Simple Random Walk in Two Dimensions

A Guided Tour

Serguei Popov

Contents

Preface

Ready by now:

- Section 1.1: Basic definitions and facts
- Section 2.1: Classical proof (of recurrence)
- Section 2.3: Lyapunov functions
- Section 2.4: Exercises (for Sections 2.1 and 2.3)
- Chapter 5: Intermezzo

At the moment, other stuff in this book is mostly copy/paste.
****

How does it look like when a mathematician explains something to a fellow mathematician (postdocs included)? Everyone knows: there are many pictures on the blackboard, there is a lot of intuition flying around, and so on. It is not surprising that mathematicians often prefer a conversation with a colleague instead of “simply” reading a book. So, the initial idea was to write a book as if I was just *explaining* things to a colleague or a research student. In such a book, there should be a lot of pictures, and plenty of detailed explanations, so that the reader would hardly have any questions left. After all, wouldn't it be nice that a person could just read in a bus (bed, park, sofa, etc.) and still learn some ideas from contemporary mathematics?

Sometimes the proof of a mathematical fact is difficult, with lots of technicalities which are hard to follow. It is not uncommon that people have troubles with understanding such proofs without first getting a “general idea” about what is going on. Also, one forgets technicalities but general ideas remain (and if the ideas are not forgotten, the technical details can usually be reconstructed with some work). So, in this book the following approach is used. The author *always* prefers to explain the intuition first. If the proof is instructive and not too long, it will be there. Otherwise, we let the interested reader to look up the details in other books and/or papers.

As one can deduce from the title, the book revolves around the two dimensional simple random walk, seemingly simple but very special and fascinating mathematical object. The purpose of this book is not to provide a complete treatment of that object, but rather make an interesting tour around it. In the end we will come to a relatively new topic of random interlacements (which can be viewed as "canonical" nearest-neighbour loops through infinity), and on the way we will take our time to digress to some related topics which are somewhat under-represented in the literature, such as, for example, Doob's h-transforms for Markov chains.

Readership and level

We expect our book to be of interest to research students and postdocs working with random walks, and to mathematicians in neighbouring fields.

It is better suited for those who want to "get the intuition first", i.e., first obtain a general idea of what is going on, and only after that pass to technicalities. The author is aware that not everybody likes this approach but still hopes that the book will find its audience.

The technical prerequisites will be rather mild. The technical material in the book will be at a level accessible to graduate students in probability, for instance, with some background in martingales and Markov chains (at the level, for instance, of [18]); the book will be largely self-contained (we also recall all necessary definitions and results in Chapter 1).

As explained above, this book is designed primarily for self-study, but it can also be used for a one-semester course in additional topics in Markov chains.

Relation to other recent books

Many topics of this book are treated at length in the literature, e.g. [17, 26, 29]; on the other hand, it contains some recent advancements (namely, soft local times and two-dimensional random interlacements) that were not covered in other books. In any case, the main distinguishing feature of this book is not its content, but rather the way it is presented.

Say that a secret plan of the author is to attract more students to the field where he works?

Special thanks are due to those whose collaborations directly relate to material presented in one or more of the chapters of this book: Francis Comets, Mikhail Menshikov, Augusto Teixeira, Gunter Schütz, Marina Vachkovskaia, and Andrew Wade. I thank ..... who read the manuscript at different stages and made many useful commments and suggestions. **(see other books for templates. . .)**

1

Introduction

Introduction: why it is such a fascinating mathematical object. The two-dimensional case is *really* critical. Blablabla.

Due to (3.23), probability to escape to $\partial \mathsf{B}(1000000)$ is approximately $1/(1.0293737 + \frac{2}{\pi} \ln 1000000) \approx 0.101785$.

Here: some funny examples for the above (step out of France starting in Paris, step out of our galaxy, all before coming back to the origin).

Let us recall the classical Polya's theorem:

Theorem 1.1 *Simple random walk in dimension d is recurrent for $d = 1, 2$ and transient for $d \geq 3$.*

A well known interpretation of this fact, attributed to Kakutani, is: "a drunken man always returns home, but a drunken bird will be eventually lost". The author cannot resist the temptation to add that this result explains why birds do not drink vodka, a fact well-known to ornithologists.

Also: write about how is difficult to simulate the two-dimensional SRW. E.g., estimate how long shall we wait until it returns to the origin, say, a hundred times.

Future side quests include: E.g., Lyapunov functions etc.

About some interesing things that we will not discuss. A little blablabla about other things, also put some pictures of DLA, they are nice :)

Level: [18] should be enough. We assume that the reader is familiar with the basic concepts of probability theory, including convergence of random variables and uniform integrability.

List literature, [26, 37] etc.

In the following, further contents of the book is described. At the end of each chapter (except for the introduction) there is a list of exercises, and at the end of the book there is a section with hints and solutions to selected exercises. A note about the exercises: they are mostly *not* meant to be easily solved during a walk in the park; their purpose is to guide an interested reader who wants to dive deeper into the subject.

1.1. *Basic definitions.* General overview of the book, and motivation. Some words about why the simple random walk in two dimensions is such a fascinating mathematical object (in some sense, the two-dimensional case is *really* critical).

Also, we recall here some basic definitions and facts for Markov chains and martingales, mainly for reference purposes.

1. *Recurrence of the walk.* First, we recall two well-known proofs of recurrence of two-dimensional simple random walk: the classical combinatorial proof, and the proof with electric networks. We then observe that the first proof heavily relies on specific combinatorics and so it is very sensitive to small changes of the model's parameters, and the second one only applies to reversible Markov chains. Then, we present a very short introduction to the Lyapunov functions method (which neither requires reversibility nor is sensitive to small perturbations of transition probabilities). Generally speaking, this method consists of finding a function (from the state space of the stochastic process to $\mathbb{R}$) such that the image under this function of the stochastic process is, in some sense, "nice". That is, this new one-dimensional process satisfies some conditions that enable one to obtain results about it and then transfer these results to the original process.
2. *Some potential theory for simple random walks.* This chapter will contain a gentle introduction to the potential theory for simple random walks, first in the transient case ($d \geq 3$), and then in two dimensions. The idea is only to recall and discuss the basic concepts (such as Green's function, potential kernel, harmonic measure) needed in the rest of the book, and then obtain explicit estimates of two-dimensional capacities and hitting probabilities, for many different kind of sets. These estimates will be heavily used in Chapters 4 and 6; also, hopefully, they may prove useful for the readers of the book in other circumstances.
3. *Simple random walk conditioned on not hitting the origin.* Here, we first recall the idea of Doob's h-transform, which permits

us to represent a conditioned (on an event of not hitting some set) Markov chain as a (unconditional) Markov chain with a different set of (possibly time-dependent) transition probabilities. We consider a few classical examples and discuss some properties of this construction. Then, we work with the Doob's transform of the simple random walk in two dimensions, with respect to its potential kernel. It turns out that this conditioned simple random walk is a fascinating object on its own right: just to cite one of its properties, the probability that a site y is ever visited by the walk started somewhere close to the origin, converges to $1/2$ as $y \to \infty$.

4. *Intermezzo: Soft local times and Poisson processes of things.* This chapter is about two topics, apparently unrelated to simple random walks. One is called soft local times; generally speaking, the method of soft local times is a way to construct an adapted stochastic process on a general space Σ, using an auxiliary Poisson point process on $\Sigma \times \mathbb{R}_+$. In Chapter 6 this method will be an important tool for dealing with *excursion processes.* Another topic we discuss is "Poisson processes of infinite objects", using as an introductory example the Poisson line process[1]. While this example *per se* is not formally necessary for the book, it helps to get some intuition about of what will happen in the next chapter.
5. *Two-dimensional random interlacements.* In this chapter we discuss random interlacements, which are Poisson processes of simple random walk trajectories. First, we review Sznitman's random interlacements model in dimension $d \geq 3$, [38]. Then, we discuss the two-dimensional case recently introduced in [12]; it is here that various plot lines of this book finally meet. This model will be constructed using the trajectories of the simple random walk conditioned on not hitting the origin, studied in Chapter 4. Using estimates on two-dimensional capacities and hitting probabilities from Chapter 3, we then prove several properties of the model, and the soft local times will enter as an important tool in some of these proofs. As stated by Sznitman in [40], "One has good decoupling properties of the excursions . . . when the boxes are sufficiently far apart. The soft

[1] judging from the author's experience, many people working with random walks do not know how Poisson line processes are constructed

local time technique of [34], especially in the form developed in the Section 2 of [11], offers a very convenient tool to express these properties".

The next section is intentionally kept dry and concise, since the author hopes that the reader would not really *read* it, but would rather occasionally use it for reference purposes.

1.1 Markov chains and martingales: basic definitions and facts

First, let us recall some basic definitions related to real-valued stochastic processes with discrete time. In the following, all random variables are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We write $\mathbb{E}$ for expectation corresponding to $\mathbb{P}$, which will be applied to real-valued random variables. Set $\mathbb{N} = \{1, 2, 3, \dots\}, \mathbb{Z}_+ = \{0, 1, 2, \dots\}, \overline{\mathbb{Z}}_+ = \mathbb{Z}_+ \cup \{+\infty\}$.

Definition 1.2 (Basic concepts for discrete-time stochastic processes)

- A discrete-time real-valued *stochastic process* is a sequence of random variables $X_n : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B})$ indexed by $n \in \mathbb{Z}_+$, where $\mathcal{B}$ is the Borel sigma-field. We write such sequences as $(X_n, n \geq 0)$, with the understanding that the time index n is always an integer.
- A *filtration* is a sequence of σ-fields $(\mathcal{F}_n, n \geq 0)$ such that $\mathcal{F}_n \subset \mathcal{F}_{n+1} \subset \mathcal{F}$ for all $n \geq 0$. Let us also define $\mathcal{F}_\infty := \sigma\big(\bigcup_{n\geq 0} \mathcal{F}_n\big) \subset \mathcal{F}$.
- A stochastic process $(X_n, n \geq 0)$ is *adapted* to a filtration $(\mathcal{F}_n, n \geq 0)$ if X_n is $\mathcal{F}_n$-measurable for all $n \in \mathbb{Z}_+$.
- For a (possibly infinite) random variable $\tau \in \overline{\mathbb{Z}}_+$, the random variable X_τ is (as the notation suggests) equal to X_n on $\{\tau = n\}$ for finite $n \in \mathbb{Z}_+$ and equal to $X_\infty := \limsup_{n\to\infty} X_n$ on $\{\tau = \infty\}$.
- A (possibly infinite) random variable $\tau \in \overline{\mathbb{Z}}_+$ is a *stopping time* with respect to a filtration $(\mathcal{F}_n, n \geq 0)$ if $\{\tau = n\} \in \mathcal{F}_n$ for all $n \geq 0$.
- If τ is a stopping time, the corresponding σ-field $\mathcal{F}_\tau$ consists of all events $A \in \mathcal{F}_\infty$ such that $A \cap \{\tau \leq n\} \in \mathcal{F}_n$ for all $n \in \overline{\mathbb{Z}}_+$.

Note that $\mathcal{F}_\tau \subset \mathcal{F}_\infty$; events in $\mathcal{F}_\tau$ include $\{\tau = \infty\}$, as well as $\{X_\tau \in B\}$ for all $B \in \mathcal{B}$.

- For $A \in \mathcal{B}$ let us define

$$\tau_A = \min\{n \geq 0 : X_n \in A\}, \tag{1.1}$$

and

$$\tau_A^+ = \min\{n \geq 1 : X_n \in A\}; \tag{1.2}$$

we may refer to either τ_A or τ_A^+ as the *hitting time* of A (also called the *passage time* into A). It is straightforward to check that both τ_A and τ_A^+ are stopping times.

Observe that, for any stochastic process $(X_n, n \geq 0)$, one can define the minimal filtration to which this process is adapted via $\mathcal{F}_n = \sigma(X_0, X_1, \ldots, X_n)$. This is the so-called *natural filtration.*

To keep the notation concise, we will frequently write X_n and $\mathcal{F}_n$ instead of $(X_n, n \geq 0)$ and $(\mathcal{F}_n, n \geq 0)$ and so on, when no confusion will arise.

Next, we need to recall some martingale-related definitions and facts.

Definition 1.3 (Martingales, submartingales, supermartingales) A real-valued stochastic process X_n adapted to a filtration $\mathcal{F}_n$ is a *martingale* (with respect to the given filtration) if, for all $n \geq 0$,

(i) $\mathbb{E}|X_n| < \infty$, and
(ii) $\mathbb{E}[X_{n+1} - X_n \mid \mathcal{F}_n] = 0$.

If in (ii) "$=$" is replaced by "$\geq$" (respectively, "$\leq$"), then X_n is called a *submartingale* (respectively, *supermartingale*).

Evidently, if X_n is a submartingale then $(-X_n)$ is a supermartingale, and vise versa; also, a martingale is both submartingale and supermartingale. Also, one can easily check the important fact that for any stopping time τ, if X_n is a (sub-, super-)martingale, then so is $X_{n\wedge\tau}$.

Martingales have a number of remarkable properties; we do not even try to elaborate on this topic here. Let us only cite the paper [32], whose title speaks for itself. In the following, we mention only the results needed in this book.

One of fundamental results is the martingale convergence theorem:

Theorem 1.4 (Martingale convergence theorem) *Assume that X_n is a submartingale such that $\sup_n \mathbb{E}[X_n^+] < \infty$. Then there is an integrable random variable X such that $X_n \to X$ a.s. as $n \to \infty$.*

Observe that, under the hypotheses of Theorem 1.4, the sequence $\mathbb{E}X_n$ is non-decreasing (by the submartingale property) and bounded above by $\sup_n \mathbb{E}[X_n^+]$, so $\lim_{n\to\infty} \mathbb{E}X_n$ exists and is finite; however, it is not necessarily equal to $\mathbb{E}X$.

Using Theorem 1.4 and Fatou's lemma, it is straightforward to obtain that the following result holds:

Theorem 1.5 (Convergence of non-negative supermartingales) *Assume that $X_n \geq 0$ is a supermartingale. Then there is an integrable random variable X such that $X_n \to X$ a.s. as $n \to \infty$, and $\mathbb{E}X \leq \mathbb{E}X_0$.*

Another fundamental result that we will frequently use is the following:

Theorem 1.6 (Optional stopping theorem) *Suppose that $\sigma \leq \tau$ are stopping times, and $X_{\tau\wedge n}$ is a uniformly integrable submartingale. Then $\mathbb{E}X_\sigma \leq \mathbb{E}X_\tau < \infty$ and $X_\sigma \leq \mathbb{E}[X_\tau \mid \mathcal{F}_\sigma]$ a.s.*

Note that, if X_n is a uniformly integrable submartingale and τ is any stopping time, then it can be shown that $X_{\tau\wedge n}$ is also uniformly integrable: see e.g. Section 5.7 of [18]. Also, observe that two applications of Theorem 1.6, one with $\sigma = 0$ and one with $\tau = \infty$, show that for any uniformly integrable submartingale X_n and any stopping time τ, it holds that $\mathbb{E}X_0 \leq \mathbb{E}X_\tau \leq \mathbb{E}X_\infty < \infty$, where $X_\infty := \limsup_{n\to\infty} X_n = \lim_{n\to\infty} X_n$ exists and is integrable, by Theorem 1.4.

Theorem 1.6 has the following corollary, obtained on setting $\sigma = 0$ and using well-known sufficient conditions for uniform integrability (see e.g. Sections 4.5 and 4.7 of [18]).

Corollary 1.7 *Let X_n be a submartingale and τ a finite stopping time. For a constant $c > 0$, suppose that at least one of the following holds:*

(i) $\tau \leq c$ *a.s.;*
(ii) $|X_{n\wedge\tau}| \leq c$ *a.s. for all* $n \geq 0$*;*
(iii) $\mathbb{E}\tau < \infty$ *and* $\mathbb{E}[|X_{n+1} - X_n| \mid \mathcal{F}_n] \leq c$ *a.s. for all* $n \geq 0$.

Then $\mathbb{E}X_\tau \geq \mathbb{E}X_0$. If X_n is a martingale and at least one of the above conditions (i)–(iii) holds, then $\mathbb{E}X_\tau = \mathbb{E}X_0$.

Next, we recall some fundamental definitions and facts on Markov processes with discrete time and countable state space, also known as *countable Markov chains*. In the following, $(X_n, n \geq 0)$ is a sequence of random variables taking values on a countable set Σ.

Definition 1.8 (Markov chains)

- A process X_n is a *Markov chain* if, for any $y \in \Sigma$, any $n \geq 0$, and any $m \geq 1$,
$$\mathbb{P}[X_{n+m} = y \mid X_0, \dots, X_n] = \mathbb{P}[X_{n+m} = y \mid X_n], \text{ a.s.}. \tag{1.3}$$
This is the *Markov property.*
- If there is no dependence on n in (1.3), the Markov chain is *homogeneous in time* (or *time-homogeneous*). Unless explicitly stated otherwise, all Markov chains considered in this book are assumed to be time-homogeneous. In this case, the Markov property (1.3) becomes
$$\mathbb{P}[X_{n+m} = y \mid \mathcal{F}_n] = p_m(X_n, y), \text{ a.s.}, \tag{1.4}$$
where $p_m : \Sigma \times \Sigma \to [0, 1]$ are the m-step Markov *transition probabilities*, for which the Chapman–Kolmogorov equation holds: $p_{n+m}(x, y) = \sum_{z \in \Sigma} p_n(x, z)p_m(z, y)$. Also, we write $p(x, y) := \mathbb{P}[X_1 = y \mid X_0 = x] = p_1(x, y)$ for the one-step transition probabilities of the Markov chain.
- We use the shorthand notation $\mathbb{P}_x[\,\cdot\,] = \mathbb{P}[\,\cdot \mid X_0 = x]$ and $\mathbb{E}_x[\,\cdot\,] = \mathbb{E}[\,\cdot \mid X_0 = x]$ for probability and expectation for the time-homogeneous Markov chain starting from initial state $x \in \Sigma$.
- A time-homogeneous, countable Markov chain is *irreducible* if for all $x, y \in \Sigma$ there exists $n_0 = n_0(x, y) \geq 1$ such that $p_{n_0}(x, y) > 0$.
- For an irreducible Markov chain, we define its *period* as the greatest common divisor of $\{n \geq 1 : p_n(x, x) > 0\}$ (it is not difficult to show that it does not depend on the choice of $x \in \Sigma$). An irreducible Markov chain with period 1 is called *aperiodic.*
- Let X_n be a Markov chain, and τ be a stopping time with respect to the natural filtration of X_n. Then, for all $x, y_1, \dots, y_k \in \Sigma$,

$n_1, \ldots, n_k \geq 1$, it holds that

$$\mathbb{P}[X_{\tau+n_j} = y_j, j = 1, \ldots, k \mid \mathcal{F}_\tau, X_\tau = x] = \mathbb{P}_x[X_{\tau+n_j} = y_j, j = 1, \ldots, k]$$

(this is the *strong Markov property*).

Suppose now that X_n is a countable Markov chain. For $A \subset \Sigma$, consistently with (1.1)–(1.2), let us define

$$\tau_A = \min\{n \geq 0 : X_n \in A\}, \text{ and } \tau_A^+ = \min\{n \geq 1 : X_n \in A\}.$$

For $x \in \Sigma$, we use the notation $\tau_x^+ := \tau_{\{x\}}^+$ and $\tau_x := \tau_{\{x\}}$ for hitting times of one-point sets. Note that for any $x \in A$ it holds that $\mathbb{P}_x[\tau_A = 0] = 1$, while $\tau_A^+ \geq 1$ is then the *return time* to A. Also note that $\mathbb{P}_x[\tau_A = \tau_A^+] = 1$ for all $x \in \Sigma \setminus A$.

Definition 1.9 For a countable Markov chain X_n, a state $x \in \Sigma$ is called

- *recurrent* if $\mathbb{P}_x[\tau_x^+ < \infty] = 1$;
- *transient* if $\mathbb{P}_x[\tau_x^+ < \infty] < 1$.

A recurrent state x is classified further as

- *positive recurrent* if $\mathbb{E}_x \tau_x^+ < \infty$;
- *null recurrent* if $\mathbb{E}_x \tau_x^+ = \infty$.

It is straightforward to see that the four properties in Definition 1.9 are *class properties*, which entails the following statement.

Proposition 1.10 *For an irreducible Markov chain, if a state $x \in \Sigma$ is recurrent (respectively, positive recurrent, null recurrent, transient) then all states in Σ are recurrent (respectively, positive recurrent, null recurrent, transient).*

By the above fact, it is legitimate to call an irreducible Markov chain itself recurrent (positive recurrent, null recurrent, transient).

Next, the following proposition is an easy consequence of the strong Markov property:

Proposition 1.11 *For an irreducible Markov chain, if a state $x \in \Sigma$ is recurrent (respectively, transient), then, regardless of the initial position of the process, it will be visited infinitely (respectively, finitely) many times almost surely.*

Finally, let us state also the following simple result which sometimes helps to prove recurrence/transience of Markov chains:

Lemma 1.12 *Let X_n be an irreducible Markov chain on a countable state space Σ.*

(i) If for some $x \in \Sigma$ and some non-empty $A \subset \Sigma$ it holds that $\mathbb{P}_x[\tau_A < \infty] < 1$, then X_n is transient.
(ii) If for some finite non-empty $A \subset \Sigma$ and all $x \in \Sigma \setminus A$ it holds that $\mathbb{P}_x[\tau_A < \infty] = 1$, then X_n is recurrent.

The reader will probably find that the above fact is evident, but we still mention that its proof can be found e.g. in [29] (cf. Lemma 2.5.1 there).

2

Recurrence of the walk

2.1 Classical proof

In this section we present the classical combinatorial proof of recurrence of the two-dimensional simple random walk.

Let us start with some general observations on recurrence and transience of random walks, which, in fact, are valid in a much broader setting. Namely, we will prove that the number of visits to the origin is a.s. finite if and only if the *expected* number of visits to the origin is finite (note that this is something which is *not* true for general random variables). This is a useful fact, because, as it frequently happens, it is easier to control the expectation than the random variable itself.

Let $p^{(d)}_m(x,y) = \mathbb{P}_x[S^{(d)}_m = y]$ be the transition probability from x to y in m steps for the simple random walk in d dimensions. Let $q_d = \mathbb{P}_0[\tau_0^+ < \infty]$ be the probability that, stating at the origin, the walk eventually returns to the origin. If $q_d < 1$, then the total number of visits (counting the initial instance $S^{(d)}_0 = 0$ as a visit) is a Geometric random variable with success probability $1 - q_d$, which has expectation $(1-q_d)^{-1} < \infty$. If $q_d = 1$, then, clearly, the walk visits the origin infinitely many times a.s.. So, the random walk is transient (i.e., $q_d < 1$) if and only if the *expected* number of visits to the origin is finite. This expected number equals (note that we can put the expectation inside the sum due to the Monotone Convergence Theorem)

$$\mathbb{E}_0 \sum_{k=0}^{\infty} \mathbf{1}\{S^{(d)}_k = 0\} = \sum_{k=0}^{\infty} \mathbb{E}_0 \mathbf{1}\{S^{(d)}_k = 0\} = \sum_{n=0}^{\infty} \mathbb{P}_0[S^{(d)}_{2n} = 0]$$

(observe that the walk can be at the starting point only after an even number of steps). We thus obtain that the recurrence of the walk is *equivalent* to

$$\sum_{n=0}^{\infty} p_{2n}^{(d)}(0,0) = \infty. \tag{2.1}$$

Before actually proving anything, let us try to *understand* why Theorem 1.1 should hold. One can represent the d-dimensional simple random walk $S^{(d)}$ as

$$S_n^{(d)} = X_1^{(d)} + \cdots + X_n^{(d)},$$

where $(X_k^{(d)}, k \geq 1)$ are i.i.d. random vectors, uniformly distributed on the set $\{\pm e_j, j = 1, \ldots, d\}$, where $e_1, \ldots, e_d$ is the canonical basis of $\mathbb{R}^d$. Since these random vectors are centered (expectation is equal to 0, component-wise), one can apply the (multivariate) Central Limit Theorem to obtain that $S_n^{(d)}/\sqrt{n}$ converges in distribution to a (multivariate) centered Normal random vector with a diagonal covariation matrix. That is, it is reasonable to expect that $S_n^{(d)}$ should be at distance of order $\sqrt{n}$ from the origin.

So, what about $p_{2n}(0,0)$? Well, if $x, y \in \mathbb{Z}^d$ are two *even* sites[1] at distance of order at most $\sqrt{n}$ from the origin, then our CLT-intuition tell us that $p_{2n}(0,x)$ and $p_{2n}(0,y)$ should be *comparable*, i.e., their ratio should be bounded away from 0 and ∞. In fact, this statement can be made rigorous by using the *local* Central Limit Theorem (e.g., Theorem 2.1.1 from [26]). Now, if there are $O(n^{d/2})$ sites where $p_{2n}(0,\cdot)$ are comparable, then the value of these probabilities (including $p_{2n}(0,0)$) should be of order $n^{-d/2}$. It remains only to observe that $\sum_{n=1}^{\infty} n^{-d/2}$ diverges only for $d =$ 1 and 2 to convince oneself that Pólya's theorem indeed holds. Notice, by the way, that for $d = 2$ we have the harmonic series which diverges just barely, its partial sums have only logarithmic growth[2].

Now, let us *prove* that (2.1) holds for the two-dimensional simple random walk. In the proof below we drop the superindex, since it is about the two-dimensional case only. For this, we first count the number of paths N_{2n} of length $2n$ that start and end at the origin. For such a path, the number of steps up must be equal to the number of steps down, and the number to the right must

[1] a site is called even if the sum of its coordinates is even; observe that the origin is even.

[2] as some physicists say, "in practice, logarithm is a constant!"

be equal to the number of steps to the left. The total number of steps up (and, also, down) can be any integer k between 0 and n; in this case, the trajectory must have $n-k$ steps to the left and $n-k$ steps to the right. So, if the number of steps up is k, the total number of trajectories starting and ending at the origin is the polynomial coefficient $\binom{2n}{k,k,n-k,n-k}$. This means that

$$N_{2n} = \sum_{k=0}^{n} \binom{2n}{k,k,n-k,n-k} = \sum_{k=0}^{n} \frac{(2n)!}{(k!)^2((n-k)!)^2}.$$

Note that

$$\frac{(2n)!}{(k!)^2((n-k)!)^2} = \binom{2n}{n}\binom{n}{k}\binom{n}{n-k};$$

the last two factors are clearly equal, but in a few lines it will become clear why we have chosen to write it this way. Since the probability of any particular trajectory of length m is 4^{-m}, we have

$$\begin{aligned} p_{2n}(0,0) &= 4^{-2n} N_{2n} \\ &= 4^{-2n}\binom{2n}{n}\sum_{k=0}^{n}\binom{n}{k}\binom{n}{n-k}. \end{aligned} \tag{2.2}$$

There is a nice combinatorial argument that allows one to deal with the sum in the right-hand side of (2.2). Consider a group of $2n$ children of which n are boys and n are girls. What is the number of ways to choose a subgroup of n children from that group? On one hand, since there are no restrictions on the gender composition of the subgroup, the answer is simply $\binom{2n}{n}$. On the other hand, the number of boys in the subgroup can vary from 0 to n, and, given that there are k boys (and, therefore, $n-k$ girls), there are $\binom{n}{k}\binom{n}{n-k}$ ways to choose the subgroup; so, the answer is precisely the above sum. This means that the above sum just equals $\binom{2n}{n}$, and we thus obtain

$$p_{2n}(0,0) = \left(2^{-2n}\binom{2n}{n}\right)^2. \tag{2.3}$$

Certainly, (2.3) is concise and beautiful; it is, however, not apriori clear which asymptotic behaviour does it have (as it frequently happens with concise and beautiful formulas). To clarify this, we use the Stirling's formula, $n! = \sqrt{2\pi n}(n/e)^n(1+o(1))$, to obtain that

$$\begin{aligned}
2^{-2n}\binom{2n}{n} &= 2^{-2n}\frac{(2n)!}{(n!)^2} \\
&= 2^{-2n}\frac{\sqrt{4\pi n}(2n/e)^{2n}}{2\pi n(n/e)^{2n}}(1+o(1)) \\
&\text{(fortunately, almost everything cancels\dots)} \\
&= \frac{1}{\sqrt{\pi n}}(1+o(1)).
\end{aligned}$$

Then, (2.3) implies that $p_{2n}(0,0) = (\pi n)^{-1}(1+o(1))$, and, using the fact that the harmonic series diverges, we obtain (2.1) and therefore recurrence. □

2.2 Electric networks

The classical book [16] is an absolute must-read.

Let $c(x,y)$ be the conductance of the edge (x,y). The transition probabilities are then defined by

$$p(x,y) = \frac{c(x,y)}{\pi(x)}, \text{ where } \pi(x) = \sum_{v:v\sim x} c(x,v). \tag{2.4}$$

Definition 2.1 Consider a Markov chain with state space Σ and transition probabilities $(p(x,y), x,y \in \Sigma)$. A function $f : \Sigma \to \mathbb{R}$ is called *harmonic* on a set $A \subset \Sigma$, if

$$f(x) = \sum_{y\in\Sigma} p(x,y)f(y),$$

for all $x \in A$.

Reversibility...

The following result is a useful *criterion* of reversibility: one can check if the Markov chain is reversible without actually calculating the reversible measure.

Theorem 2.2 *A Markov chain is reversible if and only if for any cycle $x_0, x_1, \dots, x_{n-1}, x_n = x_0$ of states it holds that:*

$$\prod_{k=0}^{n-1} p(x_k, x_{k+1}) = \prod_{k=1}^{n} p(x_k, x_{k-1}); \tag{2.5}$$

that is, the product of the transition probabilities along the cycle does not depend on the direction.

Proof It is instructive (potential fields. . .)! □

proof of recurrence with electric networks (infinite effective resistance to infinity). The harmonic series strikes again!

2.3 Lyapunov functions

The proofs of Sections 2.1 and 2.2 are simple and beautiful. This is good and bad. The problem with both proofs is that they are not robust. Assume that we change the transition probabilities of the two-dimensional simple random walk in only one site, say, $(1,1)$. For example, let the walk go from $(1,1)$ to $(1,0)$, $(1,2)$, $(0,1)$, $(2,1)$, with probabilities, say, $\frac{1}{7}$, $\frac{1}{7}$, $\frac{2}{7}$, $\frac{3}{7}$, respectively. We keep all other transition probabilities intact. Then, after this apparently innocent change, *both* proofs break down! Indeed, in the classical proof of Section 2.1 the weights of any trajectory that passes through $(1,1)$ would no longer be equal to 4^{-2n}, and so the combinatorics would be hardly manageable (instead of simple formula (2.2), a much more complicated expression will appear). The situation with the proof of Section 2.2 is even worse: the random walk is no longer reversible (cf. Exercise 2.4), so the technique of the previous section does not apply at all! It is therefore a good idea to search for a proof which is more robust, i.e., less sensible to small changes of the model's parameters.

In this section we present a very short introduction to the Lyapunov functions method[3]. Generally speaking, this method consists of finding a function (from the state space of the stochastic process to $\mathbb{R}$) such that the image under this function of the stochastic process is, in some sense, “nice”. That is, this new one-dimensional process satisfies some conditions that enable one to obtain results about it and then transfer these results to the original process.

We emphasize that this method is usually “robust”, in the sense that the underlying stochastic process need not satisfy simplifying assumptions such as the Markov property, reversibility, or time-homogeneity, for instance, and the state space of the process need

[3] it is one of the side quests that was promised.

not be necessarily countable. In particular, this approach works for *non-reversible* Markov chains.

In this section we follow mainly [29] and [19]. Other sources on the Lyapunov functions method are e.g. [2, 4, 30].

The next result is the main Lyapunov-functions-tool for proving recurrence of Markov chains.

Theorem 2.3 (Recurrence criterion) *An irreducible Markov chain X_n on a countably infinite state space Σ is recurrent if and only if there exist a function $f : \Sigma \to \mathbb{R}_+$ and a finite non-empty set $A \subset \Sigma$ such that*

$$\mathbb{E}[f(X_{n+1}) - f(X_n) \mid X_n = x] \leq 0, \text{ for all } x \in \Sigma \setminus A, \qquad (2.6)$$

and $f(x) \to \infty$ as $x \to \infty$.

Proof To prove that having a function that satisfies (2.6) is sufficient for the recurrence, let $x \in \Sigma$ be an arbitrary state, and take $X_0 = x$. Let us reason by contradiction, assuming that $\mathbb{P}_x[\tau_A = \infty] > 0$ (which would imply, in particular, that the Markov chain is transient). Set $Y_n = f(X_{n \wedge \tau_A})$ and observe that Y_n is a non-negative supermartingale. Then, by Theorem 1.5, there exists a random variable Y_∞ such that $Y_n \to Y_\infty$ a.s. and

$$\mathbb{E}_x Y_\infty \leq \mathbb{E}_x Y_0 = f(x), \qquad (2.7)$$

for any $x \in \Sigma$. On the other hand, since $f \to \infty$, it holds that the set $G_M := \{y \in \Sigma : f(y) \leq M\}$ is finite for any $M \in \mathbb{R}_+$; so, our assumption on transience implies that G_M will be visited only finitely many times, meaning that $\lim_{n\to\infty} f(X_n) = +\infty$ a.s. on $\{\tau_A = \infty\}$ (see Figure 2.1). Hence, on $\{\tau_A = \infty\}$, we must have $Y_\infty = \lim_{n\to\infty} Y_n = +\infty$, a.s.. This would contradict (2.7) under the assumption $\mathbb{P}_x[\tau_A = \infty] > 0$, because then $\mathbb{E}_x[Y_\infty] \geq \mathbb{E}_x[Y_\infty \mathbf{1}\{\tau_A = \infty\}] = \infty$. Hence $\mathbb{P}_x[\tau_A = \infty] = 0$ for all $x \in \Sigma$, which means that the Markov chain is recurrent, by Lemma 1.12 (ii).

For the "only if" part (i.e., recurrence implies that there exist f and A as above), see the proof of Theorem 2.2.1 of [19]. See also Exercise 2.10. □

As a (very simple) example of application of Theorem 2.3, consider the *one-dimensional* simple random walk $S^{(1)}$, together with

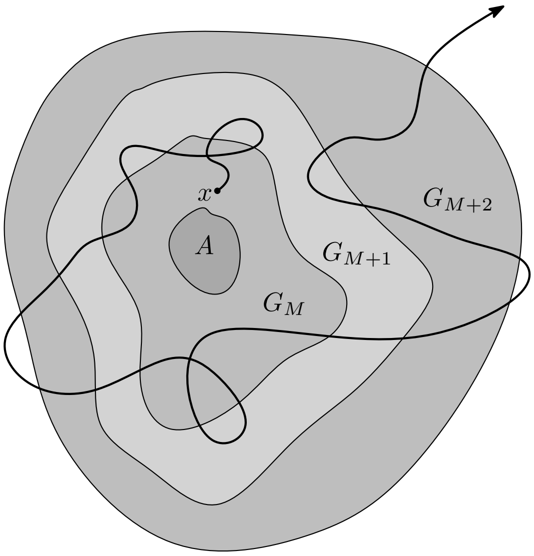

Figure 2.1 On the proof of Theorem 2.3: if the Markov chain does not hit A, it can have only finitely many visits to any finite set, and therefore goes to infinity.

the set $A = \{0\}$ and the function $f(x) = |x|$. Then (2.6) holds with equality, which shows that $S^{(1)}$ is recurrent.

Although in this chapter we are mainly interested in the recurrence, let us also formulate and prove a criterion for transience, for future reference:

Theorem 2.4 (Transience criterion) *An irreducible Markov chain X_n on a countable state space Σ is transient if and only if there exist a function $f : \Sigma \to \mathbb{R}_+$ and a non-empty set $A \subset \Sigma$ such that*

$$\mathbb{E}[f(X_{n+1}) - f(X_n) \mid X_n = x] \leq 0, \text{ for all } x \in \Sigma \setminus A, \tag{2.8}$$

and

$$f(y) < \inf_{x \in A} f(x), \text{ for at least one site } y \in \Sigma \setminus A. \tag{2.9}$$

Note that (2.8) by itself is identical to (2.6); the difference is in what we require of the nonnegative function f. Also, differently from the recurrence criterion, in the above result the set A need not be finite.

Proof of Theorem 2.4 Assume that $X_0 = y$ (where y is from (2.9)), and (similarly to the previous proof) define the process $Y_n = f(X_{n\wedge\tau_A})$. Then (2.8) implies that Y_n is a supermartingale (with respect to the filtration $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$). Since Y_n is also non-negative, Theorem 1.5 implies that there is a random variable $Y_\infty \in \mathbb{R}_+$ such that $\lim_{n\to\infty} Y_n = Y_\infty$ a.s., and $\mathbb{E}Y_\infty \le \mathbb{E}Y_0 = f(y)$. Observe that, if the Markov chain eventually hits the set A, then the value of Y_∞ equals the value of f at some random site (namely, X_{τ_A}) belonging to A; formally, we have that, a.s.,

$$\begin{aligned} Y_\infty \mathbf{1}\{\tau_A < \infty\} = \lim_{n\to\infty} Y_n \mathbf{1}\{\tau_A < \infty\} &= f(X_{\tau_A})\mathbf{1}\{\tau_A < \infty\} \\ &\ge \inf_{x\in A} f(x)\mathbf{1}\{\tau_A < \infty\}. \end{aligned}$$

So, we obtain

$$f(y) = \mathbb{E}Y_0 \ge \mathbb{E}Y_\infty \ge \mathbb{E}Y_\infty \mathbf{1}\{\tau_A < \infty\} \ge \mathbb{P}_y[\tau_A < \infty] \inf_{x\in A} f(x),$$

which implies

$$\mathbb{P}_y[\tau_A < \infty] \le \frac{f(y)}{\inf_{x\in A} f(x)} < 1,$$

proving the transience of the Markov chain X_n, by Lemma 1.12(i).

For the "only if" part, see Exercise 2.5. □

Let us now think about how to apply Theorem 2.3 to the simple random walk in two dimensions. For this, we need to find a (Lyapunov) function $f : \mathbb{Z}^2 \mapsto \mathbb{R}_+$, such that the "drift with respect to f", that is,

$$\mathbb{E}[f(S_{n+1}) - f(S_n) \mid S_n = x], \tag{2.10}$$

is nonpositive for all but finitely many $x \in \mathbb{Z}^2$, and also such that $f(x) \to \infty$ as $x \to \infty$. The reader must be warned, however, that finding a suitable Lyapunov function is a kind of an art, which usually involves a fair amount of guessing and failed attempts. Still, let us try to understand how it works. In the following, the author will do his best to explain how it *really* works, with all the failed attempts and guessing.

First of all, it is more convenient to think of f as a function of *real* arguments. Now, if there is a general rule of finding a suitable Lyapunov function for processes that live in $\mathbb{R}^d$, then it is the

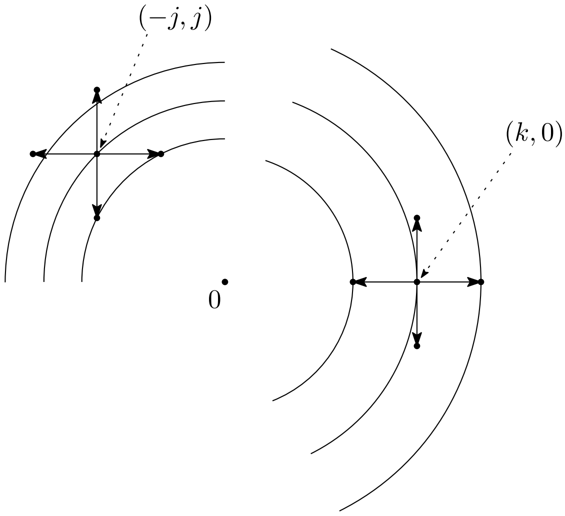

Figure 2.2 The drift with respect to $f(x) = \|x\|$ is positive

following: consider the *level sets*[4] of f and think how they should look. In the two-dimensional case we speak about the level *curves*; of course, we need the function to be sufficiently "good" to ensure that the level curves are really curves in some reasonable sense of this word.

Now, we know that the simple random walk converges to the (two-dimensional) Brownian motion, if suitably rescaled. The Brownian motion is invariant under rotations, so it seems reasonable to search for a function that only depends on the Euclidean norm of the argument, $f(x) = g(\|x\|)$ for some *increasing* function $g : \mathbb{R} \mapsto \mathbb{R}_+$. Even if we did not know about the Brownian motion, it would still be reasonable to make this assumption because, well, why not? It is easier to make calculations when there is some symmetry. Notice that, in this case, the level curves of f are just circles centered at the origin[5].

So, let us begin by looking at the level curves of a very simple function $f(x) = \|x\|$, and see what happens to the drift (2.10).

[4] for a function $f : \mathbb{R}^d \mapsto \mathbb{R}$, the level sets are sets of the form $\{x \in \mathbb{R}^d : f(x) = c\}$ for $c \in \mathbb{R}$.

[5] there are many examples where they are not circles/spheres; let us mention e.g. Section 4.3 of [29], which is based on [31].

Actually, let us just look at Figure 2.2; the level curves shown are $\{\|x\| = k-1\}$, $\{\|x\| = k\}$, $\{\|x\| = k+1\}$ on the right, and $\{\|x\| = \sqrt{j^2 + (j-1)^2}\}$, $\{\|x\| = j\sqrt{2}\}$, $\{\|x\| = 2\sqrt{2}j - \sqrt{j^2 + (j-1)^2}\}$ on the left[6]. It is quite clear then that the drift with respect to $f(x) = \|x\|$ is strictly positive in both cases. Indeed, one sees that, in the first case, the jumps to the left and to the right "compensate" each other, while the jumps up and down both slightly increase the norm. In the second case, jumps up and to the left change the norm by a larger amount than the jumps down and to the right. In fact, it is even possible to prove that the drift is positive for all $x \in \mathbb{Z}^2$, but the above examples show that, for proving the recurrence, the function $f(x) = \|x\|$ will not work anyway.

Now, think e.g. about the "diagonal case": if we move the third level curve a little bit out, then the drift with respect to the function would become nonpositive, look at Figure 2.3. It seems to be clear that, to produce such level curves, the function g should have a sublinear growth. Recall that we are "guessing" the form that g may have, so such nonrigorous reasoning is perfectly acceptable; we just need to find a function that works, and the way how we arrived to it is totally unimportant from the formal point of view. A natural first candidate would be then $g(s) = s^\alpha$, where $\alpha \in (0, 1)$. So, let us try it! Let $x \in \mathbb{Z}^2$ be such that $\|x\|$ is large, and let e be a unit vector (actually, it is $\pm e_1$ or $\pm e_2$). Write (being (y, z) the scalar product of $y, z \in \mathbb{Z}^2$)

$$
\begin{aligned}
\|x+e\|^\alpha - \|x\|^\alpha &= \|x\|^\alpha \Big(\Big(\frac{\|x+e\|}{\|x\|} \Big)^\alpha - 1 \Big) \\
&= \|x\|^\alpha \Big(\Big(\frac{(x+e, x+e)}{\|x\|^2} \Big)^{\alpha/2} - 1 \Big) \\
&= \|x\|^\alpha \Big(\Big(\frac{\|x\|^2 + 2(x,e) + 1}{\|x\|^2} \Big)^{\alpha/2} - 1 \Big) \\
&= \|x\|^\alpha \Big(\Big(1 + \frac{2(x,e)+1}{\|x\|^2} \Big)^{\alpha/2} - 1 \Big).
\end{aligned}
$$

Now, observe that $|(x, e)| \le \|x\|$, so the term $\frac{2(x,e)+1}{\|x\|^2}$ should be small (at most $O(\|x\|^{-1})$); let us also recall the Taylor expansion

[6] observe that, similarly to the previous case, these level curves have form $\{\|x\| = a - b\}$, $\{\|x\| = a\}$, $\{\|x\| = a + b\}$ with $a = j\sqrt{2}, b = j\sqrt{2} - \sqrt{j^2 + (j-1)^2}$.

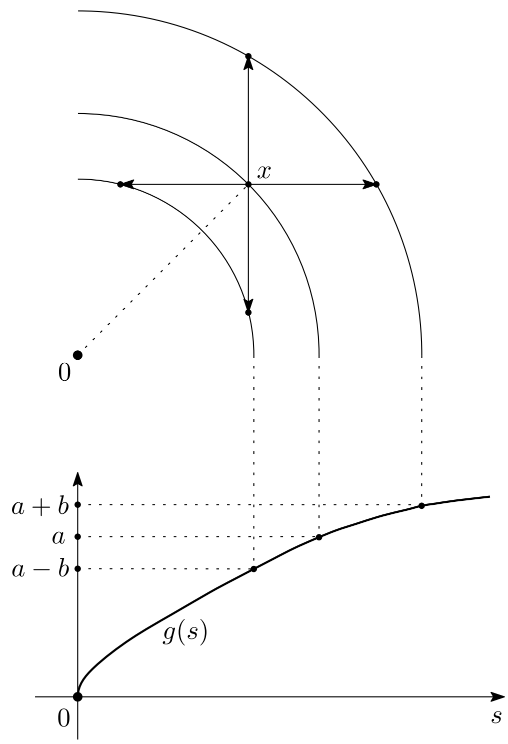

Figure 2.3 What should the function g look like? (We have $a = g(\|x\|)$,
$b = g(\|x + e_1\|) - g(\|x\|) = g(\|x\|) - g(\|x - e_1\|)$.)

$(1+y)^{\alpha/2} = 1 + \frac{\alpha}{2}y - \frac{\alpha}{4}\big(1 - \frac{\alpha}{2}\big)y^2 + O(y^3)$. Using that, we continue the above calculation:

$$\begin{aligned}
&\|x + e\|^\alpha - \|x\|^\alpha \\
&= \|x\|^\alpha \Big(\frac{\alpha(x,e)}{\|x\|^2} + \frac{\alpha}{2\|x\|^2} - \frac{\alpha}{4}\Big(1 - \frac{\alpha}{2}\Big)\frac{(2(x,e)+1)^2}{\|x\|^4} + O(\|x\|^{-3})\Big) \\
&= \alpha\|x\|^{\alpha-2}\Big((x,e) + \frac{1}{2} - \Big(1 - \frac{\alpha}{2}\Big)\frac{(x,e)^2}{\|x\|^2} + O(\|x\|^{-1})\Big). \qquad (2.11)
\end{aligned}$$

Observe that in the above display the $O(\cdot)$'s actually depend also on the direction of x (that is, the unit vector $x/\|x\|$), but this is not a problem since they are clearly uniformly bounded from above. Now, notice that, with $x = (x_1, x_2) \in \mathbb{Z}^2$,

$$\sum_{e\in\{\pm e_1,\pm e_2\}} (x,e) = 0, \quad \text{and} \qquad \sum_{e\in\{\pm e_1,\pm e_2\}} (x,e)^2 = 2x_1^2 + 2x_2^2 = 2\|x\|^2. \tag{2.12}$$

Using (2.11) and (2.12), we then obtain for $f(x) = \|x\|^\alpha$, as $\|x\| \to \infty$,

$$\begin{aligned}
&\mathbb{E}[f(S_{n+1}) - f(S_n) \mid S_n = x] \\
&= \frac{1}{4} \sum_{e\in\{\pm e_1,\pm e_2\}} \big(\|x+e\|^\alpha - \|x\|^\alpha\big) \\
&= \alpha\|x\|^{\alpha-2}\Big(\frac{1}{2} - \Big(1 - \frac{\alpha}{2}\Big)\frac{\|x\|^2}{2\|x\|^2} + O(\|x\|^{-1})\Big) \\
&= \alpha\|x\|^{\alpha-2}\Big(\frac{\alpha}{4} + O(\|x\|^{-1})\Big),
\end{aligned} \tag{2.13}$$

which, for all $\alpha \in (0,1)$, is *positive* for all sufficiently large x. So, unfortunately, we had no luck with the function $g(s) = s^\alpha$. That does not mean, however, that the above calculation was in vain; with some small changes, it will be useful for one of the exercises in the end of this chapter.

Since $g(s) = s^\alpha$ is still "too much", the next natural guess is $g(s) = \ln s$ then[7]. Well, let us try it now (more precisely, we set $f(x) = \ln\|x\|$ for $x \neq 0$ and $f(0) = 0$, but in the calculation below x is supposed to be far from the origin in any case). Using the Taylor expansion $\ln(1+y) = y - \frac{1}{2}y^2 + O(y^3)$, we write

$$\begin{aligned}
\ln\|x+e\| - \ln\|x\| &= \ln\frac{(x+e,x+e)}{\|x\|^2} \\
&= \ln\Big(1 + \frac{2(x,e)+1}{\|x\|^2}\Big) \\
&= \frac{2(x,e)}{\|x\|^2} + \frac{1}{\|x\|^2} - \frac{2(x,e)^2}{\|x\|^4} + O(\|x\|^{-3}),
\end{aligned} \tag{2.14}$$

[7] the reader may have recalled that $\ln\|x\|$ is a harmonic function in two dimensions, so $\ln\|B_t\|$ is a (local) martingale, where B is a two-dimensional standard Brownian motion. But, lattice effects may introduce some corrections...

so, using (2.12) again, we obtain (as $x \to \infty$)

$$\begin{aligned}\mathbb{E}[f(S_{n+1}) - f(S_n) \mid S_n = x] &= \frac{1}{4} \sum_{e\in\{\pm e_1,\pm e_2\}} \big(\ln \|x+e\| - \ln \|x\|\big)\\ &= \frac{1}{\|x\|^2} - \frac{1}{4} \times \frac{4\|x\|^2}{\|x\|^4} + O(\|x\|^{-3})\\ &= O(\|x\|^{-3}),\end{aligned}$$

which gives us absolutely nothing. Apparently, we need more terms in the Taylor expansion, so let us do the work: with $\ln(1+y) = y - \frac{1}{2}y^2 + \frac{1}{3}y^3 - \frac{1}{4}y^4 + O(y^5)$ we have[8]

$$\begin{aligned}\ln\|x+e\| - \ln\|x\| &= \ln\Big(1 + \frac{2(x,e)+1}{\|x\|^2}\Big)\\ &= \frac{2(x,e)}{\|x\|^2} + \frac{1}{\|x\|^2} - \frac{2(x,e)^2}{\|x\|^4} - \frac{2(x,e)}{\|x\|^4} - \frac{1}{2\|x\|^4}\\ &\quad + \frac{8(x,e)^3}{3\|x\|^6} + \frac{4(x,e)^2}{\|x\|^6} - \frac{4(x,e)^4}{\|x\|^8} + O(\|x\|^{-5}).\end{aligned}$$

Then, using (2.12) together with the fact that

$$\sum_{e\in\{\pm e_1,\pm e_2\}} (x,e)^3 = 0, \quad \text{and} \quad \sum_{e\in\{\pm e_1,\pm e_2\}} (x,e)^4 = 2(x_1^4 + x_2^4),$$

we obtain

$$\begin{aligned}&\mathbb{E}[f(S_{n+1}) - f(S_n) \mid S_n = x]\\ &= \frac{1}{\|x\|^2} - \frac{1}{\|x\|^2} - \frac{1}{2\|x\|^4} + \frac{2\|x\|^2}{\|x\|^6} - \frac{2(x_1^4+x_2^4)}{\|x\|^8} + O(\|x\|^{-5})\\ &= \|x\|^{-4}\Big(\frac{3}{2} - \frac{2(x_1^4+x_2^4)}{\|x\|^4} + O(\|x\|^{-1})\Big). \qquad (2.15)\end{aligned}$$

We want the right-hand side of (2.15) to be nonpositive for all x large enough, and it is indeed so if x is on the axes or close enough to them (for $x = (a,0)$ or $(0,a)$ the expression in the parentheses becomes $\frac{3}{2} - 2 + O(\|x\|^{-1}) < 0$ for all large enough x). Unfortunately, when we check it for the "diagonal" sites (i.e., $x = (\pm a, \pm a)$, so that $\frac{2(x_1^4+x_2^4)}{\|x\|^4} = \frac{2(a^4+a^4)}{4a^4} = 1$), we obtain that the expression in the parentheses is $\frac{3}{2} - 1 + O(\|x\|^{-1})$, which is *strictly positive* for all large enough x.

[8] the reader is invited to check that only one extra term is not enough.

So, this time we were quite close, but still missed the target. A next natural candidate would be a function that grows even slower than the logarithm; so, let us try the function $f(x) = \ln^\alpha \|x\|$ with $\alpha \in (0,1)$. Hoping for the best, we write (using $(1+y)^\alpha = 1 + \alpha y - \frac{\alpha(1-\alpha)}{2} y^2 + O(y^3)$ in the last passage)

$$
\begin{aligned}
&\ln^\alpha \|x+e\| - \ln^\alpha \|x\| \\
&= \ln^\alpha \|x\| \Big(\frac{\ln^\alpha \|x+e\|}{\ln^\alpha \|x\|} - 1 \Big) \\
&= \ln^\alpha \|x\| \left(\left(\frac{\ln \big(\|x\|^2 \big(1 + \frac{2(x,e)+1}{\|x\|^2}\big)\big)}{\ln \|x\|^2} \right)^\alpha - 1 \right) \\
&= \ln^\alpha \|x\| \left(\left(1 + (\ln \|x\|^2)^{-1} \ln \Big(1 + \frac{2(x,e)+1}{\|x\|^2} \Big) \right)^\alpha - 1 \right) \\
&\overset{(2.14)}{=} \ln^\alpha \|x\| \Bigg(\bigg(1 + (\ln \|x\|^2)^{-1} \Big(\frac{2(x,e)}{\|x\|^2} + \frac{1}{\|x\|^2} \\
&\qquad\qquad - \frac{2(x,e)^2}{\|x\|^4} + O(\|x\|^{-3}) \Big) \bigg)^\alpha - 1 \Bigg) \\
&= \ln^\alpha \|x\| \Bigg(\alpha (\ln \|x\|^2)^{-1} \Big(\frac{2(x,e)}{\|x\|^2} + \frac{1}{\|x\|^2} - \frac{2(x,e)^2}{\|x\|^4} + O(\|x\|^{-3}) \Big) \\
&\qquad\qquad - \frac{\alpha(1-\alpha)}{2} (\ln \|x\|^2)^{-2} \frac{4(x,e)^2}{\|x\|^4} + O\big((\|x\| \ln \|x\|)^{-3} \big) \Bigg).
\end{aligned}
$$

Then, using (2.12) we obtain

$$
\begin{aligned}
&\mathbb{E}[f(S_{n+1}) - f(S_n) \mid S_n = x] \\
&= \alpha \ln^{\alpha-1} \|x\| \Big(\frac{1}{\|x\|^2} - \frac{\|x\|^2}{\|x\|^4} + O(\|x\|^{-3}) \\
&\qquad\qquad - \frac{(1-\alpha)}{2} (\ln \|x\|^2)^{-1} \frac{2\|x\|^2}{\|x\|^4} + O\big((\|x\| \ln \|x\|)^{-2} \big) \Big) \\
&= - \frac{\alpha}{\|x\|^2 \ln^{2-\alpha} \|x\|} \Big(\frac{(1-\alpha)}{2} + O\big((\ln \|x\|)^{-1} \big) \Big),
\end{aligned}
$$

which is[9] negative for all sufficiently large x. Thus Theorem 2.3

[9] finally!

shows that SRW on $\mathbb{Z}^2$ is recurrent, thus proving Pólya's theorem (Theorem 1.1) in the two-dimensional case.

Now, it is time to explain why the author likes this method of proving recurrence (and many other things) of countable Markov chains. First, observe that the above proof does not use any trajectory-counting arguments (as in Section 2.1) or reversibility (as in Section 2.2), recall the example in the beginning of this section. Moreover, consider *any* Markov chain X_n on the two-dimensional integer lattice with asymptotically zero drift, and let us abbreviate $D_x = X_1 - x$. Analogously to the above, we can obtain (still using $f(x) = \ln^\alpha \|x\|$ with $\alpha \in (0,1)$)

$$\begin{aligned}
&\mathbb{E}[f(X_{n+1}) - f(X_n) \mid X_n = x] \\
&= -\frac{\alpha}{\|x\|^2 \ln^{2-\alpha}\|x\|}\Big(- \ln\|x\|^2 \mathbb{E}_x(x, D_x) - \ln\|x\|^2 \mathbb{E}_x\|D_x\|^2 \\
&\qquad + \ln\|x\|^2 \frac{2\mathbb{E}_x(x, D_x)^2}{\|x\|^2} + 2(1-\alpha)\frac{\mathbb{E}_x(x, D_x)^2}{\|x\|^2} + O\big((\ln\|x\|)^{-1}\big)\Big).
\end{aligned}$$

Now, if we can prove that the expression in the parentheses is positive for all large enough x, then this would imply the recurrence. It seems to be clear that it will be the case if the transitions probabilities at x are sufficiently close to those of the simple random walk (and the difference converges to 0 sufficiently fast as $x \to \infty$). This is what we meant when saying that the method of Lyapunov functions is robust: if it works for a particular model (the simple random walk in two dimensions, in our case), then one may expect that the same (or almost the same) Lyapunov function will also work for "close" models. See also Exercise 2.9 for some further ideas.

2.4 Exercises

Combinatorial proofs (Section 2.1):

Exercise 2.1 Understand the original proof of Polya [33]. (Warning: it uses generating functions, and it is in German.)

Exercise 2.2 Check that $p_{2n}^{(d)}(0,0) = \big(p_{2n}^{(1)}(0,0)\big)^d$ for $d = 2$. Should this be true for all $d \geq 3$ as well?

Exercise 2.3 Find a direct (combinatorial) proof of the recurrence of simple random walk on some other regular lattice (e.g., triangular, hexagonal, etc.) in two dimensions.

Electrics networks (Section 2.2):

Exercise 2.4 Show that the random walk described in the beginning of Section 2.3 (the one where we changed the transition probabilities at the site $(1,1)$) is not reversible.

Lyapunov functions (Section 2.3):

Exercise 2.5 Prove the "only if" part of the transience criterion (Theorem 2.4).

Exercise 2.6 Find a Lyapunov-function proof of transience of one-dimensional nearest-neighbor random walk with drift.

Exercise 2.7 Using Theorem 2.4, prove that simple random walk in dimensions $d \geq 3$ is transient. Hint: use $f(x) = \|x\|^{-\alpha}$ for some $\alpha > 0$.

Exercise 2.8 Prove that $f(x) = \ln \ln \|x\|$ (suitably redefined at the origin and the sites at distance at most 2.71828 from it) would also work for proving the recurrence of the two-dimensional simple random walk.

Exercise 2.9 Using Lyapunov functions, prove the recurrence of a two-dimensional spatially homogeneous zero-mean random walk with bounded jumps.

Exercise 2.10 Understand the proof of the "only if" part of the recurrence criterion (Theorem 2.3) see the proof of Theorem 2.2.1 of [19]. Can you find[10] a simpler proof?

Exercise 2.11 The following result (also known as *Foster's criterion*) provides a criterion for the positive recurrence of an irreducible Markov chain:

An irreducible Markov chain X_n on a countable state space Σ is positive recurrent if and only if there exist a positive function $f : \Sigma \to \mathbb{R}_+$, a finite non-empty set $A \subset \Sigma$, and $\varepsilon > 0$ such that

$$\mathbb{E}[f(X_{n+1}) - f(X_n) \mid X_n = x] \leq -\varepsilon, \text{ for all } x \in \Sigma \setminus A, \quad (2.16)$$

$$\mathbb{E}[f(X_{n+1}) \mid X_n = x] < \infty, \text{ for all } x \in A. \quad (2.17)$$

[10] if you find it, please, let me know!

(a) Prove the "only if" part.
(b) Understand the proof of the "if" part (see e.g. Theorems 2.6.2 and 2.6.4 of [29]).

Exercise 2.12 For the d-dimensional simple random walk, show that the first and the second moments of $\Delta_x := \|S_1^{(d)}\| - \|x\|$ under $\mathbb{E}_x$ are given by

$$\mathbb{E}_x \Delta_x = \frac{d-1}{2d\|x\|} + O(\|x\|^{-2}), \tag{2.18}$$

$$\mathbb{E}_x \Delta_x^2 = \frac{1}{d} + O(\|x\|^{-1}). \tag{2.19}$$

Exercise 2.13 Suppose now that $(X_n, n \geq 0)$ is a time-homogeneous Markov chain on an unbounded subset Σ of $\mathbb{R}_+$. Assume that X_n has uniformly bounded increments, so that

$$\mathbb{P}[|X_{n+1} - X_n| \leq B] = 1 \tag{2.20}$$

for some $B \in \mathbb{R}_+$. For $k = 1, 2$ define

$$\mu_k(x) := \mathbb{E}[(X_{n+1} - X_n)^k \mid X_n = x].$$

The first moment function, $\mu_1(x)$, is also called the one-step *mean drift* of X_n at x.

Lamperti [23, 24, 25] investigated the extent to which the asymptotic behaviour of such a process is determined by $\mu_{1,2}(\cdot)$, in a typical situation when $\mu_1(x) = O(x^{-1})$ and $\mu_2(x) = O(1)$. The following three statements are particular cases of Lamperti's fundamental results on recurrence classification:

(a) If $2x\mu_1(x) + \mu_2(x) < -\varepsilon$ for some $\varepsilon > 0$ and all large enough x, then X_n is positive recurrent;
(b) If $2x\mu_1(x) - \mu_2(x) < -\varepsilon$ for some $\varepsilon > 0$ and all large enough x, then X_n is recurrent;
(c) If $2x\mu_1(x) - \mu_2(x) > \varepsilon$ for some $\varepsilon > 0$ and all large enough x, then X_n is transient.

Prove (a), (b), and (c).

Exercise 2.14 Let $d \geq 3$. Prove that for any $\varepsilon > 0$ there exists large enough $C_d = C_d(\varepsilon)$ such that $\|S^{(d)}_{n \wedge \tau_1(\mathsf{B}(C_d))}\|^{-(d-2)+\varepsilon}$ is a supermartingale and $\|S^{(d)}_{n \wedge \tau_1(\mathsf{B}(C_d))}\|^{-(d-2)-\varepsilon}$ is a submartingale. What happens in the case $\varepsilon = 0$?

3

Some potential theory for simple random walks

Disclaimer: this chapter is by no means a systematic treatment of the subject. If the reader is looking for one, the author can recommend e.g. Chapter 6 of [26], or **(smth else?)**. Here we rather recall some general notions and tools that permit us to obtain estimates on the really interesting things — probabilities related to simple random walks.

3.1 Transient case

First, let us go back to dimensions $d \geq 3$, where the simple random walk is transient.

Three pillars: Green's function, capacity, harmonic measure.

....

We need first to recall some basic definitions related to simple random walks in higher dimensions. We shall mainly abbreviate $S := S^{(d)}$, since it always either will be clear in which dimension we are, or the argument is dimension-independent. For $d \geq 3$ let

$$G(x,y) = \mathbb{E}_x \sum_{k=0}^{\infty} \mathbf{1}\{S_k = y\} = \sum_{k=0}^{\infty} \mathbb{P}[S_k = y] \qquad (3.1)$$

denote the Green's function (i.e., the mean number of visits to y starting from x). It is important to note that in the case $x = y$ we do count this as one "initial" visit. By symmetry, it holds that $G(x,y) = G(y,x) = G(0, y-x)$, so let us abbreviate $G(y) := G(0,y)$. Now, a very important property of $G(\cdot)$ is that it is harmonic outside the origin, that is

$$G(x) = \frac{1}{2d} \sum_{y \sim x} G(y) \quad \text{for all } x \in \mathbb{Z}^d \setminus \{0\}. \qquad (3.2)$$

One readily obtains the above from the total expectation formula, with only a little bit of thinking in the case when x is a neighbour of the origin. Due to (3.2), the process $G(S_{n\wedge\tau_0(0)})$ is a martingale for *any* initial distribution of S_0; note that with τ_1 on the place of τ_0 it would not be true in the case when the walk starts from the origin.

Now, how should $G(x)$ behave as $x \to \infty$? It is (almost) clear that it converges to 0 by transience, but how fast? It is not difficult to see that it must be of order $\|x\|^{-(d-2)}$, due to the following heuristic argument. First, the simple random walk needs time of order $\|x\|^2$ to be able to deviate from its initial position by distance $\|x\|$ (which is a necessary condition if it wants to go to x). Then, at time $m > \|x\|^2$, the walk can be anywhere[1] in a ball of radius roughly $m^{1/2}$, which has volume of order $m^{d/2}$. So, the chance that the walk is in x should be of order $m^{-d/2}$; therefore, the Green's function's value in x is roughly

$$\sum_{m=\|x\|^2}^{\infty} m^{-d/2} \asymp \left(\|x\|^2\right)^{-d/2+1} = \|x\|^{-(d-2)}.$$

Note that $G(x) = \mathbb{P}_x[\tau_0(0) > \infty]G(0)$ (starting from 0, the mean number of visits to x is the mean number of visits to x counted from the time when x is first visited). **(unclear)** This implies that the probability of ever visiting y starting from x is also of order $\|x-y\|^{-(d-2)}$. **Have to state local CLT for RW's somewhere?** In fact, it is possible to obtain that

$$G(x) = \frac{C_d}{\|x\|^{d-2}} + O\big(\|x\|^{-d}\big), \tag{3.3}$$

with $C_d = \frac{\Gamma(d/2)d}{\pi^{d/2}(d-2)}$, see Theorem 4.3.1 of [26].

Let us move on. **(or another blabla)** For $x \in A$ denote the *escape probability* from A by $\mathrm{Es}_A(x) = \mathbb{P}_x[\tau_1(A) = \infty]$. The capacity of a finite set $A \subset \mathbb{Z}^d$ is defined by

$$\mathrm{cap}(A) = \sum_{x\in A} \mathrm{Es}_A(x).$$

What capacity is good for? To get some insight about this,

[1] well, not really (since the simple random walk has period two), but you understand what I mean

consider a finite $A \subset \mathbb{Z}^d$, and let $y \in \mathbb{Z}^d \setminus A$. Let us prove that

$$\mathbb{P}_y[\tau_0(A) < \infty] = \sum_{x\in A} G(y,x)\,\mathrm{Es}_A(x). \tag{3.4}$$

For the proof, we use an important idea called *the last-visit decomposition.* On the event $\{\tau_0(A) < \infty\}$, let

$$\sigma = \max\{n : S_n \in A\}$$

be the moment of the *last* visit to A (if the walk did not hit A at all, the reader is free to define σ in any convenient way, e.g., just set it to be 0). By transience (recall that A is finite!), it is clear that σ is a.s. finite. It is important to note that σ is *not* a stopping time, which actually turns out to be good! By the strong Markov property, the walk's trajectory after any stopping time is "free", that is, informally, it walks in the same way **(rewrite)**. Now, if we know that σ happened at a given time, we know something about the future, namely, that the walk must not return to A anymore. In other words, after σ the walk's law is the *conditioned* (on $\tau_1(\cdot) = \infty$) one. Now, look at Figure 3.1: what is the probability that the walker visits $x \in A$ exactly k times (on the picture, $k = 2$), and then escapes to infinity being x the last visited point of A? Clearly, for that, the walker must first visit x exactly k times (so, the piece of the trajectory until σ corresponds to the event $\{x$ is visited *at least* k times$\}$), and then escape to infinity! **(or maybe "We are interested in trajectories such that. . .)** This means that for any $x \in A$ and $k \geq 1$ it holds that

$$\mathbb{P}_y[\text{exactly } k \text{ visits to } x, S_\sigma = x] = \mathbb{P}_y[\text{at least } k \text{ visits to } x]\,\mathrm{Es}_A(x). \tag{3.5}$$

Then, summing (3.5) in k from 1 to ∞ we obtain

$$\mathbb{P}_y[\tau_0(A) < \infty, S_\sigma = x] = G(y,x)\,\mathrm{Es}_A(x), \tag{3.6}$$

and summing the above in $x \in A$ we obtain (3.4).

Next, using also (3.3), we obtain from (3.4) that

$$\mathbb{P}_y[\tau_1(A) < \infty] = \frac{C_d \,\mathrm{cap}(A)}{(\mathrm{dist}(y,A))^{d-2}}\big(1 + O(\mathrm{dist}(y,A)^{-2})\big).$$

The above means that, informally, the capacity measures how large is the set from the point of view of the random walk. **(more explanation here)**

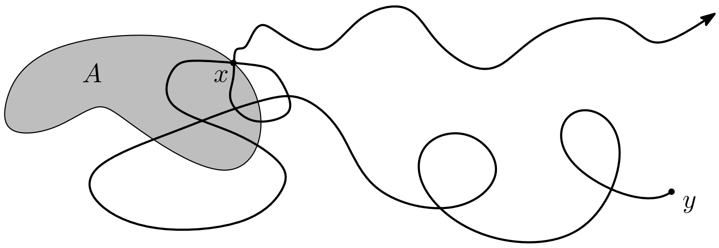

Figure 3.1 A picture useful for some proofs here

Next, exact expressions (in terms of G) for capacities of one- and two-point sets.

As for the capacity of a d-dimensional ball, observe that Proposition 6.5.2 of [26] implies (recall that $d \geq 3$)

$$\operatorname{cap}(\mathsf{B}(n)) = \frac{n^{d-2}}{C_d} + O(n^{d-3}). \tag{3.7}$$

Next, let us define the harmonic measure on A by

$$\operatorname{hm}_A(x) = \frac{\operatorname{Es}_A(x)}{\operatorname{cap}(A)}, \quad x \in A. \tag{3.8}$$

Remarkably, it holds also that hm_A is the "entrance measure to A from infinity", that is

$$\operatorname{hm}_A(x) = \lim_{y\to\infty} \mathbb{P}[S_{\tau_1(A)} = x \mid \tau_1(A) < \infty]. \tag{3.9}$$

Why this? Let us give an informal explanation. As we just saw in (3.6), the probability that the walk *leaves* the set A at x is (almost) proportional to $\operatorname{Es}_A(x)$ when the starting point in very far away. Now, look again at Figure 3.1 and reverse the direction of the trajectory (i.e., exchange y with the little arrow): "to leave A at x" now becomes "to enter A at x". Clearly, this time reversal does not change the "weight" of the trajectory[2], and this, hopefully, convinces the reader about the validity of (3.9). **(rewrite...)**

Here: explain also why the $O(\frac{r}{s})$ error term in the "entrance measure" estimate. Maybe also formulate the corresponding theorem from [26]

....

[2] formally, for infinite trajectories this only means that $0 = 0$, but you understand what I wanted to say

Estimates for hitting probabilities in terms of capacities. "min-max" lemma e.g. from Cerny-Popov [5]: **(move to exercises?)** For $n \geq 0, x \in \mathbb{Z}^d, A \subset \mathbb{Z}^d$, let

$$q_x(A; n) = \mathbb{P}_x[\tau_1(A) \leq n]$$

The following lemma will be used repeatedly to estimate the hitting probabilities:

Lemma 3.1 *For all $x \in \mathbb{Z}^d$, finite $A \subset \mathbb{Z}^d$, and $0 \leq n \leq \infty$*

$$\frac{G(x,A;n)}{\max_{y \in A} G(y,A;n)} \leq q_x(A;n) \leq \frac{G(x,A)}{\min_{y \in A} G(y,A)}. \tag{3.10}$$

Proof Using the definition of G and the strong Markov property,

$$\begin{aligned} G(x,A) &= \sum_{y \in A} \mathbb{P}_x[\tau_A < \infty, S_{\tau_A} = y] G(y,A) \\ &\geq \min_{y \in A} G(y,A) \sum_{y \in A} \mathbb{P}_x[\tau_A < \infty, S_{\tau_A} = y]. \end{aligned}$$

Since $q_x(A;n) \leq q_x(A) = \sum_{y \in A} \mathbb{P}_x[\tau_A < \infty, S_{\tau_A} = y]$, the second inequality in (3.10) follows. The first inequality is then implied by

$$\begin{aligned} G(x,A;n) &= \sum_{k=0}^{n} \sum_{y \in A} \mathbb{P}_x[\tau_A = k, S_{\tau_A} = y] G(y,A;n-k) \\ &\leq \sum_{k=0}^{n} \sum_{y \in A} \mathbb{P}_x[\tau_A = k, S_{\tau_A} = y] G(y,A;n) \\ &\leq \max_{y \in A} G(y,A;n) \sum_{k=1}^{n} \sum_{y \in A} \mathbb{P}_x[\tau_A = k, S_{\tau_A} = y], \end{aligned}$$

together with $q_x(A;n) = \sum_{k=0}^{n} \sum_{y \in A} \mathbb{P}_x[\tau_A = k, S_{\tau_A} = y]$. □

Let me stress that Lemma 3.1 is really useful, I've used it many times myself.

Proposition about capacity of a ball. Exercises: estimate (order of) capacity of the ball using Lyapunov functions. Estimate better using the asymptotic expression for the Green's function.

3.2 Potential theory in two dimensions

....

First, there is one big difference between the dimension two and higher dimensions: as shown in Chapter 2, the two-dimensional simple random walk is recurrent. This means that the mean number of visits from any site to any other site equals infinity; this prevents us from defining the Green's function in the same way as in Chapter 3.1. In spite of this unfortunate circumstance, we still would like to use martingale arguments, so a "substitute" of the Green's function is needed. Now comes the key observation: while the mean number of visits to the origin is infinite, the *difference* between the mean number of visits to the origin starting from 0 and starting from x is finite, *if suitably defined.* Namely, let us define the *potential kernel* by

$$a(x) = \sum_{k=0}^{\infty} \big(\mathbb{P}_0[S_k = 0] - \mathbb{P}_x[S_k = 0]\big). \tag{3.11}$$

Let us first see that the series converges and then figure out how big $a(x)$ should be. The "normal" approach would be using the Local CLT for this, but we prefer using another interesting and very useful tool called *coupling.* Assume that both coordinates of $x \neq 0$ are even, so, in particular, two random walks simultaneously started at x and at the origin can meet. Next, we construct these two random walks *together*, that is, on the same probability space. We do this in the following way: say, at a given moment n the positions of the walks are S'_n and S''_n; we have then $S'_0 = 0$, $S''_0 = x$. Let ζ_n and Z_n be independent random variables assuming values in $\{1, 2\}$ and $\{-1, 1\}$ respectively, with equal (to $\frac{1}{2}$) probabilities. Set

$$(S'_{n+1}, S''_{n+1}) = \begin{cases} (S'_n + Z_n e_{\zeta_n}, S''_n - Z_n e_{\zeta_n}), & \text{if } S'_n \cdot e_{\zeta_n} \neq S''_n \cdot e_{\zeta_n}, \\ (S'_n + Z_n e_{\zeta_n}, S''_n + Z_n e_{\zeta_n}), & \text{if } S'_n \cdot e_{\zeta_n} = S''_n \cdot e_{\zeta_n}; \end{cases}$$

in words, we first choose one of the two coordinates at random, and then make the walks jump in the opposite directions if the values of the chosen coordinates are different, and in the same direction in case they are equal, see Figure 3.2. Note that if the first (second) coordinates of the two walks are equal at some moment, then they will remain so forever. Let us assume, for definiteness, that x belongs to the first quadrant, that is, $x = (2a_1, 2a_2)$ for $a_{1,2} \geq 0$. Let

$$T_j = \min\{n \geq 0 : S'_n \cdot e_j = S'_n \cdot e_j\}$$

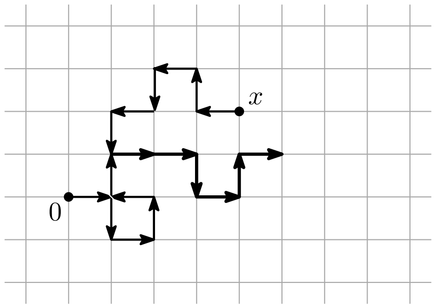

Figure 3.2 The coupling of two random walks starting at 0 and x. Note that their first coordinate become equal at time 3, and the walks meet at time 6.

for $j = 1, 2$; that is, T_j is the moment when the jth coodrinates of S' and S'' coincide for the first time. Notice that, alternatively, one can express them in the following way:

$$T_j = \min\{n \geq 0 : S'_n \cdot e_j = a_j\} = \min\{n \geq 0 : S''_n \cdot e_j = a_j\} \quad (3.12)$$

(they have to meet exactly in the middle). Let also $T = T_1 \vee T_2$ be the *coupling time*, i.e., the moment when the two walks definitely meet.

Now, we go back to (3.11) and use the strategy usually called "divide and conquer": write

$$\begin{aligned} a(x) = \sum_{k<\|x\|} \big(\mathbb{P}_0[S_k = 0] - \mathbb{P}_x[S_k = 0]\big) \\ + \sum_{k\in[\|x\|,\|x\|^2)} \big(\mathbb{P}_0[S_k = 0] - \mathbb{P}_x[S_k = 0]\big) \\ + \sum_{k\geq\|x\|^2} \big(\mathbb{P}_0[S_k = 0] - \mathbb{P}_x[S_k = 0]\big) \\ =: M_1 + M_2 + M_3, \end{aligned}$$

and then deal with the terms $M_{1,2,3}$ separately.

First, let us recall that the calculations from Section 2.1: we have obtained there that

$$\mathbb{P}_0[S_{2k} = 0] \asymp \frac{1}{k}. \quad (3.13)$$

Also, it is easy to obtain that, for all $n \geq 1$

$$\mathbb{P}_0[S_{2k} = 0] > \mathbb{P}_x[S_{2k} = 0]. \tag{3.14}$$

To see this, first consider trajectories that make $2m$ horizontal steps and $2(k-m)$ vertical ones. The number of such trajectories that go from 0 to 0 is $\binom{2k}{2m}\binom{2m}{m}\binom{2(k-m)}{k-m}$, and the number of such trajectories that go from x to 0 is $\binom{2k}{2m}\binom{2m}{m-a_1}\binom{2(k-m)}{k-m-a_2}$. It remains only to use the simple fact that $\binom{2u}{u} > \binom{2u}{v}$ for $v \neq u$. Note that (3.14) already implies that $a(x) > 0$ (but we didn't yet prove it is finite).

To deal with the term M_1, just observe that $\mathbb{P}_x[S_k = 0] = 0$ for $k < \|x\|$ — there simply will not be enough time for the walker to get from x to 0. The relation (3.13) then implies that

$$M_1 \asymp \ln \|x\|. \tag{3.15}$$

For the second term, (3.13)–(3.14) imply that

$$0 \leq M_2 \lesssim \ln \|x\|, \tag{3.16}$$

and it remains to deal with the third term. It is here that we use the coupling idea: let us write

$$\sum_{k \geq \|x\|^2} \big(\mathbb{P}_0[S_k = 0] - \mathbb{P}_x[S_k = 0]\big)$$
$$= \mathbb{E} \sum_{k \geq \|x\|^2} \big(\mathbf{1}\{S'_k = 0\} - \mathbf{1}\{S''_k = 0\}\big)$$

(write $1 = \mathbf{1}\{T \leq k\} + \mathbf{1}\{T > k\}$, and note that the kth term equals 0 on $\{T \leq k\}$)

$$= \mathbb{E} \sum_{k \geq \|x\|^2} \big(\mathbf{1}\{S'_k = 0\} - \mathbf{1}\{S''_k = 0\}\big)\mathbf{1}\{T > k\}$$

(if $T > k$ then S''_k can't be at the origin, recall (3.12))

$$= \mathbb{E} \sum_{k \geq \|x\|^2} \mathbf{1}\{S'_k = 0\}\mathbf{1}\{T > k\}$$
$$= \sum_{k \geq \|x\|^2} \mathbb{P}[S'_k = 0, T > k]$$

(since $\{T > k\} = \{T_1 > k\} \cup \{T_2 > k\}$)

$$\leq \sum_{k \geq \|x\|^2} \mathbb{P}[S'_k = 0, T_1 > k] + \sum_{k \geq \|x\|^2} \mathbb{P}[S'_k = 0, T_2 > k]$$

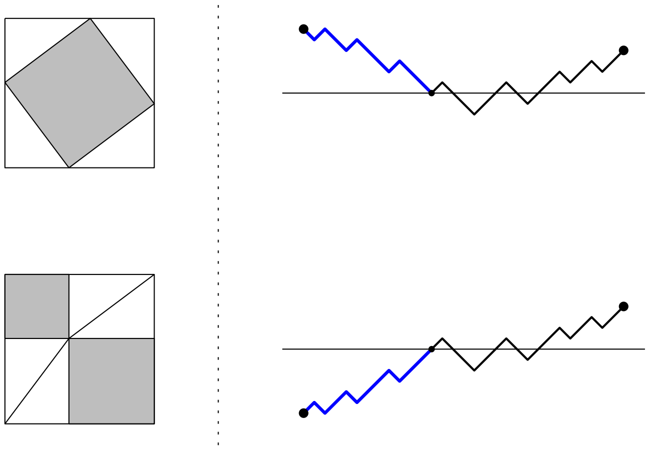

Figure 3.3 Two famous visual proofs

$$\text{(by symmetry)} \quad = 2 \sum_{k \geq \|x\|^2} \mathbb{P}[S'_k = 0, T_1 > k]. \tag{3.17}$$

We are now going to prove that the terms in the above sum are of order $k^{-3/2}$. For this, we first prove that, for all $m \geq a_1^2$,

$$\mathbb{P}_0[X_{2m} = 0, \widehat{T}(a_1) > 2m] \asymp \frac{1}{m}, \tag{3.18}$$

where X is a one-dimensional simple random walk, and $\widehat{T}(s) = \min\{\ell > 0 : X_\ell = s\}$. To show (3.18), we use the following well-known fact:

Proposition 3.2 (The Reflection Principle) *Let us consider oriented paths*[3] *in* $\mathbb{Z}^2$*, such that from* x *the path can go to either to* $x + e_1 - e_2$ *or to* $x + e_1 + e_2$*. Let two sites* $x = (x_1, x_2)$ *and* $y = (y_1, y_2)$ *be such that* $x_1 < y_1$*,* $x_2 > 0$*,* $y_2 > 0$*. Then the number of paths that go from* x *to* y *and have at least one common point with the horizontal axis is equal to the* total *number of paths that go from* $\bar{x} = (x_1, -x_2)$ *to* y*.*

Proof Just look at Figure 3.3 (on the right). □

[3] these are the *space-time* paths of one-dimensional simple random walk, the horizontal axis represents time and the vertical axis represents space

Now, we write

$$\mathbb{P}_0[X_{2m} = 0, \widehat{T}(a_1) \le 2m] = \mathbb{P}_{a_1}[X_{2m} = a_1, \widehat{T}(0) \le 2m]$$

(by the Reflection Principle)

$$= \mathbb{P}_{-a_1}[X_{2m} = a_1]$$
$$= 2^{-2m}\binom{2m}{m-a_1}.$$

So, we have for $m \ge a_1^2$,

$$\begin{aligned}
&\mathbb{P}_0[X_{2m} = 0, \widehat{T}(a_1) > 2m] \\
&= 2^{-2m}\Big(\binom{2m}{m} - \binom{2m}{m-a_1}\Big) \\
&= 2^{-2m}\binom{2m}{m}\Big(1 - \frac{(k-a_1+1)\cdots(k-1)k}{(k+1)\cdots(k+a_1)}\Big) \\
&= 2^{-2m}\binom{2m}{m}\Big(1 - \Big(1-\frac{a_1}{k+1}\Big)\cdots\Big(1-\frac{a_1}{k+a_1}\Big)\Big) \\
&= 2^{-2m}\binom{2m}{m}\Big(1 - \Big(1-\frac{a_1}{k}\Big)^{a_1}\Big)
\end{aligned}$$

(with some simple calculus)

$$\lesssim 2^{-2m}\binom{2m}{m} \times \frac{a_1^2}{m}$$

(recall Section 2.1)

$$\lesssim \frac{a_1^2}{m^{3/2}}.$$

Note also that, if N_1 is a Binomial$(2k, \frac{1}{2})$ random variable (think of the number of steps in vertical direction of two-dimensional simple random walk up to time $2k$), then it holds that[4]

$$\mathbb{P}\big[\tfrac{2}{3}k \le N_1 \le \tfrac{4}{3}k\big] \ge 1 - e^{-ck}.$$

So, using this, we have

$$\mathbb{P}[S'_{2k} = 0, T_1 > 2k]$$

[4] use e.g. the Chernoff's bound

$$= \sum_{m=0}^{k} \mathbb{P}[N_1 = m]\mathbb{P}_0[X_{2m} = 0, \widehat{T}(a_1) > 2m]\mathbb{P}_0[X_{2(k-m)} = 0]$$
$$\lesssim \frac{a_1^2}{m^{3/2}} \times \frac{1}{\sqrt{m}};$$

going back to (3.17) we find that the term M_3 is bounded above by a constant, and this finally shows that $a(x)$ exists and is of order $\ln \|x\|$.

It is possible to prove that **(more bla-bla-bla?)** as $x \to \infty$,

$$a(x) = \frac{2}{\pi} \ln \|x\| + \gamma' + O(\|x\|^{-2}), \tag{3.19}$$

where, being $\gamma = 0.5772156\ldots$ the Euler-Mascheroni constant,

$$\gamma' = \frac{2\gamma + \ln 8}{\pi} = 1.0293737\ldots, \tag{3.20}$$

cf. Theorem 4.4.4 of [26].

Also, the function a is harmonic outside the origin, i.e.,

$$\frac{1}{4} \sum_{y:y\sim x} a(y) = a(x) \quad \text{for all } x \neq 0. \tag{3.21}$$

explain why harmonic: intuitively, if a Green's function is harmonic, then the difference between two Green's functions is harmonic as well.

Observe that (3.21) immediately implies that $a(S_{k\wedge\tau_0(0)})$ is a martingale, we will repeatedly use this fact in the sequel. With some abuse of notation, we also consider the function

$$a(r) = \frac{2}{\pi} \ln r + \gamma'$$

of a *real* argument $r \geq 1$ (note that, in general, $a(x)$ need not be equal to $a(\|x\|)$). The advantage of using this notation is e.g. that, due to (3.19), we may write, as $r \to \infty$,

$$\sum_{y\in\partial B(x,r)} \nu(y)a(y) = a(r) + O\Big(\frac{\|x\| \vee 1}{r}\Big) \tag{3.22}$$

for *any* probability measure ν on $\partial B(x,r)$.

Say that $a(y) = 1$ if $y \sim 0$ (cite [37]). Here: the Optional Stopping Theorem together with (3.19) imply that

$$\mathbb{P}_0[\tau_1(\partial \mathsf{B}(r))] = \frac{1}{\gamma' + \frac{2}{\pi} \ln r}. \tag{3.23}$$

The *harmonic measure* of a finite $A \subset \mathbb{Z}^2$ is the entrance law "starting at infinity",

$$\mathrm{hm}_A(x) = \lim_{\|y\| \to \infty} \mathbb{P}_y[S_{\tau_1(A)} = x]. \tag{3.24}$$

The existence of the above limit follows from Proposition 6.6.1 of [26]; also, this proposition together with (6.44) implies that

$$\mathrm{hm}_A(x) = \frac{2}{\pi} \lim_{R \to \infty} \mathbb{P}_x\big[\tau_1(A) > \tau_1\big(\partial B(R)\big)\big] \ln R. \tag{3.25}$$

Intuitively, (3.25) means that the harmonic measure at $x \in \partial A$ is proportional to the probability of escaping from x to a large sphere. Observe also that, by recurrence of the walk, hm_A is a probability measure on ∂A. Now, for a finite set A containing the origin, we define its capacity by

$$\mathrm{cap}(A) = \sum_{x \in A} a(x)\, \mathrm{hm}_A(x); \tag{3.26}$$

in particular, $\mathrm{cap}\big(\{0\}\big) = 0$ since $a(0) = 0$. For a set not containing the origin, its capacity is defined as the capacity of a translate of this set that does contain the origin. Indeed, it can be shown that the capacity does not depend on the choice of the translation. A number of alternative definitions are available, cf. Section 6.6 of [26]. Observe that, by symmetry, the harmonic measure of any two-point set is uniform, so $\mathrm{cap}\big(\{x, y\}\big) = \frac{1}{2} a(y-x)$ for any $x, y \in \mathbb{Z}^2$. Also, (3.22) implies that

$$\mathrm{cap}\big(B(r)\big) = a(r) + O(r^{-1}). \tag{3.27}$$

Formula for the capacity via escape probabilities to large circumference.

....

comment on the exact values of $a(e_1), a(e_1 + e_2)$ etc., [37]

capacity,

tricky calculations of hitting probabilities and (as a consequence) capacities of special kinds of sets.

....

Here, we formulate several basic facts about simple random walks on annuli.

Lemma 3.3 *(i) For all $x \in \mathbb{Z}^2$ and $R > r > 0$ such that $x \in \mathsf{B}(R) \setminus \mathsf{B}(r)$ we have*

$$\mathbb{P}_x\big[\tau_1(\partial\mathsf{B}(R)) < \tau_1(\mathsf{B}(r))\big] = \frac{\ln\|x\| - \ln r + O(r^{-1})}{\ln R - \ln r}, \quad (3.28)$$

as $r, R \to \infty$.

(ii) For all $x \in \mathbb{Z}^d$, $d \geq 3$, and $R > r > 0$ such that $x \in \mathsf{B}(R) \setminus \mathsf{B}(r)$ we have

$$\mathbb{P}_x\big[\tau_1(\partial\mathsf{B}(R)) < \tau_1(\mathsf{B}(r))\big] = \frac{r^{-(d-2)} - \|x\|^{-(d-2)} + O(r^{-(d-1)})}{r^{-(d-2)} - R^{-(d-2)}}, \quad (3.29)$$

as $r, R \to \infty$.

Proof Essentially, this comes out of an application of the Optional Stopping Theorem to the martingales $a(S_{n\wedge\tau_0(0)})$ (in two dimensions) or $G(S_{n\wedge\tau_0(0)})$ (in higher dimensions). See Lemma 3.1 of [12] for the part (i). As for the part (2), apply the same kind of argument and use the expression for the Green's function e.g. from Theorem 4.3.1 of [26]. □

Lemma 3.4 *Let $c > 1$ be fixed. Then for all large enough n we have for all $v \in (\mathsf{B}(cn) \setminus \mathsf{B}(n)) \cup \partial\mathsf{B}(n)$*

$$c_1\frac{\|v\| - n + 1}{n} \leq \mathbb{P}_v\big[\tau_1(\partial\mathsf{B}(cn)) < \tau_1(\mathsf{B}(n))\big] \leq c_2\frac{\|v\| - n + 1}{n}. \quad (3.30)$$

with $c_{1,2}$ depending on c.

Proof This follows from Lemma 3.3 together with the observation that (3.28)–(3.29) start working when $\|x\| - n$ become larger than a constant (and, if x is too close to $\mathsf{B}(n)$, we just pay a constant price to force the walk out). See also Lemma 8.5 of [34] (for $d \geq 3$) and Lemma 6.3.4 together with Proposition 6.4.1 of [26] (for $d = 2$). □

do we need the following lemma?

Lemma 3.5 *Fix $c > 1$ and $\delta > 0$ such that $1 + \delta < c - \delta$, and abbreviate $A_n = (\mathsf{B}(cn) \setminus \mathsf{B}(n)) \cup \partial\mathsf{B}(n)$. Then, there exist positive constants c_3, c_4 (depending only on c, δ, and the dimension) such* *that for all* $u_{1,2} \in \mathbb{Z}^d$ *with* $(1+\delta)n < \|u_{1,2}\| < (c-\delta)n$ *and* $\|u_1 - u_2\| \geq \delta n$ *it holds that* $c_3 n^{-(d-2)} \leq G_{A_n}(u_1, u_2) \leq c_4 n^{-(d-2)}$.

Proof Indeed, we first notice that Proposition 4.6.2 of [26] (together with the estimates on the Green's function and the potential kernel, Theorems 4.3.1 of [26] and (3.19)) imply that $G_{A_n}(v, u_2) \asymp n^{-(d-2)}$ for all $d \geq 2$, where $\delta' n - 1 < \|v - u_2\| \leq \delta' n$, and $\delta' \leq \delta$ is a small enough constant. Then, use the fact that from any u_1 as above, the simple random walk comes from u_1 to $\mathsf{B}(u_2, \delta' n)$ without touching ∂A_n with uniformly positive probability. $\square$

....

3.3 Exercises

General theory in the transient case (Section 3.1):

For a finite set $A \subset \mathbb{Z}^d$ and $x, y \in A \setminus \partial A$ define

$$G_A(x,y) = \mathbb{E}_x \sum_{k=0}^{\tau_1(\partial A)-1} \mathbf{1}\{S_k = y\}$$

to be the mean number of visits to y starting from x before hitting ∂A (since A is finite, this definition makes sense for all dimensions).

Exercise 3.1 Prove that $G_A(x,y) = G_A(y,x)$ for any $x, y \in A \setminus \partial A$.

Now, there is a useful fact that follows from a reversibility calculation:

Lemma 3.6 *For all* $y \notin A$ *we have*

$$\mathbb{P}_y[\tau_1(A) < \infty, S_{\tau_1(A)} = x] = \mathbb{E}_x \sum_{j=1}^{\tau_1(\partial A_n)} \mathbf{1}\{S_j = x\} = \widehat{G}_A(x,y) \tag{3.31}$$

(that is, the probability of entering A *through* x *is equal to the mean number of visits to* y *before hitting* ∂A_n*, starting from* x*) for all* $x \in \partial A$.

Proof This follows from a standard reversibility argument. Indeed, write (the sums below are over all nearest-neighbor trajectories ϱ beginning in y and ending in x that do not touch A before hitting x; ϱ^* stands for ϱ reversed, $|\varrho|$ is the number of edges in ϱ, and $k(\varrho)$ is the number of times ϱ was in y)

$$\begin{aligned}\mathbb{P}_y[\tau_1(A) < \infty, S_{\tau_1(A)} = x] &= \sum_{\varrho} (2d)^{-|\varrho|} \\ &= \sum_{\varrho} (2d)^{-|\varrho^*|} \\ &= \sum_{j=1}^{\infty} \sum_{\varrho : k(\varrho)=j} (2d)^{-|\varrho^*|},\end{aligned}$$

and observe that the jth term in the last line is equal to the probability that y is visited at least j times (starting from x) before coming back to A. This implies (3.31). □

variational expressions for the capacity

capacity of some sets (segment, plaquette, cilinder) using various methods

capacity of a ball using Lyapunov functions (use Exercise 2.14?)

tricky calculations of hitting probabilities (also using the Möbius tranform)

complete the proof of existence of $a(x)$ for all x (hint: first for site with even sum of coodrinates, and then for the rest)

4

Simple random walk conditioned on not hitting the origin

4.1 Doob's h-transform

Let us start with a *one-dimensional* example. Let S be the simple random walk in dimension 1, starting at some site $b > 0$.

Etc.

4.2 Conditional simple random walk

....

Let us define another random walk $(\widehat{S}_n, n \geq 0)$ on $\mathbb{Z}^2$ (in fact, on $\mathbb{Z}^2 \setminus \{0\}$) in the following way: the transition probability from x to y equals $\frac{a(y)}{4a(x)}$ for all $x \sim y$ (this definition does not make sense for $x = 0$, but this is not a problem since the walk $\widehat{S}$ can never enter the origin anyway). The walk $\widehat{S}$ can be thought of as the Doob h-transform of the simple random walk, under condition of not hitting the origin (see Lemma 4.1 and its proof). Note that (3.21) implies that the random walk $\widehat{S}$ is indeed well defined, and, clearly, it is an irreducible Markov chain on $\mathbb{Z}^2 \setminus \{0\}$. We denote by $\widehat{P}_x, \widehat{E}_x$ the probability and expectation for the random walk $\widehat{S}$ started from $x \neq 0$. Let $\widehat{\tau}_0, \widehat{\tau}_1$ be defined as in (**??**)–(**??**), but with $\widehat{S}$ in the place of S. Then, it is straightforward to observe that

- the walk $\widehat{S}$ is reversible, with the reversible measure $\mu_x := a^2(x)$;
- in fact, it can be represented as a random walk on the two-dimensional lattice with conductances (or weights) $\big(a(x)a(y), x, y \in \mathbb{Z}^2, x \sim y\big)$;
- let $\mathcal{N}$ be the set of the four neighbours of the origin. Then, a direct calculation shows that $1/a(\widehat{S}_{k\wedge\widehat{\tau}_0(\mathcal{N})})$ is a martingale **(write it!)**. Theorem 2.4 then implies that the random walk $\widehat{S}$ is transient.

....

Next, we relate the probabilities of certain events for the walks S and $\widehat{S}$. For $M \subset \mathbb{Z}^2$, let $\Gamma_M^{(x)}$ be the set of all nearest-neighbour finite trajectories that start at $x \in M \setminus \{0\}$ and end when entering ∂M for the first time; denote also $\Gamma_{y,R}^{(x)} = \Gamma_{B(y,R)}^{(x)}$. For $A \subset \Gamma_M^{(x)}$ write $S \in A$ if there exists k such that $(S_0, \dots, S_k) \in A$ (and the same for the conditional walk $\widehat{S}$). In the next result we show that $\mathbb{P}_x\big[\,\cdot \mid \tau_1(0) > \tau_1\big(\partial B(R)\big)\big]$ and $\widehat{P}_x[\,\cdot\,]$ are almost indistinguishable on $\Gamma_{0,R}^{(x)}$ (that is, the conditional law of S almost coincides with the unconditional law of $\widehat{S}$). A similar result holds for excursions on a "distant" (from the origin) set.

Lemma 4.1 *(i) Assume $A \subset \Gamma_{0,R}^{(x)}$. We have*

$$P_x\big[S \in A \mid \tau_1(0) > \tau_1\big(\partial B(R)\big)\big] = \widehat{P}_x\big[\widehat{S} \in A\big]\big(1+O((R\ln R)^{-1})\big). \tag{4.1}$$

(ii) Assume that $A \subset \Gamma_M^{(x)}$ and suppose that $0 \notin M$, and denote $s = \operatorname{dist}(0, M)$, $r = \operatorname{diam}(M)$. Then, for $x \in M$,

$$P_x[S \in A] = \widehat{P}_x[\widehat{S} \in A]\Big(1 + O\Big(\frac{r}{s \ln s}\Big)\Big). \tag{4.2}$$

Proof Let us prove part (i). Assume without loss of generality that no trajectory from A passes through the origin. Then, it holds that

$$\widehat{P}_x[\widehat{S} \in A] = \sum_{\varrho \in A} \frac{a(\varrho_{\text{end}})}{a(x)} \Big(\frac{1}{4}\Big)^{|\varrho|},$$

with $|\varrho|$ the length of ϱ. On the other hand, by (**??**)

$$P_x[S \in A \mid \tau_1(0) > \tau_1\big(\partial B(R)\big)] = \frac{a(R) + O(R^{-1})}{a(x)} \sum_{\varrho \in A} \Big(\frac{1}{4}\Big)^{|\varrho|}.$$

Since $\varrho_{\text{end}} \in \partial B(R)$, we have $a(\varrho_{\text{end}}) = a(R) + O(R^{-1})$, and so (4.1) follows.

The proof of part (ii) is analogous (observe that $a(y_1)/a(y_2) = 1 + O\big(\frac{r}{s \ln s}\big)$ for any $y_1, y_2 \in M$). □

Next, we estimate the probability that the $\widehat{S}$-walk avoids a ball centered at the origin:

Lemma 4.2 *Assume $r \geq 1$ and $\|x\| \geq r+1$. We have*

$$\widehat{P}_x\big[\widehat{\tau}_1\big(B(r)\big) = \infty\big] = 1 - \frac{a(r) + O(r^{-1})}{a(x)}.$$

Proof By Lemma 4.1 (i) we have

$$\widehat{P}_x\big[\widehat{\tau}_1\big(B(r)\big) = \infty\big] = \lim_{R\to\infty} P_x\big[\tau_1(\partial B(r)) > \tau_1\big(\partial B(R)\big) \mid \tau_1(0) > \tau_1\big(\partial B(R)\big)\big].$$

The claim then follows from (**??**)–(**??**). □

Remark 4.3 Alternatively, one can deduce the proof of Lemma 4.2 from the fact that $1/a(\widehat{S}_{k\wedge\widehat{\tau}_0(\mathcal{N})})$ is a martingale, together with the Optional Stopping Theorem.

We will need also an expression for the probability of avoiding *any* finite set containing the origin:

Lemma 4.4 *Assume that $0 \in A \subset B(r)$, and $\|x\| \geq r+1$. Then*

$$\widehat{P}_x[\widehat{\tau}_1(A) = \infty] = 1 - \frac{\operatorname{cap}(A)}{a(x)} + O\Big(\frac{r \ln r \ln \|x\|}{\|x\|}\Big). \tag{4.3}$$

Proof Indeed, using Lemmas **??** and 4.1 (i) together with (**??**), we write

$$\begin{aligned}
&\widehat{P}_x[\widehat{\tau}_1(A) = \infty] \\
&= \lim_{R\to\infty} P_x\big[\tau_1(A) > \tau_1\big(\partial B(R)\big) \mid \tau_1(0) > \tau_1\big(\partial B(R)\big)\big] \\
&= \lim_{R\to\infty} \frac{a(R) + O(R^{-1})}{a(x)} \times \frac{a(x) - \operatorname{cap}(A) + O\big(\frac{r\ln r \ln\|x\|}{\|x\|}\big)}{a(R) - \operatorname{cap}(A) + O\big(R^{-1} + \frac{r\ln r\ln\|x\|}{\|x\|}\big)},
\end{aligned}$$

thus obtaining (4.3). □

It is also possible to obtain exact expressions for one-site escape probabilities, and probabilities of (not) hitting a given site:

$$\widehat{P}_x[\widehat{\tau}_1(y) < \infty] = \frac{a(x) + a(y) - a(x-y)}{2a(x)}, \tag{4.4}$$

for $x \neq y$, $x, y \neq 0$ and

$$\widehat{P}_x[\widehat{\tau}_1(x) < \infty] = 1 - \frac{1}{2a(x)} \tag{4.5}$$

for $x \neq 0$. We temporarily postpone the proof of (4.4)–(4.5). Observe that, in particular, we recover from (4.5) the transience of $\widehat{S}$.

Also, observe that (4.4) implies the following surprising fact: for any $x \neq 0$,

$$\lim_{y\to\infty} \widehat{P}_x[\widehat{\tau}_1(y) < \infty] = \frac{1}{2}.$$

The above relation leads to the following heuristic explanation for Theorem 6.2 (iii) (in the case when A is fixed and $\|x\| \to \infty$). Since the probability of hitting a distant site is about 1/2, by conditioning that this distant site is vacant, we essentially throw away three quarters of the trajectories that pass through a neighbourhood of the origin: indeed, the double-infinite trajectory has to avoid this distant site two times, before and after reaching that neighbourhood.

Let us state several other general estimates, for the probability of (not) hitting a given set (which is, typically, far away from the origin), or, more specifically, a disk:

Lemma 4.5 *Assume that $x \notin B(y,r)$ and $\|y\| > 2r \geq 1$. Abbreviate also $\Psi_1 = \|y\|^{-1}r$, $\Psi_2 = \frac{r\ln r \ln\|y\|}{\|y\|}$, $\Psi_3 = r\ln r\big(\frac{\ln\|x-y\|}{\|x-y\|} + \frac{\ln\|y\|}{\|y\|}\big)$.*

(i) We have

$$\widehat{P}_x\big[\widehat{\tau}_1(B(y,r)) < \infty\big] = \frac{\big(a(y)+O(\Psi_1)\big)\big(a(y)+a(x)-a(x-y)+O(r^{-1})\big)}{a(x)\big(2a(y)-a(r)+O(r^{-1}+\Psi_1)\big)}. \tag{4.6}$$

(ii) Consider now any nonempty set $A \subset B(y,r)$. Then, it holds that

$$\widehat{P}_x\big[\widehat{\tau}_1(A) < \infty\big] = \frac{\big(a(y)+O(\Psi_1)\big)\big(a(y)+a(x)-a(x-y)+O(r^{-1}+\Psi_3)\big)}{a(x)\big(2a(y)-\mathrm{cap}(A)+O(\Psi_2)\big)}. \tag{4.7}$$

Observe that (4.6) is *not* a particular case of (4.7); this is because (**??**) typically provides a more precise estimate than (**??**).

Proof Fix a (large) $R > 0$, such that $R > \max\{\|x\|, \|y\|+r\}+1$. Denote

$$\begin{aligned}
h_1 &= P_x\big[\tau_1(0) < \tau_1\big(\partial B(R)\big)\big],\\
h_2 &= P_x\big[\tau_1\big(B(y,r)\big) < \tau_1\big(\partial B(R)\big)\big],\\
p_1 &= P_x\big[\tau_1(0) < \tau_1\big(\partial B(R)\big) \wedge \tau_1(B(y,r))\big],
\end{aligned}$$

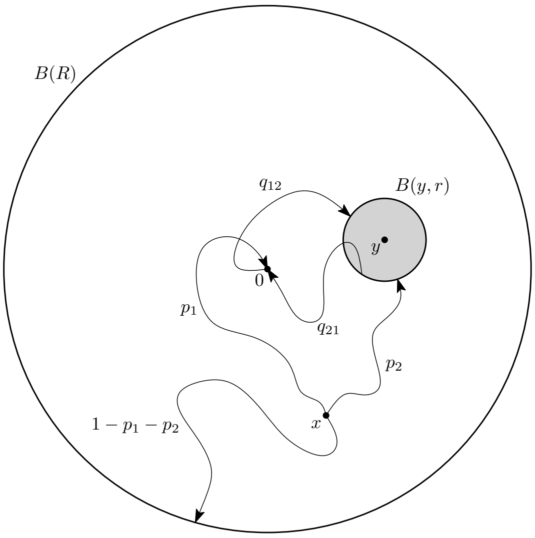

Figure 4.1 On the proof of Lemma 4.5

$$\begin{aligned}
p_2 &= P_x\big[\tau_1\big(B(y,r)\big) < \tau_1\big(\partial B(R)\big) \wedge \tau_1(0)\big],\\
q_{12} &= P_0\big[\tau_1\big(B(y,r)\big) < \tau_1\big(\partial B(R)\big)\big],\\
q_{21} &= P_\nu\big[\tau_1(0) < \tau_1\big(\partial B(R)\big)\big],
\end{aligned}$$

where ν is the entrance measure to $B(y,r)$ starting from x conditioned on the event $\big\{\tau_1\big(B(y,r)\big) < \tau_1\big(\partial B(R)\big) \wedge \tau_1(0)\big\}$, see Figure 4.1. Using Lemma 3.3, we obtain

$$h_1 = 1 - \frac{a(x)}{a(R) + O(R^{-1})}, \tag{4.8}$$

$$h_2 = 1 - \frac{a(x-y) - a(r) + O(r^{-1})}{a(R) - a(r) + O(R^{-1}\|y\| + r^{-1})}, \tag{4.9}$$

and

$$q_{12} = 1 - \frac{a(y) - a(r) + O(r^{-1})}{a(R) - a(r) + O(R^{-1}\|y\| + r^{-1})}, \tag{4.10}$$

$$q_{21} = 1 - \frac{a(y) + O(\|y\|^{-1}r)}{a(R) + O(R^{-1}\|y\|)}. \tag{4.11}$$

Then, as a general fact, it holds that

$$h_1 = p_1 + p_2 q_{21},$$

$$h_2 = p_2 + p_1 q_{12}.$$

Solving these equations with respect to p_1, p_2, we obtain

$$p_1 = \frac{h_1 - h_2 q_{21}}{1 - q_{12}q_{21}}, \tag{4.12}$$

$$p_2 = \frac{h_2 - h_1 q_{12}}{1 - q_{12}q_{21}}, \tag{4.13}$$

and so, using (4.8)–(4.11), we write

$$\begin{aligned}
&P_x\big[\tau_1(B(y,r)) < \tau_1\big(\partial B(R)\big) \mid \tau_1(0) > \tau_1\big(\partial B(R)\big)\big] \\
&= \frac{p_2(1-q_{21})}{1-h_1} \\
&= \frac{(h_2 - h_1 q_{12})(1-q_{21})}{(1-h_1)(1-q_{12}q_{21})} \\
&= \Bigg(\frac{a(x)}{a(R)+O(R^{-1})} + \frac{a(y)-a(r)+O(r^{-1})}{a(R)-a(r)+O(R^{-1}\|y\|+r^{-1})} \\
&\qquad - \frac{a(x-y)-a(r)+O(r^{-1})}{a(R)-a(r)+O(R^{-1}\|y\|+r^{-1})} + O\Big(\frac{\ln\|x-y\|\ln\|y\|}{\ln^2 R}\Big)\Bigg) \\
&\quad \times \frac{a(y)+O(\|y\|^{-1}r)}{a(R)+O(R^{-1}\|y\|)} \times \Big(\frac{a(x)}{a(R)+O(R^{-1})}\Big)^{-1} \\
&\quad \times \Bigg(\frac{a(y)-a(r)+O(r^{-1})}{a(R)-a(r)+O(R^{-1}\|y\|+r^{-1})} + \frac{a(y)+O(\|y\|^{-1}r)}{a(R)+O(R^{-1}\|y\|)} + O\Big(\frac{\ln^2\|y\|}{\ln^2 R}\Big)\Bigg)^{-1}.
\end{aligned}$$

Sending R to infinity, we obtain the proof of (4.6).

To prove (4.7), we use the same procedure. Define $h'_{1,2}$, $p'_{1,2}$, q'_{12}, q'_{21} in the same way but with A in place of $B(y,r)$. It holds that $h'_1 = h_1$, q'_{21} is expressed in the same way as q_{21} (although q'_{21} and q_{21} are not necessarily equal, the difference is only in the error terms $O(\cdot)$) and, by Lemma **??**,

$$\begin{aligned}
h'_2 &= 1 - \frac{a(x-y) - \operatorname{cap}(A) + O\big(\frac{r\ln r\ln\|x-y\|}{\|x-y\|}\big)}{a(R) - \operatorname{cap}(A) + O\big(R^{-1}\|x-y\| + \frac{r\ln r\ln\|x-y\|}{\|x-y\|}\big)}, \\
q'_{12} &= 1 - \frac{a(y) - \operatorname{cap}(A) + O\big(\frac{r\ln r\ln\|y\|}{\|y\|}\big)}{a(R) - \operatorname{cap}(A) + O\big(R^{-1}\|y\| + \frac{r\ln r\ln\|y\|}{\|y\|}\big)}.
\end{aligned}$$

After the analogous calculations, we obtain (4.7). □

Proof of relations (4.4)–(4.5) Formula (4.5) rephrases (6.5) with $A = \{0, x\}$. Identity (4.4) follows from the same proof as in Lemma 4.5 (i), using (**??**) instead of (**??**). $\square$

....

4.3 Exercises

Doob's h transforms of Lamperti's RWs?

Exercise 4.1 For one-dimensional random walk X_n with drift (i.e., it jumps to the left with probability $p \in \left(0, \frac{1}{2}\right)$ and to the right with probability $1 - p$), prove that (somewhat surprisingly) $|X_n|$ is a Markov chain, and calculate its transition probabilities.

5

Intermezzo: Soft local times and Poisson processes

5.1 Soft local times

Let us start with the following elementary question. Assume that X and Y are two random variables with the same support[1] but different distributions. Let $X_1, X_2, X_3, \ldots$ be a sequence of independent copies of X. Does there exist an infinite permutation (i.e., a bijection $\mathbb{N} \mapsto \mathbb{N}$) $\sigma = (\sigma(1), \sigma(2), \sigma(3), \ldots)$ such that the sequence $X_{\sigma(1)}, X_{\sigma(2)}, X_{\sigma(3)}, \ldots$ has the same law as the sequence $Y_1, Y_2, Y_3, \ldots$, a sequence of independent copies of Y? Of course, such a permutation should be random: if it is deterministic, then the permuted sequence would simply have the original law[2]. For constructing σ, one is allowed to use additional random variables (independent of the X-sequence) besides the realization of the X-sequence itself. As far as the author knows, constructing the permutation without using additional randomness (i.e., when the permutation is a *deterministic* function of the *random* sequence $X_1, X_2, X_3, \ldots$) is still an open problem, a rather interesting one.

As usual, when faced with such a question, one tries a "simple" case first, to see if it gives any insight on the general problem. For example, take X to be Binomial$(n, \frac{1}{2})$ and Y to be discrete Uniform$[0, n]$. One may even consider the case when X and Y are Bernoullis, with different probabilities of success. How can one obtain σ in these cases?

After some thought, one will come with the following solution, simple and straightforward: just generate the i.i.d. sequnce $Y_1, Y_2, Y_3, \ldots$ independently, then there is a permutation that sends

[1] Informally, the support of a random variable Z is the (minimal) set where it lives. Formally, it is the intersection of all *closed* sets F such that $\mathbb{P}[Z \in F] = 1$.

[2] even more, the permutation σ should depend on $X_1, X_2, X_3, \ldots$; if it is independent of the X-sequence, it is still easy to check that the permuted sequence has the original law.

X-sequence to the Y-sequence. Indeed (this argument works for any pair of discrete random variables with the same support), almost surely any possible value of X (and Y) occurs infinitely many times both in the X-sequence and the Y-sequence. It is then quite straightforward to see that there is a permutation that sends one sequence to the other.

Now, let us be honest with ourselves: this solution looks like cheating. In a way, it is simply *too easy.* Common wisdom tells us, however, that there ain't no such thing as a free solution; in this case, the problem is that the above construction does not work at all when the random variables are continuous. Indeed, if we generate the two sequences independently, then, almost surely, no element of the first sequence will be even present in the second one. So, a different approach is needed.

Later in this section, we will see how to solve the above problem using a sequence of i.i.d. Exponential random variables as additional randomness. The solution will come out as an elementary application of the method of *soft local times*, the main subject of this section. Generally speaking, the method of soft local times is a way to construct an adapted stochastic process on a general space Σ, using an auxiliary Poisson point process on $\Sigma \times \mathbb{R}_+$.

Naturally, we assume that the reader knows what is a Poisson point process in $\mathbb{R}^d$ with (not necessarily constant) rate λ. If one needs to consider a Poisson process on, say, $\mathbb{Z} \times \mathbb{R}$, then it is still easy to understand what exactly it should be (a union of Poisson processes on the straight lines indexed by the sites of $\mathbb{Z}$). In any case, all this fits into the Poissonian paradigm: what happens in a domain does not affect what is going on in a disjoint domain, the probability that there is exactly one point in a "small" domain of volume δ located "around" x is $\delta\lambda(x)$ (up to terms of smaller order), and the probability that there are at least two points in that small domain is $o(\delta)$. Here, the tradition dictates that the author cites a comprehensive book on the subject, so, [35].

Coming back to the soft local times method, we mention that, in full generality, it was introduced in [34]; see also [11, 12] which contain short surways of this method applied to constructions of excursion processes[3]. The idea of using projections of Poisson processes for constructions of other (point) processes is not new, see

[3] but this will be treated in a detailed way in Section 6.3.1 below.

e.g. [27, 20]. The key tool of this method (Lemma 5.1) appears in [42] in a simpler form, and the motivating example we gave in the beginning of this section is also from that paper.

Next, we are going to present the key result that makes the soft local times possible. Over here, we call it "the magic lemma". Assume that we have a space Σ, which has enough structure[4] that permits us to construct a Poisson point process on Σ of rate μ, where μ is a measure on Σ.

Now, we main object we need is the Poisson point process on $\Sigma \times \mathbb{R}_+$, with rate $\mu \otimes dv$, where dv is the Lebesgue measure on $\mathbb{R}_+$. At this point we have to write some formalities. In the next display, Ξ is a countable index set. We prefer not to use $\mathbb{Z}_+$ for the indexing, because we are not willing to fix any particular ordering of the points of the Poisson process for the reason that will become clear in a few lines. Let

$$\mathcal{M} = \Big\{\eta = \sum_{\varrho \in \Xi} \delta_{(z_\varrho, v_\varrho)}; z_\varrho \in \Sigma, v_\varrho \in \mathbb{R}_+, \\ \text{and } \eta(K) < \infty \text{ for all compact } K\Big\}, \quad (5.1)$$

be the set[5] of point configurations of this process. It is a general fact that one can canonically construct a Poisson point process η as above; see e.g. Proposition 3.6 on p.130 of [35] for details of this construction.

The result below is our "magic lemma": it provides us with a way to simulate a random element of Σ with law absolutely continuous with respect to μ, using the Poisson point process η. We first write it formally, and then explain, what does it mean.

Lemma 5.1 *Let $g : \Sigma \to \mathbb{R}_+$ be a measurable function with $\int g(z)\mu(dz) = 1$. For $\eta = \sum_{\varrho\in\Xi} \delta_{(z_\varrho, v_\varrho)} \in \mathcal{M}$, we define*

$$\xi = \inf\big\{t \geq 0;\ \textit{there exists } \varrho \in \Xi \textit{ such that } tg(z_\varrho) \geq v_\varrho\big\}. \quad (5.2)$$

Then, under the law $\mathbb{Q}$ of the Poisson point process η,

1. there exists a.s. a unique $\hat{\varrho} \in \Xi$ such that $\xi g(z_{\hat{\varrho}}) = v_{\hat{\varrho}}$,

[4] for example, the following is enough: let Σ be a locally compact and Polish metric space, and μ is a Radon measure (i.e., every compact set has finite μ-measure) on the measurable space $(\Sigma, \mathcal{B})$, where $\mathcal{B}$ is the Borel σ-algebra on Σ.

[5] endowed with sigma-algebra $\mathcal{D}$ generated by the evaluation maps $\eta \mapsto \eta(A)$, $A \in \mathcal{B} \otimes \mathcal{B}(\mathbb{R})$.

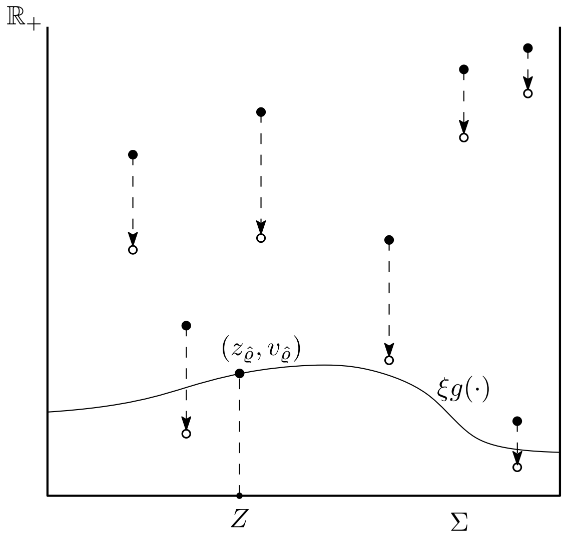

Figure 5.1 The magic lemma. The random variable Z has density g, $\bullet$'s are points of η, and $\circ$'s are points of the new Poisson process η'.

2. $(z_{\hat{\varrho}}, \xi)$ *is distributed as* $g(z)\mu(dz) \otimes \mathrm{Exp}(1)$,

3. $\eta' := \sum_{\varrho \neq \hat{\varrho}} \delta_{(z_\varrho, v_\varrho - \xi g(z_\varrho))}$ *has the same law as* η *and is independent of* $(\xi, \hat{\varrho})$.

That is, in plain words (see Figure 5.1):

- In (5.2) we define ξ as the smallest positive number such that there is exactly one point $(z_{\hat{\varrho}}, v_{\hat{\varrho}})$ of the Poisson process on the graph of $\xi g(\cdot)$, and nothing below this graph.
- The first coordinate Z of the chosen point is a random variable with density g (with respect to μ). Also, ξ is Exponential with parameter 1, and it is independent of Z.
- Remove the point that was chosen, and translate every other point (z, v) of η down by amount $\xi g(z)$. Call this new configuration η'. Then, η' is also a Poisson point process on $\Sigma \times \mathbb{R}_+$ with rate $\mu \otimes dv$, and it is independent of ξ and Z.

Sketch of the proof of Lemma 5.1. The formal proof can be found in [34] (Lemma 5.1 is Proposition 4.1 of [34]), and here we give only an informal argument to convince the reader that the above lemma is not only magic, but also true. In fact, this result is one

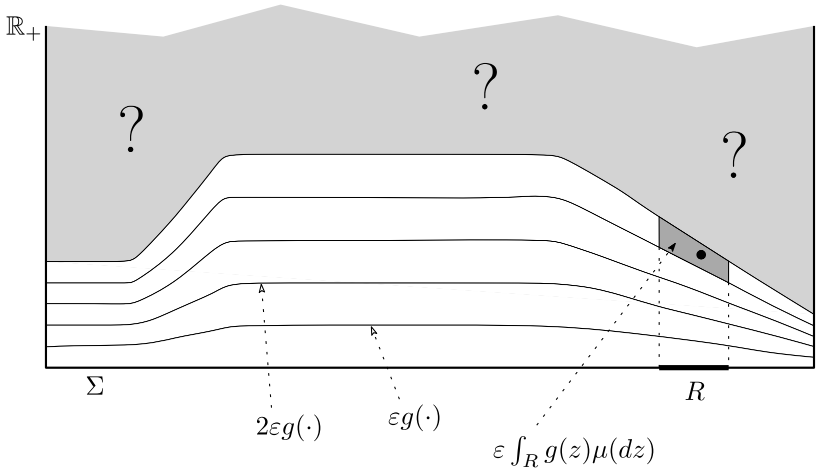

Figure 5.2 A slow exploration of the space: why Lemma 5.1 is valid

of those statements that become evident after one thinks about it for a couple of minutes; so, it may be a good idea for the reader to ponder on it for some time before going further.

So, one may convince oneself that the result holds e.g. in the following way. Fix a very small $\varepsilon > 0$ and let us *explore* the space as shown on Figure 5.2. That is, first look at the domain $\{(z, u) : u \leq \varepsilon g(z)\}$ and see if we find a point of the Poisson process there (observe that finding *two* points is highly improbable). If we don't, then we look at the domain $\{(z, u) : \varepsilon g(z) < u \leq 2\varepsilon g(z)\}$, and so on.

How many steps do we need to discover the first point? First, observe that g is a density, so it integrates to 1 with respect to μ, and therefore the area[6] below εg equals ε. So the number of points below εg is Exponential with rate ε, which means that on the first step (as well as on each subsequent one) we are successful with probability $1 - e^{-\varepsilon}$. Hence the number of steps N_ε until the first success is Geometric$(1 - e^{-\varepsilon})$. It is then quite straightforward to see that $\varepsilon N_\varepsilon$ converges in law to an Exponential random variable with parameter 1 as $\varepsilon \to 0$ (note that $1 - e^{-\varepsilon} = \varepsilon + o(\varepsilon)$ as $\varepsilon \to 0$). Therefore, ξ should indeed be Exponential(1).

The above fact could have been established in a direct way

[6] with respect to $\mu \otimes dv$

(note that $\mathbb{Q}[\xi > t]$ equals the probability that the set $\{(z,u) : u \leq tg(z)\}$ is empty, and the "volume" of that set is exactly t), but with an argument as above the questions about Z become more clear. Indeed, consider an arbitrary (measurable) set $R \subset \Sigma$. Then, on each step, we find a point with first coordinate in R with probability $1 - \exp\big(-\varepsilon \int_R g(z)\mu(dz)\big) = \varepsilon \int_R g(z)\mu(dz) + o(\varepsilon)$. Note that this probability does not depend on the number of steps already taken; that is, independently of the past, the conditional probability of finding a point with first coordinate in R given that something is found on the current step[7] is roughly $\int_R g(z)\mu(dz)$. This shows that ξ and Z are independent random variables.

As for the third part, simply observe that, at the time we discovered the first point, the shaded part on Figure 5.2 is still completely unexplored, and so its contents is independent of the pair (ξ, Z). In other words, we have a Poisson process on the set $\{(z,v) : v > \xi g(z)\}$ with the same rate, which can be transformed to the Poisson process in $\Sigma \times \mathbb{R}_+$ by subtracting $\xi g(\cdot)$ from the second coordinate (observe that such transformation is volume-preserving). □

Now, the key observation is that Lemma 5.1 allows us to construct virtually *any* discrete-time adapted stochastic process! Moreover, one can effectively couple two or more stochastic processes using the same realization of the Poisson process. One can better visualize the picture in a continuous space, so, to give a clear idea on how the method works, assume that we desire to obtain a realization of a sequence of (not necessarily independent nor Markovian) random variables $X_1, X_2, X_3, \ldots$ taking values in the interval $[0,1]$. Let us also construct *simultaneously* the sequence $Y_1, Y_2, Y_3, \ldots$, where (Y_k) are i.i.d. Uniform$[0,1]$ random variables, thus effectively obtaining a *coupling* of the X- and Y-sequences. We assume that the law of X_k conditioned on $\mathcal{F}_{k-1}$ is a.s. absolutely continuous with respect to the Lebesgue measure on $[0,1]$, where $\mathcal{F}_{k-1}$ is the sigma-algebra generated by $X_1, \ldots X_{k-1}$.

This idea of using the method of soft local times to couple (possibly complicated) stochastic processes with independent sequences already proved to be useful in many situations; for this book it will be useful as well, as we will see in Chapter 6.

[7] that is, we effectively condition on ξ, and show that the conditional law of Z does not depend on the value of ξ.

Our method for constructing a coupling of the X- and Y-sequences is illustrated on Figure 5.3. Consider a Poisson point process in $[0,1]\times\mathbb{R}_+$ with rate 1. Then, one can obtain a realization of the Y-sequence by simply ordering the points according to their second coodrinates, and then taking $Y_1, Y_2, Y_3, \dots$ to be the first coodrinates of these points. Now, to obtain a realization of the X-sequence using the same Poisson point process, one proceeds as follows.

- First, take the density $g(\cdot)$ of X_1 and multiply it by the unique positive number ξ_1 so that there is exactly one point of the Poisson process lying on the graph of $\xi_1 g$ and nothing strictly below it; X_1 is then the first coordinate of that point.
- Using Lemma 5.1, we see that, if we remove the point chosen on the previous step[8] and then translate every other point (z,u) of the Poisson process to $(z, u-\xi_1 g(z))$, then we obtain a Poisson process in $[0,1]\times\mathbb{R}_+$ which is independent of the pair (ξ_1, X_1).
- Thus, we are ready to use Lemma 5.1 again in order to construct X_2.
- So, consider the conditional density $g(\,\cdot \mid \mathcal{F}_1)$ of X_2 given $\mathcal{F}_1$ and find the smallest positive number ξ_2 in such a way that exactly one point lies on the graph of $\xi_2 g(\,\cdot \mid \mathcal{F}_1) + \xi_1 g(\cdot)$ and exactly one (the point we picked first) below it; again, X_2 is the first coordinate of the point that lies on the graph.
- Continue with $g(\,\cdot \mid \mathcal{F}_2)$, and so on.

The fact that the X-sequence obtained in this way has the prescribed law is readily justified by the subsequent application of Lemma 5.1. Now, let us state the formal result (it corresponds to Proposition 4.3 of [34]); here it is only a bit more general since we formulate it for general adapted processes.

Formally, for a general stochastic process $(Z_n, n\ge 0)$ adapted to a filtration $(\mathcal{F}_n, n\ge 0)$ we define

$$\begin{aligned}\xi_1 &= \inf\big\{t\ge 0:\ \text{there exists } \varrho\in\Xi \text{ such that } tg(z_\varrho)\ge v_\varrho\big\},\\ G_1(z) &= \xi_1 g(z\mid \mathcal{F}_0),\ \text{ for } z\in\Sigma,\end{aligned}$$

where $g(\,\cdot\mid\mathcal{F}_0)$ is the density of Z_1 given $\mathcal{F}_0$, and

$$(z_1, v_1) \text{ is the unique pair in } \{(z_\varrho, v_\varrho)\}_{\varrho\in\Xi} \text{ with } \xi_1 G_1(z_1) = v_1.$$

[8] this point has coordinates $(X_1, \xi_1 g(X_1))$

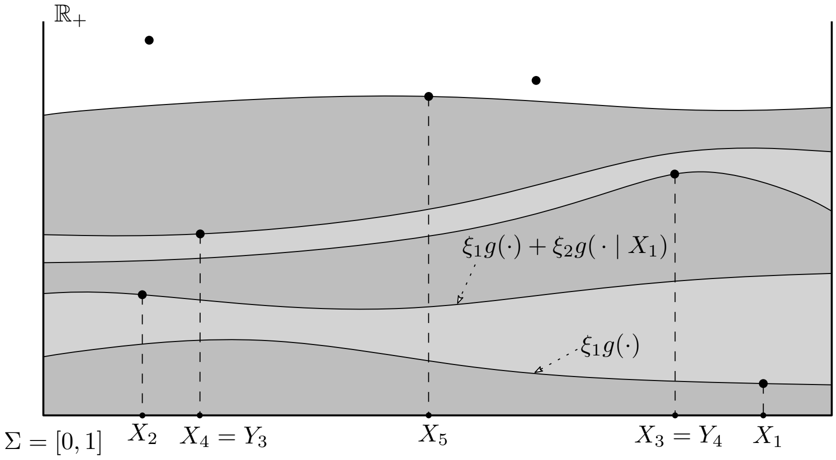

Figure 5.3 Soft local times: the simultaneous construction of the processes X and Y (here, $X_k = Y_k$ for $k = 1, 2, 5$); it is very important to observe that the points of the two processes need not necessarily appear in the same order with respect to the vertical axis.

(that is, we call 1 the corresponding element of Θ). Denote also $R_1 = \{(z_1, v_1)\}$. Then, for $n \geq 2$ we proceed inductively,

$$\begin{aligned} \xi_n &= \inf \big\{ t \geq 0 : \text{there exists } (z_\varrho, v_\varrho) \notin R_{n-1} \\ &\qquad\qquad \text{such that } G_{n-1}(z_\varrho) + t g(z_\varrho \mid \mathcal{F}_{n-1}) \geq v_\varrho \big\}, \\ G_n(z) &= G_{n-1}(z) + \xi_n g(z \mid \mathcal{F}_{n-1}), \end{aligned} \tag{5.3}$$

and

$$(z_n, v_n) \text{ is the unique pair } (z_\varrho, v_\varrho) \notin R_{n-1} \text{ with } G_n(z_\varrho) = v_\varrho;$$

also, set $R_n = R_{n-1} \cup \{(z_n, v_n)\}$. Then, the previous discussion implies that the following result holds:

Proposition 5.2 *It holds that*

(i) $(z_1, \ldots, z_n) \stackrel{law}{=} (Z_1, \ldots, Z_n)$ *and they are independent from* $\xi_1, \ldots, \xi_n$,

(ii) the point process

$$\sum_{(z_\varrho, v_\varrho) \notin R_n} \delta_{(z_\varrho, v_\varrho - G_n(z_\varrho))}$$

is distributed as η *and independent of the above,*

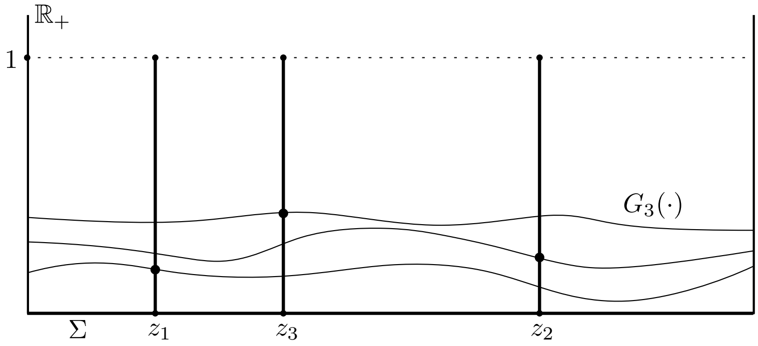

Figure 5.4 Comparison of soft and usual local times: the usual local time $L_3(\cdot)$ has three peaks (of size 1) at points z_1, z_2, z_3, and equals 0 in all other points. The soft one looks much softer.

for all $n \geq 1$.

We call G_n the *soft local time* of the process, at time n, with respect to the reference measure μ. To justify the choice of this name, consider a stochastic process in a finite or countable state space, and define the "usual" local time of the process by

$$L_n(z) = \sum_{k=1}^{n} \mathbf{1}\{X_k = z\}. \tag{5.4}$$

Now, just look at Figure 5.4.

Next, we establish a very important relation between these two different local times: their expectations are equal.

Proposition 5.3 *For all* $z \in \Sigma$ *it holds that*

$$\mathbb{E}G_n(z) = \mathbb{E}L_n(z) = \sum_{k=1}^{n} \mathbb{P}[X_k = z]. \tag{5.5}$$

Notice that in continuous space we cannot expect the above result to be true, since typically $\mathbb{E}L_n(z)$ would be just 0 for any z. Nevertheless, an analogous result holds in the general setting as well (cf. Theorem 4.6 of [34]), but, to formulate it properly, one would need to define the so-called *expected local time density* first (cf. (4.16) of [34]), which we prefer not to do here.

Proof of Proposition 5.3. It is an easy calculation that uses con-

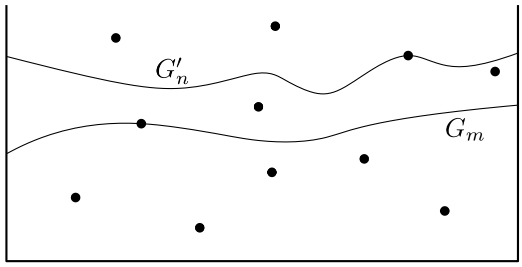

Figure 5.5 The set of Ys (with soft local time $G'_n(\cdot)$) contains all the Xs (with soft local time $G_m(\cdot)$) and three other points.

ditioning and induction. First, observe that $g(z \mid \mathcal{F}_{n-1}) = \mathbb{P}[X_n = z \mid \mathcal{F}_{n-1}]$, so we have

$$\mathbb{E}G_1(z) = \mathbb{E}(\mathbb{P}[X_1 = z \mid \mathcal{F}_0]) = \mathbb{P}[X_1 = z] = \mathbb{E}L_1(z).$$

Then, we proceed by induction: note that $G_{n-1}(z)$ is $\mathcal{F}_{n-1}$-measurable, and ξ_n is a mean-1 random variable which is independent of $\mathcal{F}_{n-1}$. Recall also (5.3) and write

$$\begin{aligned}
\mathbb{E}G_n(z) &= \mathbb{E}\big(\mathbb{E}(G_n(z) \mid \mathcal{F}_{n-1})\big)\\
&= G_{n-1}(z) + \mathbb{E}\big(\mathbb{E}(\xi_n g(z \mid \mathcal{F}_{n-1}) \mid \mathcal{F}_{n-1})\big)\\
&= G_{n-1}(z) + \mathbb{E}\big(g(z \mid \mathcal{F}_{n-1})\mathbb{E}(\xi_n \mid \mathcal{F}_{n-1})\big)\\
&= G_{n-1}(z) + \mathbb{E}\big(\mathbb{P}[X_n = z \mid \mathcal{F}_{n-1})\big)\\
&= G_{n-1}(z) + \mathbb{P}[X_n = z].
\end{aligned}$$

This concludes the proof. $\square$

As mentioned before, soft local times work really well for couplings of stochastic processes: indeed, just construct them in the way described above using *the same* realization of the Poisson point process. Observe that, for this coupling of the processes (X_n) and (Y_n) it holds that

$$\mathbb{P}\big[\{X_1, \dots, X_m\} \subset \{Y_1, \dots, Y_n\}\big] \geq \mathbb{P}\big[G_m(z) \leq G'_n(z) \text{ for all } z \in \Sigma\big], \tag{5.6}$$

where G' is the soft local time of Y, see Figure 5.5. Then, in principle, one may use large deviations tools to estimate the right-hand side of (5.6). One have to pay attention to the following, though: it is easy to see that the random variables $(\xi_1, \dots, \xi_m)$ are not independent of $(\xi'_1, \ldots, \xi'_n)$ (which enter to G'_n). This can be usually circumvented in the following way: we find a *deterministic* function $\varphi : \Sigma \mapsto \mathbb{R}$ which should typically be "between" G_m and G'_n, and then write

$$\mathbb{P}\big[G_m(z) \leq G'_n(z) \text{ for all } z \in \Sigma\big] \geq \mathbb{P}\big[G_m(z) \leq \varphi(z) \text{ for all } z \in \Sigma\big] \\ + \mathbb{P}\big[\varphi(z) \leq G'_n(z) \text{ for all } z \in \Sigma\big] - 1.$$

Note that, in the right-hand side of the above relation we do not have this "conflict of ξ's" anymore. Let us also mention that in the above large deviation estimates one has to deal with sequences of *random functions* (not just real-valued random variables). When the state space Σ is finite, this difficulty can be usually circumvented by considering the values of the functions separately in each point of Σ and then using the union bound, hoping that this last step would not cost too much. Otherwise, one has to do the large deviations for random functions directly using some advanced tools from the theory of empirical processes[9]; see e.g. Section 6 of [14] and Lemma 2.9 of [7] for examples of how large deviations for soft local times may be treated.

Now, finally, let us go back to the example from the beginning of this section: recall that we had a realization of an i.i.d. sequence $X_1, X_2, X_3, \ldots$, and we wanted to find an infinite permutation σ such that $X_{\sigma(1)}, X_{\sigma(2)}, X_{\sigma(3)}, \ldots$ is also an i.i.d. sequence, however, sampled from another distibution (with the same support). With Proposition 5.2 to hand, the solution is relatively simple. Take a sequence of i.i.d. Exponential(1) random variables $\xi_1, \xi_2, \xi_3, \ldots$; this sequence will serve as an additional randomness. As an example, let us consider the case when X is Uniform on $[-1, 1]$, and Y has the "triangular" density $f(y) = (1 - |y|)\mathbf{1}\{y \in [-1, 1]\}$. The first step is to *reconstruct* a Poisson process in $[-1, 1] \times \mathbb{R}_+$, using X's and ξ's. This can be done in the following way (see Figure 5.6): for all $n \geq 1$, put a point to $\big(X_n, \frac{1}{2}(\xi_1 + \cdots + \xi_n)\big)$. Then, using this Poisson process, we obtain the sequence $Y_1, Y_2, Y_3, \ldots$ of i.i.d. triangular random variables in the way described above; look at Figure 5.6 which speaks for itself. Clearly, one sequence is a permutation of the other: they use the same points! We leave as

[9] note that they have to be more advanced than the Talagrand's inequality (see e.g. ??? of [3]) since, because of these i.i.d. Exponential ξ's, the terms are not a.s. bounded

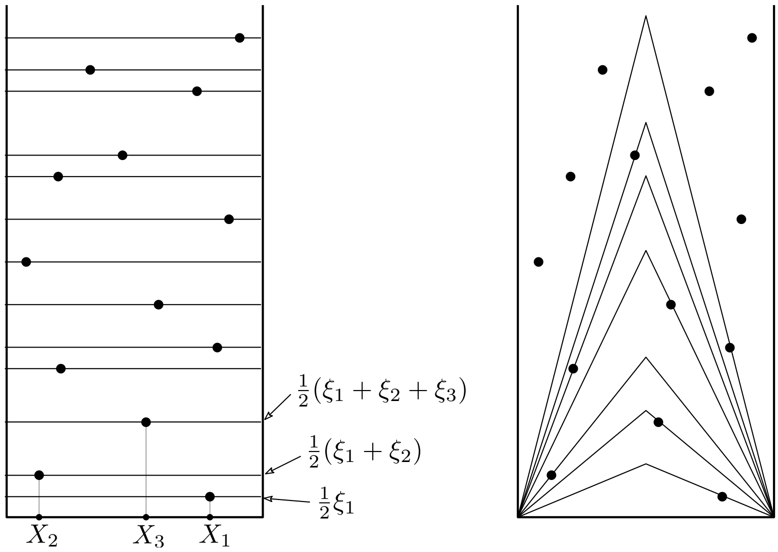

Figure 5.6 Making uniforms triangular. We first obtain a particular instance of the Poisson process in $[-1, 1] \times \mathbb{R}_+$ using the X-sequence, and then use the same collection of points to build the Y-sequence. It holds that $\sigma(1) = 1$, $\sigma(2) = 3$, $\sigma(3) = 2$, $\sigma(4) = 6$, $\sigma(5) = 4$, $\sigma(6) = 10$, $\sigma(7) = 5$.

an exercise for the reader to check that, *this time*, essentially the same solution works in the general case.

5.2 Poisson processes of things

Of course, all people know what is a Poisson process of points in $\mathbb{R}^d$. But what if we need a Poisson process of more complicated things, which still live in $\mathbb{R}^d$? What is the *right* way to define it? Naturally, we need the picture to be invariant with respect to isometries[10]. Also, it should be, well, *as independent as possible*, whatever it may mean.

Observe that, if those things are bounded (not necessarily uniformly), one can use the following natural procedure: take a d-dimensional Poisson *point* process of rate $\lambda > 0$, and "attach" the things to the points independently (as e.g. on Figure 5.7). A broad example of this is the Poisson Boolean model, cf. e.g. [28].

[10] translations, rotations, reflections, and combinations of them.

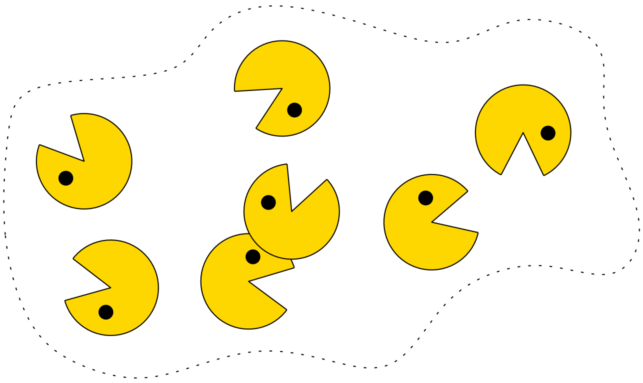

Figure 5.7 A Poisson process of *finite* things

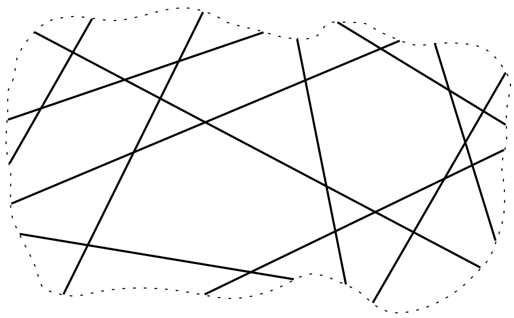

Figure 5.8 A Poisson line process (observed in the dotted domain) in $\mathbb{R}^2$

However, the situation becomes much more complicated if we need to build a Poisson process of *infinite* things. For example, what about a two-dimensional Poisson process of lines, which should look like as shown on Figure 5.8?

An idea that first comes to mind is simply to take a two-dimensional Poisson point process, and draw independent lines in random uniform directions through each point. One quickly realises, however, that this way we would rather see what is shown on Figure 5.9: there will be *too many* lines, one would obtain a dense set on the plane instead of the nice picture above. Another idea can be the following: first, fix a straight line on the plane (it can be the horizontal axis or just anything; it is the thicker line on

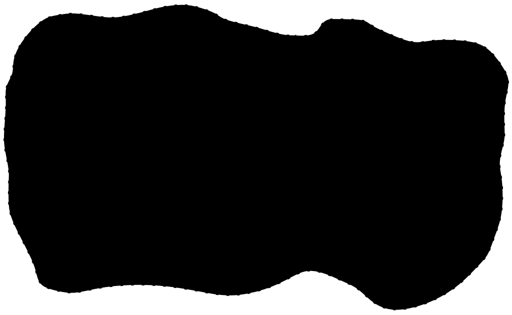

Figure 5.9 Too many lines!

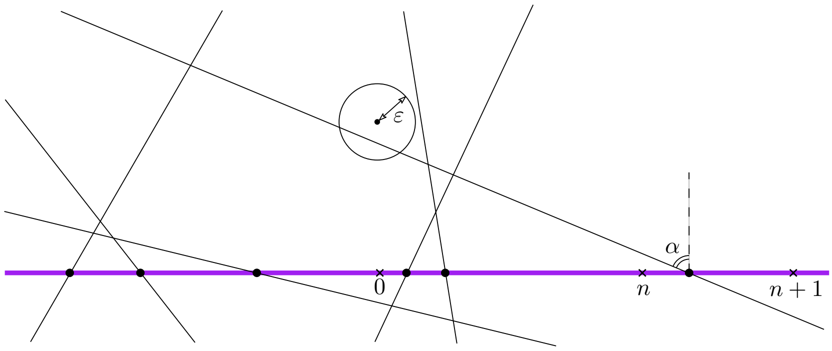

Figure 5.10 Constructing a Poisson line process using the reference line (here, $\alpha \in [-\pi/2, \pi/2]$ is the angle between the line and the normal vector)

Figure 5.10), and then consider a one-dimensional Poisson point process on this line. Then, through each of these points, draw a line with uniformly distributed direction (that is, α on Figure 5.10 is uniform in $[-\pi/2, \pi/2]$; for definiteness, think that the positive values of α are on the left side with respect to the normal vector pointing up) independently, thus obtaining the "process of lines" (not including the "reference" line) in $\mathbb{R}^2$.

Well, this looks as a reasonable procedure, but, in fact, it is not. Let us show that, as a result, we obtain a dense set again. Assume without loss of generality that the reference line is the horizontal axis, and consider a disk of radius $\varepsilon > 0$ situated somewhere above

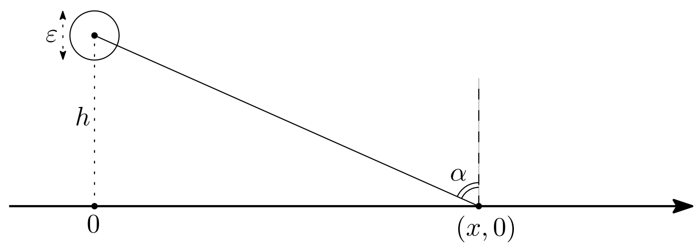

Figure 5.11 Calculating the expected number of lines that intersect the small ball. It holds that $\alpha = \arccos \frac{h}{\sqrt{x^2+h^2}}$.

the origin (as on Figure 5.10). For all $n \in \mathbb{Z}$, consider the events

$$H_n = \big\{\text{there is at least one line attached to a point in } [n, n+1), \\ \text{which intersects the disk}\big\}.$$

The events $(H_n, n \in \mathbb{Z})$ are independent by construction, and it not difficult to see that $\mathbb{P}[H_n] \asymp \varepsilon n^{-1}$ (indeed, for each point of $[n, n+1)$, the "angular size" of the disk as seen from that point is just of that order). Therefore, the divergence of the harmonic series[11] implies that a.s. this disk is crossed infinitely many times, and from this it is straightforward to obtain that the set of lines is dense.

Can this procedure be "repaired"? Well, examining the above argument, we see that the problem was that we gave "too much weight" to the angles α close to $\pm\pi/2$. Therefore, choosing the direction uniformly does not work, and hence we need to choose it with some other density $\varphi(\cdot)$ on $[-\pi/2, \pi/2]$ (of course, it should be symmetric with respect to 0, i.e., the direction of the normal).

What should be this φ? Consider a small disk of diameter ε situated at distance h above the origin, as on Figure 5.11. Consider a point $(x, 0)$ on the reference line (horizontal axis), with $x > 0$. Then, clearly, to intersect the disk, the direction of a straight line passing through x must be in $[\alpha - \frac{\delta}{2}, \alpha + \frac{\delta}{2}]$, where $\alpha = \arccos \frac{h}{\sqrt{x^2+h^2}}$ and (up to terms of smaller order) $\delta = \frac{\varepsilon}{\sqrt{x^2+h^2}}$. So, if λ is the rate of the Poisson point process on the reference line and $N(h, \varepsilon)$ is the mean number of lines intersecting the small

[11] this again! Why do we meet the harmonic series so frequently in two dimensions?...

ball, we have

$$\mathbb{E}N(h,\varepsilon) = \lambda\varepsilon \int_{-\infty}^{+\infty} \frac{\varphi\big(\arccos \frac{h}{\sqrt{x^2+h^2}}\big)}{\sqrt{x^2+h^2}}\, dx + o(\varepsilon). \tag{5.7}$$

This does not look very nice, but notice that, if we just erase "φ" and "arccos" from (5.7)[12], the integral would become

$$\int_{-\infty}^{+\infty} \frac{h}{x^2+h^2}\, dx = \int_{-\infty}^{+\infty} \frac{1}{\big(\frac{x}{h}\big)^2+1}\, d\Big(\frac{x}{h}\Big) = \int_{-\infty}^{+\infty} \frac{du}{u^2+1} = \pi,$$

so the parameter h disappears. And, actually, it is easy to get rid of φ and arccos at once: just choose $\varphi(\alpha) = \frac{1}{2}\cos\alpha$. So, we obtain from (5.7) that $\mathbb{E}N(h,\varepsilon) = \frac{1}{2}\lambda\varepsilon + o(\varepsilon)$, which is a good sign that $\varphi(\alpha) = \frac{1}{2}\cos\alpha$ may indeed work for defining the Poisson line process.

The above construction is obviously invariant with respect to translations *in the direction of the reference line*, and, *apparently*, in the other directions too (there is no dependence on h for the expectations, but still some formalities are missing), but what about the rotational invariance? This can be proved directly[13], but, instead of doing this now, let us consider another (in fact, more general[14]) approach to defining Poisson processes of things. The idea is to represent these things as *points in the parameter space*; i.e., each possible "thing" is described by a (unique) set of parameters, chosen in some convenient (and clever!) way. Then, we just take a Poisson *point* process in that parameter space, which a process of things naturally.

So, how can one carry this out in our case? Remember that we already constructed something translationally invariant, so let us try to find a parameter space where the rotational invariance would naturally appear. Note that any straight line that does not pass through the origin can be uniquely determined by two parameters: the distance r from the line to the origin, and the angle θ between the horizontal axis and the shortest segment linking the line to the origin. So, the idea is to take a realization of a Poisson point process (with some constant rate) in the parameter space

[12] and the two parentheses as well, although it not strictly necessary.

[13] please, try to do it!

[14] in fact, it is *the* approach.

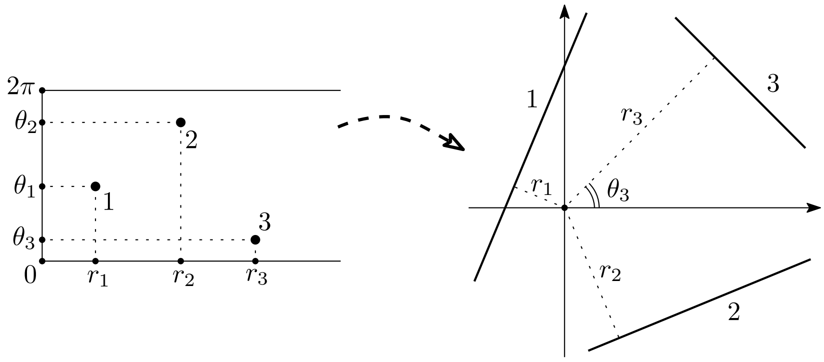

Figure 5.12 Constructing a Poisson line process as a point process in the parameter space.

$\mathbb{R}_+ \times [0, 2\pi)$, and translate it to a set of lines in $\mathbb{R}^2$, as shown on Figure 5.12.

Now, what kind of process do we obtain? First, it is clearly invariant under rotations. Secondly, it is not so obvious that it should be invariant with respect to translations. Instead of trying to prove it directly, we prefer to show that this construction is equivalent to the one with reference line (and hence get the translational invariance for free). Indeed, assume again that the reference line is the horizontal axis. Then (look at Figure 5.13) we have $\theta = \alpha$ and $dx = \cos\alpha\, dr$, so the probability that there is a line of the process crossing the reference line in the interval $[x, x + dx]$ (with respect to the first coordinate) and having the direction in the interval $[\alpha, \alpha + d\alpha]$ is proportional to $\cos\alpha\, dr d\alpha$, as required.

At this point we prefer to end this discussion and recommend the beautiful book [22] to an interested reader; in particular, that book contains a lot of information about Poisson processes of lines and (hyper)planes.

Finally, here is the general message of this section: it may be possible to construct something which can be naturally called a Poisson process of things, but the construction may be quite non-trivial. As for the Poisson line process itself, it serves as a supporting example for the previous sentence and as a “get-some-

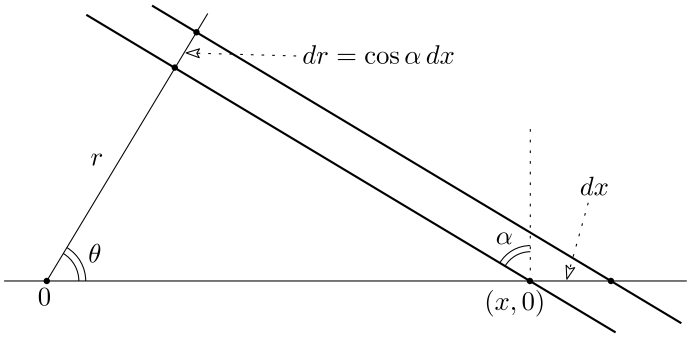

Figure 5.13 Equivalence of the two constructions.

intuition" example for the next chapter[15], but it is not *directly* connected to anything else in the rest of this book. There is one more reason, however, for its presence here: it is beautiful. As an additional argument in favor of the last affirmation, let us consider the following question: what is the distribution of the direction of a "typical" line from the Poisson line process? Well, it should obviously be uniform (the process is invariant under rotations, after all). Now, what is the distribution of the direction of a "typical" line intersecting the reference line? This time, it should obviously obey the cosine law. And here comes the paradox: a.s. *all* lines of the Poisson line process intersect the reference line, so we are talking about *the same sets of lines*! So, what is "the direction of a typical line", after all?

5.3 Exercises

Soft local times (Section 5.1):

Exercise 5.1 Look again at Figure 5.6. Can you find the value of $\sigma(8)$?

Exercise 5.2 Let $(X_i)_{i\geq 1}$ be a Markov chain on a finite set Σ, with transition probabilities $p(x, x')$, initial distribution π_0, and stationary measure π. Let A be a subset of Σ. Prove that for any

[15] in particular, the reader is invited to pay special attention to Exercises 5.11 and 5.12 below

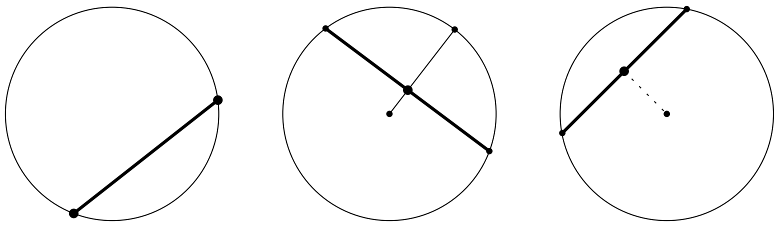

Figure 5.14 Constructing a random chord

$n \geq 1$ and $\lambda > 0$ it holds that

$$\mathbb{P}_{\pi_0}[\tau_A \leq n] \geq \mathbb{P}_{\pi_0}\Big[\xi_0 \pi_0(x) + \sum_{j=1}^{n-1} \xi_j p(X_j, x) \geq \lambda \pi(x), \text{ for all } x \in \Sigma\Big] - e^{-\lambda \pi(A)}, \tag{5.8}$$

where ξ_i are i.i.d. Exp(1) random variables, also independent of the Markov chain X.

Exercise 5.3 Find a nontrivial application of (5.8).

Exercise 5.4 Give a *rigorous* proof of Lemma 5.1 in case Σ is discrete (i.e., finite or countably infinite set).

Poisson processes of things (Section 5.2):

Exercise 5.5 Let us recall the Bertrand paradox: "what is the probability that a random chord of a circle is longer than the side of the inscribed equilateral triangle?".

The answer, of course, depends on how exactly we decide to choose the random chord. One may consider (at least) *three* apparently natural ways, see Figure 5.14 (from left to right):

1. choose two points uniformly and independently, and draw a chord between them;
2. first choose a radius[16] at random, then choose a random point on it (all that uniformly), and then draw the chord perpendicular to the radius through that point;
3. choose a random point *inside the disk* (note that, almost surely, that point will not be the center), and then draw the unique chord perpendicular to the corresponding radius;

[16] i.e., a straight line segment linking the center to a boundary point.

The author does not ask you to prove that the probability of the above event will be $\frac{1}{3}$, $\frac{1}{2}$, and $\frac{1}{4}$ respectively for the three above methods, since it is very easy. Instead, let me ask the following question: how to find *the right* way to choose a random chord (and therefore resolve the paradox)? One reasonable idea is to consider a Poisson line process and *condition* on the fact that only one line intersects the circle, so that this intersection generates the chord. To which of the three above ways it corresponds?

Exercise 5.6 Note that the uniform distribution on a finite set (or a subset of $\mathbb{R}^d$ with finite Lebesgue measure) has the following *characteristic property*: if we condition that the chosen point belongs to a fixed subset, then this conditional distribution is uniform again (on that subset).

Now, consider a (smaller) circle which lies fully inside the original circle, and *condition* that the random chord (that *you* defined above) of the bigger circle intersects the smaller one, thus generating a chord in it as well. Does this induced random chord have the *right* distribution?

Exercise 5.7 The above method of defining a random chord works for any convex domain. What do you think, is there a *right* way of defining a random chord for nonconvex (even nonconnected) domains?

Note that for such a domain one straight line can generate several chords at once.

Exercise 5.8 Explain the paradox in the end of Section 5.2.

Exercise 5.9 Argue that the above paradox has a lot to do with the motivating example of Section 5.1; in fact, show how one can generate the Poisson line process using the "strip" representation in two ways (with reference line, and without).

Exercise 5.10 (Random billiards) A particle moves with constant speed inside some (connected, but not necessarily simply connected) domain $\mathcal{D}$. When it hits the boundary, it is reflected in random direction according to the *cosine law*[17] (i.e., with density proportional to the cosine of the angle with the normal vector), independently of the incoming direction, and keeping the absolute value of its speed. Let $X_t \in \mathcal{D}$ be the location of the process at

[17] recall the construction of the Poisson line process that used the reference line.

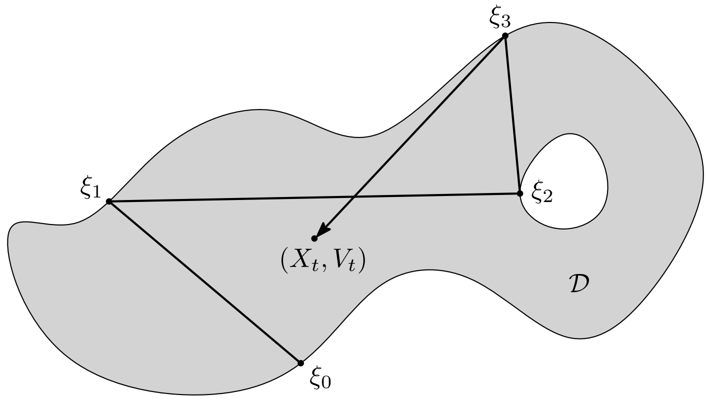

Figure 5.15 Random billiard (starting on the boundary of $\mathcal{D}$)

time t, and $V_t \in [0, 2\pi)$ be the corresponding direction; $\xi_n \in \partial\mathcal{D}$, $n = 0, 1, 2, \ldots$ are the points where the process hits the boundary, as shown on Figure 5.15.

Prove that

- the stationary measure of the random walk ξ_n is uniform on $\partial\mathcal{D}$;
- the stationary measure of the process (X_t, V_t) is the product of uniform measures on $\mathcal{D}$ and $[0, 2\pi)$.

Observe that this result holds for any (reasonable) domain $\mathcal{D}$!

The d-dimensional version of this process appeared in [36] under the name of "running shake-and-bake algorithm", and was subsequently studied in [8, 9, 10]. For some physical motivation for the cosine reflection law see e.g. [13] and references therein.

Exercise 5.11 Sometimes, instead of defining a Poisson process of infinite objects "as a whole", it is easier to define its image inside a finite "window". This is not the case for the Poisson line processes[18], but one can still do it. Let $A \subset \mathbb{R}^2$ be a convex domain. Prove that the following procedure defines a Poisson line process as seen in A: take a Poisson point process on ∂A, and then, independently for each of its points, trace a ray (pointing inside the domain) according to the cosine law.

Then, prove *directly* (i.e., forget about the Poisson line process

[18] I mean, it is not *easier*

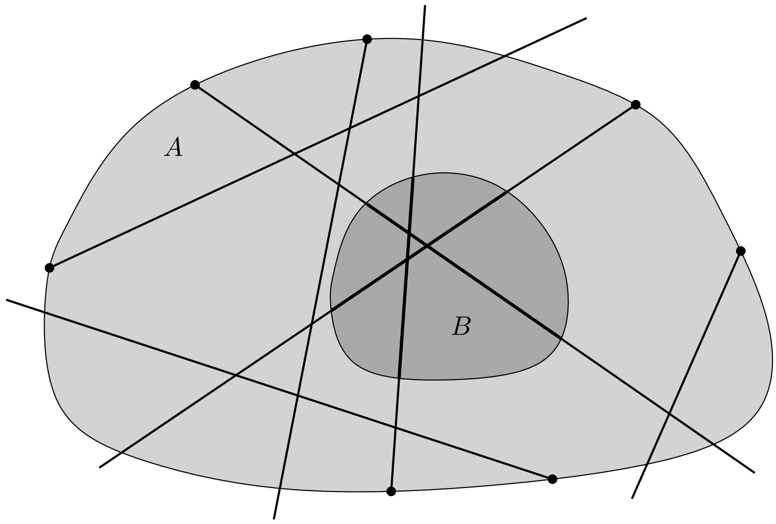

Figure 5.16 Poisson line process seen in a finite convex set

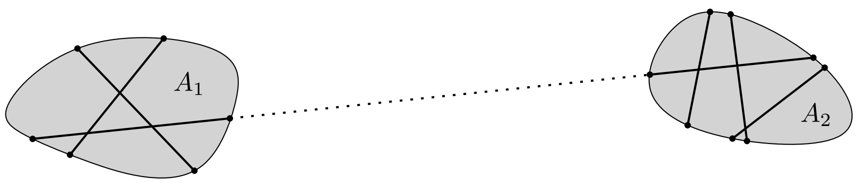

Figure 5.17 Poisson line process seen in two disjoint sets; if a straight line intersects A_1 the way shown on the picture, it *must* intersect A_2.

in $\mathbb{R}^2$ for now) that the above procedure is *consistent*: if $B \subset A$ is convex too, then the restriction of the process in A to B has the *right* law (i.e., the same as if we took a Poisson point process on ∂B with the same intensity, and traced rays from each of its points, traced independent rays with the cosine law), see Figure 5.16.

Exercise 5.12 Now, let us consider two nonintersecting domains $A_1, A_2 \subset \mathbb{R}^2$, and abbreviate $r = \max\{\text{diam}(A_1), \text{diam}(A_1)\}$, $s = \text{dist}(A_1, A_2)$. Consider a two-dimensional Poisson line process with rate λ. It is quite clear that the restrictions of this process on A_1 and A_2 are *not* independent, just look at Figure 5.17. However, in the case $s \gg r$ one still can *decouple* them. Let G_1 and G_2 be two events *supported* on A_1 and A_2. This means that, informally speaking, the occurrence of the event G_k is determined by the configuration seen on A_k, for $k = 1, 2$. Prove that, for some positive constant C we have

$$\big|\mathbb{P}[G_1 \cap G_2] - \mathbb{P}[G_1]\mathbb{P}[G_2]\big| \leq \frac{C\lambda r}{s}. \tag{5.9}$$

Exercise 5.13 Find the expected value of the orthogonal projection of the unit cube on a randomly oriented plane.

6

Two-dimensional random interlacements

6.1 Introduction: random interlacements in dimension $d \geq 3$

Explain about RI for $d \geq 3$.

Mention that *all* one-dimensional Poisson processes with constant rate can be constructed at once, as projections of a Poisson process with rate 1 in $\mathbb{R} \times \mathbb{R}_+$, as on Figure 6.1.

....

Random interlacements were introduced by Sznitman in [38], motivated by the problem of disconnection of the discrete torus $\mathbb{Z}_n^d := \mathbb{Z}^d / n\mathbb{Z}^d$ by the trace of simple random walk, in dimension 3 or higher. Detailed accounts can be found in the survey [6] and the recent books [17, 39]. Loosely speaking, the model of random interlacements in $\mathbb{Z}^d$, $d \geq 3$, is a stationary Poissonian soup of (transient) doubly infinite simple random walk trajectories on the

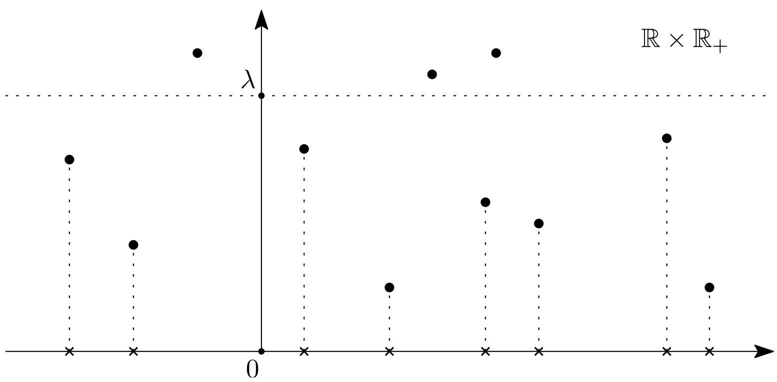

Figure 6.1 A simultaneous construction of one-dimensional Poisson processes

integer lattice. There is an additional parameter $u > 0$ entering the intensity measure of the Poisson process, the larger u is the more trajectories are thrown in. The sites of $\mathbb{Z}^d$ that are not touched by the trajectories constitute the *vacant set* $\mathcal{V}^u$. The random interlacements are constructed simultaneously for all $u > 0$ in such a way that $\mathcal{V}^{u_1} \subset \mathcal{V}^{u_2}$ if $u_1 > u_2$. In fact, the law of the vacant set at level u can be uniquely characterized by the following identity:

$$\mathbb{P}[A \subset \mathcal{V}^u] = \exp\big(-u\,\mathrm{cap}(A)\big) \quad \text{for all finite } A \subset \mathbb{Z}^d, \tag{6.1}$$

where $\mathrm{cap}(A)$ is the capacity of A, recall ... (where it was defined).

....

6.2 The two-dimensional case

....

At first glance, the title of this section seems to be meaningless, just because even a single trajectory of two-dimensional simple random walk a.s. visits all sites of $\mathbb{Z}^2$, so the vacant set would be always empty. Nevertheless, there is also a natural notion of capacity in two dimensions (cf. Section 6.6 of [26]), so one may wonder if there is a way to construct a decreasing family $(\mathcal{V}^\alpha, \alpha > 0)$ of random subsets of $\mathbb{Z}^2$ in such a way that a formula analogous to (6.1) holds for every finite A. This is, however, clearly not possible since the two-dimensional capacity of one-point sets equals 0. On the other hand, it turns out to be possible to construct such a family so that

$$\mathbb{P}[A \subset \mathcal{V}^\alpha] = \exp\big(-\pi\alpha\,\mathrm{cap}(A)\big) \tag{6.2}$$

holds *for all sets containing the origin* (the factor π in the exponent is just for convenience, as explained below). We present this construction in Section **??**. To build the interlacements, we use trajectories of simple random walks conditioned on never hitting the origin. Of course, the law of the vacant set is no longer translationally invariant, but we show that it has the property of *conditional* translation invariance, cf. Theorem 6.2 below. In addition, we will see that (similarly to the $d \geq 3$ case) the random object we construct has strong connections to random walks on two-dimensional torus. All this makes us believe that “two-dimensional random interlacements” is the right term for the object we introduce in this paper.

....

Our next definitions are appropriate for the transient case. For a finite $A \subset \mathbb{Z}^2$, we define the *equilibrium measure*

$$\widehat{e}_A(x) = \mathbf{1}\{x \in A\}\widehat{P}_x[\widehat{\tau}_1(A) = \infty]\mu_x, \tag{6.3}$$

and the capacity (with respect to $\widehat{S}$)

$$\widehat{\mathrm{cap}}(A) = \sum_{x \in A} \widehat{e}_A(x). \tag{6.4}$$

Observe that, since $\mu_0 = 0$, it holds that $\widehat{\mathrm{cap}}(A) = \widehat{\mathrm{cap}}(A \cup \{0\})$ for any set $A \subset \mathbb{Z}^2$.

Now, we use the general construction of random interlacements on a transient weighted graph introduced in [41]. In the following few lines we briefly summarize this construction. Let W be the space of all doubly infinite nearest-neighbour transient trajectories in $\mathbb{Z}^2$,

$$\begin{aligned} W = \big\{ &\varrho = (\varrho_k)_{k \in \mathbb{Z}} : \varrho_k \sim \varrho_{k+1} \text{ for all } k; \\ &\qquad \text{the set } \{m : \varrho_m = y\} \text{ is finite for all } y \in \mathbb{Z}^2 \big\}. \end{aligned}$$

We say that ϱ and ϱ' are equivalent if they coincide after a time shift, i.e., $\varrho \sim \varrho'$ when there exists k such that $\varrho_{m+k} = \varrho_m$ for all m. Then, let $W^* = W/\sim$ be the space of trajectories modulo time shift, and define χ^* to be the canonical projection from W to W^*. For a finite $A \subset \mathbb{Z}^2$, let W_A be the set of trajectories in W that intersect A, and we write W_A^* for the image of W_A under χ^*. One then constructs the random interlacements as Poisson point process on $W^* \times \mathbb{R}^+$ with the intensity measure $\nu \otimes du$, where ν is described in the following way. It is the unique sigma-finite measure on W^* such that for every finite A

$$\mathbf{1}_{W_A^*} \cdot \nu = \chi^* \circ Q_A,$$

where the finite measure Q_A on W_A is determined by the following equality:

$$Q_A\big[(\varrho_k)_{k \geq 1} \in F, \varrho_0 = x, (\varrho_{-k})_{k \geq 1} \in G\big] = \widehat{e}_A(x) \cdot \widehat{P}_x[F] \cdot \widehat{P}_x[G \mid \widehat{\tau}_1(A) = \infty].$$

The existence and uniqueness of ν was shown in Theorem 2.1 of [41].

For a configuration $\sum_\lambda \delta_{(w_\lambda^*, u_\lambda)}$ of the above Poisson process, the process of random interlacements at level α (which will be referred to as RI(α)) is defined as the set of trajectories with label less than or equal to $\pi\alpha$, i.e.,

$$\sum_{\lambda: u_\lambda \le \pi\alpha} \delta_{w^*_\lambda} .$$

Observe that this definition is somewhat unconventional (we used $\pi\alpha$ instead of just α, as one would normally do), but we will see below that it is quite reasonable in two dimensions, since the formulas become generally cleaner.

It is important to have in mind the following "constructive" description of random interlacements at level α "observed" on a finite set $A \subset \mathbb{Z}^2$. Namely,

- take a Poisson($\pi\alpha\,\widehat{\mathrm{cap}}(A)$) number of particles;
- place these particles on the boundary of A independently, with distribution $\bar{e}_A = \big((\widehat{\mathrm{cap}}\,A)^{-1}\widehat{e}_A(x), x \in A\big)$;
- let the particles perform independent $\widehat{S}$-random walks (since $\widehat{S}$ is transient, each walk only leaves a finite trace on A).

It is also worth mentioning that the FKG inequality holds for random interlacements, cf. Theorem 3.1 of [41].

The *vacant set* at level α,

$$\mathcal{V}^\alpha = \mathbb{Z}^2 \setminus \bigcup_{\lambda: u_\lambda \le \pi\alpha} \omega^*_\lambda(\mathbb{Z}),$$

is the set of lattice points not covered by the random interlacement. It contains the origin by definition. In Figure 6.2 we present a simulation of the vacant set for different values of the parameter.

As a last step, we need to show that we have indeed constructed the object for which (6.2) is verified. For this, we need to prove the following fact:

Proposition 6.1 *For any finite set $A \subset \mathbb{Z}^2$ such that $0 \in A$ it holds that* $\mathrm{cap}(A) = \widehat{\mathrm{cap}}(A)$.

Proof Indeed, consider an arbitrary $x \in \partial A$, $x \neq 0$, and (large) r such that $A \subset B(r-2)$. Write using (3.19)

$$\widehat{P}_x\big[\widehat{\tau}_1(A) > \widehat{\tau}_1\big(\partial B(r)\big)\big] = \sum_{\varrho} \frac{a(\varrho_{\mathrm{end}})}{a(x)} \Big(\frac{1}{4}\Big)^{|\varrho|}$$

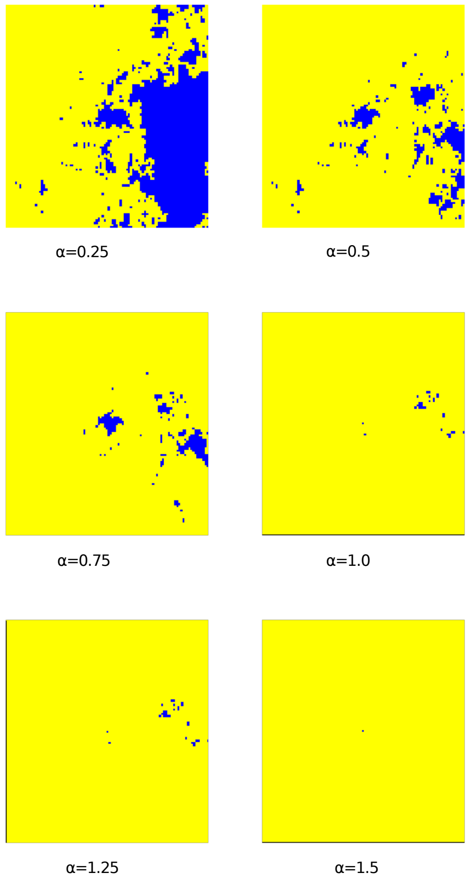

Figure 6.2 A realization of the vacant set (dark blue) of $\mathrm{RI}(\alpha)$ for different values of α. For $\alpha = 1.5$ the only vacant site is the origin. Also, note that we see the same neighbourhoods of the origin for $\alpha = 1$ and $\alpha = 1.25$; this is not surprising since just a few new walks enter the picture when increasing the rate by a small amount.

$$= \big(1 + o(1)\big)\frac{\frac{2}{\pi}\ln r}{a(x)}\sum_{\varrho}\Big(\frac{1}{4}\Big)^{|\varrho|}$$

$$= \big(1+o(1)\big)\frac{\frac{2}{\pi}\ln r}{a(x)}\,P_x\big[\tau_1(A) > \tau_1\big(\partial B(r)\big)\big],$$

where the sums are taken over all trajectories ϱ that start at x, end at $\partial B(r)$, and avoid $A\cup\partial B(r)$ in between; $\varrho_{\text{end}}\in\partial B(r)$ stands for the ending point of the trajectory, and $|\varrho|$ is the trajectory's length. Now, we send r to infinity and use (3.25) to obtain that, if $0\in A$,

$$a(x)\widehat{P}_x[\widehat{\tau}_1(A)=\infty] = \mathrm{hm}_A(x). \tag{6.5}$$

Multiplying by $a(x)$ and summing over $x\in A$ (recall that $\mu_x = a^2(x)$) we obtain the expressions in (3.26) and (6.4) and thus conclude the proof. □

Together with formula (1.1) of [41], Proposition 6.1 shows the fundamental relation (6.2) announced in introduction: for all finite subsets A of $\mathbb{Z}^2$ containing the origin,

$$\mathbb{P}[A\subset\mathcal{V}^\alpha] = \exp\big(-\pi\alpha\,\mathrm{cap}(A)\big).$$

As mentioned before, the law of two-dimensional random interlacements is not translationally invariant, although it is of course invariant with respect to reflections/rotations of $\mathbb{Z}^2$ that preserve the origin. Let us describe some other basic properties of two-dimensional random interlacements:

Theorem 6.2 *(i) For any $\alpha>0$, $x\in\mathbb{Z}^2$, $A\subset\mathbb{Z}^2$, it holds that*

$$\mathbb{P}[A\subset\mathcal{V}^\alpha \mid x\in\mathcal{V}^\alpha] = \mathbb{P}[-A+x\subset\mathcal{V}^\alpha \mid x\in\mathcal{V}^\alpha]. \tag{6.6}$$

More generally, for all $\alpha>0$, $x\in\mathbb{Z}^2\setminus\{0\}$, $A\subset\mathbb{Z}^2$, and any lattice isometry M exchanging 0 and x, we have

$$\mathbb{P}[A\subset\mathcal{V}^\alpha \mid x\in\mathcal{V}^\alpha] = \mathbb{P}[MA\subset\mathcal{V}^\alpha \mid x\in\mathcal{V}^\alpha]. \tag{6.7}$$

(ii) With γ' from (3.19) we have

$$\mathbb{P}[x\in\mathcal{V}^\alpha] = \exp\Big(-\pi\alpha\frac{a(x)}{2}\Big) = e^{-\gamma'\pi\alpha/2}\|x\|^{-\alpha}\big(1+O(\|x\|^{-2})\big). \tag{6.8}$$

(iii) For A such that $0\in A\subset B(r)$ and $x\in\mathbb{Z}^2$ such that $\|x\|\geq 2r$ we have

$$\mathbb{P}[A\subset\mathcal{V}^\alpha \mid x\in\mathcal{V}^\alpha] = \exp\left(-\frac{\pi\alpha}{4}\,\mathrm{cap}(A)\frac{1+O\big(\frac{r\ln r\ln\|x\|}{\|x\|}\big)}{1-\frac{\mathrm{cap}(A)}{2a(x)}+O\big(\frac{r\ln r}{\|x\|}\big)}\right). \tag{6.9}$$

(iv) For $x, y \neq 0$, $x \neq y$, we have $\mathbb{P}\big[\{x,y\} \subset \mathcal{V}^\alpha\big] = \exp\big(-\pi\alpha\Psi\big)$, *where*

$$\Psi = \frac{a(x)a(y)a(x-y)}{a(x)a(y) + a(x)a(x-y) + a(y)a(x-y) - \frac{1}{2}\big(a^2(x) + a^2(y) + a^2(x-y)\big)}.$$

Moreover, as $s := \|x\| \to \infty$, $\ln\|y\| \sim \ln s$ *and* $\ln\|x-y\| \sim \beta\ln s$ *with some* $\beta \in [0,1]$, *we have*

$$\mathbb{P}\big[\{x,y\} \subset \mathcal{V}^\alpha\big] = s^{-\frac{4\alpha}{4-\beta}+o(1)}. \tag{6.10}$$

(v) Assume that $\ln\|x\| \sim \ln s$, $\ln r \sim \beta \ln s$ *with* $\beta < 1$. *Then, as* $s \to \infty$,

$$\mathbb{P}\big[B(x,r) \subset \mathcal{V}^\alpha\big] = s^{-\frac{2\alpha}{2-\beta}+o(1)}. \tag{6.11}$$

These results invite a few comments.

Remark 6.3 1. The statement in (i) describes an invariance property given that a point is vacant. We refer to it as the conditional stationarity of two-dimensional random interlacements.

2. We can interpret (iii) as follows: the conditional law of RI(α) given that a distant site x is vacant, is similar – near the origin – to the unconditional law of RI($\alpha/4$). Combined with (i), the similarity holds near x as well. Moreover, one can also estimate the "local rate" away from the origin, see Figure 6.3. More specifically, observe from Lemma **??** (ii) that $\mathrm{cap}(A_2) \ll \ln s$ with $s = \mathrm{dist}(0, A_2)$ large implies $\mathrm{cap}\big(\{0\}\cup A_2\big) = \frac{a(s)}{2}(1+o(1))$. If x is at a much larger distance from the origin than A_2, say $\ln\|x\| \sim \ln(s^2)$, then (6.9) reveals a "local rate" equal to $\frac{2}{7}\alpha$, that is, $\mathbb{P}[A_2 \subset \mathcal{V}^\alpha \mid x \in \mathcal{V}^\alpha] = \exp\big(-\frac{2}{7}\pi\alpha\,\mathrm{cap}\big(\{0\}\cup A_2\big)(1+o(1))\big)$; indeed, the expression in the denominator in (6.9) equals approximately $1 - \frac{\mathrm{cap}(\{0\}\cup A_2)}{2a(x)} \approx 1 - \frac{a(s)/2}{2a(s^2)} \approx \frac{7}{8}$.

3. By symmetry, the conclusion of (iv) remains the same in the situation when $\ln\|x\|, \ln\|x-y\| \sim \ln s$ and $\ln\|y\| \sim \beta\ln s$.

Proof of (i) and (ii) To prove (i), observe that

$$\mathrm{cap}\big(\{0,x\}\cup A\big) = \mathrm{cap}\big(\{0,x\}\cup(-A+x)\big)$$

by symmetry. For the second statement in (i), note that, for $A' = \{0,x\}\cup A$, it holds that $\mathrm{cap}\big(A'\big) = \mathrm{cap}\big(MA'\big) = \mathrm{cap}\big(\{0,x\}\cup MA\big)$. Item (ii) follows from the above mentioned fact that $\mathrm{cap}\big(\{0,x\}\big) = \frac{1}{2}a(x)$ together with (3.19). □

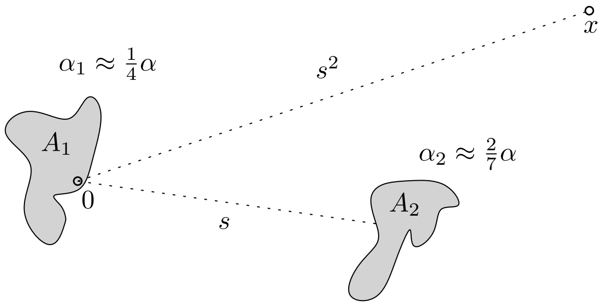

Figure 6.3 How the "local rate" looks like if we condition on the event that a "distant" site is vacant.

We postpone the proof of other parts of Theorem 6.2, since it requires some estimates for capacities of various kinds of sets. We now turn to estimates on the cardinality of the vacant set.

Theorem 6.4 *(i) We have*

$$\mathbb{E}\big(|\mathcal{V}^\alpha \cap B(r)|\big) \sim \begin{cases} \frac{2\pi}{2-\alpha} e^{-\gamma'\pi\alpha/2} \times r^{2-\alpha}, & \textit{for } \alpha < 2, \\ 2\pi e^{-\gamma'\pi\alpha/2} \times \ln r, & \textit{for } \alpha = 2, \\ \textit{const}, & \textit{for } \alpha > 2. \end{cases}$$

(ii) For $\alpha > 1$ it holds that $\mathcal{V}^\alpha$ is finite a.s. Moreover, $\mathbb{P}\big[\mathcal{V}^\alpha = \{0\}\big] > 0$ and $\mathbb{P}\big[\mathcal{V}^\alpha = \{0\}\big] \to 1$ as $\alpha \to \infty$.

(iii) For $\alpha \in (0,1)$, we have $|\mathcal{V}^\alpha| = \infty$ a.s. Moreover,

$$\mathbb{P}\big[\mathcal{V}^\alpha \cap \big(B(r) \setminus B(r/2)\big) = \emptyset\big] \le r^{-2(1-\sqrt{\alpha})^2 + o(1)}. \tag{6.12}$$

It is worth noting that the "phase transition" at $\alpha = 1$ in (ii) corresponds to the cover time of the torus, as shown in Theorem **??** below.

Proof of (i) and (ii) (incomplete, in the latter case) Part (i) immediately follows from Theorem 6.2 (ii).

The proof of the part (ii) is easy in the case $\alpha > 2$. Indeed, observe first that $\mathbb{E}|\mathcal{V}^\alpha| < \infty$ implies that $\mathcal{V}^\alpha$ itself is a.s. finite. Also, Theorem 6.2 (ii) actually implies that $\mathbb{E}|\mathcal{V}^\alpha \setminus \{0\}| \to 0$ as $\alpha \to \infty$, so $\mathbb{P}\big[\mathcal{V}^\alpha = \{0\}\big] \to 1$ by the Chebyshev inequality.

Now, let us prove that, in general, $\mathbb{P}\big[|\mathcal{V}^\alpha| < \infty\big] = 1$ implies that $\mathbb{P}\big[\mathcal{V}^\alpha = \{0\}\big] > 0$. Indeed, if $\mathcal{V}^\alpha$ is a.s. finite, then one can find a sufficiently large R such that $\mathbb{P}\big[|\mathcal{V}^\alpha \cap (\mathbb{Z}^2 \setminus B(R))| = 0\big] > 0$.

Since $\mathbb{P}[x \notin \mathcal{V}^\alpha] > 0$ for any $x \neq 0$, the claim that $\mathbb{P}\big[\mathcal{V}^\alpha = \{0\}\big] > 0$ follows from the FKG inequality applied to events $\{x \notin \mathcal{V}^\alpha\}$, $x \in B(R)$ together with $\big\{|\mathcal{V}^\alpha \cap (\mathbb{Z}^2 \setminus B(R))| = 0\big\}$. $\square$

As before, we postpone the proof of part (iii) and the rest of part (ii) of Theorem 6.4. Let us remark that we believe that the right-hand side of (6.12) gives the correct order of decay of the above probability; we, however, do not have a rigorous argument at the moment. Also, note that the question whether $\mathcal{V}^1$ is a.s. finite or not, is open.

Let us now give a heuristic explanation about the unusual behaviour of the model for $\alpha \in (1,2)$: in this non-trivial interval, the vacant set is a.s. finite but its expected size is infinite. The reason is the following: the number of $\widehat{S}$-walks that hit $B(r)$ has Poisson law with rate of order $\ln r$ (recall (3.27)). Thus, decreasing this number by a constant factor (with respect to the expectation) has only a polynomial cost. On the other hand, by doing so, we increase the probability that a site $x \in B(r)$ is vacant for all $x \in B(r)$ at once, which increases the expected size of $\mathcal{V}^\alpha \cap B(r)$ by a polynomial factor. It turns out that this effect causes the actual number of uncovered sites in $B(r)$ to be typically of much smaller order then the expected number of uncovered sites there.

....

6.3 Proofs for random interlacements

6.3.1 Excursions and soft local times

In this section we will develop some tools for dealing with excursions of two-dimensional random interlacements and random walks on tori; in particular, one of our goals is to construct a coupling between the set of RI's excursions and the set of excursions of the simple random walk X on the torus $\mathbb{Z}_n^2 = \mathbb{Z}^2/n\mathbb{Z}^2$.

First, if $A \subset A'$ are (finite) subsets of $\mathbb{Z}^2$ or $\mathbb{Z}_n^2$, then the excursions between ∂A and $\partial A'$ are pieces of nearest-neighbour trajectories that begin on ∂A and end on $\partial A'$, see Figure 6.4, which is, hopefully, self-explanatory. We refer to Section 3.4 of [12] for formal definitions. Here and in the sequel we denote by $(Z^{(i)}, i \geq 1)$ the (complete) excursions of the walk X between ∂A and $\partial A'$, and by $(\widehat{Z}^{(i)}, i \geq 1)$ the RI's excursions between ∂A and $\partial A'$ (depen-

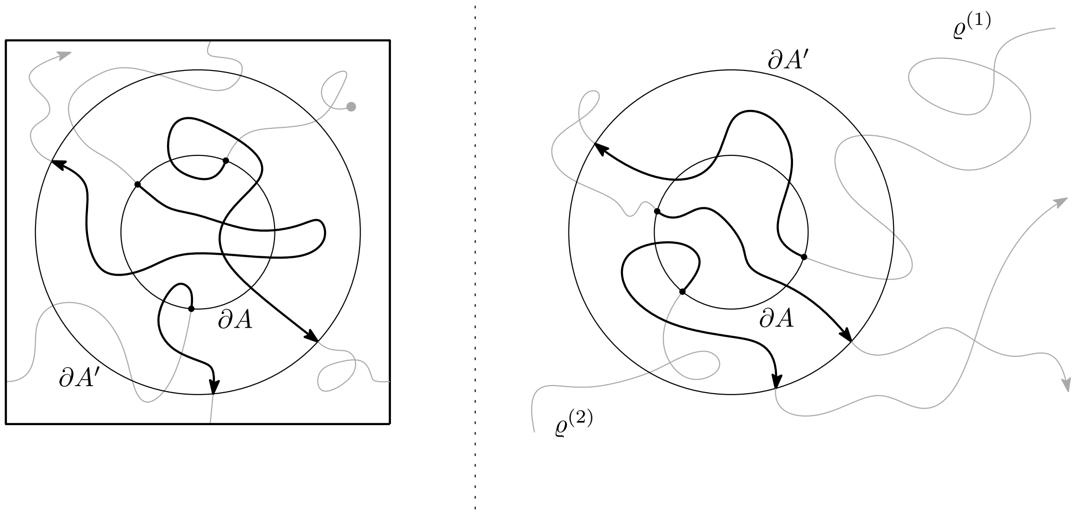

Figure 6.4 picture way too wide! Excursions (pictured as bold pieces of trajectories) for simple random walk on the torus (on the left), and random interlacements (on the right). Note the walk "jumping" from right side of the square to the left one, and from the bottom one to the top one (the torus is pictured as a square). For random interlacements, two trajectories, $\varrho^{1,2}$, intersect the set A; the first trajectory produces two excursions, and the second only one.

dence on n, A, A' is not indicated in these notations when there is no risk of confusion).

Now, assume that we want to construct the excursions of $\mathrm{RI}(\alpha)$, say, between $\partial\mathsf{B}(y_0, n)$ and $\partial\mathsf{B}(y_0, cn)$ for some $c > 0$ and $y_0 \in \mathbb{Z}^2$. Also, (let us identify the torus $\mathbb{Z}^2_{n_1}$ with the square of size n_1 centered in the origin of $\mathbb{Z}^2$) we want to construct the excursions of the simple random walk on the torus $\mathbb{Z}^2_{n_1}$ between $\partial\mathsf{B}(y_0, n)$ and $\partial\mathsf{B}(y_0, cn)$, where $n_1 > n+1$. It turns out that one may build both sets of excursions simultaneously on the same probability space, in such a way that, typically, most of the excursions are present in both sets (obviously, after a translation by y_0). This is done using the *soft local times* method; we refer to Section 4 of [34] for the general theory (see also Figure 1 of [34] which gives some quick insight on what is going on), and also to Section 2 of [11]. Here, we describe the soft local times approach in a less formal way. Assume, for definiteness, that we want to construct the simple random walk's excursions on $\mathbb{Z}^2_{n_1}$, between ∂A and $\partial A'$, and suppose that the starting point x_0 of the walk X does not belong to A.

We first describe our approach for the case of the torus. For $x \notin A$ and $y \in \partial A$ let us denote $\varphi(x, y) = \mathbb{P}_x[X_{\tau_1(A)} = y]$. For an excursion Z let $\iota(Z)$ be the first point of this excursion, and $\ell(Z)$ be the last one; by definition, $\iota(Z) \in \partial A$ and $\ell(Z) \in \partial A'$. Clearly, for the random walk on the torus, the sequence $\big((\iota(Z^{(j)}), \ell(Z^{(j)})), j \geq 1\big)$ is a Markov chain with transition probabilities

$$P_{(y,z),(y',z')} = \varphi(z, y')\mathbb{P}_{y'}[X_{\tau_1(\partial A')} = z'].$$

Now, consider a *marked* Poisson point process on $\partial A \times \mathbb{R}_+$ with rate 1. The (independent) marks are the simple random walk trajectories started from the first coordinate of the Poisson points (i.e., started at the corresponding site of ∂A) and run until hitting $\partial A'$. Then (see Figure 6.5; observe that A and A' need not be necessarily connected, as shown on the picture)

- let ξ_1 be the a.s. unique positive number such that there is only one point of the Poisson process on the graph of $\xi_1\varphi(x_0, \cdot)$ and nothing below;
- the mark of the chosen point is the first excursion (call it $Z^{(1)}$) that we obtain;
- then, let ξ_2 be the a.s. unique positive number such that the graph of $\xi_1\varphi(x_0, \cdot) + \xi_2\varphi(\ell(Z^{(1)}), \cdot)$ contains only one point of the Poisson process, and there is nothing between this graph and the previous one;
- the mark $Z^{(2)}$ of this point is our second excursion;
- and so on.

It is possible to show that the sequence of excursions obtained in this way indeed has the same law as the simple random walk's excursions (in particular, conditional on $\ell(Z^{(k-1)})$, the starting point of kth excursion is indeed distributed according to $\varphi(\ell(Z^{(k-1)}), \cdot)$); moreover, the ξ's are i.i.d. random variables with Exponential(1) distribution.

So, let us denote by $\xi_1, \xi_2, \xi_3, \ldots$ a sequence of i.i.d. random variables with Exponential distribution with parameter 1. According

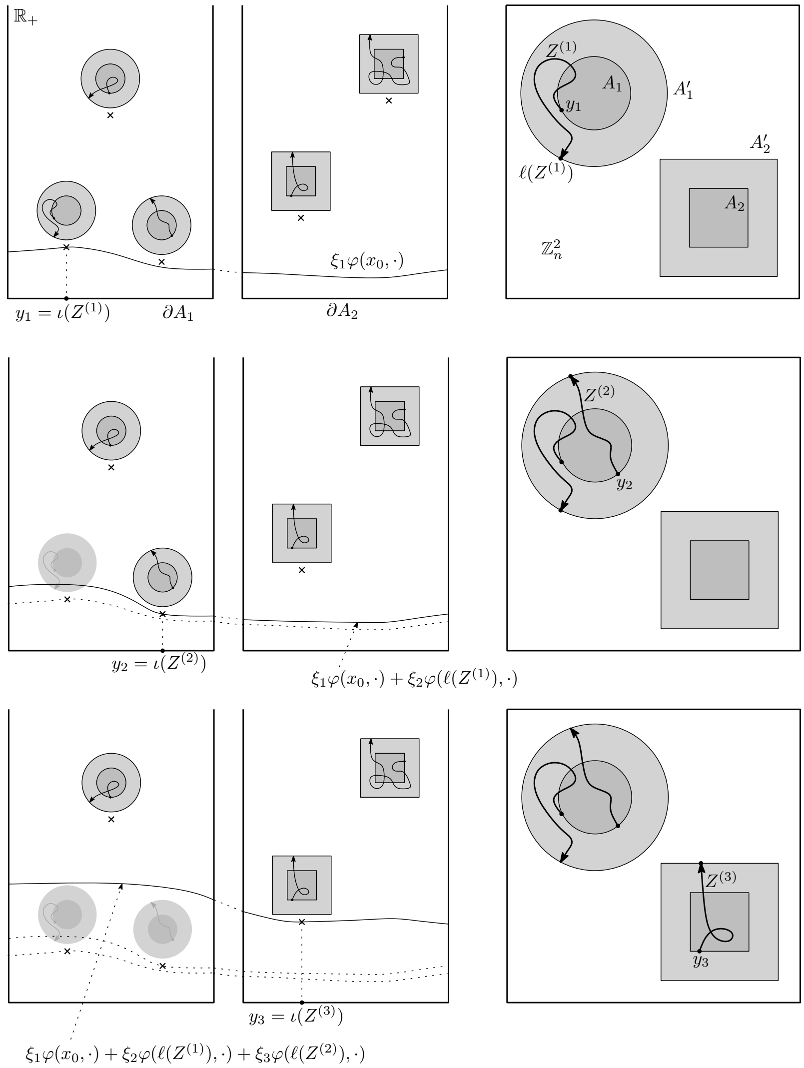

Figure 6.5 Construction of the first three excursions between ∂A and $\partial A'$ on the torus $\mathbb{Z}_n^2$ using the soft local times (here, $A = A_1 \cup A_2$ and $A' = A'_1 \cup A'_2$)

to the above informal description, the soft local time of kth excursion is a random vector indexed by $y \in \partial A$, defined as follows:

$$L_k(y) = \xi_1 \varphi(x_0, y) + \sum_{j=2}^{k} \xi_j \varphi(\ell(Z^{(j-1)}), y). \tag{6.13}$$

For the random interlacements, the soft local times are defined analogously. Recall that $\widehat{\mathrm{hm}}_A$ defines the (normalized) harmonic measure on A with respect to the $\widehat{S}$-walk. For $x \notin A$ and $y \in \partial A$ let

$$\widehat{\varphi}(x, y) = \mathbb{P}_x[\widehat{S}_{\widehat{\tau}_1(A)} = y, \widehat{\tau}_1(A) < \infty] + \mathbb{P}_x[\widehat{\tau}_1(A) = \infty]\,\widehat{\mathrm{hm}}_A(y). \tag{6.14}$$

Analogously, for the random interlacements, the sequence $\big((\iota(\widehat{Z}^{(j)}), \ell(\widehat{Z}^{(j)})), j \geq 1\big)$ is also a Markov chain, with transition probabilities

$$\widehat{P}_{(y,z),(y',z')} = \widehat{\varphi}(z, y')\mathbb{P}_{y'}[\widehat{S}_{\widehat{\tau}_1(\partial A')} = z'].$$

The process of picking the excursions for the random interlacements is quite analogous: if the last excursion was $\widehat{Z}$, we use the probability distribution $\widehat{\varphi}(\ell(\widehat{Z}), \cdot)$ to choose the starting point of the next excursion. Clearly, the last term in (6.14) is needed for $\widehat{\varphi}$ to have total mass 1; informally, if the $\widehat{S}$-walk from x does not ever hit A, we just take the "next" trajectory of the random interlacements that does hit A, and extract the excursion from it (see also (4.10) of [43]). Again, let $\hat{\xi}_1, \hat{\xi}_2, \hat{\xi}_3, \ldots$ be a sequence of i.i.d. random variables with Exponential distribution with parameter 1. Then, define the soft local time of random interlacement of kth excursion as

$$\widehat{L}_k(y) = \hat{\xi}_1 \widehat{\varphi}(x_0, y) + \sum_{j=2}^{k} \hat{\xi}_j \widehat{\varphi}(\ell(\widehat{Z}^{(j-1)}), y). \tag{6.15}$$

6.4 Exercises

add Lawler's book, papers about RI (2dim, and more), Kendall-Moran, Polya's original paper

Hints and solutions to selected exercises

Exercise 2.3.

You may find it useful to look at [15].

Exercise 2.4.

Use the cycle criterion (Theorem 2.2) with e.g. the cycle $(0,0) \to (0,1) \to (1,1) \to (1,0) \to (0,0)$.

Exercise 2.5.

Fix an arbitrary $x_0 \in \Sigma$, set $A = \{x_0\}$, and

$$f(x) = \mathbb{P}_x[\tau_{x_0} < \infty] \text{ for } x \in \Sigma$$

(so, in particular, $f(x_0) = 1$). Then (2.8) holds with equality for all $x \neq x_0$, and, by transience, one can find $y \in \Sigma$ such that $f(y) < 1 = f(x_0)$.

Exercise 2.6.

Let $p = p(n, n+1)$ (for all n), and assume for definiteness that $p > \frac{1}{2}$. Consider the function $f(x) = \bigl(\frac{1-p}{p}\bigr)^x$ and the set $A = (-\infty, 0]$; then use Theorem 2.4.

Note also that, for proving that this random walk is transient, one may also use Theorem 2.5.15 of [29] (which we did not consider in this book) together with a simpler function $f(x) = x$. There are many different Lyapunov function tools that one may use!

Exercise 2.7.

Quite analogously to (2.11)–(2.13), it is elementary to obtain for $f(x) = \|x\|^{-\alpha}$

$$\mathbb{E}[f(X_{n+1}) - f(X_n) \mid X_n = x] = -\alpha\|x\|^{-\alpha-2}\Bigl(\frac{1}{2} - \Bigl(1+\frac{\alpha}{2}\Bigr)\frac{1}{d} + O\bigl(\|x\|^{-1}\bigr)\Bigr).$$

The inequality $\frac{1}{2} - \bigl(1 + \frac{\alpha}{2}\bigr)\frac{1}{d} > 0$ solves to $\alpha < d - 2$, so any fixed $\alpha \in (0, d-2)$ will do the job.

By the way, what do you think, is it surprising that the "critical" value for α is $d-2$? Read Section 3.1 and think again about this question!

Exercise 2.9.

Hint: being X_n the two-dimensional walk, define first its *covariation matrix* by $M := \mathbb{E}_0\bigl((X_1)^\top X_1\bigr)$. Find a suitable linear transformation[1] of the process for which M will become the identity matrix. Then use the same Lyapunov function that worked for the simple random walk.

Exercise 2.11 (a).

Fix an arbitrary $x_0 \in \Sigma$, and set $A = \{x_0\}$. Observe that, for $x \neq x_0$,

$$\mathbb{E}_x \tau_{x_0} = \sum_{y \in \Sigma} p(x,y)\mathbb{E}_y(1 + \tau_{x_0}),$$

and that

$$\sum_{y \in \Sigma} p(x_0,y)\mathbb{E}_y \tau_{x_0} = \mathbb{E}_{x_0}\tau^+_{x_0} < \infty,$$

so the function $f(x) = \mathbb{E}_x \tau_{x_0}$ satisfies (2.16)–(2.17) with $\varepsilon = 1$.

Exercise 2.12.

Note that the calculation (2.11) is dimension-independent, and (2.12) remains valid as well, with obvious changes. Then obtaining (2.18) is straightforward (use (2.11) with $\alpha = 1$ and observe that the factor $\frac{1}{4}$ in the next display after (2.11) will become $\frac{1}{2d}$ in the general case). As for (2.19), show first that

$$\mathbb{E}_x(\|S_1^{(d)}\|^2 - \|x\|^2) = 1 \text{ for all } x \in \mathbb{Z}^d,$$

and then use (2.18) together with the identity $(b-a)^2 = b^2 - a^2 - 2a(b-a)$ with $a = \|x\|, b = \|S_1^{(d)}\|$.

Exercise 2.13.

Hint: try using the following Lyapunov functions: $f(x) = x^2$ for (a), $f(x) = x^\alpha$ for some $\alpha > 0$ for (b), and $f(x) = x^{-\alpha}$ for (c). Note that α will depend on ε in (b) and (c)!

[1] why does it exist?

Exercise 2.14.

Hint: use Exercise 2.12.

Exercise 3.1.

First, prove that

$$\mathbb{P}_x[S_n = y, \tau_1(\partial A) > n] = \mathbb{P}_y[S_n = x, \tau_1(\partial A) > n]$$

for any n (hint: which trajectories correspond to the two above events?), and then deduce the desired result.

Exercise 5.1.

Unfortunately, only with what one sees on Figure 5.6, it is not possible to find it. Think, for instance, that there may be a point just slightly above the top of the biggest triangle, which (I mean, the point) did not make it into the picture.

Exercise 5.2.

Due to Proposition 5.2, we can find a coupling $\mathbb{Q}$ between the Markov chain (X_i) and an i.i.d. collection (Y_i) (with law π), in such a way that for any $\lambda > 0$ and $t \geq 0$,

$$\begin{aligned}&\mathbb{Q}\big[\{Y_1, \ldots, Y_R\} \subset \{X_1, \ldots, X_t\}\big]\\&\geq \mathbb{P}_{\pi_0}\Big[\xi_0\pi_0(x) + \sum_{j=1}^{t-1} \xi_j p(X_j, x) \geq \lambda\pi(x), \text{ for all } x \in \Sigma\Big]\end{aligned} \tag{6.16}$$

where ξ_i are i.i.d. Exp(1) random variables, independent of R, a Poisson(λ)-distributed random variable. Then, obtain from a simple calculation the fact that $\mathbb{P}[\{Y_1, \ldots, Y_R\} \cap A = \emptyset] = e^{-\lambda\pi(A)}$.

Exercise 5.3.

If you find one, please, let me know.

Exercise 5.5.

To the second one. See also [21] for a more complete discussion of this.

Exercise 5.9.

Think, how one can generate the lines in the order corresponding

(a) to the distances from the origin to the lines;

(b) to the distances from the origin points of intersection with the horizontal axis.

Exercise 5.13.

Answer: 3/2. It is the last problem of the famous Mathematical Trivium [1] of Vladimir Arnold. Clearly, the problem reduces to finding the expected area of a projection of a square (note that a.s. only three faces of the cube contribute to the projection), and then one can calculate the answer doing a bit of integration. There is another way to solve it, however, that does not require any computations at all, and works for any convex body, not only for the cube. One may reason in the following way:

- imagine the surface of a convex body to be composed of many small plaquettes, and use the linearity of expectation to argue that the expected area of the projection *equals* the surface area of the body times a constant (that is, it does not depend on the surface's shape itself!);
- to obtain this constant, consider a certain special convex body whose projections are the same in all directions.

References

[1] Arnol′d, V. I. 1992. Trivium mathématique. *Gaz. Math.*, 87–96.

[2] Asmussen, S. 2003. *Applied probability and queues*. Second edn. Applications of Mathematics (New York), vol. 51. New York: Springer-Verlag. Stochastic Modelling and Applied Probability.

[3] Boucheron, S., Lugosi, G., Massart, P., and Ledoux, M. 2013. *Concentration inequalities : a nonasymptotic theory of independence*. Oxford university press.

[4] Brémaud, Pierre. 1999. *Markov chains*. Texts in Applied Mathematics, vol. 31. Springer-Verlag, New York. Gibbs fields, Monte Carlo simulation, and queues.

[5] Černý, J., and Popov, S. 2012. On the internal distance in the interlacement set. *Electron. J. Probab.*, **17**, no. 29, 25.

[6] Černý, Jiří, and Teixeira, Augusto. 2012. *From random walk trajectories to random interlacements*. Ensaios Matemáticos [Mathematical Surveys], vol. 23. Sociedade Brasileira de Matemática, Rio de Janeiro.

[7] Comets, F., and Popov, S. 2017. The vacant set of two-dimensional critical random interlacement is infinite. *arXiv:1606.05805*.

[8] Comets, F., Popov, S., Schütz, G. M., and Vachkovskaia, M. 2009. Billiards in a general domain with random reflections. *Arch. Ration. Mech. Anal.*, **191**(3), 497–537.

[9] Comets, F., Popov, S., Schütz, G. M., and Vachkovskaia, M. 2010a. Knudsen gas in a finite random tube: transport diffusion and first passage properties. *J. Stat. Phys.*, **140**(5), 948–984.

[10] Comets, F., Popov, S., Schütz, G. M., and Vachkovskaia, M. 2010b. Quenched invariance principle for the Knudsen stochastic billiard in a random tube. *Ann. Probab.*, **38**(3), 1019–1061.

[11] Comets, Francis, Gallesco, Christophe, Popov, Serguei, and Vachkovskaia, Marina. 2013. On large deviations for the cover time of two-dimensional torus. *Electron. J. Probab.*, **18**, no. 96, 18.

[12] Comets, Francis, Popov, Serguei, and Vachkovskaia, Marina. 2016. Two-dimensional random interlacements and late points for random walks. *Comm. Math. Phys.*, **343**(1), 129–164.

[13] Cook, Scott, and Feres, Renato. 2012. Random billiards with wall temperature and associated Markov chains. *Nonlinearity*, **25**(9), 2503–2541.

[14] de Bernardini, D.F., Gallesco, C., and Popov, S. 2017. On uniform closeness of local times of Markov chains and i.i.d. sequences. *arXiv:1610.02532*.

[15] Di Crescenzo, A., Macci, C., Martinucci, B., and Spina, S. 2017. Random walks on graphene: generating functions, state probabilities and asymptotic behavior. *arXiv:1610.09310*.

[16] Doyle, P. G., and Snell, J. L. 1984. *Random walks and electric networks*. Carus Mathematical Monographs, vol. 22. Washington, DC: Mathematical Association of America.

[17] Drewitz, Alexander, Ráth, Balázs, and Sapozhnikov, Artëm. 2014. *An introduction to random interlacements*. Springer Briefs in Mathematics. Springer, Cham.

[18] Durrett, R. 2010. *Probability: theory and examples*. Fourth edn. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge: Cambridge University Press.

[19] Fayolle, G., Malyshev, V. A., and Menshikov, M. V. 1995. *Topics in the constructive theory of countable Markov chains*. Cambridge: Cambridge University Press.

[20] Garcia, Nancy L., and Kurtz, Thomas G. 2008. Spatial point processes and the projection method. Pages 271–298 of: *In and out of equilibrium. 2*. Progr. Probab., vol. 60. Birkhäuser, Basel.

[21] Jaynes, E. T. 1973. The well-posed problem. *Found. Phys.*, **3**, 477–492.

[22] Kendall, M. G., and Moran, P. A. P. 1963. *Geometrical probability*. Griffin's Statistical Monographs & Courses, No. 10. Hafner Publishing Co., New York.

[23] Lamperti, J. 1960. Criteria for the recurrence or transience of stochastic process. I. *J. Math. Anal. Appl.*, **1**, 314–330.

[24] Lamperti, J. 1962. A new class of probability limit theorems. *J. Math. Mech.*, **11**, 749–772.

[25] Lamperti, J. 1963. Criteria for stochastic processes. II. Passage-time moments. *J. Math. Anal. Appl.*, **7**, 127–145.

[26] Lawler, G. F., and Limic, V. 2010. *Random walk: a modern introduction*. Cambridge Studies in Advanced Mathematics, vol. 123. Cambridge: Cambridge University Press.

[27] Lopes Garcia, Nancy. 1995. Birth and death processes as projections of higher-dimensional Poisson processes. *Adv. in Appl. Probab.*, **27**(4), 911–930.

[28] Meester, Ronald, and Roy, Rahul. 1996. *Continuum percolation*. Cambridge Tracts in Mathematics, vol. 119. Cambridge University Press, Cambridge.

[29] Menshikov, Mikhail, Popov, Serguei, and Wade, Andrew. 2017. *Non-homogeneous random walks – Lyapunov function methods for near-critical stochastic systems*. Cambridge: Cambridge University Press.

[30] Meyn, S., and Tweedie, R. L. 2009. *Markov chains and stochastic stability*. Second edn. Cambridge: Cambridge University Press. With a prologue by Peter W. Glynn.

[31] Peres, Y., Popov, S., and Sousi, P. 2013. On recurrence and transience of self-interacting random walks. *Bull. Braz. Math. Soc.*, **44**(4), 1–27.

[32] Peres, Yuval. 2009. The unreasonable effectiveness of martingales. Pages 997–1000 of: *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms.* SIAM, Philadelphia, PA.

[33] Pólya, Georg. 1921. Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Straßennetz. *Math. Ann.*, **84**(1-2), 149–160.

[34] Popov, Serguei, and Teixeira, Augusto. 2015. Soft local times and decoupling of random interlacements. *J. Eur. Math. Soc. (JEMS)*, **17**(10), 2545–2593.

[35] Resnick, Sidney I. 2008. *Extreme values, regular variation and point processes.* Springer Series in Operations Research and Financial Engineering. Springer, New York. Reprint of the 1987 original.

[36] Romeijn, H. E. 1998. A general framework for approximate sampling with an application to generating points on the boundary of bounded convex regions. *Statist. Neerlandica*, **52**(1), 42–59.

[37] Spitzer, F. 1976. *Principles of random walk.* Second edn. New York: Springer-Verlag. Graduate Texts in Mathematics, Vol. 34.

[38] Sznitman, Alain-Sol. 2010. Vacant set of random interlacements and percolation. *Ann. of Math. (2)*, **171**(3), 2039–2087.

[39] Sznitman, Alain-Sol. 2012. *Topics in occupation times and Gaussian free fields.* Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich.

[40] Sznitman, Alain-Sol. 2017. Disconnection, random walks, and random interlacements. *Probab. Theory Relat. Fields, to appear.*

[41] Teixeira, A. 2009. Interlacement percolation on transient weighted graphs. *Electron. J. Probab.*, **14**, no. 54, 1604–1628.

[42] Tsirelson, Boris. 2006. Brownian local minima, random dense countable sets and random equivalence classes. *Electron. J. Probab.*, **11**, no. 7, 162–198 (electronic).

[43] Černý, Jiří, and Teixeira, Augusto. 2016. Random walks on torus and random interlacements: Macroscopic coupling and phase transition. *Ann. Appl. Probab.*, **26**(5), 2883–2914.

Index