following sense.

Proposition 8.1.2 (Uniqueness of the law of BM^d) *Let $S = (\mathbb{R}^d)^{[0,\infty)}$ and let $\mathcal{B}$ be its cylindrical σ-field (cf. Definition 2.2.1).*

a) *Suppose that $\{B_t\}_{t \geq 0}$ is a BM^d. Then the map $\omega \mapsto (B_t(\omega))_{t \geq 0}$ from $(\Omega, \mathcal{F})$ to $(S, \mathcal{B})$ is measurable.*

b) *Suppose that $\{B_t\}_{t \geq 0}$ and $\{\widetilde{B}_t\}_{t \geq 0}$ are BM^d's. Then, their laws on $(S, \mathcal{B})$ 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;*

$$P((B_t)_{t \geq 0} \in A) = P((\widetilde{B}_t)_{t \geq 0} \in A) \quad \textit{for all } A \in \mathcal{B}. \tag{8.3}$$

Proof: a): This follows from a general consideration of cylindrical σ-fields (cf. Exercise 2.2.2).
b): It is not difficult to see that

$$\begin{aligned} &P(B_{t_1} \in A_1, \ldots, B_{t_n} \in A_n) \\ &= \int_{A_1} p_{t_1}(x_1)dx_1 \int_{A_2} p_{t_2 - t_1}(x_2 - x_1)dx_2 \ldots \int_{A_n} p_{t_n - t_{n-1}}(x_n - x_{n-1})dx_n \end{aligned} \tag{8.4}$$

for $n \geq 1$, $0 = t_0 < t_1 < \ldots < t_n$ and $A_1, \ldots A_n \in \mathcal{B}(\mathbb{R}^d)$ (Exercise 8.1.2). This proves (8.3) for all cylinder set A, and hence for all $A \in \mathcal{B}$ (cf. Lemma 2.2.2). $\square$

Exercise 8.1.2 Prove (8.4). Hint: Note that $\{B_{t_j} - B_{t_{j-1}}\}_{j=1}^n$ are independent and that $\{B_{t_1} \in A_1, \dots, B_{t_n} \in A_n\} = \{(B_{t_j} - B_{t_{j-1}})_{j=1}^n \in D\}$, where $D = \cap_{j=1}^n \{y \in (\mathbb{R}^d)^n \ ; \ y_1 + ... + y_j \in A_j\}$. Therefore,

$$\begin{aligned} & P(B_{t_1} \in A_1, \dots, B_{t_n} \in A_n) \\ & = \int_D p_{t_1}(y_1) p_{t_2 - t_1}(y_2) \cdots p_{t_n - t_{n-1}}(y_n) dy_1 \cdots dy_n. \end{aligned}$$

Exercise 8.1.3 Prove that $\{B_t\}_{t \geq 0}$ is a BM^d if and only if their coordinates $\{B_t^i\}_{t \geq 0}$ ($i = 1, \dots, d$) are independent BM^1's. Hint: It is not difficult to prove "if" part. Then, Proposition 8.1.2 together with the result of "if" part can be used to prove "only if" part.

Exercise 8.1.4 Suppose that $\{B_t\}_{t \geq 0}$ is a BM^d. Then, prove that $\{c^{-1/2} B_{ct}\}_{t \geq 0}$ is a BM^d for all $c > 0$ and that $\{UB_t\}_{t \geq 0}$ is a BM^d for any orthogonal $d \times d$ matrix U.

One of the most striking property of the Brownian motion is:

Proposition 8.1.3 (Nowhere non-differentiability[11]) *With probability one, $t \mapsto B_t$ is not differentiable at any $t \geq 0$.*

Remark: It is easy to prove that for *fixed* $t \geq 0$,

$$\lim_{h \searrow 0} \frac{|B_{t+h} - B_t|}{h} = \infty \quad \text{in probability}$$

in the sense that $\lim_{h \searrow 0} P\left(\left| \frac{B_{t+h} - B_t}{h} \right| \geq M \right) = 1$ for any $M > 0$. In fact,

$$P\left(\left| \frac{B_{t+h} - B_t}{h} \right| \geq M \right) \stackrel{(8.2)}{=} P\left(\left| \frac{B_1}{\sqrt{h}} \right| \geq M \right) = P\left(|B_1| \geq M\sqrt{h} \right) \stackrel{h \searrow 0}{\longrightarrow} 1.$$

Proof of Proposition 8.1.3[12]: It is enough to prove the proposition for $d = 1$. We will prove that with probability one,

$$\overline{\lim_{h \searrow 0}} \frac{|B_{t+h} - B_t|}{h} = \infty \ \text{ for any } t \geq 0,$$

It ie enough to show this with "at any $t \geq 0$" replaced by "at any $t \in [0, 1]$" (countable additivity of the measure). Thus, it is enough to prove that the following set $D \subset \Omega$ is a null set:

$$D = \left\{ \overline{\lim_{h \searrow 0}} \frac{|B_{t+h} - B_t|}{h} < \infty \ \text{ for some } t \in [0, 1] \right\}.$$

Let us write for simplicity $B(a, b] = B(b) - B(a)$ for $0 \leq a \leq b$. We will show that

1) $\qquad D \subset E \stackrel{\text{def.}}{=} \bigcup_{\ell \in \mathbb{N}^*} \bigcup_{N \in \mathbb{N}^*} \bigcap_{n \geq N} \bigcup_{i=1}^n \bigcap_{j=1}^3 E_{\ell, n, i, j},$

[11]Due to R.E.A.C.Paley, N. Wiener and A. Zygmund (1933)

[12]Due to A. Dvoretsky, P. Erd/"os, S. Kakutani (1961)

where

$$E_{\ell,n,i,j} = \left\{ \left| B\left(\frac{i+j}{n}, \ \frac{i+j+1}{n} \right] \right| \leq \frac{8\ell}{n} \right\}.$$

Let us first prove that $P(E) = 0$, by which the reason for $\bigcap_{j=1}^{3}$ in the definition of E will also be explained. Since the intervals $\left(\frac{i+j}{n}, \ \frac{i+j+1}{n}\right]$, $j = 1, 2, 3$ are disjoint,

$$B\left(\frac{i+j}{n}, \ \frac{i+j+1}{n} \right], \quad j = 1, 2, 3$$

are i.i.d. with the same distribution as $|B(1)|/\sqrt{n}$ (Recall (B2) in Definition 8.1.1 and (8.2)). We therefore have that

$$\begin{aligned} P\left(\bigcup_{i=1}^{n} \bigcap_{j=1}^{3} E_{\ell,n,i,j} \right) &\leq \sum_{i=1}^{n} P\left(\bigcap_{j=1}^{3} E_{\ell,n,i,j} \right) = nP\left(|B(1)| \leq \frac{8\ell}{\sqrt{n}} \right)^3 \\ &= n \left(\frac{1}{\sqrt{2\pi}} \int_{|x| \leq \frac{8\ell}{\sqrt{n}}} e^{-x^2/2} \, dx \right)^3 \leq \frac{C\ell^3}{n^{1/2}}. \end{aligned}$$

By this and Fatou's lemma, we get

$$P\left(\bigcup_{N \in \mathbb{N}^*} \bigcap_{n \geq N} \bigcup_{i=1}^{n} \bigcap_{j=1}^{3} E_{\ell,n,i,j} \right) \leq \varliminf_{n \to \infty} P\left(\bigcup_{i=1}^{n} \bigcap_{j=1}^{3} E_{\ell,n,i,j} \right) = 0,$$

which proves $P(E) = 0$.
We next turn to (1). On the set D, there exist $t \in [0,1]$, $\ell, N \in \mathbb{N}^*$ such that

2) $$\sup_{0 < h \leq 4/n} \frac{|B_{t+h} - B_t|}{h} \leq \ell \quad \text{for all } n \geq N$$

We set $i = \lfloor nt \rfloor \leq n$ so that

3) $$\frac{i}{n} \leq t < \frac{i+j}{n} < \frac{i+j+1}{n} \leq t + \frac{4}{n}, \quad j = 1, 2, 3.$$

Then, for $j = 1, 2, 3$,

$$\left| B\left(\frac{i+j}{n}, \ \frac{i+j+1}{n} \right] \right| \leq \left| B\left(t, \ \frac{i+j}{n} \right] \right| + \left| B\left(t, \ \frac{i+j+1}{n} \right] \right| \stackrel{(2)-(3)}{\leq} 2 \cdot \frac{4\ell}{n} = \frac{8\ell}{n}.$$

This proves (1).

□

Exercise 8.1.5 ($\star$) Prove the following refinement of Proposition 8.1.3: if $\alpha > 1/2$, then, with probability one,

$$\varlimsup_{h \searrow 0} \frac{|B_{t+h} - B_t|}{h^\alpha} = \infty \ \text{ for any } t \geq 0.$$

Hint: It is enough to prove that the following set $D \subset \Omega$ is a null set:

$$D = \left\{ \varlimsup_{h \searrow 0} \frac{|B_{t+h} - B_t|}{h^\alpha} < \infty \ \text{ for some } t \in [0,1] \right\}.$$

With $J \in \mathbb{N}^*$ such that $(\alpha - \frac{1}{2})J > 1$, show that

$$D \subset E \stackrel{\text{def.}}{=} \bigcup_{\ell \in \mathbb{N}^*} \bigcup_{N \in \mathbb{N}^*} \bigcap_{n \geq N} \bigcup_{i=1}^{n} \bigcap_{j=1}^{J+1} \left\{ \left| B\left(\frac{i+j}{n}, \ \frac{i+j+1}{n} \right] \right| \leq \frac{2(J+1)\ell}{n^\alpha} \right\}.$$

and that $P(E) = 0$.

8.2 Does a Brownian motion exist?

We present a construction of a BM^1 in this subsection. This is enough to prove the existence of BM^d for any $d \geq 1$ (cf. Exercise 8.1.3). We begin by introducing Haar functions $h_{n,k}$: $[0, \infty) \to \mathbb{R}$ $(n, k = 0, 1, \ldots)$ as follows;

$$h_{0,k} = 1_{[k,k+1)} \quad h_{n,k}(t) = \begin{cases} 2^{\frac{n-1}{2}} & \text{if } n \geq 1 \text{ and } t \in [2k/2^n, (2k+1)/2^n), \\ -2^{\frac{n-1}{2}} & \text{if } n \geq 1 \text{ and } t \in [(2k+1)/2^n, (2k+2)/2^n), \\ 0 & \text{if } n \geq 1 \text{ and } t \notin [2k/2^n, (2k+2)/2^n). \end{cases} \tag{8.5}$$

Exercise 8.2.1 Draw pictures of Haar functions.

Lemma 8.2.1 *Define*

$$\langle\, f, g \,\rangle = \int_0^\infty f(t)g(t)dt \quad f, g \in L^2[0, \infty).$$

Then $\{h_{n,k}\}_{n,k \geq 0}$ *is a complete orthnormal system of* $L^2[0, \infty)$*, i.e.,*

$$\langle\, h_{n,k}, h_{n',k'} \,\rangle = \begin{cases} 1, & \text{if } (n, k) = (n', k') \\ 0, & \text{if otherwise.} \end{cases} \tag{8.6}$$

and

$$\bigcap_{n,k \geq 0} \{g \in L^2[0, \infty) \ ; \ \langle\, h_{n,k}, g \,\rangle = 0\} = \{g \equiv 0\} \tag{8.7}$$

Proof: The proof of (8.6) is easy and is left to the readers (cf. Exercise 8.2.2 below). To prove (8.7), we take a function g from the set on the left-hand-side of (8.7) and show that

$$G(t) = G(0) \quad \text{for all } t \geq 0,$$

where $G(t) \stackrel{\text{def.}}{=} \int_0^t g(s)ds$. Since diadic rationals are dense, it is enough to prove

1) $\qquad G\left(\frac{2k+1}{2^n}\right) = G(0) \quad \text{for all } n, k \geq 0.$

We will prove (1) by induction on n. For $n = 0$, we have

2) $\qquad G(k+1) - G(k) = \int_k^{k+1} g(t)dt = \langle\, h_{0,k}, g \,\rangle = 0, \quad k = 0, 1, \ldots.$

Similarly for $n = 1, 2, \ldots$, and $k = 0, 1, \ldots$,

$$\begin{aligned} & G\left(\tfrac{2k+1}{2^n}\right) - \tfrac{1}{2}G\left(\tfrac{2k}{2^n}\right) - \tfrac{1}{2}G\left(\tfrac{2k+2}{2^n}\right) \\ = \ & \tfrac{1}{2}\left(G\left(\tfrac{2k+1}{2^n}\right) - G\left(\tfrac{2k}{2^n}\right)\right) - \tfrac{1}{2}\left(G\left(\tfrac{2k+2}{2^n}\right) - G\left(\tfrac{2k+1}{2^n}\right)\right) \\ = \ & \tfrac{1}{2}\int_{\frac{2k}{2^n}}^{\frac{2k+1}{2^n}} g(t)dt - \tfrac{1}{2}\int_{\frac{2k+1}{2^n}}^{\frac{2k+2}{2^n}} g(t)dt = \tfrac{1}{2}2^{-\frac{n-1}{2}}\langle\, h_{n,k}, g \,\rangle = 0. \end{aligned}$$

We see from this and (2) that induction works. □

Exercise 8.2.2 Prove (8.6).

Recall that $\nu_1 \in \mathcal{P}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ stands for the Gaussian distribution with mean-zero and variance one. We let $\xi = (\xi_{n,k})_{n,k\geq 0}$ denote a r.v. with $P\{\xi \in \cdot\} = \otimes_{n,k\geq 0}\nu_1$ defined on a probability space $(\Omega, \mathcal{F}, P)$. We will prove the existence of BM^1 in the following form;

Theorem 8.2.2 a) *A series*

$$B_t = \sum_{n,k\geq 0} \xi_{n,k} \int_0^t h_{n,k}(u)du, \quad t \geq 0 \tag{8.8}$$

is absolutely convergent almost surely. More precisely, for any $\alpha \in [0, 1/2)$ and $T > 0$, there is a $(0, \infty)$-valued r.v. M such that with probability one,

$$\sum_{n,k\geq 0} \left|\xi_{n,k} \int_s^t h_{n,k}(u)du\right| \leq M|t-s|^\alpha \quad \text{for any } 0 \leq s < t \leq T. \tag{8.9}$$

b) *$\{B_t\}_{t\geq 0}$ defined above is a* BM^1.

Theorem 8.2.2 tells us also that $t \mapsto B_t$ for the Brownian motion is α-Hölder continuous for any $\alpha < 1/2$.

Corollary 8.2.3 *If $\{B_t\}_{t\geq 0}$ is a* BM^1*, then for any $\alpha \in [0, 1/2)$ and $T > 0$,*

$$P\left(\sup_{0\leq s<t\leq T} \frac{|B_t - B_s|}{|t-s|^\alpha} < \infty\right) = 1.$$

Remark: $t \mapsto B_t$ is not α-Hölder continuous for any $\alpha > 1/2$ (Exercise 8.1.5).

We now turn to the proof of Theorem 8.2.2, which will be divided into some lemmas. We begin with the following

Lemma 8.2.4

$$P(S(\xi) < \infty) = 1,$$

where

$$S(\xi) = \sup_{n,k\geq 0} |\xi_{n,k}|/\sqrt{\log(2+n+k)}.$$

Proof: We will in fact prove that

$$P\left(|\xi_{n,k}| \leq \sqrt{6\log(2+n+k)} \quad \text{except finitely many } (n,k)\text{'s}\right) = 1.$$

We first compute any $y > 0$ that

$$\begin{aligned} P\{|\xi_{n,k}| > y\} &= \sqrt{2/\pi} \int_y^\infty \exp(-x^2/2)dx \\ &\leq \sqrt{2/\pi} \int_y^\infty (x/y)\exp(-x^2/2)dx = \sqrt{2/\pi}\exp(-y^2/2)/y. \end{aligned}$$

We then use this inequality as follows,

$$
\begin{aligned}
E\left[\sum_{n,k\geq 0} 1\{|\xi_{n,k}| > \sqrt{6\log(2+n+k)}\right] &= \sum_{n,k\geq 0} P\{|\xi_{n,k}| > \sqrt{6\log(2+n+k)}\} \\
&\leq \sqrt{2/\pi}\sum_{n,k\geq 0} \exp(-3\log(2+n+k)) \\
&= \sqrt{2/\pi}\sum_{n,k\geq 0} (2+n+k)^{-3} < \infty.
\end{aligned}
$$

As a consequence, $\sum_{n,k\geq 0} 1\{|\xi_{n,k}| > \sqrt{6\log(2+n+k)}\} < \infty$, P-a.s., which is equivalent to what we wanted to prove. □

Lemma 8.2.5 *Define a linear subspace $\mathcal{X}$ of $L^2([0,\infty))$ by $\mathcal{X} = \cup_{T>0,l\geq 1}\mathcal{X}_{l,T}$, where*

$$
\mathcal{X}_{l,T} = \{\sum_{i=1}^{l} \theta_i 1_{[0,t_i)} \ ; \ (\theta_i)_{i=1}^{l} \in \mathbb{R}^l, \ 0 < t_1 < \ldots < t_l \leq T\}.
$$

Then the following hold;

a) *For $g \in \mathcal{X}$, a series*

$$
B(g) = \sum_{n,k\geq 0} \xi_{n,k}\langle\, h_{n,k}, g \,\rangle, \tag{8.10}
$$

is almost surely absolutely convergent. More precisely, for any $q > 2$, $l \geq 1$ and $T > 0$; there is a constant $C = C(q,l,T) \in (0,\infty)$ such that

$$
\sum_{n,k\geq 0} |\xi_{n,k}\langle\, h_{n,k}, g \,\rangle| \leq CS(\xi)\|g\|_{L^q[0,\infty)} \quad \text{for all } g \in \mathcal{X}_{l,T}, \tag{8.11}
$$

where $S(\xi)$ is defined in Lemma 8.2.4.

b) *$\{B(g)\}_{g\in\mathcal{X}}$ is a family of a mean-zero Gaussian r.v.'s such that*

$$
E[B(g_1)B(g_2)] = \langle\, g_1, g_2 \,\rangle, \quad \text{for all } g_1, g_2 \in \mathcal{X}. \tag{8.12}
$$

c) *If $\{g_j\}_{j=1}^{n} \subset \mathcal{X}$ and $\langle\, g_i, g_j \,\rangle = 0$ for $i \neq j$, then $\{B(g_j)\}_{j=1}^{n}$ are independent.*

Proof: (**a**): It is enough to prove (8.11). Let $q > 2$, $l \geq 1$, $T > 0$ and $g = \sum_{i=1}^{l} \theta_i 1_{[0,t_i)} \in \mathcal{X}_{l,T}$. We take $p \in (1,2)$ such that $\frac{1}{p} + \frac{1}{q} = 1$ and define $\varepsilon = \frac{1}{p} - \frac{1}{2} > 0$. Note first that

1) If $\langle\, h_{n,k}, g \,\rangle \neq 0$, then $\frac{2k}{2^n} \leq T$,

2) $\displaystyle\sum_{k:\langle\, h_{n,k},g \,\rangle\neq 0} 1 \leq (1+T)l,$

3) $\|h_{n,k}\|_{L^p[0,\infty)} = 2^{\frac{n-1}{2} - \frac{n-1}{p}} = 2^{-(n-1)\varepsilon}.$

In fact, (1) and (3) are obvious from the definition of $h_{n,k}$. The second inequality (2) can be seen as follows;

$$\sum_{k:\langle\, h_{n,k}, g\,\rangle\neq 0} 1 \le \sum_{i=1}^{l} \sum_{k:\langle\, h_{n,k}, 1_{[0,t_i)}\,\rangle\neq 0} 1$$

and for any $t>0$,

$$\sum_{k:\langle\, h_{n,k}, 1_{[0,t)}\,\rangle\neq 0} 1 \le \begin{cases} 1+t, & \text{if } n=0 \\ 1 & \text{if } n\ge 1. \end{cases}$$

If $\langle\, h_{n,k}, g\,\rangle \neq 0$, then

$$\begin{aligned} |\xi_{n,k}| &\le S(\xi)\sqrt{\log(2+n+k)} \le S(\xi)\sqrt{\log(2+n+2^nT)} \quad \text{by (1)}, \\ |\langle\, h_{n,k}, g\,\rangle| &\le 2^{-(n-1)\varepsilon}\|g\|_{L^q[0,\infty)} \quad \text{by (3) and Hölder.} \end{aligned}$$

Therefore,

$$\begin{aligned} \sum_{n,k\ge 0} |\xi_{n,k}\langle\, h_{n,k}, g\,\rangle| &= \sum_{n\ge 0}\sum_{k:\langle\, h_{n,k},g\,\rangle\neq 0} |\xi_{n,k}\langle\, h_{n,k}, g\,\rangle| \\ &\le \sum_{n\ge 0}\sum_{k:\langle\, h_{n,k},g\,\rangle\neq 0} S(\xi)\sqrt{\log(2+n+2^nT)}2^{-(n-1)\varepsilon}\|g\|_{L^q[0,\infty)} \\ &\le (1+T)lS(\xi)\|g\|_{L^q[0,\infty)}\sum_{n\ge 0}\sqrt{\log(2+n+2^nT)}2^{-(n-1)\varepsilon}. \end{aligned}$$

The series in the third line converges and this proves (8.11).

(**b**): Ingredients of the proof will be Lemma 8.2.1 and some basic properties of Gaussian r.v.'s listed in Exercise 5.3.3–Exercise 5.3.6. For $g\in\mathcal{X}$, we define $B(g)$ by (8.10) and $B_N(g)$ by the partial sum;

$$B_N(g) = \sum_{n=0}^{N}\sum_{k\ge 0}\xi_{n,k}\langle\, h_{n,k}, g\,\rangle.$$

Then,

- $B_N(g)$ for each $g\in\mathcal{X}$ is a is a mean-zero Gaussian r.v..

In fact, $B_N(g)$ is a finite summation of independent mean-zero Gaussian r.v.'s (cf. (2)) and hence is a mean-zero Gaussian r.v. by Exercise 5.3.3.
Next, as a consequence of part (a),

- $\lim_{N\nearrow\infty} B_N(g) = B(g)$, P-a.s.

Moreover,

- $\lim_{N\nearrow\infty} E[B_N(g_1)B_N(g_2)] = \langle\, g_1, g_2\,\rangle$ for $g_1, g_2\in\mathcal{X}$.

This can be seen as follows;

$$\begin{aligned} E[B_N(g_1)B_N(g_2)] &= \sum_{n,n'=0}^{N}\sum_{k,k'\ge 0}\langle\, h_{n,k}, g_1\,\rangle\langle\, h_{n',k'}, g_2\,\rangle E[\xi_{n,k}\xi_{n',k'}] \\ &= \sum_{n=0}^{N}\sum_{k\ge 0}\langle\, h_{n,k}, g_1\,\rangle\langle\, h_{n,k}, g_2\,\rangle \stackrel{N\nearrow\infty}{\longrightarrow} \sum_{n\ge 0}\sum_{k\ge 0}\langle\, h_{n,k}, g_1\,\rangle\langle\, h_{n,k}, g_2\,\rangle \\ &= \langle\, g_1, g_2\,\rangle, \quad \text{by Parseval's identity.} \end{aligned}$$

These, together with Exercise 5.3.6, prove that $B(g)$ for each $g \in \mathcal{X}$ is a Gaussian r.v. and that (8.12) holds for $g_1, g_2 \in \mathcal{X}$.
(**c**): By part (**b**), $\sum_{j=1}^n \theta_j B(g_j) = B(\sum_{j=1}^n \theta_j g_j)$ is a Gaussian r.v. for $(\theta_j)_{j=1}^n \in \mathbb{R}^n$. Hence it follows from Exercise 5.3.4 that $(B(g_j))_{j=1}^n$ is an $\mathbb{R}^n$-valued Gaussian r.v. By this, (8.12) and Exercise 5.3.5, we see that $\{B(g_j)\}_{j=1}^n$ are independent. $\Box$

Proof of Theorem 8.2.2: We set $g_t = 1_{[0,t)}$. Then, we see from (8.8) and (8.10) that for $0 \le s \le t < \infty$,

1) $\quad B_t = B(g_t),$

2) $\quad B_t - B_s = B(g_t - g_s).$

Since $\|g_t - g_s\|_{L^q[0,\infty)} = |t-s|^{1/q}$, the bound (8.9) follows from (8.11) and (2). Let next us check (**B0**)–(**B3**) (with $d = 1$) for $\{B_t\}_{t \ge 0}$.
(**B0**): This is obvious by the definition (8.8).
(**B1**): This follows from (8.9).
(**B2**): If $n \ge 2$ and $0 = t_0 < t_1 < \ldots < t_n$, then $\langle\, g_{t_i} - g_{t_{i-1}}, g_{t_j} - g_{t_{j-1}} \,\rangle = 0$ for $i \ne j$ Therefore, $B_{t_i} - B_{t_{i-1}} = B(g_{t_i} - g_{t_{i-1}})$ $(i = 1, \ldots n)$ are independent by Lemma 8.2.5 **(c)** .
(**B3**): $\langle\, g_t - g_s, g_t - g_s \,\rangle = t - s$ for $0 \le s < t$. Hence it follows from Lemma 8.2.5 **(b)** that $B_t - B_s = B(g_t - g_s)$ is a Gaussian r.v. with the variance $t - s$. $\Box$

8.3 Itô's formula

In this subsection, we will explain Itô's formula for the Brownian motion and its applications without going much into proofs. For complete proofs and more details, we will refer to the following references;

[Dur84] Durrett, R. :"Brownian Motion and Martingales in Analysis", Wadsworth & Brooks/Cole Advanced Books & Software, Belmont, Calfornia.

[Fol76] Folland, G. B.: Introduction to partial differential equations, Princeton University Press, 1976.

[KS91] Karatzas, I. and Shreve, S. E.: Brownian Motion and Stochastic Calculus, Second Edition. Springer Verlag (1991).

In what follows, $B_t = (B_t^i)_{i=1}^d$, $t \ge 0$ denotes a BM^d on a probability space $(\Omega, \mathcal{F}, P)$.
For $f \in C^2(\mathbb{R}^d)$ and $i = 1, ..., d$, we write $\partial_i f = \frac{\partial}{\partial x^i} f$ and $\Delta f = \sum_{1 \le i \le d} \left(\frac{\partial}{\partial x^i}\right)^2 f$.

Theorem 8.3.1 (Itô's formula for the Brownian motion) *Suppose that $f \in C^2(\mathbb{R}^d)$. Then, P-a.s.,*

$$f(B_t) - f(0) = \sum_{1 \le i \le d} \int_0^t \partial_i f(B_s) dB_s^i + \frac{1}{2} \int_0^t \Delta f(B_s) ds, \quad \textit{for all } t \ge 0. \tag{8.13}$$

Here the integral $\int_0^t \partial_i f(B_s) dB_s^i$ is called the stochasic integral with respect to the Brownian motion explained below.

Remark: The function $s \mapsto B_s^i$ is *not* of bounded variation in any interval. Therefore, the integral $\int_0^t \partial_i f(B_s) dB_s^i$ cannot be defined as a Lebesgue-Stieltjes integral.

We now turn to the explanation of the stochasic integral with respect to the Brownian motion. Unfortunately, we have to begin with a technicality.

We define the *filteration* $(\mathcal{F}_t)_{t \geq 0}$ as follows;

$$\begin{aligned} \mathcal{F}_t^0 &= \sigma[(B_s)_{s \leq t}] \quad t \geq 0, \\ \mathcal{F}_t &= \cap_{\varepsilon > 0} \sigma[\mathcal{F}_{t+\varepsilon}^0, \ \mathcal{N}] \quad t \geq 0. \end{aligned}$$

where $\mathcal{N} = \{N \in \mathcal{F} \quad P(N) = 0\}$. It would be natural to introduce $\mathcal{F}_t^0$, which is all the information available up to time t. The technical advantage of introducing $\mathcal{F}_t$ ("an infinitesimal peeping in the future plus null sets") is to enlarge $\mathcal{F}_t^0$ to get the following properties;

- Right continuity; $\cap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon} = \mathcal{F}_t$, $t \geq 0$,
- Completeness; $\mathcal{N} \subset \mathcal{F}_t$, $t \geq 0$.

On the other hand, enlargement is only a little so that we still have that

- For $t \geq 0$, $\sigma[B_{t+s} - B_t \ ; \ s \geq 0]$ and $\mathcal{F}_t$ are independent under P.

See [KS91,page 93] for the proof of above mentioned properties of $\mathcal{F}_t$.

Definition 8.3.2 (**Stopping times**) A r.v. $\tau : \Omega \to [0, \infty]$ is called a *stopping time* if

$$\{\tau \leq t\} \in \mathcal{F}_t \text{ for all } t \geq 0. \tag{8.14}$$

Example 8.3.3 Let $\Gamma \subset \mathbb{R}^d$ and define

$$\tau(\Gamma) = \inf\{t > 0 \ ; \ B_t \in \Gamma\}.$$

It is known that $\tau(\Gamma)$ is a stopping time if $\Gamma \in \mathcal{B}(\mathbb{R}^d)$. This is not difficult to prove if Γ is either open or closed. Here, in the proof, one sees how the right continuity of $\mathcal{F}_t$ is used. See the remark below and Exercise 8.3.1.

Remark : Consider the following condition[13] for a r.v. $\tau : \Omega \to [0, \infty]$;

$$\{\tau < t\} \in \mathcal{F}_t \text{ for all } t \geq 0. \tag{8.15}$$

Then, this is equivalent to (8.14). In fact, we have

1) $\qquad \{\tau < t\} = \cup_{n \geq 1} \{\tau \leq t - \frac{1}{n}\}$,

2) $\qquad \{\tau > t\} = \cap_{m \geq 1} \cup_{n \geq m} \{\tau \geq t - \frac{1}{n}\}$.

[13] A r.v. τ with this condition is called an *optional time*. We see from the argument of this remark that a stopping time is always an optional time, and that the converse is true when the filteration is right continuous.

We see from (1) that (8.14) implies (8.15), while the converse can be seen from (2) and the right continuity of $\mathcal{F}_t$.

The observation above can be used to prove that $\tau(\Gamma)$ defined in Example 8.3.3 is a stopping time for an open set Γ. We prove that $\tau(\Gamma)$ satisfies (8.15) as follows:

$$\{\tau(\Gamma) < t\} = \bigcup_{s \in (0,t)} \{B_s \in \Gamma\} = \bigcup_{s \in \mathbb{Q} \cap (0,t)} \{B_s \in \Gamma\} \in \mathcal{F}_t,$$

where, to get the second equality, we have used that Γ is open and that $s \mapsto B_s$ is continuous.

Exercise 8.3.1 Prove that $\tau(\Gamma)$ defined in Example 8.3.3 is a stopping time if Γ is closed. Hint: There is a sequence of open sets $G_1 \supset G_2 \supset ...$ such that $\Gamma = \cap_{m \geq 1} G_m$.

We now define some classes of integrands for the stochastic integral.

Definition 8.3.4 (**Integrands for stochastic integral**) We define a function space Φ as the totality of $\varphi : [0,\infty) \times \Omega \to \mathbb{R}$ such that[14]

$$\varphi|_{[0,t]\times\Omega} \text{ is } \mathcal{B}([0,t]) \otimes \mathcal{F}_t \text{ measurable for all } t \geq 0.$$

We also define

$$\begin{aligned} \Phi_2 &= \{\varphi \in \Phi \ ; \ E \textstyle\int_0^t |\varphi(s,\omega)|^2 ds < \infty \text{ for all } t > 0\} \\ \Phi_2^{\text{loc.}} &= \{\varphi \in \Phi \ ; \ \textstyle\int_0^t |\varphi(s,\omega)|^2 ds < \infty, \ P\text{-a.s. for all } t > 0\}. \end{aligned}$$

Clearly, $\Phi_2 \subset \Phi_2^{\text{loc.}} \subset \Phi$.

Example 8.3.5 Let $g : \mathbb{R}^d \to \mathbb{R}$ be Borel measurable and

$$\varphi(s,\omega) = g(B_s(\omega)).$$

Then,

- If g is bounded, then $\varphi \in \Phi_2$.
- If $\sup_K |g| < \infty$ for any bounded set $K \subset \mathbb{R}^d$ (in particular, if $g \in C(\mathbb{R}^d)$), then $\varphi \in \Phi_2^{\text{loc.}}$.

Theorem 8.3.6 *For $t \geq 0$ and $\varphi \in \Phi_2^{\text{loc.}}$, there are r.v.'s (called the stochastic integral with respect to the Brownian motion)*

$$\int_0^t \varphi(s,\omega) dB_s^i \quad i = 1,..,d \tag{8.16}$$

with the following properties;

(a) *If*

$$\varphi(s,\omega) = \xi_a(\omega) 1_{[a,b)}(s) \tag{8.17}$$

where $0 \leq a < b$ and ξ_a is an $\mathcal{F}_a$-measurable r.v., then

$$\int_0^t \varphi(s,\omega) dB_s^i = \xi_a(\omega)(B_{t\wedge b}^i - B_{t\wedge a}^i). \tag{8.18}$$

[14] This property is called *progressive measurability*

(b) *For $t \geq 0$, $\alpha, \beta \in \mathbb{R}$ and $\varphi, \psi \in \Phi$*

$$\int_0^t (\alpha\varphi(s,\omega) + \beta\psi(s,\omega))dB_s^i = \alpha \int_0^t \varphi(s,\omega)dB_s^i + \beta \int_0^t \psi(s,\omega)dB_s^i, \tag{8.19}$$

(c) *If $\varphi \in \Phi_2$ and $t \geq 0$, then,*

$$E\left[\left|\int_0^t \varphi(s,\omega)dB_s^i\right|^2\right] = E\int_0^t |\varphi(s,\omega)|^2 ds < \infty, \tag{8.20}$$

$$E\left[\int_0^t \varphi(s,\omega)dB_s^i\right] = 0 \tag{8.21}$$

Moreover, (8.21) and (8.20) remain true if t is replaced by $t \wedge \tau$, with a stopping time τ.

We now indicate how the construction of the integrals (8.16) goes (See [KS91, Section 3.2] dor details).

- If φ is a linear combination of the functions of the form (8.17), we then defiene the integrals by (8.18) and (8.19).
- For $\varphi \in \Phi_2$, we approximate φ by a sequence of linear combinations explained above (in a certain L^2 sense), and define the integral by a limiting procedure.
- For $\varphi \in \Phi_2^{\text{loc.}}$, We consider

$$\begin{aligned}\tau_n &= n \wedge \inf\left\{t > 0\ ;\ \int_0^t |\varphi(s,\omega)|^2 ds \geq n\right\}\\ \varphi_n(s,\omega) &= \varphi(s,\omega)1_{[0,\tau_n]}(s).\end{aligned}$$

Then, $\tau_n \nearrow \infty$ and $\varphi_n \in \Phi_2$. We then define the integrals (8.16) by

$$\int_0^t \varphi(s,\omega)dB_s^i = \int_0^t \varphi_n(s,\omega)dB_s^i \quad \text{for } t \leq \tau_n.$$

Remark: In advanced texts in probability, properties (8.20) and (8.21) are usually stated in the following stronger form. For $\varphi \in \Phi_2$, we set $M_t = \int_0^t \varphi(s,\omega)dB_s^i$. Then,

$$E[M_t : A] = E[M_u : A] \quad \text{if } 0 \leq u < t \text{ and } A \in \mathcal{F}_u. \tag{8.22}$$

The same is true for $M_t = \left|\int_0^t \varphi(s,\omega)dB_s^i\right|^2 - \int_0^t |\varphi(s,\omega)|^2 ds$. In general, a sequence $(M_t)_{t\geq 0} \subset L^1(P)$ is called a *martingale*, if $M_t : \Omega \to \mathbb{R}$ is $\mathcal{F}_t$-measurable for all $t \geq 0$, and (8.22) holds. Martingale is in fact one of the most useful tools in modern probability theory. See again [Dur84] or [KS91] for details.

Example 8.3.7 (Application to the Dirichlet problem) Let a bounded open set $G \subset \mathbb{R}^d$, $f \in C(\partial G)$, and $g \in C(\overline{G})$ be given. A classical problem in the theory of partial differential equations is to show the existence and uniqueness of $u \in C(\overline{G}) \cap C^2(G)$ such that

a) $\frac{1}{2}\Delta u = -g$ in G,

b) $u|_{\partial G} = f$.

A special case where $g \equiv 0$ is especially famous as *Dirichlet problem*. Here, we will prove the uniqueness by running a Brownian motion. We will in fact represent the solution of (a) and (b) as follows. Let

$$\tau(x, \partial G) = \inf\{t > 0 \; ; \; x + B_t \in \partial G\}, \quad x \in G,$$

which is the first time for a Brownian motion (starting from $x \in G$) to reach the boundary ∂G. By Exercise 8.3.1, $\tau(x, \partial G)$ is a stopping time. We will then prove that a solution u to (a) and (b) is represented as

1) $$u(x) = E[f(x + B_{\tau(x,\partial G)})] + E \int_0^{\tau(x,\partial G)} g(x + B_s) ds$$

hence is unique.

Let us first prove that[15]

2) $$E[t(x, \partial G)] \leq \sup_{y \in G} |y - x|^2 < \infty.$$

Since $B_t^1 = \int_0^t dB_s^1$, we have by (8.20) that

$$\begin{aligned} E[\tau(x, \partial G) \wedge t] &= E\left[|B^1_{\tau(x,\partial G)\wedge t}|^2\right] \\ &= E\left[|(x + B^1_{\tau(x,\partial G)\wedge t}) - x|^2\right] \leq \sup_{y \in G} |y - x|^2, \end{aligned}$$

from which we obtain (2) by letting $t \nearrow \infty$.

We now turn to (1). Suppose that $x \in G$ and $u \in C(\overline{G}) \cap C^2(G)$ satisfies (a) and (b). We have by Itô's formula applied to $y \mapsto u(x + y)$ that,

$$\begin{aligned} &u(x + B_{t\wedge\tau(x,\partial G)}) - u(x) \\ &= \sum_{1\leq i\leq d} \int_0^{t\wedge\tau(x,\partial G)} \partial_i u(x + B_s) dB_s^i + \frac{1}{2} \int_0^{t\wedge\tau(x,\partial G)} \Delta u(x + B_s) ds \\ &= \sum_{1\leq i\leq d} \int_0^{t\wedge\tau(x,\partial G)} \partial_i u(x + B_s) dB_s^i - \int_0^{t\wedge\tau(x,\partial G)} g(x + B_s) ds. \end{aligned}$$

We then take expectation and use (8.21) to see that

$$E[u(x + B_{t\wedge\tau(x,\partial G)})] - u(x) = -E \int_0^{t\wedge\tau(x,\partial G)} g(x + B_s) ds.$$

By (2), we can use the dominated convergence thorem in the limit $t \nearrow \infty$ 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 a certain regularity condition on ∂D that

$$P(\tau(x, \partial G) = 0) = 1 \text{ for all } x \in \partial G$$

to show the continuity of u at the boundary. See [Dur84, Sections 8.5,8.6], [KS91, Section 4.2] for the proofs and details.

[15] This fact is not needed when $g \equiv 0$.

Remark: Of course, the existence and uniqueness of u discussed in Example 8.3.7 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 G assuming that G has a smooth boundary [Fol76,pages 112, 343].

9 Appendix to Section 2

9.1 Proof of Lemma 2.1.1

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

$$\mathcal{D} \stackrel{\text{def.}}{=} \{B \in \mathcal{B} \ ; \ \mu(B) = \nu(B)\}. \tag{9.1}$$

If the class $\mathcal{D}$ defined by (9.1) happens to be closed under intersection, it is then not difficult to prove that $\mathcal{D}$ is a σ-field[16] and hence that $\sigma[\mathcal{A}] \subset \mathcal{D}$. Unfortunately, $\mathcal{D}$ is *not* closed under intersection in general. In fact, we see in Exercise 9.6.2 an example where

- the family $\mathcal{D}$ in (9.1) is not a σ-field and hence is not closed under intersection (Exercise 9.1.1).
- "$\mu(A) = \nu(A)$ for all $A \in \mathcal{A}$" does not imply "$\mu(A) = \nu(A)$ for all $A \in \sigma(\mathcal{A})$" if $\mathcal{A}$.

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

Definition 9.1.1 Suppose that S is a set. A subset $\mathcal{A}$ of 2^S is called a π-*system* if $A_1 \cap A_2 \in \mathcal{A}$ for any $A_1, A_2 \in \mathcal{A}$. A subset $\mathcal{D}$ of 2^S is called[17] a δ-*system* or a *Dynkin class* if the following conditions are satisfied;

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

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

Exercise 9.1.1 Prove that a δ-system $\mathcal{D}$ is a σ-field if and only if $\mathcal{D}$ is a π-system.

Lemma 9.1.2 (Dynkin's π-δ lemma) *Suppose that S is a set and that $\mathcal{A} \subset \mathcal{D} \subset 2^S$, where $\mathcal{A}$ is a π-system and $\mathcal{D}$ is a δ-system. Then, $\sigma[\mathcal{A}] \subset \mathcal{D}$.*

Proof: Define $\delta[\mathcal{A}] = \cap \mathcal{D}$, where the intersection is taken over all δ-system $\mathcal{D}$ which contains $\mathcal{A}$. Then, it is easy to prove that $\delta[\mathcal{A}]$ is the smallest δ-system that contains $\mathcal{A}$ and therefore that

$$\delta[\mathcal{A}] \subset \mathcal{D} \tag{9.2}$$

To complete the proof of Lemma 9.1.2, it is thus enough to show that

$$\sigma[\mathcal{A}] = \delta[\mathcal{A}] \text{ if } \mathcal{A} \text{ is a } \pi\text{-system.} \tag{9.3}$$

It is clear that $\sigma[\mathcal{A}] \supset \delta[\mathcal{A}]$, since $\sigma[\mathcal{A}]$ is a δ-system which contains $\mathcal{A}$. To prove the other inclusion, it is enough to prove that $\delta[\mathcal{A}]$ is a π-system, i.e. ,

$$A \cap B \in \delta[\mathcal{A}] \tag{9.4}$$

for all $\{A, B\} \subset \delta[\mathcal{A}]$ (Exercise 9.1.1). We now introduce

$$\begin{aligned} \mathcal{B}_1 &= \{B \in \delta[\mathcal{A}] \ ; \ (9.4) \text{ holds for all } A \in \mathcal{A}\}, \\ \mathcal{B}_2 &= \{B \in \delta[\mathcal{A}] \ ; \ (9.4) \text{ holds for all } A \in \delta[\mathcal{A}]\}. \end{aligned}$$

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

[17] Here, $\mathcal{D}$ is general and is not necessarily the one defined by (9.1)

What we want to prove is paraphrased as $\mathcal{B}_2 = \delta[\mathcal{A}]$. But we begin by proving that $\mathcal{B}_1 = \delta[\mathcal{A}]$. Since $\mathcal{A} \subset \mathcal{B}_1 \subset \delta[\mathcal{A}]$, the equality $\mathcal{B}_1 = \delta[\mathcal{A}]$ can easily be verified by showing that $\mathcal{B}_1$ is a δ-system. By the equality $\mathcal{B}_1 = \delta[\mathcal{A}]$, we now know that $\mathcal{A} \subset \mathcal{B}_2 \subset \delta[\mathcal{A}]$. Therefore, it is also easy to prove the desired equality $\mathcal{B}_2 = \delta[\mathcal{A}]$ by showing that $\mathcal{B}_2$ is a δ-system . □

Proof of Lemma 2.1.1: It is easy so see that $\mathcal{D}$ defined by (9.1) is a δ-system . Since $\mathcal{A} \subset \mathcal{D}$ and $\mathcal{A}$ is a π-system, we see by Lemma 9.1.2 that $\sigma[\mathcal{A}] \subset \mathcal{D}$. □

9.2 Uniform distribution and an existence theorem for independent r.v.'s

To difine a random walk (cf. Definition 4.2.1 below), we will need countably many independent r.v's. A question[18] 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 9.2.1 *Consider a sequence of probability spaces $\{(S_n, \mathcal{B}_n, \mu_n)\}_{n \geq 1}$ where for each n, S_n is a complete separable metric space and $\mathcal{B}_n$ is the Borel σ-field. Then, there is a sequence of independent r.v.'s $\{X_n : \Omega \to S_n\}_{n \geq 1}$ such that $\mu_n(B) = P(X_n \in B)$ for all $n \geq 1$ and $B \in \mathcal{B}_n$.*

Remark: Proposition 9.2.1 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 9.2.1 without relying on any big theorem from measure theory.

We begin with examples:

Example 9.2.2 Let us now construct an i.i.d. sequence $\{U_n\}_{n \geq 1}$ of $[0, 1)$-valued r.v.'s with the uniform distribution. By Example 3.3.2, there is an i.i.d. sequence $\{X_{n,k}\}_{n,k \geq 1}$ of $\{0, 1\}$-valued r.v.'s with $P\{X_{n,k} = 1\} = 1/2$. We define $\{U_n\}_{n \geq 1}$ by

$$U_n = \sum_{k \geq 1} 2^{-k} X_{n,k}.$$

Then, each U_n is uniformly distributed by Lemma 9.4.1. Moreover, $\{U_n\}_{n \geq 1}$ are independent by Exercise 2.3.9.

To prove Proposition 9.2.1, we will use Example 3.3.2, Example 9.2.2 and the following lemma.

Lemma 9.2.3 *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-field. Then, there is a measurable map $\varphi : [0, 1) \to S$ such that*

$$P\{\varphi(U) \in B\} = \mu(B), \quad \textit{for all } B \in \mathcal{B}, \tag{9.5}$$

where $U : \Omega \to [0, 1)$ is a uniformly distributed r.v.

[18] This may be a question which a physicist would not care about. Those who do not worry about this question can skip this subsection.

Lemma 9.2.3 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 9.2.3 will be presented in subsection 9.3.

We now prove Proposition 9.2.1.

Proof of Proposition 9.2.1: Let $\{U_n\}_{n\geq 1}$ be $[0,1)$-valued r.v.'s with the uniform distribution constructed in Example 9.2.2. 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 9.2.3. 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 9.2.1. □

Exercise 9.2.1 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 9.2.2 Let $\mu_n \in \mathcal{P}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ $(n = 1, \ldots)$ be a sequence of probability measures. Use Example 9.2.2 and Exercise 9.2.1 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 2.1.1.

9.3 Proof of Lemma 9.2.3

The proof of Lemma 9.2.3 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 spaecial 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 9.2.3 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

$$\begin{aligned} Q_{s_1} &\ni s_1 \quad \text{for all } s_1 \in C, \\ \mu(N) &= 0, \quad \text{where } N \stackrel{\text{def.}}{=} S \setminus \cup_{s_1 \in C} Q_{s_1}, \\ \mu(Q_{s_1}) &> 0, \quad \text{for all } s_1 \in C. \end{aligned} \tag{9.6}$$

In fact, this can be done just by setting

$$
\begin{aligned}
Q_{s_1} &= \textstyle\prod_{j=1}^d [s_1^j, s_1^j + 2^{-1}), \quad \text{for } s_1 = (s_1^j)_{j=1}^d \in 2^{-1}\mathbb{Z}^d, \\
C &= \{s_1 \in 2^{-1}\mathbb{Z}^d \ ; \ \mu(Q_{s_1}) > 0\}.
\end{aligned} \tag{9.7}
$$

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

$$
\begin{aligned}
Q_{s_1} &\supset Q_{s_1 s_2} \ni s_2 \quad \text{for all } s_2 \in C(s_1), \\
\mu(N_{s_1}) &= 0, \quad \text{where } N_{s_1} \stackrel{\text{def.}}{=} Q_{s_1} \setminus \cup_{s_2 \in C(s_1)} Q_{s_1,s_2}, \\
\mu(Q_{s_1 s_2}) &> 0 \quad \text{for all } s_2 \in C(s_1).
\end{aligned}
$$

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 $s_1 \in nnC, \ldots, s_n \in C(s_1 \cdots s_{n-1})$. 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

$$
Q_{s_1 \cdots s_n} \supset Q_{s_1 \cdots s_{n+1}} \ni s_{n+1} \quad \text{for all } s_{n+1} \in C(s_1, \ldots, s_n). \tag{9.8}
$$

$$
\mu(N_{s_1 \cdots s_n}) = 0, \quad \text{where } N_{s_1 \cdots s_n} \stackrel{\text{def.}}{=} Q_{s_1 \cdots s_n} \setminus \cup_{s_{n+1} \in C(s_1, \ldots, s_n)} Q_{s_1 \cdots s_{n+1}}, \tag{9.9}
$$

$$
\mu(Q_{s_1 \cdots s_{n+1}}) > 0 \quad \text{for all } s_{n+1} \in C(s_1, \ldots, s_n).
$$

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

$$
[0, 1) = \cup_{s_1 \in C} I_{s_1}, \tag{9.10}
$$

$$
I_{s_1 \cdots s_{n-1}} = \cup_{s_n \in C(s_1, \ldots, s_{n-1})} I_{s_1 \cdots s_n}, \tag{9.11}
$$

$$
|I_{s_1 \cdots s_n}| = \mu(Q_{s_1 \cdots s_n}). \tag{9.12}
$$

Step 3: We now define $\varphi_n : [0, 1) \to S$ by

$$
\varphi_n(\theta) = s_n \quad \text{if } \theta \in I_{s_1 \cdots s_n}. \tag{9.13}
$$

Let us check the following;

$$
\varphi_n : [0, 1) \to S \text{ is well defined and measurable for all } n \geq 1. \tag{9.14}
$$

$$
(\varphi_n(\theta))_{n \geq 1} \text{ is a Cauchy sequence for for all } \theta \in [0, 1). \tag{9.15}
$$

By (9.10) and (9.11), 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 (9.15), 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

$$|\varphi_{m+n}(\theta) - \varphi_n(\theta)| \le 2^{-n}\sqrt{d}.$$

Step 4: By, (9.15) and (9.15), we can define a measurable map $\varphi : [0,1) \to \mathbb{R}^d$ by $\varphi(\theta) = \lim_{n\nearrow\infty} \varphi_n(\theta)$ for all $\theta \in [0,1)$. Let us see that φ satisfies (9.5). To do so, define a set

$$N_0 = \cup_{n\ge 1} \cup_{s_1\in C} \cup_{s_2\in C(s_1)} \ldots \cup_{s_n\in C(s_1,\ldots,s_{n-1})} N \cup N_{s_1} \cup N_{s_1 s_2} \cup \ldots \cup N_{s_1\ldots s_n}$$

which is μ-measure zero by (9.6) and (9.9). Moreover, for each $x \in \mathbb{R}^d \backslash N_0$ and $n \ge 1$, 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

$$f_n(x) = \sum_{s_1\in C}\sum_{s_2\in C(s_1)} \cdots \sum_{s_n\in C(s_1,\ldots,s_{n-1})} f(s_n) 1\{x \in Q_{s_1\ldots s_n}\}.$$

We see that

$$\lim_{n\nearrow\infty} f_n(x) = f(x) \quad \text{for all } x \in S\backslash N_0, \tag{9.16}$$

since $|x - s_n| \le 2^{-n}\sqrt{d}$ if $x \in Q_{s_1\ldots s_n}$. Therefore,

$$\begin{aligned}
Ef(\varphi(U)) &= \lim_{n\nearrow\infty} Ef(\varphi_n(U)) \quad \text{by definition of } \varphi, \\
&= \lim_{n\nearrow\infty} \sum_{s_1\in C}\sum_{s_2\in C(s_1)} \cdots \sum_{s_n\in C(s_1,\ldots,s_{n-1})} f(s_n)|I_{s_1,\ldots,s_n}| \quad \text{by definition of } \varphi_n, \\
&= \lim_{n\nearrow\infty} \sum_{s_1\in C}\sum_{s_2\in C(s_1)} \cdots \sum_{s_n\in C(s_1,\ldots,s_{n-1})} f(s_n)\mu(Q_{s_1,\ldots,s_n}) \quad \text{by (9.23)}, \\
&= \lim_{n\nearrow\infty} \int f_n d\mu \quad \text{by definition of } f_n, \\
&= \int f d\mu \quad \text{by (9.16)}.
\end{aligned}$$

This proves (9.5) (cf. Lemma 2.1.2). □

Proof of Lemma 9.2.3 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

$$\begin{aligned}
Q_{s_1} &\ni s_1 \quad \text{for all } s_1 \in C, \\
\mu(N) &= 0, \quad \text{where } N \stackrel{\text{def.}}{=} S\backslash \cup_{s_1\in C} Q_{s_1}, \\
\mathrm{diam}(Q_{s_1}) &\le 2^{-1}, \\
\mu(Q_{s_1}) &> 0,
\end{aligned} \tag{9.17}$$

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 $\mathrm{diam}(\underline{B}_n) \leq 2^{-1}$. 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

$$\begin{aligned}
Q_{s_1} &\supset Q_{s_1 s_2} \ni s_2 \quad \text{for all } s_2 \in C(s_1).\\
\mu(N_{s_1}) &= 0, \quad \text{where } N_{s_1} \stackrel{\text{def.}}{=} Q_{s_1} \setminus \cup_{s_2\in C(s_1)} Q_{s_1,s_2},\\
\mathrm{diam}(Q_{s_1 s_2}) &\leq 2^{-2},\\
\mu(Q_{s_1 s_2}) &> 0.
\end{aligned}$$

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 $s_1 \in C$, ..., $s_n \in C(s_1\cdots s_{n-1})$. 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

$$Q_{s_1\cdots s_n} \supset Q_{s_1\cdots s_{n+1}} \ni s_{n+1} \quad \text{for all } s_{n+1} \in C(s_1,\ldots,s_n). \tag{9.18}$$

$$\mu(N_{s_1\cdots s_n}) = 0, \quad \text{where } N_{s_1\cdots s_n} \stackrel{\text{def.}}{=} Q_{s_1\cdots s_n} \setminus \cup_{s_{n+1}\in C(s_1,\ldots,s_n)} Q_{s_1\cdots s_{n+1}}, \tag{9.19}$$

$$\mathrm{diam}(Q_{s_1\cdots s_n}) \leq 2^{-n}, \tag{9.20}$$

$$\mu(Q_{s_1\cdots s_{n+1}}) > 0 \quad \text{for all } s_{n+1} \in C(s_1,\ldots,s_n).$$

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

$$[0,1) = \cup_{s_1\in C} I_{s_1}, \tag{9.21}$$

$$I_{s_1\cdots s_{n-1}} = \cup_{s_n\in C(s_1,\ldots,s_{n-1})} I_{s_1\cdots s_n}, \tag{9.22}$$

$$|I_{s_1\cdots s_n}| = \mu(Q_{s_1\cdots s_n}). \tag{9.23}$$

Step 3: We now define $\varphi_n : [0,1) \to S$ by

$$\varphi_n(\theta) = s_n \quad \text{if } \theta \in I_{s_1\cdots s_n}. \tag{9.24}$$

Let us check the following;

$$\varphi_n : [0,1) \to S \text{ is well defined and measurable for all } n \geq 1. \tag{9.25}$$

$$(\varphi_n(\theta))_{n\geq 1} \text{ is a Cauchy sequence for for all } \theta \in [0,1). \tag{9.26}$$

By (9.21) and (9.22), 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 (9.26), 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 (9.20) that

$$\text{dist.}(\varphi_{m+n}(\theta), \varphi_n(\theta)) \le 2^{-n}.$$

Step 4: By, (9.26) and (9.26), we can define a measurable map $\varphi : [0,1) \to S$ by $\varphi(\theta) = \lim_{n\nearrow\infty} \varphi_n(\theta)$ for all $\theta \in [0,1)$. Let us see that φ satisfies (9.5). To do so, take $f \in C_{\rm b}(S)$ and define a set

$$N_0 = \cup_{n\ge1} \cup_{s_1\in C} \cup_{s_2\in C(s_1)} \ldots \cup_{s_n\in C(s_1,\ldots,s_{n-1})} N \cup N_{s_1} \cup N_{s_1 s_2} \cup \ldots \cup N_{s_1\ldots s_n}$$

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 (9.17) and (9.19). Moreover, we see from (9.20) that

$$\lim_{n\nearrow\infty} f_n(x) = f(x) \quad \text{for all } x \in S\backslash N_0. \tag{9.27}$$

Therefore,

$$\begin{aligned}
Ef(\varphi(U)) &= \lim_{n\nearrow\infty} Ef(\varphi_n(U)) \quad \text{by definition of } \varphi, \\
&= \lim_{n\nearrow\infty} \sum_{s_1\in C} \sum_{s_2\in C(s_1)} \cdots \sum_{s_n\in C(s_1,\ldots,s_{n-1})} f(s_n)|I_{s_1,\ldots,s_n}| \quad \text{by definition of } \varphi_n, \\
&= \lim_{n\nearrow\infty} \sum_{s_1\in C} \sum_{s_2\in C(s_1)} \cdots \sum_{s_n\in C(s_1,\ldots,s_{n-1})} f(s_n)\mu(Q_{s_1,\ldots,s_n}) \quad \text{by (9.23)}, \\
&= \lim_{n\nearrow\infty} \int f_n d\mu \quad \text{by definition of } f_n, \\
&= \int f d\mu \quad \text{by (9.27)}.
\end{aligned}$$

This proves (9.5) (cf. Lemma 2.1.2). □

9.4 Complement to section 3.3

Lemma 9.4.1 *Suppose that $q \ge 2$ is an integer and that $V = \sum_{k\ge1} q^{-k} Y_k$, where $\{Y_k\}_{k\ge1}$ 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) $\Longleftrightarrow$ (b)";*

a1) *$\{Y_k\}_{k\ge1}$ 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, $(X_n)_{n\ge 1}$ in Example 3.3.1 and $(Y_n)_{n\ge 1}$ 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

$$\cup_{n\ge 1} \cup_{0\le s\le q^n-1} \{V = sq^{-n}\},$$

and therefore for P-almost all $\omega \in \Omega$, $Y_k(\omega)$ is uniqly determied as the k^{th} digit of the q-adic expansion of the number $V(\omega)$. We therefore see from (3.3) that $(X_n)_{n\ge 1}$ in Example 3.3.1 and $(Y_n)_{n\ge 1}$ have the same distribution, which proves (a1) & (a2). $\Box$

Exercise 9.4.1 Check an alternative proof of Lemma 9.4.1, (a1) & (a2) $\Rightarrow$ (b) presented below. It is enough to prove that for any $t \in [0,1)$

$$P\{V \le t\} = t \tag{9.28}$$

(cf. Exercise 2.1.1). 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 (9.28) by $f(s_1,s_2,\ldots)$. We will prove that

$$f(s_1,s_2,\ldots) = q^{-1}s_1 + q^{-1}f(s_2,s_3,\ldots). \tag{9.29}$$

We have that

$$\begin{aligned} \{U \le t\} &= \{Y_1 < s_1\} \cup \left\{ Y_1 = s_1,\ \sum_{k=2}^{\infty} q^{-k}Y_k \le \sum_{k=2}^{\infty} q^{-k}s_k \right\} \\ &= \{Y_1 < s_1\} \cup \left\{ Y_1 = s_1,\ \sum_{k=1}^{\infty} q^{-k}Y_{k+1} \le \sum_{k=1}^{\infty} q^{-k}s_{k+1} \right\}. \end{aligned} \tag{9.30}$$

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 (9.30) imply that

$$\begin{aligned} P\{V \le t\} &= P\{Y_1 < s_1\} + P\{Y_1 = s_1\}P\left\{ \sum_{k=1}^{\infty} q^{-k}Y_{k+1} \le \sum_{k=1}^{\infty} q^{-k}s_{k+1} \right\} \quad \text{by (i)} \\ &= s_1 q^{-1} + q^{-1}P\left\{ \sum_{k=1}^{\infty} q^{-k}Y_k \le \sum_{k=1}^{\infty} q^{-k}s_{k+1} \right\} \quad \text{by (ii)} \\ &= s_1 q^{-1} + q^{-1}f(s_2,s_3,\ldots), \end{aligned} \tag{9.31}$$

which proves (9.29).

With (9.29) in hand, proof of (9.28) is easy. In fact, we have for any $n = 1,2,\ldots$

$$f(s_1,s_2,\ldots) = \sum_{k=1}^{n} q^{-k}s_k + q^{-n}f(s_{n+1},s_{n+2},\ldots) \tag{9.32}$$

by induction. Then (9.29) follows by letting $n \nearrow \infty$. $\Box$

9.5 Convolution

Definition 9.5.1 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

$$(\mu_1 * \cdots * \mu_n)(B) = \left(\otimes_{j=1}^n \mu_j\right) \left\{(x_j)_{j=1}^n \in (\mathbb{R}^d)^n \ ; \ x_1 + \ldots + x_n \in B\right\}, \quad B \in \mathcal{B}(\mathbb{R}^d). \quad (9.33)$$

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 (2.9) that

$$P(X_1 + \ldots + X_n \in \cdot) = \mu_1 * \cdots * \mu_n. \quad (9.34)$$

Lemma 9.5.2 (a) *For Borel finite measures μ_1, μ_2 on $\mathbb{R}^d$,*

$$(\mu_1 * \mu_2)^\wedge(\theta) = \widehat{\mu_1}(\theta)\widehat{\mu_2}(\theta) \quad \text{for all } \theta \in \mathbb{R}^d. \quad (9.35)$$

(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*

$$(f_1 * f_2)(x) = \int f_1(x - y) f_2(y) dy \quad (9.36)$$

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*

$$(f_1 * f_2)(x) = \sum_{y \in \mathbb{Z}^d} f_1(x - y) f_2(y) dy. \quad (9.37)$$

Proof: It is easy to see (9.35). (9.36) can be seen as follows;

$$\begin{aligned} \mu_1 * \mu_2(B) &= \int \mu_1 \otimes \mu_2(dzdy) 1_B(z + y) \\ &= \int f_1(z) f_2(y) dzdy 1_B(z + y) \\ &= \int f_1(x - y) f_2(y) dxdy 1_B(x) \\ &= \int_B (f_1 * f_2)(x) dx. \end{aligned} \quad (9.38)$$

The proof of (9.37) is similar to that of (9.36). □

Example 9.5.3 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 5.3.3,

$$P(\chi_1 + \chi_2 \in \cdot) = \nu_{V_1} * \nu_{V_2} = \nu_{V_1 + V_2}. \quad (9.39)$$

Example 9.5.4 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 (3.1),

$$P(X + Y \in \cdot) = \gamma_{r,a} * \gamma_{r,b} = \gamma_{r,a+b}. \quad (9.40)$$

Example 9.5.5 Then, by Exercise 5.3.9,

$$P(N_1 + N_2 \in \cdot) = \pi_{r_1} * \pi_{r_2} = \pi_{r_1+r_2}. \tag{9.41}$$

Exercise 9.5.1 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.,

$$P\{U_1 + U_2 \in B\} = \int_B v(t)dt, \tag{9.42}$$

where

$$v(t) = (u * u)(t) = \frac{t - 2a}{(b - a)^2} 1_{[2a,a+b]}(t) + \frac{2b - t}{(b - a)^2} 1_{[a+b,2b]}(t).$$

Then, conclude from (5.8) and (9.42) that

$$\widehat{v}(\theta) = \widehat{u}(\theta)^2 = \left(\frac{\mathbf{e}(\theta b)-\mathbf{e}(\theta a)}{(b-a)\theta}\right)^2. \tag{9.43}$$

Exercise 9.5.2 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. (1.22)). Suppose also that $\{N, X_1, X_2, \ldots\}$ are independent. Prove then that

$$P\{X_1 + \ldots + X_N \in \cdot\} = \sum_{n \geq 1} e^{-r} r^n (\mu_1 * \cdots * \mu_n)/n! \tag{9.44}$$

The distribution on the right-hand-side of (9.44) is called the *compound Poisson distribution*. Poisson distribution is a compound Poisson distribution with $X_j \equiv 1$.

9.6 Independent families of random variables

Definition 9.6.1 a) Independent events: Suppose that $\mathcal{A} \subset \mathcal{F}$. Then, $\mathcal{A}$ said to be *independent*, if

$$P\left(\cap_{A \in \mathcal{A}_0} A\right) = \prod_{A \in \mathcal{A}_0} P(A) \quad \text{for any finite subset } \mathcal{A}_0 \text{ in } \mathcal{A}. \tag{9.45}$$

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*[19], if

$$\{A_\lambda\}_{\lambda \in \Lambda} \subset \mathcal{F} \text{ is independent in the sense of (a) for any } A_\lambda \in \mathcal{A}_\lambda \ (\lambda \in \Lambda). \tag{9.46}$$

The families $\{\mathcal{A}_\lambda\}_{\lambda \in \Lambda}$ are said to be *independent* if the σ-fields $\{\sigma[\mathcal{A}_\lambda]\}_{\lambda \in \Lambda}$ are quasi-independent.

Remark: The condition (9.46) does not imply that $\{\sigma[\mathcal{A}_\lambda]\}_{\lambda \in \Lambda}$ are independent σ-fields (cf. Exercise 9.6.2). 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-fields, then the notion of indpendence and quasi-indpendence coincide. See also Proposition 9.7.1 below.

[19]This terminology does not appear in standard text books in probability theory. It is introduced by the author of this notes for the convenience.

Exercise 9.6.1 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-fields.

Exercise 9.6.2 In the setting of Definition 9.6.1(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 9.6.1 (a).
ii) $\mathcal{A}_1 = \{A_1\}$ and $\mathcal{A}_{23} = \{A_2, A_3\}$ are quasi-independent in the sense of Definition 9.6.1 (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 (cf. Proposition 9.7.1).

Remark: In Exercise 9.6.2, $P(B|A_1) = P(B)$ 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-field. cf. Lemma 2.1.1.

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 9.6.2 A σ-field:

$$\sigma\left[X_\lambda^{-1}(B_\lambda) \ ; \ B_\lambda \in \mathcal{B}_\lambda, \ \lambda \in \Lambda\right] \tag{9.47}$$

is called the σ-field 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-field $\sigma\left[\{X_\lambda\}_{\lambda \in \Lambda}\right]$ (cf. (9.47)) 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 9.6.3 *For a disjoint decomposition $\Lambda = \cup_{\gamma \in \Gamma} \Lambda(\gamma)$ of the index set Λ, the following conditions are equivalent:*

a) *The σ-fields*

$$\sigma[X_\lambda \ ; \ \lambda \in \Lambda(\gamma)], \quad \gamma \in \Gamma$$

are independent (cf. Definition 9.6.1(b)).

b) *R.v.'s $\{\widetilde{X}\}_{\gamma \in \Gamma}$ defined by*

$$\widetilde{X}_\gamma : \omega \mapsto (X_\lambda(\omega))_{\lambda \in \Lambda(\gamma)} \in \prod_{\lambda \in \Lambda(\gamma)} S_\lambda, \quad \gamma \in \Gamma. \tag{9.48}$$

are independent.

Definition 9.6.4 Families of r.v.'s

$$\{X_\lambda \ ; \ \lambda \in \Lambda(\gamma)\}, \quad \gamma \in \Gamma \tag{9.49}$$

in Proposition 9.6.3 are said to be *independent* if they satisfy one of (therefore all of) conditions in the corollary.

Proof of Proposition 9.6.3: The equivalence is a consequence of Proposition 2.3.1 and an identity $\sigma[\widetilde{X}_\gamma] = \sigma[X_\lambda \ ; \ \lambda \in \Lambda(\gamma)]$, which can be seen from Exercise 2.2.2. □

Remarks:

1) The independence of the families of r.v.'s (Definition 9.6.4) can be considered as is a special case of the independence of r.v.'s (Proposition 2.3.1), if we consider r.v.'s $\{\widetilde{X}\}_{\gamma \in \Gamma}$ defined by (9.48).

2) In the setting of Proposition 9.6.3, let us consider the following condition:

$$\{X_{\lambda(\gamma)}\}_{\gamma \in \Gamma} \text{ are independent r.v.'s for any choice of } \lambda(\gamma) \in \Lambda(\gamma) \ (\gamma \in \Gamma). \tag{9.50}$$

This condition follows from the independence of the families (9.49). However, the converse is not true. A counterexample is again provided by Exercise 9.6.2. Consider $\{1_{A_1}\}$ and $\{1_{A_2}, 1_{A_3}\}$ there. Since, $\{A_i\}_{i=1}^3$ are pairwise independent, we have (9.50) by Exercise 9.6.1. However, $\{1_{A_1}\}$ and $\{1_{A_2}, 1_{A_3}\}$ are not independent, since $\sigma[\{A_1, A_2\}] = 2^\Omega$.

Exercise 9.6.3 Suppose that $(X_n)_{n \geq 1}$ are $\mathbb{R}^d$-valued independent r.v.'s and let $S_n = X_1 + \cdots + X_n$. Prove then that, for each fixed $m \geq 1$, two families of r.v.'s

$$\{S_n\}_{n=1}^m, \quad \{S_{n+m} - S_m\}_{n \geq 1}$$

are independent. Hint: Note that $\sigma(\{S_n\}_{n=1}^m) = \sigma(\{X_n\}_{n=1}^m)$ and that $\sigma(\{S_{n+m} - S_m\}_{n \geq 1}) = \sigma(\{X_{n+m}\}_{n \geq 1})$. Then, use Exercise 2.3.9.

9.7 Kolmogorov's 0-1 law

Results in this subsection will not be used in the other part of this notes.[20] We however include this subsection since Kolmogorov's 0-1 law is an important aspect of independent r.v.'s. We start with the following abstract statement.

Proposition 9.7.1 *Suppose that $\{\mathcal{A}_\lambda\}_{\lambda \in \Lambda}$ are quasi-independent π-systems on $(\Omega, \mathcal{F}, P)$ indexed by a set Λ. Then, $\{\sigma[\mathcal{A}_\lambda]\}_{\lambda \in \Lambda}$ are independent σ-fields.*

Proof: We may assume that $\Omega \in \mathcal{A}_\lambda$ for all λ. We have to prove that $\{\sigma[\mathcal{A}_{\lambda_j}]\}_{j=1}^n$ are independent for any $n \geq 1$ and for any distinct $\lambda_1, \ldots, \lambda_n \in \Lambda$. We set $\mathcal{A}_j = \mathcal{A}_{\lambda_j}$ for simplicity. We begin by proving that

$$\text{if } \{\mathcal{A}_j\}_{j=1}^n \text{ are independent } \pi\text{-systems, then so } \{\sigma[\mathcal{A}_1], \mathcal{A}_2, \ldots \mathcal{A}_n\} \text{ are.} \tag{9.51}$$

To show (9.51), we fix $A_j \in \mathcal{A}_j$ $(2 \leq n \leq n)$ arbitrarily and prove by Lemma 2.1.1 that

$$P\left(A \cap \left(\cap_{j=2}^n A_j\right)\right) = P(A)\, P\left(\cap_{j=2}^n A_j\right) \quad A \in \sigma[\mathcal{A}_1].$$

[20]The reader can skip this subsection for this reason.

This, together with the identity $P\left(\cap_{j=2}^n A_j\right) = \prod_{j=2}^n P(A_j)$, implies (9.51). Once (9.51) is obtained, then the proof of the proposition can be finished by induction. Since $\{\sigma[\mathcal{A}_1], \mathcal{A}_2, \ldots, \mathcal{A}_n\}$ are independent π-system again, we can apply (9.51) to see that $\{\sigma[\mathcal{A}_1], \sigma[\mathcal{A}_2], \mathcal{A}_3, \ldots, \mathcal{A}_n\}$ are independent π-system. By proceeding inductively, we conclude that $\{\sigma[\mathcal{A}_j]\}_{j=1}^n$ are independent. □

We have following corollaries to Proposition 9.7.1, first of which is in fact even stronger than the Proposition.

Corollary 9.7.2 *Let $\{\mathcal{A}_\lambda\}_{\lambda\in\Lambda}$ be quasi-independent π-systems on $(\Omega, \mathcal{F}, P)$. Suppose that the set Λ of indices is decomposed into a disjoint union as $\Lambda = \cup_{\gamma\in\Gamma}\Lambda(\gamma)$. Then the σ-fields*

$$\sigma\left[\mathcal{A}_\lambda\ ;\ \lambda \in \Lambda(\gamma)\right], \quad \gamma \in \Gamma$$

are independent.

Proof: For each $\gamma \in \Gamma$, we define $\widetilde{\mathcal{A}}_\gamma$ as a collection of sets of the form $\cap_{\lambda\in I} A_\lambda$ where I is a finite subset of $\Lambda(\gamma)$ and $A_\lambda \in \mathcal{A}_\lambda$ for $\lambda \in I$. It is then easy to verify that

$$\sigma\left[\mathcal{A}_\lambda\ ;\ \lambda \in \Lambda(\gamma)\right] = \sigma\left[\widetilde{\mathcal{A}}_\gamma\right]. \tag{9.52}$$

We will prove that

$$\widetilde{\mathcal{A}}_\gamma \text{ is a } \pi\text{-system for each } \gamma \in \Gamma, \tag{9.53}$$

and that

$$\{\widetilde{\mathcal{A}}_\gamma\}_{\gamma\in\Gamma} \text{ are independent.} \tag{9.54}$$

The corollary follows from (2.1), (9.53), (9.54) and Proposition 9.7.1.

To prove (9.53), take two sets $A = \cap_{\lambda\in I} A_\lambda$ and $A' = \cap_{\lambda\in I'} A'_\lambda$ from $\widetilde{\mathcal{A}}_\gamma$. We then have that

$$\begin{aligned} A \cap A' &= \left(\cap_{\lambda\in I} A_\lambda\right) \cap \left(\cap_{\lambda\in I'} A'_\lambda\right) \\ &= \cap_{\lambda\in I\cup I'} A''_\lambda, \end{aligned}$$

where

$$A''_\lambda = \begin{cases} A_\lambda & \text{if } \lambda \in I \backslash I' \\ A'_\lambda & \text{if } \lambda \in I' \backslash I \\ A_\lambda \cap A'_\lambda & \text{if } \lambda \in I \cap I' \end{cases}$$

Since each $\mathcal{A}_\lambda$ is a π-system, $A''_\lambda \in \mathcal{A}_\lambda$ and therefore, $A \cap A' \in \widetilde{\mathcal{A}}_\gamma$.

To prove (9.54), we have to show that $\{A_j\}_{j=1}^n$ are independent for any $n \geq 1$, for any distinct $\gamma_1, \ldots, \gamma_n \in \Gamma$ and for any $A_j \in \widetilde{\mathcal{A}}_{\gamma_j}$. We write $A_j = \cap_{\lambda\in I(j)} A_\lambda$ with $I(j) \subset \Lambda(\gamma_j)$. Since $\{A_\lambda\ ;\ \lambda \in \cup_{j=1}^n I(j)\}$ are independent events, we have

$$\begin{aligned} P\left(\cap_{j=1}^n A_j\right) &= P\left(\cap_{j=1}^n \cap_{\lambda\in I(j)} A_\lambda\right) \\ &= \prod_{j=1}^n \prod_{\lambda\in I(j)} P\left(A_\lambda\right) \\ &= \prod_{j=1}^n P\left(\cap_{\lambda\in I(j)} A_\lambda\right) \\ &= \prod_{j=1}^n P\left(A_j\right), \end{aligned}$$

which proves (9.54). □

Corollary 9.7.3 *Suppose that $\mathcal{A}_1, \mathcal{A}_2, \ldots$ are quasi-independent π-systems on $(\Omega, \mathcal{F}, P)$ and that $A \in \cap_{n\geq 1}\sigma[\mathcal{A}_m \ ; \ m \geq n]$. Then, $P(A) = 0$ or 1.*

Proof: We define $\mathcal{F}_n = \sigma[\mathcal{A}_m \ ; \ m \leq n]$ and $\mathcal{T}_n = \sigma[\mathcal{A}_m \ ; \ m \geq n]$. Then, by Corollary 9.7.2, $\mathcal{F}_n$ and $\mathcal{T}_{n+1}$ are independent for all $n \geq 1$. From this, it is easy to see that $\cup_{n\geq 1}\mathcal{F}_n$ and $\cap_{n\geq 1}\mathcal{T}_n$ are independent. Since each of $\cup_{n\geq 1}\mathcal{F}_n$ and $\cap_{n\geq 1}\mathcal{T}_n$ is a π-system, we conclude from Proposition 9.7.1 that

$$\sigma[\cup_{n\geq 1}\mathcal{F}_n] \text{ and } \cap_{n\geq 1}\mathcal{T}_n \text{ are independent.} \tag{9.55}$$

On the other hand, it is easy to see that

$$\sigma[\cup_{n\geq 1}\mathcal{F}_n] \supset \cap_{n\geq 1}\mathcal{T}_n. \tag{9.56}$$

For $A \in \cap_{n\geq 1}\mathcal{T}_n$, we see from (9.55) and (9.56) that $A \in \cap_{n\geq 1}\mathcal{T}_n$ and $A \in \sigma[\cup_{n\geq 1}\mathcal{F}_n]$ are independent. Therefore, $P(A) = P(A \cap A) = P(A)^2$, and hence $P(A) = 0$ or 1. □

Exercise 9.7.1 Suppose that r.v.'s X_n indexed by $n = 1, 2, \ldots$ are independent, where each X_n takes values in some measurable space $(S_n, \mathcal{B}_n)$. Then, apply Corollary 9.7.3 with $\mathcal{A}_n = \sigma[X_n]$ to prove that $P(A) = 0$ or 1 for all A in a σ-field defined by;

$$\mathcal{T} = \cap_{n\geq 1}\sigma[\{X_m\}_{m\geq n}]. \tag{9.57}$$

This result is known as *Kolmogorov's 0-1 law.* The σ-field $\mathcal{T}$ above is called the *tail σ-field* of the r.v.'s $\{X_n\}_{n\geq 1}$.

Exercise 9.7.2 Apply Corollary 9.7.2 to prove that the families of r.v.'s (9.49) are independent if r.v.'s $\{X_\lambda\}_{\lambda\in\Lambda}$ are independent. Hint: What we assume is that $\{\sigma(X_\lambda)\}_{\lambda\in\Lambda}$ are independent σ-fields.

10 Appendix to Section 4

10.1 Reflection principle and its applications

Reflection principle (Exercise 10.1.2) is a important tool to study symmetric, nearest neighbor random walk in $\mathbb{Z}$. In this subsection, we will forcus on the reflection principle and its applications in a series of exercises.

Exercise 10.1.1 Let $\{S_n\}_{n\geq 1}$ be a symmetric random walk (cf. Exercise 4.3.2)and define $\{S_n^{(m)}\ ;\ m,n \geq 0\}$ 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 10.1.2 *Reflection principle;* Consider a $\mathbb{Z}$-valued random walk such that $P\{X_1 = 0\} = r$ and $P\{X_1 = \pm 1\} = \frac{1-r}{2}$. For positive integers x and y and n, prove that

$$P\left(x + S_n = y, \tau < n\right) = P(x + S_n = -y) = P(-x + S_n = y), \tag{10.1}$$

where $\tau = \inf\{n \geq 1\ ;\ x + S_n = 0\}$. Hint: The second equality is easy. To prove the first one, observe the following;

- By Exercise 10.1.1, $(S_n)_{n\geq 0}$ and $(S_n^{(m)})_{n\geq 0}$ have the same distribution for each $m \geq 0$.
- If $1 \leq m < n$, then $\{x + S_n^{(m)} = y, \tau = m\} = \{x + S_n = -y, \tau = m\}$.
- $\{x + S_n = -y\} \subset \{1 \leq \tau < n\}$, since $x + S_0 = x > 0$ and $x + S_n = -y < 0$.

We see from these observations that

$$\begin{aligned}
P\left(x + S_n = y, \tau < n\right) &= \sum_{m=1}^{n-1} P\left(x + S_n = y, \tau = m\right) \\
&= \sum_{m=1}^{n-1} P\left(x + S_n^{(m)} = y, \tau = m\right) \\
&= \sum_{m=1}^{n-1} P\left(x + S_n = -y, \tau = m\right) \\
&= P\left(x + S_n = -y\right).
\end{aligned}$$

Exercise 10.1.3 Consider a $\mathbb{Z}$-valued random walk such that $P\{X_1 = 0\} = r$ and $P\{X_1 = \pm 1\} = \frac{1-r}{2}$. For integers $x \geq 0$, $y \geq 0$ and $n \geq 1$, prove that

$$P(S_1 \geq -x, \ldots, S_{n-1} \geq -x, S_n = y - x) = P(S_n = y - x) - P(S_n = y + x + 2). \tag{10.2}$$

Then, use (10.2) to show that

$$\begin{aligned}
P(S_1 \geq 0, \ldots, S_n \geq 0) &= P(S_n \in \{0, 1\}) && (10.3) \\
&= \tfrac{1-2r}{1-r} P(S_n = 0) + \tfrac{1}{1-r} P(S_{n+1} = 0). && (10.4)
\end{aligned}$$

Hint: The equality (10.2) can be seen as follows;

$$\begin{aligned}
&P(S_1 \geq -x, \ldots, S_{n-1} \geq -x, S_n = y - x) \\
&= P(1 + x + S_1 > 0, \ldots, 1 + x + S_{n-1} > 0, 1 + x + S_n = 1 + y) \\
&= P(1 + x + S_n = 1 + y) - P(-1 - x + S_n = 1 + y) \quad \text{by (10.1)}.
\end{aligned}$$

The equality (10.3) follows from (10.2). The equality (10.4) is easy.

Exercise 10.1.4 Consider a $\mathbb{Z}$-valued random walk such that $P\{X_1 = 0\} = r$ and $P\{X_1 = \pm 1\} = \frac{1-r}{2}$. For positive integers n and x, prove that

$$\begin{aligned} P\{S_1 < x, \dots, S_{n-1} < x, S_n = x\} &= P\{S_1 > -x, \dots, S_{n-1} > -x, S_n = -x\} \\ &= \tfrac{1-r}{2} P\{S_1 \geq 1-x, \dots, S_{n-2} \geq 1-x, S_{n-1} = 1-x\} \\ &= \tfrac{1-r}{2} \left(P\{S_{n-1} = 1-x\} - P\{S_{n-1} = 1+x\}\right). \qquad (10.5) \end{aligned}$$

Hint: The first equality comes from th symmetry. The second one can be seen by using the independence of $\{S_k\}_{k=1}^{n-1}$ and $\{X_n\}$. The last equality follows from (10.2).

Exercise 10.1.5 Consider a $\mathbb{Z}$-valued random walk such that $P\{X_1 = 0\} = r$ and $P\{X_1 = \pm 1\} = \frac{1-r}{2}$. For integers $n \geq 1$ and $y \geq 0$, prove that

$$\begin{aligned} P(S_1 > 0, \dots, S_{n-1} > 0, S_n = 1+y) &= \tfrac{1-r}{2} P(S_1 \geq 0, \dots, S_{n-2} \geq 0, S_{n-1} = y) \qquad (10.6) \\ &= \tfrac{1-r}{2} \left(P\{S_{n-1} = y\} - P\{S_{n-1} = y+2\}\right), \quad (10.7) \\ P(S_1 \neq 0, \dots, S_n \neq 0) &= 2P(S_1 > 0, \dots, S_n > 0) \\ &= (1-r)P(S_1 \geq 0, \dots, S_{n-1} \geq 0) \\ &= (1-r)P(S_{n-1} \in \{0,1\}) \\ &= (1-2r)P(S_{n-1} = 0) + P(S_n = 0). \end{aligned}$$

Hint: The equalities (10.6) and (10.7) can be seen as follows;

$$\begin{aligned} & 2P(S_1 > 0, \dots, S_{n-1} > 0, S_n = 1+y) \\ = \; & 2P(S_1 = 1, 1+S_2-S_1 > 0, \dots, 1+S_{n-1}-S_1 > 0, 1+S_n-S_1 = 1+y) \\ = \; & 2P(S_1 = 1)P(1+S_2-S_1 > 0, \dots, 1+S_{n-1}-S_1 > 0, 1+S_n-S_1 = 1+y) \\ = \; & (1-r)P(1+S_1 > 0, \dots, 1+S_{n-2} > 0, 1+S_{n-1} = 1+y) \\ = \; & (1-r)P(S_1 \geq 0, \dots, S_{n-2} \geq 0, S_{n-1} = y) \\ = \; & (1-r)\left(P\{S_{n-1} = y\} - P\{S_{n-1} = y+2\}\right) \quad \text{by (10.2)} \end{aligned}$$

Exercise 10.1.6 Consider a $\mathbb{Z}$-valued random walk such that $P\{X_1 = 0\} = r$ and $P\{X_1 = \pm 1\} = \frac{1-r}{2}$. Prove that $E[S_n 1_{\Omega_n}] = \frac{1-r}{2}$ for any $n \geq 1$, where $\Omega_n = \{S_1 > 0, \dots, S_n > 0\}$. Hint: Note first that $E[(x+X_1)1\{x+X_1 > 0\}] = x$ for any positive integer x. We then see that

$$\begin{aligned} E[S_n 1_{\Omega_n}] &= \sum_{x \geq 1} E[1_{\Omega_{n-1}} 1\{S_{n-1} = x, x+X_n > 0\}(x+X_n)] \\ &= \sum_{x \geq 1} P(\Omega_{n-1} \cap \{S_{n-1} = x\}) E[(x+X_n)1\{x+X_n > 0\}] \\ &= \sum_{x \geq 1} P(\Omega_{n-1} \cap \{S_{n-1} = x\}) x \\ &= E[S_{n-1} 1_{\Omega_{n-1}}]. \end{aligned}$$

10.2 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) $$\sum_{n\ge 1} 1\{X_n \neq Y_n\} < \infty \text{ a.s.}$$

This can be seen as follows;

$$\begin{aligned}
E\sum_{n\ge 1} 1\{X_n \neq Y_n\} &\overset{\text{Fubini}}{=} \sum_{n\ge 1} P\{X_n \neq Y_n\} \\
&\le \sum_{n\ge 1} P\{X_n > n\} = \sum_{n\ge 1} P\{X_1 > n\} \\
&\le \sum_{n\ge 1} \int_{n-1}^{n} dt P\{X_1 > t\} = \int_0^\infty dt P\{X_1 > t\} \\
&\overset{(1.11)}{=} EX_1 < \infty,
\end{aligned}$$

which in particular implies (1).

We see from (1) that Theorem 4.1.2 follows from:

2) $$\lim_{n\nearrow\infty} \frac{T_n}{n} = E[X_1] \text{ a.s.}$$

We first prove (2) along the subsequence $l(n) = \lfloor q^n \rfloor$, where $q > 1$:

3) $$\lim_{n\nearrow\infty} \frac{T_{l(n)}}{l(n)} = E[X_1] \text{ a.s.}$$

Since

$$EY_n = EX_n 1\{X_n \le n\} = EX_1 1\{X_1 \le n\} \to EX_1,$$

we have

$$\lim_{n\nearrow\infty} \frac{E[T_n]}{n} = EX_1.$$

Thus, (3) follows from:

4) $$\lim_{n\nearrow\infty} \frac{T_{l(n)} - E[T_{l(n)}]}{l(n)} = E[X_1] \text{ a.s.}$$

To show (4), we prepare the following estimate:

5) $$\text{var}(T_n) \le nE[X_1^2 \mathbf{1}\{X_1 \le n\}]$$

Indeed,

$$\begin{aligned}
\text{var}(T_n) &\overset{(2.11)}{=} \sum_{j=1}^{n} \text{var}(Y_j) \le \sum_{j=1}^{n} E[Y_j^2] \\
&= \sum_{j=1}^{n} E[X_1^2 \mathbf{1}\{X_1 \le j\}] \le nE[X_1^2 \mathbf{1}\{X_1 \le n\}].
\end{aligned}$$

We next observe that

6) $$\sum_{n:l(n)\geq x} \frac{1}{l(n)} \leq \frac{2q}{(q-1)x} \text{ for any } x > 0.$$

In fact, let M be the smallest $n \in \mathbb{N}$ such that $l(n) \geq x$. Then, $q^M \geq x$. Note also that $l(n) \geq q^n/2$ for all $n \in \mathbb{N}$. Thus,

$$\sum_{n:l(n)\geq x} \frac{1}{l(n)} \leq 2 \sum_{n\geq M} q^{-n} = 2q^{-M} \sum_{n\geq 0} q^{-n} \leq \frac{2q}{(q-1)x}.$$

With (5) and (6), we proceed as follows:

$$\begin{aligned} E\sum_{n\geq 1} \left| \frac{T_{l(n)} - E[T_{l(n)}]}{l(n)} \right|^2 &= \sum_{n\geq 1} l(n)^{-2} \text{var}(S_{l(n)}) \overset{(5)}{\leq} E\left[X_1^2 \sum_{n\geq 1} l(n)^{-1} \mathbf{1}\{X_1 \leq n\} \right] \\ &\overset{(6)}{\leq} \frac{2q}{q-1} E[X_1] < \infty. \end{aligned}$$

This implies that $\sum_{n\geq 1} \left| \frac{T_{l(n)} - E[T_{l(n)}]}{l(n)} \right|^2 < \infty$, 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

$$l(k) \leq n < l(k+1).$$

We have by the positivity of $\{X_m\}$ that

$$l(k+1)^{-1} T_{l(k)} \leq n^{-1} T_n \leq l(k)^{-1} T_{l(k+1)}.$$

By letting $n \nearrow \infty$, we see from (3) that

$$q^{-1} EX_1 \leq \varliminf_{n\nearrow\infty} n^{-1} T_n \leq \varlimsup_{n\nearrow\infty} n^{-1} T_n, \leq qEX_1,$$

which conclude the proof, since $q > 1$ is arbitrary. □

11 Appendix to Sections 5

11.1 Proof of Lemma 5.1.4

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) $\qquad \int_{\mathbb{R}^d} |h_t * f - f|(x)dx \le \int_{\mathbb{R}^d} h_1(y)g_t(y)dy \quad \text{where} \quad g_t(y) = \int_{\mathbb{R}^d} |f(x-\sqrt{t}y) - f(x)|dx.$

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 (5.9),

4) $\qquad h_t^\wedge(\theta) = \exp(-t|\theta|^2/2)$.

Using (5.9) again, we see that $h_t = h_t^{\wedge\vee}$. Therefore,

$$\begin{aligned}
f * h_t(x) &= f * h_t^{\wedge\vee}(x) \\
&= (2\pi)^{-d}\int f(x-y)dy \int \underbrace{\mathbf{e}(-\theta\cdot y)}_{=\mathbf{e}(-\theta\cdot x)\mathbf{e}((\theta\cdot(x-y)))} h_t^\wedge(\theta)d\theta \\
&\overset{\text{Fubini}}{=} (2\pi)^{-d}\int \mathbf{e}(-\theta\cdot x)h_t^\wedge(\theta)d\theta \underbrace{\int f(x-y)\mathbf{e}((\theta\cdot(x-y)))dy}_{=f^\wedge(\theta)} \\
&= (f^\wedge h_t^\wedge)^\vee(x).
\end{aligned}$$

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 (5.6).
b):

$$\begin{aligned}
\int f d\mu &\overset{(5.6)}{=} \int_{\mathbb{R}^d} d\mu(x)(2\pi)^{-d}\int_{\mathbb{R}^d} d\theta\mathbf{e}(-\theta\cdot x)\widehat{f}(\theta) \\
&\overset{\text{Fubini}}{=} (2\pi)^{-d}\int_{\mathbb{R}^d} d\theta\widehat{f}(\theta)\underbrace{\int \mathbf{e}(-\theta\cdot x)d\mu_n(x)}_{=\widehat{\mu}(-\theta)}
\end{aligned}$$

□

11.2 Some results from Fourier transform

Theorem 11.2.1 (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* $\lim_{n \nearrow \infty} \mu_n^{\wedge}(\theta) = f(\theta)$ *for all* $\theta \in \mathbb{R}^d$ *and that the convergence is uniform in* $|\theta| \le \delta$ *for some* $\delta > 0$*. Then, there exists a* $\mu \in \mathcal{P}(\mathbb{R}^d)$ *such that* $f = \mu^{\wedge}$*.*

Theorem 11.2.2 (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* $N \in \mathbb{N}^*$ *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.*

12 Appendix to Section 7

12.1 Integral formula for the Green function

Lemma 12.1.1 *For $x \in \mathbb{Z}^d$,*

$$g_s(x) = \frac{1}{(2\pi)^d}\int_{\pi I} d\theta \frac{\mathbf{e}(-\theta \cdot x)}{1 - s\widehat{\nu}(\theta)}, \quad 0 \le s < 1. \tag{12.1}$$

Moreover,

$$g_1(0) \ \ge \ \frac{1}{(2\pi)^d}\int_{\pi I} d\theta \frac{1 - \mathrm{Re}\widehat{\nu}(\theta)}{|1 - \widehat{\nu}(\theta)|^2}. \tag{12.2}$$

$$g_1(x) \ \le \ \frac{1}{(2\pi)^d}\int_{\pi I} \frac{d\theta}{1 - \mathrm{Re}\widehat{\nu}(\theta)}, \quad x \in \mathbb{Z}^d, \tag{12.3}$$

Proof:

$$\begin{aligned} g_s(x) &= \sum_{n \ge 0} s^n P(S_n = x) \stackrel{(7.13)}{=} \frac{1}{(2\pi)^d}\sum_{n\ge 0} s^n \int_{\pi I} d\theta \mathbf{e}(-\theta \cdot x)\widehat{\nu}(\theta)^n \\ &\stackrel{\text{Fubini}}{=} \frac{1}{(2\pi)^d}\int_{\pi I} d\theta \mathbf{e}(-\theta \cdot x)\sum_{n \ge 0} s^n \widehat{\nu}(\theta)^n = \frac{1}{(2\pi)^d}\int_{\pi I} d\theta \frac{\mathbf{e}(-\theta \cdot x)}{1 - s\widehat{\nu}(\theta)}. \end{aligned}$$

(12.2) can be seen as follows. Since the left-hand-side of (12.1) is a real number, we may replace the integrand in the right-hand-side by its real part. We therefore see that

3) $$g_s(0) = (2\pi)^{-d}\int_I d\theta \frac{1 - s\mathrm{Re}\widehat{\nu}(\theta)}{|1 - s\widehat{\nu}(\theta)|^2}$$

and that

$$g_1(0) \stackrel{\text{MCT}}{=} \lim_{s \nearrow 1} g_s(0) \stackrel{(3),\ \text{Fatou}}{\ge} \frac{1}{(2\pi)^d}\int_I d\theta \lim_{s\nearrow 1} \frac{1 - s\mathrm{Re}\widehat{\nu}(\theta)}{|1 - s\widehat{\nu}(\theta)|^2} = \frac{1}{(2\pi)^d}\int_{\pi I} d\theta \frac{1 - \mathrm{Re}\widehat{\nu}(\theta)}{|1 - \widehat{\nu}(\theta)|^2}.$$

(12.3) can be seen as follows. Note that

4) $$|1 - s\widehat{\nu}(\theta)| \ge 1 - s\mathrm{Re}\widehat{\nu}(\theta) \ge s(1 - \mathrm{Re}\widehat{\nu}(\theta)).$$

Hence,

$$g_1(x) \stackrel{\text{MCT}}{=} \lim_{s \nearrow 1} g_s(x) \stackrel{(12.1)}{\le} \lim_{s\nearrow 1} \frac{1}{(2\pi)^d}\int_{\pi I} \frac{d\theta}{|1 - s\widehat{\nu}(\theta)|} \stackrel{(4)}{\le} \frac{1}{(2\pi)^d}\int_{\pi I} \frac{d\theta}{1 - \mathrm{Re}\widehat{\nu}(\theta)}.$$

□

Exercise 12.1.1 Consider a $\mathbb{Z}$-valued random walk such that

$$P(X_1 = 1) = p > 0,\ P(X_1 = -1) = q > 0 \text{ and } P(X_1 = 0) = r = 1 - p - q.$$

Use residue theorem to compute the integral (12.1) and conclude that

$$g_s(x) = \begin{cases} \gamma(s)^{-1/2}\left(\frac{1 - rs - \gamma(s)^{1/2}}{2qs}\right)^{|x|} & \text{if } x \ge 0, \\ \gamma(s)^{-1/2}\left(\frac{1 - rs - \gamma(s)^{1/2}}{2ps}\right)^{|x|} & \text{if } x \le 0, \end{cases}$$

where $\gamma(s) = (1 - rs)^2 - 4pqs^2$.

We next relate the recurrence/transience with the bound on the ch. f. $\widehat{\nu}(\theta)$ near $\theta = 0$.

Lemma 12.1.2 a) *Suppose that $d \le 2$ and that there exist $\delta_i > 0$ $(i = 0, 1, 2)$ such that*

$$\delta_1|\theta|^2 \le 1 - \mathrm{Re}\widehat{\nu}(\theta) \le |1 - \widehat{\nu}(\theta)| \le \delta_2|\theta|^2, \quad \textit{for } |\theta| < \delta_0. \tag{12.4}$$

Then the random walk is recurrent.

b) *Suppose that $d \ge 3$ and that there exists $\delta > 0$ such that*

$$1 - \mathrm{Re}\widehat{\nu}(\theta) \ge \delta|\theta|^2 \quad \textit{if } \theta \in [-\pi, \pi]^d. \tag{12.5}$$

Then the random walk is transient.

Proof: Note first that

$$\int_{\theta \in \delta I} \frac{d\theta}{|\theta|^2} \begin{cases} = \infty & \text{if } d \le 2, \\ < \infty & \text{if } d \ge 3, \end{cases} \tag{12.6}$$

a): We see from (12.4) that there exists $C > 0$ such that

1) $$\frac{1 - \mathrm{Re}\widehat{\nu}(\theta)}{|1 - \widehat{\nu}(\theta)|^2} \ge \frac{C}{|\theta|^2}, \quad |\theta| \le \delta_0$$

and hence we get (R3) in Proposition 7.1.2:

$$g_1(0) \overset{(12.2)}{\ge} \frac{1}{(2\pi)^d} \int_{\pi I} d\theta \frac{1 - \mathrm{Re}\widehat{\nu}(\theta)}{|1 - \widehat{\nu}(\theta)|^2} \overset{(1)}{\ge} \frac{C}{(2\pi)^d} \int_{\delta_0 I} \frac{d\theta}{|\theta|^2} \overset{(12.6)}{=} \infty.$$

b): We get (T5) in Proposition 7.1.2 as follows:

$$g_1(x) \overset{(12.3)}{\le} \frac{1}{(2\pi)^d} \int_{\pi I} d\theta \frac{1}{1 - \mathrm{Re}\widehat{\nu}(\theta)} \overset{(12.5)}{\le} \frac{1}{(2\pi)^d \delta} \int_{\pi I} \frac{d\theta}{|\theta|^2} \overset{(12.6)}{<} \infty, \quad x \in \mathbb{Z}^d.$$

□

Exercise 12.1.2 Prove (12.6).

Exercise 12.1.3 Suppose that $\int_{\pi I} \frac{d\theta}{1 - \mathrm{Re}\widehat{\nu}(\theta)} < \infty$, which is true for the simple random walk with $d \ge 3$. Prove then the following.
i) $\frac{1}{1-\widehat{\nu}} \in L^1(\pi I)$, $\quad g_1(x) = (2\pi)^{-d} \int_{\pi I} d\theta \frac{\mathrm{e}(-\theta \cdot x)}{1 - \widehat{\nu}(\theta)}$, $\quad x \in \mathbb{Z}^d$.
ii) $g_1(x) \to 0$ and $h(x) \to 0$ as $|x| \to \infty$. Hint: The Riemann-Lebesgue lemma.
iii) $P(H \not\subset \{S_n\}_{n \ge 1}) = 1$, where $H = \{x \in \mathbb{Z}^d \ ; \ h(x) > 0\}$. This is in contrast with Exercise 7.1.3. Hint: $P(H \subset \{S_n\}_{n \ge 1}) \le h(x)$ for any $x \in H$.

12.2 When is the random walk "truely d-dimensional"?

In Section 12.3, we generalize Theorem 7.1.1 to more general class of random walks in $\mathbb{Z}^d$. Now, let us define these classes.

Definition 12.2.1 A random walk in $\mathbb{R}^d$ is said to be *truly d-dimensional* if

$$\Theta_\perp \overset{\text{def.}}{=} \{\theta \in \mathbb{R}^d \ ; \ \theta \cdot X_1 = 0, \ P\text{-a.s.}\} = \{0\}. \tag{12.7}$$

Condition (12.7) says that the random walk is not confined in a subspace with positive codimension.

Definition 12.2.2 • For $\mathbb{Z}^d$-valued random walk, we define

$$\mathcal{R}_n = \{z \in \mathbb{Z}^d \ ; \ P\{S_n = z\} > 0\}. \tag{12.8}$$

• A $\mathbb{Z}^d$-valued random walk is said to be *aperiodic*[21] if

$$\{x - y \ ; \ x, y \in \cup_{n\geq 1}\mathcal{R}_n\} = \mathbb{Z}^d. \tag{12.9}$$

If otherwise, the random walk is called *periodic*.

Remark The left-hand-side of (12.9) is nothing but the Abelian subgroup of $\mathbb{Z}^d$ generated by $\mathcal{R}_1$.

Lemma 12.2.3 *Let $(S_n)_{n\geq 0}$ be a $\mathbb{Z}^d$-valued random walk.*

(a)

$$\mathcal{R}_n = \{x_1 + ... + x_n \ ; \ x_i \in \mathcal{R}_1\}. \tag{12.10}$$

(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 12.2.4 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 (12.10) that the random walk is aperiodic. In particular, the simple random walk is aperiodic.

12.3 A generalization of Theorem 7.1.1

In this subsection, we prove the following theorem.

Theorem 12.3.1 *Suppose that $(S_n)_{n\geq 0}$ is a $\mathbb{Z}^d$-valued, truly d-dimensional random walk with the mean velocity $v = 0$.*

(a) *If $d \leq 2$ and $X_1 \in L^2(P)$, then the random walk is recurrent.*

(b) *If $d \geq 3$, then the random walk is transient.*

Remark 12.3.2 For part (b) we will give a proof only in the aperiodic case. The proof for the periodic case, which is more intricate, is presented in Section 12.5.

Theorem 12.3.1 is a consequence of Lemma 12.1.2 and the following lemmas.

Lemma 12.3.3 *For a truly d-dimensional random walk with $\nu = P\{X_1 \in \cdot\}$, there exist $\delta_i > 0$, $i = 1, 2$ such that*

$$1 - \mathrm{Re}\widehat{\nu}(\theta) \geq \delta_1 |\theta|^2 \quad \textit{if } |\theta| \leq \delta_2. \tag{12.11}$$

[21] 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].

Proof: The proof is based on the observation that the expectation $E[|\sigma \cdot X_1|^2]$ (can be $+\infty$, but) can never be zero for $\sigma \neq 0$. We now use (7.16) and (7.17) 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

$$\begin{aligned}
F(\delta, \sigma) &= E\left[|\sigma \cdot X_1|^2 \ : \ |\sigma \cdot X_1| \leq \pi/\delta\right], \\
& \qquad \delta > 0, \ \sigma \in S^{d-1} = \{y \in \mathbb{R}^d \ ; \ |y| = 1\}.
\end{aligned}$$

Hence it is enough to show that there exists $\delta_2 > 0$ such that

$$\inf\{F(\delta, \sigma) \ ; \ \delta < \delta_2, \sigma \in S^{d-1}\} > 0. \tag{12.12}$$

Since $F(\delta, \sigma)$ is decreasing in δ, (12.12) is equivalent to;

$$\inf\{F(\delta, \sigma) \ ; \ \sigma \in S^{d-1}\} > 0 \ \text{ for some } \delta > 0. \tag{12.13}$$

We prove (12.13) by contradiction. Suppose that (12.13) is false. Then, there is $\delta_n \searrow 0$ and $\{\sigma_n\}_{n \geq 1} \subset S^{d-1}$ such that $\lim_{n \nearrow \infty} F(\delta_n, \sigma_n) = 0$. By the compactness of S^{d-1} and by taking a subsequence, we may assume that $\lim_{n \nearrow \infty} \sigma_n = \sigma$ for some $\sigma \in S^{d-1}$. Then, by Fatou's lemma,

$$\lim_{n \nearrow \infty} F(\delta_n, \sigma_n) \geq E\left[|\sigma \cdot X_1|^2\right] \neq 0,$$

which is a contradiction. □

Lemma 12.3.4 *Let $(S_n)_{n \geq 0}$ be an aperiodic random walk with with $\nu = P\{X_1 \in \cdot\}$. Then,*

$$\{\theta \in \mathbb{R}^d \ ; \ \widehat{\nu}(\theta) = 1\} = \{2\pi m \ ; \ m \in \mathbb{Z}^d\}. \tag{12.14}$$

Moreover, there exists $\delta > 0$ such that

$$1 - \mathrm{Re}\widehat{\nu}(\theta) \geq \delta|\theta|^2 \quad \text{if } \theta \in [-\pi, \pi]^d. \tag{12.15}$$

Proof: Let $(S'_n)_{n \geq 0}$ be an independent copy of $(S_n)_{n \geq 0}$. We first observe that

$$\cup_{n,n' \geq 0}\{x \in \mathbb{Z}^d \ ; \ P\{S_n - S'_{n'} = x\} > 0\} = \mathbb{Z}^d. \tag{12.16}$$

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,

$$E\mathbf{e}(X) = \mathbf{e}(t) \iff E\cos(X-t) = 1 \iff X \in \{t + 2\pi m\}_{m\in\mathbb{Z}},\ P\text{-a.s.} \tag{12.17}$$

Let $S'_n = X'_1 + ... + X'_n$. We then have that

$$\begin{aligned}
\widehat{\nu}(\theta) = 1 &\iff E\mathbf{e}(\theta\cdot X_1) = E\mathbf{e}(\theta\cdot X'_1) = 1 \\
&\iff E\mathbf{e}(\theta\cdot S_n) = E\mathbf{e}(\theta\cdot S'_{n'}) = 1, \quad \text{for all } n, n' \geq 1, \\
&\Rightarrow E\mathbf{e}(\theta\cdot (S_n - S'_{n'})) = 1, \quad \text{for all } n, n' \geq 1, \\
&\iff \theta\cdot (S_n - S'_{n'}) \in \{2\pi m\}_{m\in\mathbb{Z}}, \quad P\text{-a.s. for all } n, n' \geq 1, \quad \text{by (12.17)} \\
&\iff \theta\cdot x \in \{2\pi m\}_{m\in\mathbb{Z}}, \quad \text{for all } x \in \mathbb{Z}^d, \quad \text{by (12.16)} \\
&\iff \theta \in \{2\pi m\}_{m\in\mathbb{Z}^d}
\end{aligned}$$

We see from (12.11) that (12.15) is valid for $|\theta| \leq \delta_2$. We next prove (12.11) for the case $|\theta| \geq \delta_2$. By (12.14), $\{\theta \in \pi I \ ; \ \widehat{\nu}(\theta) = 1\ \} = \{0\}$. 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 $|\theta| \geq \delta_2$,

$$1 - \mathrm{Re}\widehat{\nu}(\theta) \geq \delta_3 \geq \delta_3\delta_2^{-1}|\theta|^2.$$

□

Proof of Theorem 12.3.1: We use Lemma 12.1.2.
(a): We see (12.4) from (6.4) and (12.11).
(b): (12.15) implies (12.5). □

12.4 More on the true d dimensionality

Lemma 12.4.1 *Consider an L^2-random walk with the mean vlocity v and the covariance matrix Γ.*

(a) *If $\det(\Gamma) > 0$, then the random walk is truly d-dimensional.*

(b) *If the random walk is truly d-dimensional and $v = 0$, then $\det(\Gamma) > 0$.*

Proof: **(a)**: It is enough to prove that $\Theta_\perp \subset \{\theta \in \mathbb{R}^d \ ; \ \theta\cdot\Gamma\theta = 0\}$. It is easy to see that for $\theta \in \mathbb{R}^d$,

$$\theta\cdot\Gamma\theta = E\,|(X_1 - v)\cdot\theta|^2 = \tfrac{1}{2}E\,|(X_1 - X_2)\cdot\theta|^2 \tag{12.18}$$

If $\theta \in \Theta_\perp$, we then have that $(X_1 - X_2)\cdot\theta = 0$, P-a.s. and hence that $\theta\cdot\Gamma\theta = 0$.
(b): If $v = 0$, then (12.18) proves $\Theta_\perp \supset \{\theta \in \mathbb{R}^d \ ; \ \theta\cdot\Gamma\theta = 0\}$. □

Exercise 12.4.1 Find a truely d-dimensional random walk with degenerate covariance matrix. Hint: Consider a $\mathbb{Z}^d$-valued random walk such that $P(S_1 = (\delta_{jk})_{k=1}^d) = 1/d$ for all $j = 1, \ldots, d$.

12.5 Proof of Theorem 12.3.1(b): periodic case

Proof is a consequence of (12.6) and the following

Lemma 12.5.1 *Consider a $\mathbb{Z}^d$-valued, truly d-dimensional L^2-random walk with the mean velocity $v = 0$. If $\delta_0 > 0$ is small enough, then for all $0 < \delta \leq \delta_0$, there is positive constant B such that*

$$g_1(0) \leq B\int_{\delta I} |\theta|^{-2}d\theta. \tag{12.19}$$

The proof of (12.19) is based on the following technical lemma

Lemma 12.5.2 *Define* $w_\delta : \mathbb{R}^d \to [0,\infty)$ *(*$\delta > 0$*) by*

$$w_\delta(\theta) = \prod_{j=1}^d \tfrac{\delta - |\theta_j|}{\delta^2} 1\{|\theta_j| \le \delta\}. \tag{12.20}$$

Then, for $x \in \mathbb{R}^d$*,*

$$\begin{aligned} \widehat{w_\delta}(x) &\stackrel{\text{def.}}{=} \int_{\mathbb{R}^d} \mathbf{e}(x\cdot\theta) w_\delta(\theta) d\theta \\ &= \prod_{j=1}^d \left(\tfrac{\sin(\delta x_j/2)}{\delta x_j/2}\right)^2, \end{aligned} \tag{12.21}$$

$$w_\delta(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} \mathbf{e}(x\cdot\theta)\widehat{w_\delta}(\theta) d\theta. \tag{12.22}$$

Proof: $w_\delta(\theta) = \prod_{j=1}^d v_\delta(\theta_j)$, where $v_\delta(\theta_j) = \frac{\delta-|\theta_j|}{\delta^2} 1\{|\theta_j| \le \delta\}$. On the other hand, it follows from (9.43) that

$$\int_{\mathbb{R}} \mathbf{e}(x\theta_j) v_\delta(\theta_j) d\theta_j = \left(\tfrac{\sin(\delta x/2)}{\delta x/2}\right)^2, \quad x \in \mathbb{R}. \tag{12.23}$$

Therefore,

$$\begin{aligned} \widehat{w_\delta}(x) &= \int_{\mathbb{R}^d} \prod_{j=1}^d \mathbf{e}(x_j\theta_j) v_\delta(\theta_j) d\theta \\ &= \prod_{j=1}^d \int_{\mathbb{R}} \mathbf{e}(x_j\theta_j) v_\delta(\theta_j) d\theta_j \\ &= \prod_{j=1}^d \left(\tfrac{\sin(\delta x_j/2)}{\delta x_j/2}\right)^2. \end{aligned}$$

The last equality (12.22) follows from (12.21) and the Fourier inversion formula. □

Proof of Lemma 12.5.1:

$$\begin{aligned} P\{S_n = 0\} &= \widehat{w_\delta}(0) P\{S_n = 0\} \quad \text{since } \widehat{w_\delta}(0) = 1, \\ &\le \sum_{x\in\mathbb{Z}^d} \widehat{w_\delta}(x) P\{S_n = x\} \\ &= \sum_{x\in\mathbb{Z}^d} \int_{\mathbb{R}^d} \mathbf{e}(x\cdot\theta) w_\delta(\theta) d\theta P\{S_n = x\} \quad \text{by (12.21)} \\ &= \int_{\mathbb{R}^d} d\theta w_\delta(\theta) \sum_{x\in\mathbb{Z}^d} \mathbf{e}(x\cdot\theta) P\{S_n = x\} \quad \text{by Fubini's theorem,} \\ &= \int_{\mathbb{R}^d} d\theta w_\delta(\theta) \widehat{\nu}(\theta)^n \quad \text{by (7.13).} \end{aligned} \tag{12.24}$$

If $0 \le s < 1$, then $s|\nu(\theta)| \le s < 1$. Note also that

$$0 \le w_\delta \le \delta^{-d} 1_{\delta I}. \tag{12.25}$$

Therefore,

$$\begin{aligned}
g_s(0) &= \sum_{n\geq 0} s^n P\{S_n = 0\} \\
&\leq \sum_{n\geq 0} \int_{\mathbb{R}^d} d\theta w_\delta(\theta)(s\widehat{\nu}(\theta))^n \quad \text{by (12.24),} \\
&= \int_{\mathbb{R}^d} d\theta w_\delta(\theta) \sum_{n\geq 0} (s\widehat{\nu}(\theta))^n \quad \text{by Fubini's theorem,} \\
&= \int_{\mathbb{R}^d} d\theta w_\delta(\theta)(1 - s\widehat{\nu}(\theta))^{-1} \\
&\leq \int_{\mathbb{R}^d} d\theta w_\delta(\theta)|1 - s\widehat{\nu}(\theta)|^{-1} \\
&\leq \delta^{-d} \int_{\delta I} d\theta |1 - s\widehat{\nu}(\theta)|^{-1} \quad \text{by (12.25),} \\
&\leq \delta^{-d} C \int_{\delta I} |\theta|^{-2} d\theta \quad \text{by (12.11).}
\end{aligned}$$

We see the desired upper bound for $g_1(0)$ by letting $s \nearrow 1$.
□

12.6 When to stop walking?

Definition 12.6.1 Consider a random walk $S_n = X_1 + \ldots + X_n$ ($n \geq 1$) defined on a probability space $(\Omega, \mathcal{F}, P)$. We define a sequence $(\mathcal{F}_n)_{n\geq 1}$ of sub σ-fields of $\mathcal{F}$ by

$$\mathcal{F}_n = \sigma(X_1, \ldots, X_n). \tag{12.26}$$

The σ-fields $\mathcal{F}_n$ can be considered as all information that the random walker knows *up to time* n. A r.v. $\tau : \Omega \to \{0, 1, \ldots, \} \cup \{\infty\}$ is called a *stopping time* if

$$\{\tau = n\} \in \mathcal{F}_n \quad \text{for all } n = 1, 2, \ldots. \tag{12.27}$$

Condition (12.27) says that the event $\{\tau = n\}$ depends only on $(X_k)_{k\geq 1}$ up to time n. For a stopping time τ, we define

$$\mathcal{F}_\tau = \{A \in \mathcal{F} \ ; \ A \cap \{\tau = n\} \in \mathcal{F}_n \text{ for all } n = 1, 2, \cdots\}. \tag{12.28}$$

Exercise 12.6.1 For a random walk, define its m^{th}-hitting time of a Borel set $A \subset \mathbb{R}^d$ by

$$\eta_A^{(m)} = \inf \left\{ n \geq 1 \mid \sum_{k=1}^n 1\{S_k \in A\} = m \right\}. \tag{12.29}$$

Prove that $\eta_A^{(m)}$ is a stopping time.

Exercise 12.6.2 Prove that $\mathcal{F}_\tau$ defined by (12.28) is a sub σ-field of $\mathcal{F}$.

Lemma 12.6.2 (Strong Markov Property) *Suppose that* $\tau : \Omega \to \{0, 1, \ldots, \} \cup \{\infty\}$ *is a stopping time such that*

$$P\{\tau < \infty\} > 0. \tag{12.30}$$

Then, under the measure $P(\cdot|\tau < \infty)$,

a) $\mathcal{F}_\tau$ *and* $\sigma(\{X_{\tau+n}\}_{n\geq 1})$ *are independent,*

b) $(X_{\tau+n})_{n\geq 1}$ *is an i.i.d. with* $P(X_{\tau+n} \in \cdot|\tau < \infty) = P(X_1 \in \cdot)$.

Proof: It is enough to prove that

$$P(A \cap \{(X_{\tau+k})_{k=1}^n \in B\} \mid \tau < \infty) = P(A \mid \tau < \infty)P\{(X_k)_{k=1}^n \in B\} \tag{12.31}$$

for all $A \in \mathcal{F}_\tau$, $n \geq 1$ and $B \in \mathcal{B}\left((\mathbb{R}^d)^n\right)$. This can be seen as follows,

$$\begin{aligned}
&P(\{\tau < \infty\} \cap A \cap \{(X_{\tau+k})_{k=1}^n \in B\}) \\
&= \sum_{l\geq 1} P(\{\tau = l\} \cap A \cap \{(X_{l+k})_{k=1}^n \in B\}) \\
&= \sum_{l\geq 1} P(\{\tau = l\} \cap A)P\{(X_{l+k})_{k=1}^n \in B\} \\
&= P(\{\tau < \infty\} \cap A)P\{(X_k)_{k=1}^n \in B\}.
\end{aligned}$$

which is equivalent to (12.31). □

Exercise 12.6.3 Let σ and τ be stopping times such that $\sigma \leq \tau$ a.s. Prove then that $\mathcal{F}_\sigma \subset \mathcal{F}_\tau$

Exercise 12.6.4 The purpose of this exercise is to illustrate that property **(a)** in Lemma 12.6.2 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 $\tau = \inf\{n \geq 1 \; |S_n \geq 2 \}$. Prove then that two events $\{\tau = 1\}$ and $\{X_{\tau+1} = 1\}$ are independent if and only if $p = 1/2$.

12.7 Green function and hitting times

We now define

$$h_s^{(m)}(x) = \begin{cases} E[s^{\eta_x^{(m)}}], & \text{if } 0 \leq s < 1, \\ h^{(m)}(x), & \text{if } s = 1. \end{cases} \tag{12.32}$$

and

$$h_s(x) = h_s^{(1)}(x), \quad 0 \leq s \leq 1. \tag{12.33}$$

Note that, by the monotone convergence theorem,

$$\begin{aligned}
h_1^{(m)}(x) &= \lim_{s\nearrow 1} h_s^{(m)}(x), && (12.34) \\
h^{(\infty)}(x) &= \lim_{m\nearrow\infty} h^{(m)}(x). && (12.35)
\end{aligned}$$

Exercise 12.7.1 Prove that

$$E[\eta_x^{(m)} 1\{\eta_x^{(m)} < \infty\}] = \lim_{s \nearrow 1} \tfrac{\partial}{\partial s} h_s^{(m)}(x). \tag{12.36}$$

We now prove (7.8) in the following generalized form.

Lemma 12.7.1 *Consider a random walk $(S_n)_{n \geq 0}$. For all $s \in [0,1]$, $x \in \mathbb{R}^d$ and $m \geq 1$,*

$$\begin{aligned} h_s^{(m)}(x) &= h_s(x) h_s^{(m-1)}(0) && (12.37)\\ &= h_s(x) h_s(0)^{m-1} && (12.38)\\ g_s(x) &= \delta_{0,x} + \frac{h_s(x)}{1 - h_s(0)}, && (12.39)\\ & && (12.40) \end{aligned}$$

Proof of Lemma 12.7.1: It is enough to prove (12.38) and (12.39) for $s < 1$. The results for $s = 1$ can be obtained by passing to the limit $s \nearrow 1$. We begin by proving (12.38) for $s < 1$. To do so, we may assume that $P\{\eta_x < \infty\} > 0$. In fact, (12.38) is just "0=0" if otherwise. Note that

$$\eta_x^{(m)} = \eta_x + \widetilde{\eta}_0^{(m-1)},$$

if $\eta_x < \infty$, where

$$\widetilde{\eta}_0^{(m-1)} = \inf\left\{ n \geq 1 \mid \sum_{k=1}^{n} 1\{X_{\eta_x+1} + \ldots + X_{\eta_x+n} = 0\} = m-1 \right\}.$$

By Lemma 12.6.2, the r.v. $\widetilde{\eta}_0^{(m-1)}$ on $(\Omega, \mathcal{F}, P(\ \cdot\ |\eta_x < \infty))$ is independent of η_x and has the same distribution as the r.v. $\eta_0^{(m-1)}$ on $(\Omega, \mathcal{F}, P)$. Note also that

$$s^{\eta_x^{(m)}} = s^{\eta_x^{(m)}} 1\{\eta_x < \infty\}.$$

We therefore have that

$$\begin{aligned} E\left[s^{\eta_x^{(m)}}\right] &= P\{\eta_x < \infty\} E\left[s^{\eta_x + \widetilde{\eta}_0^{(m-1)}} | \eta_x < \infty\right] \\ &= P\{\eta_x < \infty\} E\left[s^{\eta_x} | \eta_x < \infty\right] E\left[s^{\eta_0^{(m-1)}}\right] \\ &= E\left[s^{\eta_x}\right] E\left[s^{\eta_0^{(m-1)}}\right]. \end{aligned} \tag{12.41}$$

By applying (12.41) for $x = 0$ inductively, we see that

$$E\left[s^{\eta_0^{(m-1)}}\right] = E\left[s^{\eta_0}\right]^{m-1},$$

which, in conjunction with (12.41), proves (12.38). We next prove (12.39) for $s < 1$ as follows:

$$\begin{aligned} g_s(x) &= \delta_{0,x} + \sum_{n \geq 1} s^n P\{S_n = x\}, \\ \sum_{n \geq 1} s^n P\{S_n = x\} &= \sum_{n \geq 1} s^n \sum_{m \geq 1} P\{\eta_x^{(m)} = n\} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{m\ge1} E\left[\sum_{n\ge1} s^{\eta_x^{(m)}} 1\{\eta_x^{(m)} = n\}\right] \\
&= \sum_{m\ge1} E\left[s^{\eta_x^{(m)}}\right] \\
&= \sum_{m\ge1} h_s(x)h_s(0)^{m-1} \quad \text{; by (12.38)} \\
&= \frac{h_s(x)}{1-h_s(0)}.
\end{aligned}
$$

□

Exercise 12.7.2 Prove that for any $x, y \in \mathbb{R}^d$,

$$
1 - h(x+y) \ge \max\{P\{\eta_x < \eta_{x+y}\}(1-h(y)),\ P\{\eta_y < \eta_{x+y}\}(1-h(x))\}. \tag{12.42}
$$

Hint: Let us prove that $1 - h(x+y) \ge P\{\eta_x < \eta_{x+y}\}(1-h(y))$. To do so, we may assume that $h(x) > 0$ ($P\{\eta_x < \eta_{x+y}\} = 0$ if otherwise). Since $h(x) = P\{\eta_x < \infty\}$, we have

$$
\begin{aligned}
1 - h(x+y) &= P\{\eta_{x+y} = \infty\} \\
&\ge P\{\eta_x < \eta_{x+y},\ \widetilde{\eta}_y = \infty\},
\end{aligned}
$$

where

$$
\widetilde{\eta}_y = \inf\{n \ge 1\ ;\ X_{\eta_x+1} + \ldots + X_{\eta_x+n} = y\}.
$$

Therefore, by Lemma 12.6.2,

$$
\begin{aligned}
P\{\eta_x < \infty,\ \widetilde{\eta}_y = \infty\} &= P\{\eta_x < \eta_{x+y}\}P\{\widetilde{\eta}_y = \infty \mid \eta_x < \infty\} \\
&= P\{\eta_x < \eta_{x+y}\}P\{\eta_y = \infty\} \\
&= P\{\eta_x < \eta_{x+y}\}(1-h(y)).
\end{aligned}
$$

By exchanging the role of x and y, we also see that $1 - h(x+y) \ge P\{\eta_y < \eta_{x+y}\}(1-h(x))$.

Exercise 12.7.3 Use a similar argument in the proof (12.42) to show that

$$
h(x+y) \ge h(x)h(y) \quad \text{for any } x, y \in \mathbb{R}^d. \tag{12.43}
$$

Exercise 12.7.4 Generalize (7.6) by showing

$$
h_s(z) = s(1-h_s(0))P\{X_1 = z\} + sPh_s(z - X_1), \quad z \in \mathbb{R}^d,\ 0 \le s < 1. \tag{12.44}
$$

Exercise 12.7.5 Consider a $\mathbb{Z}$-valued random walk defined by

$$
P\{X_1 = 1\} = p > 0,\ P\{X_1 = -1\} = q > 0 \text{ and } P\{X_1 = 0\} = r = 1 - p - q.
$$

i) Use Exercise 12.1.1 and (12.39) to prove that

$$
h_s(x) = \begin{cases}
\left(\frac{1-rs-\gamma(s)^{1/2}}{2qs}\right)^{|x|} & \text{if } x > 0, \\
1 - \gamma(s)^{1/2} & \text{if } x = 0, \\
\left(\frac{1-rs-\gamma(s)^{1/2}}{2ps}\right)^{|x|} & \text{if } x < 0,
\end{cases} \tag{12.45}
$$

where $\gamma(s) = (1-rs)^2 - 4pqs^2$.

ii) Use (12.34), (12.36), (12.38), (7.8) and (12.45) to prove that if $q \le p$ and $m \ge 1$,

$$P[\eta_x^{(m)} < \infty] = \begin{cases} (1-|p-q|)^{m-1} & \text{if } x > 0, \\ (1-|p-q|)^m & \text{if } x = 0, \\ (q/p)^{|x|}(1-|p-q|)^{m-1} & \text{if } x < 0. \end{cases} \tag{12.46}$$

$$E[\eta_x^{(m)}] = \begin{cases} x/(p-q) & \text{if } x > 0,\ q < p \text{ and } m = 1, \\ \infty & \text{if otherwise.} \end{cases} \tag{12.47}$$

Exercise 12.7.6 Consider a $\mathbb{Z}^d$-valued random walk. For $x, y \in \mathbb{Z}^d$ and $A \subset \mathbb{Z}^d$, define

$$\begin{aligned} \eta(x, A) &= \inf\{n \ge 1 \,;\, x + S_n \in A\}, & (12.48) \\ \eta(x, y) &= \eta(x, \{y\}), \\ g_s^A(x, y) &= E\left[\sum_{n=0}^{\eta(x,A^c)-1} s^n 1\{x + S_n = y\}\right], & (12.49) \\ h_s^A(x, y) &= E\left[s^{\eta(x,y)} 1\{\eta(x, y) < \eta(x, A^c)\}\right], \quad 0 \le s < 1 & (12.50) \\ h_1^A(x, y) &= P\{\eta(x, y) < \eta(x, A^c)\}. & (12.51) \\ H_s^A(x, y) &= E\left[s^{\eta(x,A^c)} 1\{x + S_{\eta(x,A^c)} = y\}\right], \quad 0 \le s < 1 & (12.52) \\ H_1^A(x, y) &= P\{\eta(x, A^c) < \infty,\ x + S_{\eta(x,A^c)} = y\}. & (12.53) \end{aligned}$$

Prove then that for $x, y \in A$,

$$\begin{aligned} g_s^A(x, y) &= \delta_{x,y} + \frac{h_s^A(x, y)}{1 - h_s^A(x, x)}, \quad 0 < s \le 1, & (12.54) \\ g_s(y - x) &= g_s^A(x, y) + \sum_{z \in \mathbb{Z}^d \setminus A} H_s^A(x, z) g_s(y - z), \quad 0 < s \le 1. & (12.55) \end{aligned}$$

Special case of these identities can be found in [Law91]; See Exercise 1.5.7 and Proposition 1.5.8. of that book.

Exercise 12.7.7 Consider a symmetric, $\mathbb{Z}^d$-valued, aperiodic L^2-random walk.

i) Use (7.13) to prove that

$$P\{S_n = x\} = (2\pi)^{-d} \int_{\pi I} d\theta \cos(\theta \cdot x) \widehat{\nu}(\theta)^n \tag{12.56}$$

Hint: $P\{S_n = x\} = \frac{1}{2} P\{S_n = x\} + \frac{1}{2} P\{S_n = -x\}$ by symmetry.

ii) Use (12.56) to show that the following for any $d \ge 1$;

$$\begin{aligned} a(x) &\stackrel{\text{def.}}{=} \lim_{n \nearrow \infty} \sum_{k=0}^{n} \{P(S_k = 0) - P(S_k = x)\} & (12.57) \\ &= (2\pi)^{-d} \int_{\pi I} d\theta \frac{1 - \cos(\theta \cdot x)}{1 - \widehat{\nu}(\theta)}, & (12.58) \\ &= \lim_{s \nearrow 1} (g_s(0) - g_s(x)). & (12.59) \end{aligned}$$

The function $a(x)$ is called the *potential kernel* of the random walk. Hint: Use (12.15) and an inequality $1 - \cos(\theta \cdot x) \le (\theta \cdot x)^2/2$ to prove

$$\int_{\pi I} d\theta \sup_{0 \le s \le 1} \left| \frac{1 - \cos(\theta \cdot x)}{1 - s\widehat{\nu}(\theta)} \right| < \infty. \tag{12.60}$$

Then, use (12.56), (12.60) and the dominated convergence theorem to prove (12.58) and (12.59).

Remark 12.7.2 i) We will see in (12.62) that $a(z)$ has the following probabilistic meaning;

$$a(z) = E\left[\sum_{n=0}^{\eta_z - 1} 1\{S_n = 0\}\right] / (1 + h(z)).$$

ii) The symmetry we have assumed to prove the existence of the limit (12.57) is not essential, but to simplify the discussion for $d = 1$. In fact, for $d \ge 2$, we can prove the existence of the limit (12.57) and (12.59) without symmetry by (7.13), since $|1 - \mathbf{e}(\theta \cdot x)| \le |\theta \cdot x|$. Even for $d = 1$, it is known that the limit (12.57) exists without symmetry [Spi76, page 352].

Exercise 12.7.8 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 12.1.1 and (12.59) to compute $a(x)$ in Exercise 12.7.7 explicitly;

$$a(x) = |x|/(1 - r).$$

Exercise 12.7.9 Consider a symmetric, $\mathbb{Z}^d$-valued, aperiodic L^2-random walk. Use (12.59) and (12.55) to prove that

$$g_1^{\mathbb{Z}^d \setminus \{z\}}(x, y) = a(z - x) + h(z - x)a(y - z) - a(y - x). \tag{12.61}$$

and in particular ($x = y = 0 \ne z$) that

$$a(z) = g_1^{\mathbb{Z}^d \setminus \{z\}}(0, 0)/(1 + h(z)). \tag{12.62}$$

Exercise 12.7.10 Consider a $\mathbb{Z}^d$-valued random walk such that $P(X_1 = 0) \ne 1$. For a finite set $A \subset \mathbb{Z}^d$, there is an $\varepsilon > 0$ such that

$$P \exp\left(\varepsilon \eta(x, A^{\mathrm{c}})\right) < \infty \quad \text{for all } x \in A \tag{12.63}$$

(cf. (12.48)). Check the proof of (12.63) presented below.

Proof of (12.63); We first pick $z \ne 0$ such that $\nu(z) \stackrel{\text{def.}}{=} P(X_1 = z) > 0$. Since $A - A = \{x - x' \; ; \; x, x' \in A\}$ is a finite set, $Nz \notin A - A$ if N is large enough. We set $\delta = 1 - \nu(z)^N < 1$. We will prove by induction that

$$\sup_{x \in A} P\{\eta(x, A^{\mathrm{c}}) > kN\} \le \delta^k, \quad k = 1, 2, \ldots. \tag{12.64}$$

We begin with $k = 1$.

$$\begin{aligned} P\{\eta(x, A^{\mathsf{c}}) \le N\} &\ge P\{x + S_N \notin A\} \\ &\ge P\{S_N \notin A - A\} \\ &\ge P\{S_N = Nz\} \\ &\ge P\{X_1 = \ldots, = X_N = z\} = \nu(z)^N. \end{aligned}$$

This proves (12.64) for $k = 1$. We now suppose (12.64) for some k. Then,

$$P\{\eta(x, A^{\mathsf{c}}) > (k+1)N\} \le \sum_{y \in A} P\{\eta(x, A^{\mathsf{c}}) > kN,\ x + S_{kN} = y,\ \widetilde{\eta}(y, A^{\mathsf{c}}) > N\},$$

where

$$\widetilde{\eta}(y, A^{\mathsf{c}}) = \inf\{n \ge 1\ ;\ y + S_{n+kN} - S_{kN} \in A^{\mathsf{c}}\}.$$

Since

$$\{\eta(x, A^{\mathsf{c}}) > kN,\ x + S_{kN} = y\} \in \mathcal{F}_{kN}, \tag{12.65}$$

$$\widetilde{\eta}(y, A^{\mathsf{c}}) \text{ is independent of } \mathcal{F}_{kN}, \tag{12.66}$$

$$\widetilde{\eta}(y, A^{\mathsf{c}}) \text{ has the same distribution as } \eta(y, A^{\mathsf{c}}), \tag{12.67}$$

we have

$$\begin{aligned} &\sum_{y \in A} P\{\eta(x, A^{\mathsf{c}}) > kN,\ x + S_{kN} = y\ \widetilde{\eta}(y, A^{\mathsf{c}}) > N\} \\ &= \sum_{y \in A} P\{\eta(x, A^{\mathsf{c}}) > kN,\ x + S_{kN} = y\} P\{\eta(y, A^{\mathsf{c}}) > N\} \quad \text{by (12.65), (12.66) and (12.67),} \\ &\le \delta \sum_{y \in A} P\{\eta(x, A^{\mathsf{c}}) > kN,\ x + S_{kN} = y\} \quad \text{by (12.64) for } k = 1, \\ &= \delta P\{\eta(x, A^{\mathsf{c}}) > kN\} \\ &\le \delta^{k+1} \quad \text{by the induction hypothesis.} \end{aligned}$$

This completes the induction and proves (12.64).

Now, (12.64) can be used to prove that there are $C > 0$ and $\varepsilon > 0$ such that

$$P\{\eta(x, A^{\mathsf{c}}) > t\} \le C \exp(-\varepsilon t), \quad \text{for all } t \ge 1,$$

which proves (12.63) (cf. Exercise 1.2.5).

Exercise 12.7.11 Consider a symmetric, $\mathbb{Z}^d$-valued, aperiodic L^2-random walk. Use (12.59) and (12.55) to prove that, if $A \subset \mathbb{Z}^d$ is finite, then

$$a(y - x) = -g_1^A(x, y) + \sum_{z \in \mathbb{Z}^d \setminus A} H_1^A(x, z) a(y - z), \quad x, y \in A. \tag{12.68}$$

cf. [Law91, Proposition 1.6.3] for the simple random walk case.