Simple Random Walk in Two Dimensions

A Guided Tour

Serguei Popov
Contents

Preface v

Preface v

Notations viii

Notations viii

1 Introduction 1
1.1 Markov chains and martingales: basic definitions and facts 4

2 Recurrence of the walk 10
2.1 Classical proof 10
2.2 Electric networks 14
2.3 Lyapunov functions 19
2.4 Exercises 30

3 Some potential theory for simple random walks 34
3.1 Transient case 35
3.2 Potential theory in two dimensions 47
3.3 Exercises 67

4 Simple random walk conditioned on not hitting the origin 74
4.1 Doob’s h-transforms 74
4.2 Conditioned SRW: basic properties 78
4.3 Green’s function and capacity 82
4.4 Harmonic measure 94
4.5 Range of the conditioned SRW 103
4.6 Exercises 115

5 Intermezzo: Soft local times and Poisson processes 118
5.1 Soft local times 118
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2</td>
<td>Poisson processes of objects</td>
<td>135</td>
</tr>
<tr>
<td>5.3</td>
<td>Exercises</td>
<td>141</td>
</tr>
<tr>
<td>6</td>
<td><strong>Random interlacements</strong></td>
<td>147</td>
</tr>
<tr>
<td>6.1</td>
<td>Random interlacements in higher dimensions</td>
<td>147</td>
</tr>
<tr>
<td>6.2</td>
<td>The two-dimensional case</td>
<td>157</td>
</tr>
<tr>
<td>6.3</td>
<td>Proofs for random interlacements</td>
<td>166</td>
</tr>
<tr>
<td>6.4</td>
<td>Exercises</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td><em>Hints and solutions to selected exercises</em></td>
<td>181</td>
</tr>
<tr>
<td></td>
<td><em>References</em></td>
<td>195</td>
</tr>
<tr>
<td></td>
<td><em>Index</em></td>
<td>200</td>
</tr>
</tbody>
</table>
Preface

Ready by now:

- Section 1.1: Basic definitions and facts
- Chapter 2: Recurrence of the walk
- Chapter 3: Potential theory
- Chapter 4: Conditioned SRW
- Chapter 5: Intermezzo (except the “exact coupling” for Markov chains part)

At the moment, other things in this book are mostly copy/paste (or not, but need polishing).

********************************************

How does it look like when a mathematician explains something to a fellow mathematician (postdocs included)? (rewrite it a bit)

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? **Write something like: the author has to confess that, unfortunately, as attested by many early readers, he ultimately failed in this endeavour. Still, hope that at least some pieces etc.**

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.

To explain it further, our approach can be characterized in the following way: we want to understand all things through a direct probabilistic intuition. Yes, the author is aware that the latter is generally not possible. On the other hand, it is still easier to “claim the territory” using the tools that one controls quite well and then extend one’s dominion over is by some additional tools, than to try claiming that territory with unfamiliar tools.

This 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, with some background in martingales and Markov chains (at the level of [32], for example); 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 on additional topics in Markov chains.

The author would like to warn the reader already at this point that, in order to make the book more accessible, he prefers to sometimes exaggerate on explaining things and writing explicitly what a considerable proportion of the readers may consider too evident. Therefore, if you find that some arguments are simply too evident and/or too well-known, that’s probably because that they are so indeed. Please forgive the author for that and be guaranteed that it is certainly not a sign of disrespect.

Relation to other recent books

Many topics of this book are treated at length in the literature, e.g. [31, 46, 52]; 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.

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, Nina Gantert, Mikhail Menshikov, Augusto Teixeira, Marina Vachkovskaia, and Andrew Wade. I also thank Daniel Ungaretti Borges, Gustavo Henrique Tasca, Victor Freguglia Souza, Thainá Soares Silva, Matheus Gorito de Paula, who read the manuscript at different stages and made many useful comments and suggestions.
Notations

Here we list the notations recurrently used in this book.

- $f(x) \asymp g(x)$ means that there exist $0 < C_1 < C_2 < \infty$ such that $C_1 g(x) \leq f(x) \leq C_2 g(x)$ for all $x$; $f(x) \lesssim g(x)$ means that there is $C_3 > 0$ such that $f(x) \leq C_3 g(x)$ for all $x$.
- $f(x) = \mathcal{O}(g(x))$ as $x \to a$ means that $\limsup_{x \to a} |f(x)/g(x)| < \infty$.
- $\mathcal{N} = \partial_c \{0\} = \{ \pm e_1, 2 \} \subset \mathbb{Z}^2$ is the set of the four neighbours of the origin (in two dimensions).
- $\mathcal{B}(x,r) = \{ y : ||y-x|| \leq r \}$ is the ball (disk) in $\mathbb{R}^d$ or $\mathbb{Z}^d$; $\mathcal{B}(r)$ stands for $\mathcal{B}(0,r)$.
- $(\mathcal{F}_n, n \geq 0)$ is a filtration (a nondecreasing sequence of sigma-algebras).
• \((p(x, y), x, y \in \Sigma)\) are transition probabilities of a Markov chain on a state space \(\Sigma\), and \((p_n(x, y), x, y \in \Sigma)\) are the \(n\)-step transition probabilities
• \((S_n, n \geq 0)\) is the simple random walk (also abbreviated as SRW) in \(\mathbb{Z}^d\)
• \((\hat{S}_n, n \geq 0)\) is the simple random walk in \(\mathbb{Z}^2\) conditioned on never hitting the origin
• \(P_x\) and \(E_x\) are probability and expectation for a random walk (the one that we are considering at the moment) starting from \(x\)
• \(\tau_A \geq 0\) and \(\tau^+_A \geq 1\) are entrance and hitting times of \(A\) by the SRW
• \(\hat{\tau}_A \geq 0\) and \(\hat{\tau}^+_A \geq 1\) are entrance and hitting times of \(A\) by the conditioned SRW
• \(G(\cdot, \cdot)\) is the Green’s function for the SRW in three or more dimensions, \(G_A(\cdot, \cdot)\) is the Green’s function restricted on \(A\), and \(a(\cdot)\) is the potential kernel for the two-dimensional SRW
• \(E_s_A(x) = P_x[\tau^+_A = \infty]1\{x \in A\}\) is the escape probability from \(x\) for SRW in dimensions \(d \geq 3\), and \(\hat{E}_s_A(x) = P_x[\hat{\tau}^+_A = \infty]1\{x \in A\}\) is the escape probability from \(x\) for the conditioned SRW in two dimensions
• for SRW, \(\text{hm}_A(x)\) is the value of the harmonic measure in \(x \in A\), and \(\text{cap}(A)\) is the capacity of \(A\); \(\hat{\text{hm}}_A(x)\) and \(\hat{\text{cap}}(A) = \text{cap}(A \cup \{0\})\) are the corresponding quantities for the conditioned SRW
• \(N_x = \sum_{j=0}^{\infty} 1\{S_j = x\}\) is the total number of visits to \(x\), and \(N^{(k)}_x = \sum_{j=0}^{k} 1\{S_j = x\}\) is the total number of visits to \(x\) up to time \(k\)
• let \(A\) be a fixed subset of \(\mathbb{Z}^2\); then \(N^+_x = \sum_{j=0}^{\tau^+_A - 1} 1\{S_j = x\}\) is the number of visits to \(x\) before the first return to \(A\), and by \(N^+_x = \sum_{j=0}^{\infty} 1\{S_j = x\}\) the number of visits to \(x\) after the first return to \(A\) (with \(N^+_x = 0\) on \(\{\tau^+_A = \infty\}\))
• \(\hat{N}_x, \hat{N}^{(k)}_x, \hat{N}_x^+, \hat{N}^{(k)}_x^+\) are the corresponding quantities for the conditioned walk \(\hat{S}\)
• \(\hat{G}(x, y) = a(y)(a(x) + a(y) - a(x - y))/a(x)\) is the Green’s function of the conditioned walk
• \(\hat{\ell}(x, y) = 1 + \frac{a(y) - a(x-y)}{a(x)} = \frac{\hat{G}(x,y)}{a(y)}, \text{ and } \hat{g}(x, y) = \frac{\hat{G}(x,y)}{a^2(y)} = \hat{g}(y, x)\) is the “symmetrized” \(\hat{G}\)
• \(G_n(z)\) is the soft local time of the process at time \(n\) at site \(z\)
Notations

- $\mathbb{Z}^d_n = \mathbb{Z}^d/n\mathbb{Z}^d$ is the $d$-dimensional torus of (linear) size $n$ (with the graph structure inherited from $\mathbb{Z}^d$)
- $\mathcal{I}^u$ and $\mathcal{V}^u$ are the interlacement and vacant sets for the random interlacement model on level $u > 0$ in dimensions $d \geq 3$; in two dimensions, we usually denote the level by $\alpha$, so these become $\mathcal{I}^\alpha$ and $\mathcal{V}^\alpha$


1

Introduction

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

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”. This result may explain why birds do not drink vodka; still, inspite of the recurrence, the drunken man’s life is not easy as well: as we will see below, it may take him *quite* some time to return home.

Due to (3.42), probability to escape to $\partial B(n)$ is approximately $(1.0293737 + \frac{2}{n} \ln n)^{-1}$.

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).

Here: some funny examples for the above (step out of Paris starting in Louvre, step out of our galaxy, all before coming back to the origin, with step size 1m).

Radius of Paris is around 5000m, and $(1.0293737 + \frac{2}{5000} \ln 5000)^{-1} \approx 0.155$. Radius of the Milky Way galaxy is around $10^{21}$m, and $(1.0293737 + \frac{2}{10^{21}} \ln 10^{21})^{-1} \approx 0.031$. (Yes, when one starts in the center of our galaxy, there is a risk that the starting location happens to be in a big black hole; we restrict ourselves to purely mathematical aspects of the above question, though.)

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. And maybe that Kesten’s result about filling a disk ([25]).
Future side quests include: E.g., Lyapunov functions etc.

About some interesting things that we will not discuss. A little blablabla about other things, also put some pictures of DLA, they are nice :) cite Sava-Huss article arxiv:1902.03800

Level: [32] 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, [46, 66] 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; the purpose of at least some of them is to guide an interested reader who wants to dive deeper into the subject.

1. Basic definitions. In the remaining part of this chapter, we recall here some basic definitions and facts for Markov chains and martingales, mainly for reference purposes.

2. 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.

3. 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 different kind of sets. **will we do it there?**

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. (?)

4. **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 (not conditioned) 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 $y \to \infty$. Perhaps even more surprisingly, the proportion of visited sites of “typical” large sets approaches in distribution a Uniform[0, 1] random variable.

5. **Intermezzo: Soft local times and Poisson processes of objects.** 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 $\Sigma$, 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.

6. **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 [68] in dimension $d \geq 3$. Then,

$^1$ judging from the author’s experience, many people working with random walks do not know how Poisson line processes are constructed
we discuss the two-dimensional case recently introduced in [18]; 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 [72], “One has good decoupling properties of the excursions . . . when the boxes are sufficiently far apart. The soft local time technique . . . 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 $\mathbb{N} = \{1, 2, 3, \ldots\}, \mathbb{Z}_+ = \{0, 1, 2, \ldots\}, \mathbb{Z}_+^+ = \mathbb{Z}_+ \cup \{+\infty\}$.

**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 $\sigma$-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 $\mathcal{F}_\infty := \sigma(\bigcup_{n \geq 0} \mathcal{F}_n) \subset \mathcal{F}$.
- 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 $n \in \mathbb{Z}_+$.
1.1 Basic definitions

- For a (possibly infinite) random variable $\tau \in \mathbb{Z}_+$, the random variable $X_\tau$ is (as the notation suggests) equal to $X_n$ on $\{\tau = n\}$ for finite $n \in \mathbb{Z}_+$ and equal to $X_\infty := \lim \sup_{n \to \infty} X_n$ on $\{\tau = \infty\}$.

- A (possibly infinite) random variable $\tau \in \mathbb{Z}_+$ is a stopping time with respect to a filtration $\mathcal{F}_n, n \geq 0$ if $\{\tau \leq n\} \in \mathcal{F}_n$ for all $n \geq 0$.

- If $\tau$ is a stopping time, the corresponding $\sigma$-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 $n \in \mathbb{Z}_+$.

- 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
  \[ \tau_A = \min\{n \geq 0 : X_n \in A\}, \]  
  (1.1)  
  and
  \[ \tau_A^+ = \min\{n \geq 1 : X_n \in A\}; \]  
  (1.2)
  we may refer to either $\tau_A$ or $\tau_A^+$ as the hitting time of $A$ (also called the passage time into $A$). It is straightforward to check that both $\tau_A$ and $\tau_A^+$ 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 $\mathcal{F}_n$) if, for all $n \geq 0$,

(i) $\mathbb{E}|X_n| < \infty$, and

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

If in (ii) “=” is replaced by “$\geq$” (respectively, “$\leq$”), then $X_n$ is called a submartingale (respectively, supermartingale). If the filtration is not specified, that means that the natural filtration is used.
Evidently, if $X_n$ is a submartingale then $(-X_n)$ is a supermartingale, and vice versa; also, a martingale is both submartingale and supermartingale. Also, one can easily check the important fact that, if $X_n$ is a (sub-, super-)martingale, then so is $X_n \wedge \tau$ for any stopping time $\tau$.

Martingales have a number of remarkable properties; we do not even try to elaborate on this topic here. Let us only cite the paper [55], 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 $\sup_n E X_n^+ < \infty$. 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 $E X_n$ is non-decreasing (by the submartingale property) and bounded above by $\sup_n E [X_n^+]$; so $\lim_{n \to \infty} E X_n$ exists and is finite. However, it is not necessarily equal to $E X$.

Using Theorem 1.4 and the 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 $E X \leq E X_0$.

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 $E X_\sigma \leq E X_\tau < \infty$ and $X_\sigma \leq E [X_\tau \mid F_\sigma]$ a.s.

Note that, if $X_n$ is a uniformly integrable submartingale and $\tau$ 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 [32]. 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 $\tau$, it holds that $E X_0 \leq E X_\tau \leq E X_\infty < \infty$, where $X_\infty := \limsup_{n \to \infty} X_n = \lim_{n \to \infty} X_n$ exists and is integrable, by Theorem 1.4.
1.1 Basic definitions

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 [32]).

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

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

**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 $m \geq 1$,
  $$\mathbb{P}[X_{n+m} = y \mid X_0, \ldots, X_n] = \mathbb{P}[X_{n+m} = y \mid X_n], \text{ a.s.} \quad (1.3)$$
  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
  $$\mathbb{P}[X_{n+m} = y \mid \mathcal{F}_n] = p_m(X_n, y), \text{ a.s.,} \quad (1.4)$$
  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 $x \in \Sigma$. 
Introduction

• 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 \( \tau \) be a stopping time with respect to the natural filtration of \( X_n \). Then, for all \( x, y_1, \ldots, y_k \in \Sigma \), \( n_1, \ldots, n_k \geq 1 \), it holds that

\[
P_{x}[X_{\tau+n_j} = y_j, j = 1, \ldots, k | F_{\tau}, X_{\tau} = x] = P_{x}[X_{\tau+n_j} = y_j, j = 1, \ldots, k]
\]

(this is the strong Markov property).

• For a Markov chain, a probability measure \( (\pi(x), x \in \Sigma) \) is called an invariant measure if \( \sum_{x \in \Sigma} \pi(x)p(x, y) = \pi(y) \) for all \( y \in \Sigma \). It then holds that \( P_{\pi}[X_n = y] = \pi(y) \) for all \( n \) and \( y \) (where \( P_{\pi} \) means that the initial state of the process is chosen according to \( \pi \)).

Suppose now that \( X_n \) is a countable Markov chain. Recall the definitions of hitting times \( \tau_A \) and \( \tau_{A}^{+} \) from (1.1)–(1.2). 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 \( P_x[\tau_A = 0] = 1 \), while \( \tau_A^{+} \geq 1 \) is then the return time to \( A \). Also note that \( P_x[\tau_A = \tau_{A}^{+}] = 1 \) 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 \( P_x[\tau_{x}^{+} < \infty] = 1 \);
• transient if \( P_x[\tau_{x}^{+} < \infty] < 1 \).

A recurrent state \( x \) is classified further as

• positive recurrent if \( E_x\tau_{x}^{+} < \infty \);
• null recurrent if \( E_x\tau_{x}^{+} = \infty \).

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 $\Sigma$ 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 $\Sigma$.

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

(For the proof, cf. e.g. Lemma 2.5.1 of [52].)
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 $p_m(x, y) = \mathbb{P}_x[S_m = y]$ 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, starting at the origin, the walk eventually returns to the origin. If $q_d < 1$, then the total number of visits (counting the initial instance $S_0 = 0$ 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\(^1\)

$$\mathbb{E}_0 \sum_{k=0}^{\infty} \mathbf{1}\{S_k = 0\} = \sum_{k=0}^{\infty} \mathbb{E}_0 \mathbf{1}\{S_k = 0\} = \sum_{n=0}^{\infty} \mathbb{P}_0[S_{2n} = 0]$$

(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

\(^1\) note that we can put the expectation inside the sum because of the Monotone Convergence Theorem
walk is equivalent to

\[ \sum_{n=0}^{\infty} p_{2n}(0,0) = \infty. \] (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 \(S\) as

\[ S_n = X_1 + \cdots + X_n, \]

where \((X_k, k \geq 1)\) 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 \(S_n/\sqrt{n}\) converges in distribution to a (multivariate) centered Normal random vector with a diagonal covariance matrix. That is, it is reasonable to expect that \(S_n\) 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\(^2\) 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 \(\infty\). In fact, this statement can be made rigorous by using the local Central Limit Theorem (e.g., Theorem 2.1.1 from [46]). 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\(^3\).

Now, let us prove that (2.1) holds for the one- and two-dimensional simple random walks. In the one-dimensional case, it is quite simple to calculate \(p_{2n}(0,0)\): it is the probability that a \(\text{Binomial}(2n, \frac{1}{2})\)-random variable equals 0, so it is \(2^{-2n}{\binom{2n}{n}}\). Certainly, this expression is concise and beautiful; it is, however, not a priori clear which asymptotic behaviour it has (as it frequently

---

\(^2\) a site is called even if the sum of its coordinates is even; observe that the origin is even.

\(^3\) as some physicists say, “in practice, logarithm is a constant!”
Recurrence

happens with concise and beautiful formulas). To clarify this, we use the Stirling’s approximation

\[ n! = \sqrt{2\pi n} (n/e)^n (1 + o(1)) \]

to obtain that

\[
2^{-2n} \left( \frac{2n}{n} \right) = 2^{-2n} \frac{(2n)!}{(n!)^2} = 2^{-2n} \frac{\sqrt{4\pi n}(2n/e)^{2n}}{2\pi n(n/e)^{2n}} (1 + o(1))
\]

(remarkably, almost everything cancels)

\[
= \frac{1}{\sqrt{\pi n}} (1 + o(1)).
\]

The series \( \sum_{k=1}^{\infty} k^{-1/2} \) diverges, so (2.1) holds, and this implies recurrence in dimension 1.

Let us now deal with the two-dimensional case. 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 of steps to the right must 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

\[
p_{2n}(0,0) = 4^{-2n} N_{2n}
\]

\(^{4}\text{see e.g. http://mathworld.wolfram.com/StirlingsApproximation.html}\)
2.1 Classical proof

\[ p_{2n}(0,0) = \left( 2^{-2n} \binom{2n}{n} \right)^2. \]  

(2.2)

There is a nice combinatorial argument that allows one to deal with the sum in the right-hand side of (2.3). 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 that, in two dimensions,

\[ p_{2n}(0,0) = \left( 2^{-2n} \binom{2n}{n} \right)^2. \]  

(2.3)

The calculation (2.2) then implies that

\[ p_{2n}(0,0) = \frac{1}{\pi n} (1 + o(1)) \]  

(2.4)

for the two-dimensional SRW, and, using the fact that the harmonic series diverges, we obtain (2.1) and therefore recurrence.

\[ \square \]

It is curious to notice that (2.4) means that the probability of being at the origin at time 2n for the two-dimensional SRW is exactly the square of corresponding probability in one dimension. Such coincidences usually happen for a reason, and this case is no exception: in fact, it is possible to decouple the “one-dimensional components” of the two-dimensional SRW by considering its projections on the axes rotated \( \pi/4 \) anticlockwise; these projections are independent\(^5\). Indeed, it is straightforward to verify\(^6\) that \( S_n \cdot (e_1 + e_2) \) and \( S_n \cdot (e_1 - e_2) \) are independent one-dimensional SRWs.

\(^5\) this fact is folklore, but the author thanks Alejandro Ramirez for making him aware of it

\(^6\) this is left as an exercise
2.2 Electric networks

First, we recall the following

**Definition 2.1** A Markov chain with transition probabilities \((p(x,y), x, y \in \Sigma)\) is called **reversible** with the reversible measure \(\pi = (\pi(x), x \in \Sigma)\) when \(\pi(x)p(x,y) = \pi(y)p(y,x)\) for all \(x, y \in \Sigma\).

It is important to note that, in the above definition, we do not assume that \(\pi\) is a probability measure; its total mass can be any positive number or even infinity (in the case when the state space \(\Sigma\) is infinite). For example, note that \(\pi \equiv 1\) is reversible for the SRW (in any dimension). However, when the total mass of the reversible measure is 1, it is also invariant for the Markov chain, as shown by the following simple calculation:

\[
\sum_{x \in \Sigma} \pi(x)p(x,y) = \sum_{x \in \Sigma} \pi(y)p(y,x) = \pi(y)\sum_{x \in \Sigma} p(y,x) = \pi(y).
\]

Also, a reversible measure cannot be unique: if \(\pi\) is a reversible measure, then so is \(c\pi\) for any \(c > 0\).

In this book we will not dive deep into the theory of reversible Markov chains; let us only note that the latter is rich and useful. It is good to know the following 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 \(x_0, x_1, \ldots, x_{n-1}, x_n = x_0\) of states it holds that:

\[
\prod_{k=0}^{n-1} p(x_k, x_{k+1}) = \prod_{k=1}^{n} p(x_k, x_{k-1}); \tag{2.6}
\]

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

**Proof** First, assume that the Markov chain is reversible, and let us do the proof for \(n = 2\) (i.e., cycle of size 3); the reader will easily see that the same argument works for all \(n\). The idea is to multiply (2.6) by \(\pi(x_0)\) and then “let it go through the cycle”:

\[
\pi(x_0)p(x_0, x_1)p(x_1, x_2)p(x_2, x_0) = p(x_1, x_0)\pi(x_1)p(x_1, x_2)p(x_2, x_0) \\
= p(x_1, x_0)p(x_2, x_1)\pi(x_2)p(x_2, x_0) \\
= p(x_1, x_0)p(x_2, x_1)p(x_0, x_2)\pi(x_0),
\]
then we cancel $\pi(x_0)$ and obtain the claim.

Now, let us assume that (2.6) holds, and prove that the Markov chain is reversible. The main difficulty is to find a good candidate for the reversible measure – in principle, \textit{a priori} there is none. But if one recalls how the reversible measure for one-dimensional nearest-neighbor random walk looks like (see Exercise 2.6), one may come out with following procedure:

• fix some site $x_0$ and put $\pi(x_0) = 1$;
• for $x \neq x_0$, fix a path $(x_0, x_1, \ldots, x_k, x)$ which goes from $x_0$ to $x$; then, set

$$
\pi(x) = \frac{p(x_0, x_1)p(x_1, x_2)\cdots p(x_k, x)}{p(x_1, x_0)} \frac{p(x_1, x_2)\cdots p(x, x_k)}{p(x_2, x_1)\cdots p(x, x_k)} \cdot (2.7)
$$

We then need to check that this definition is correct in the sense that if there is another path $(x_0, z_1, \ldots, z_\ell, x)$ which goes from $x_0$ to $x$, then

$$
p(x_0, x_1)p(x_1, x_2)\cdots p(x_k, x) = p(x_0, z_1)p(z_1, z_2)\cdots p(z_\ell, x)\cdot (2.7)
$$

this, however, immediately follows from the condition (2.6) with the cycle $(x_0, x_1, \ldots, x_k, x, z_1, \ldots, z_\ell, x_0)$. It remains to take some $y$ adjacent to $x$ (as on Figure 2.1) and check that $\pi(x)p(x, y) = \pi(y)p(y, x)$; this is because, by (2.7), $\pi(y) = \pi(x)\frac{p(x, y)}{p(y, x)}$ (note that $(x_0, x_1, \ldots, x_k, x, y)$ is a path from $x_0$ to $y$).

Now, one of the advantages of reversibility is that, somewhat unexpectedly, it permit us to use \textit{physical intuition} for analysing the Markov chain. More concretely, it is possible to represent a reversible Markov chain via an \textit{electric network} (think of the
Recurrence

edges of its transition graph as wires that have some conductance/resistance). In the following, we assume that $p(x, x) = 0$ for all $x \in \Sigma$, this will permit us to avoid loops which do not make much sense in electricity.

**Definition 2.3** An electric network is a (non-oriented) graph $(V, E)$ without loops, with positive weights $(c(e), e \in E)$ assigned to its edges. The quantity $c(e)$ is thought of as conductance of $e$, and $r(e) := 1/c(e)$ stands for the resistance of the edge $e$.

An important observation is that we can actually permit the conductances (as well as resistances) to assign values in $[0, +\infty]$ (however, the graph will have to be modified): if $c(e) = 0$ (or $r(e) = \infty$), that means that the edge $e$ is simply removed (no wire at all), and if $c(e) = \infty$ (or $r(e) = 0$) for $e = (x, y)$, that means that the vertices $x$ and $y$ are “glued together” (short-circuited by wires of zero resistance).

The correspondance between reversible Markov chains and electric networks is described in the following way. Given an electric network, the transition probabilities of the corresponding Markov chain are then defined by

$$p(x, y) = \frac{c(x, y)}{C(x)}, \quad \text{where } C(x) = \sum_{v : v \sim x} c(x, v). \quad (2.8)$$

It is then straightforward to verify that it is reversible with the reversible measure $C$. Conversely, consider a reversible Markov chain with the reversible measure $\pi$ and with $p(x, x) = 0$ for all $x$; then define $c(x, y) = \pi(x)p(x, y)$. Also, clearly, the SRW (on any graph) corresponds to the electric network with all conductances being equal to 1.

We will now take the following route. Instead of proving our way to the desired tools, we will just formulate the necessary facts and explain informally why they should be valid; this is because the author feels that there are already many great sources for learning this theory. First of all, the classical book [30] is an absolute must-read. I can recommend also Chapter 9 of [47] for a modern short introduction to the subject, and Chapters 2, 3, and 9 of [49] for an in-depth treatment.

The central notion of this theory is that of effective resistance. The “physical” definition of effective resistance between $x$ and $y$ (denoted by $R_{\text{eff}}(x, y)$) is simple: attach a 1 volt battery to $x$ and $y,
2.2 Electric networks

measure the (outgoing) current $I$ at $x$, and define, in accordance
with the Ohm’s law, $R_{\text{eff}}(x, y) = 1/I$. We can also recall how to
calculate resistances of several wires which are put together in the
simplest cases of serial and parallel connections, see Figure 2.2.

Next, for a subset $B \subset V$ such that $x \in B \setminus \partial B$, define $R_{\text{eff}}(x, \partial B)$
to be the effective resistance between $x$ and all vertices of $\partial B$ glued
together. Then, being $B(x, n)$ the ball of radius $n$ with respect to
the graph distance, define effective resistance to infinity $R_{\text{eff}}(x, \infty)$ as

$$R_{\text{eff}}(x, \infty) = \lim_{n \to \infty} R_{\text{eff}}(x, \partial B(x, n))$$

(the limit exists because that sequence is monotonously increasing,
as it is not difficult to see). Intuitively, we just measure the
resistance between $x$ and the infinity, see Figure 2.3.

Now, it is possible to prove a remarkable fact that relates the
escape probability from $x$ to the effective resistance:

$$P_x[\tau_x^+ = \infty] = \frac{1}{C(x)R_{\text{eff}}(x, \infty)}; \quad (2.9)$$

this implies that the effective resistance to infinity is infinite if and
only if the corresponding Markov chain is recurrent.

The following properties of effective resistance are physically
evident (but have a non-trivial proof; they are consequences of
the Rayleigh’s Monotonicity Law):

intuitively: if the particle cannot escape to infinity, then the “current” does not
flow to infinity, et voilà the infinite resistance
(i) if we cut certain edges of the network (i.e., set the conductances of these edges to 0), then the effective resistance cannot decrease;

(ii) if we “glue” (short-circuit) certain vertices together (i.e., connect them with edges of infinite conductance), then the effective resistance cannot increase.

A curious consequence of (i) is that the SRW on any subgraph of \( \mathbb{Z}^2 \) is recurrent (just imagine proving this using the approach of the previous section!).

Let us now obtain another proof of recurrence of the two-dimensional SRW, this time with electric networks. As we noted before, we need to prove that the effective resistance to infinity is infinite. Let us use the above observation (ii): if we short-circuit some vertices together and prove that the effective resistance is still infinite, then we are done. Now, look at Figure 2.4: let us glue together the vertices that lie on (graph) distance \( n \) from the origin, for \( n = 1, 2, 3, \ldots \). We use the two rules depicted on Figure 2.2: there are \( 4(2n+1) \) edges from level \( n - 1 \) to level \( n \), so the overall conductance of the “superedge” between \( n - 1 \) and \( n \) is \( 4(2n + 1) \), and so its resistance is \( \frac{1}{4(2n+1)} \). The resistance to infinity of this

\[ \text{Figure 2.3} \text{ On the definition of effective resistance to infinity.} \]
new graph is therefore

$$\sum_{n=0}^{\infty} \frac{1}{4(2n + 1)} = \infty,$$

as required. The harmonic series strikes again!

2.3 Lyapunov functions

The proofs of Sections 2.1 and 2.2 are simple and beautiful. In principle, this is good, but: not infrequently, such proofs are not very robust, in the sense that they do not work anymore if the setup is changed even a little bit. Indeed, assume that we modify 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 (2.3), a much more complicated expression will appear). The situation with the proof of Section 2.2 is not very satisfactory as well: the random walk is no longer reversible (cf. Exercise 2.7), 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. 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 not be necessarily countable. In particular, this approach works for non-reversible Markov chains.

In this section we follow mainly [52] and [33]. Other sources on the Lyapunov functions method are e.g. [3, 8, 53]; see also [67] for a take on Lyapunov functions from a more applied perspective.

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

**Theorem 2.4 (Recurrence criterion)** An irreducible Markov chain $X_n$ on a countably infinite state space $\Sigma$ 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

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

and $f(x) \to \infty$ as $x \to \infty$. (2.10)

The quantity in (2.10) can be thought of as the drift vector at $x$ with respect to the function $f$. To understand the meaning of Theorem 2.4, recall first that, for a function $f : \mathbb{R}^d \to \mathbb{R}$, the level sets are sets of the form $\{x \in \mathbb{R}^d : f(x) = c\}$ for $c \in \mathbb{R}$; in the following heuristics, we think of $f$ as a function of continuous argument, just to be able to visualize the things better. If $f$ converges to infinity as $x \to \infty$, then the level sets will look as depicted on Figure 2.5, and Theorem 2.4 says that, to prove the recurrence, it

---

8 It is one of the side quests that was promised.

9 The reader may wonder what “$x \to \infty$” might mean on an arbitrary countable set, with no particular enumeration fixed. Notice, however, that if a sequence of sites converges to infinity with respect to one enumeration, it will do so with respect to any other one; so, writing “$x \to \infty$” is legitimate in any case.
2.3 Lyapunov functions

is enough to find a function as above such that the drift vectors look inside its level sets. In fact, just by observing Figure 2.5 it is easy to believe that the Markov chain should be recurrent, since it has a “tendency” to “go inside”.

The term “Lyapunov function” comes from Differential Equations: there, a similar (in spirit) construction is used to prove stability\(^\text{10}\) of the solutions; see e.g. [45].

**Proof of Theorem 2.4** To prove that having a function that satisfies (2.10) 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 \(P_x[\tau_A = \infty] > 0\) (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 \(Y_\infty\) such that \(Y_n \to Y_\infty\) a.s. and

\[
E_x Y_\infty \leq E_x Y_0 = f(x),
\]

(2.11)

for any \(x \in \Sigma\). On the other hand, since \(f \to \infty\), it holds that the set \(V_M := \{y \in \Sigma : f(y) \leq M\}\) is finite for any \(M \in \mathbb{R}_+\); so, our assumption on transience implies that \(V_M\) 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 2.6). Hence, on \(\{\tau_A = \infty\}\), we must have \(Y_\infty = \lim_{n \to \infty} Y_n = +\infty\), a.s.. This would contra-

\(^{10}\) which can be seen as deterministic analogue of recurrence
Recurrence

Figure 2.6 On the proof of Theorem 2.4: a transient Markov chain can visit any finite set only finitely many times.

dict (2.11) under the assumption $P_x[\tau_A = \infty] > 0$, because then
$E_x[Y_\infty] \geq E_x[Y_\infty 1\{\tau_A = \infty\}] = \infty$. Hence $P_x[\tau_A = \infty] = 0$ 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 [33]. See also Exercise 2.15.

As a (very simple) example of application of Theorem 2.4, consider the one-dimensional simple random walk $S^{(1)}$, together with the set $A = \{0\}$ and the function $f(x) = |x|$. Then (2.10) 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 2.5 (Transience criterion)** An irreducible Markov chain $X_n$ on a countable state space $\Sigma$ 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
\begin{equation}
E[f(X_{n+1}) - f(X_n) \mid X_n = x] \leq 0, \text{ for all } x \in \Sigma \setminus A, \quad (2.12)
\end{equation}
and
\begin{equation}
f(y) < \inf_{x \in A} f(x), \text{ for at least one site } y \in \Sigma \setminus A. \quad (2.13)
\end{equation}

Note that (2.12) by itself is identical to (2.10); the difference is in what we require of the nonnegative function $f$ (we need (2.13) instead of convergence to infinity; in most applications of Theorem 2.5, the function $f$ will converge to 0). Also, differently from the recurrence criterion, in the above result the set $A$ need not be finite.

Proof of Theorem 2.5 Assume that $X_0 = y$, where $y$ is from (2.13), and (similarly to the previous proof) define the process $Y_n = f(X_{n\wedge \tau_A})$. Then (2.12) implies that $Y_n$ is a supermartingale. Since $Y_n$ is also non-negative, Theorem 1.5 implies that there is a random variable $Y_\infty \geq 0$ such that $\lim_{n \to \infty} Y_n = Y_\infty$ a.s., and $E Y_\infty \leq E Y_0 = f(y)$. Observe that, if the Markov chain eventually hits the set $A$, then the value of $Y_\infty$ equals the value of $f$ at some random site (namely, $X_{\tau_A}$) belonging to $A$; formally, we have that, a.s.,
\begin{equation}
Y_\infty 1\{\tau_A < \infty\} = \lim_{n \to \infty} Y_n 1\{\tau_A < \infty\} = f(X_{\tau_A}) 1\{\tau_A < \infty\} \\
\geq \inf_{x \in A} f(x) 1\{\tau_A < \infty\}.
\end{equation}
So, we obtain
\begin{equation}
f(y) = EY_0 \geq E Y_\infty \geq E Y_\infty 1\{\tau_A < \infty\} \geq P_y[\tau_A < \infty] \inf_{x \in A} f(x),
\end{equation}
which implies
\begin{equation}
P_y[\tau_A < \infty] \leq \frac{f(y)}{\inf_{x \in A} f(x)} < 1,
\end{equation}
proving the transience of the Markov chain $X_n$, by Lemma 1.12(i).

For the “only if” part, see Exercise 2.9.

Let us now think about how to apply Theorem 2.4 to the simple random walk in two dimensions. For this, we need to find a (Lyapunov) function $f : \mathbb{Z}^2 \to \mathbb{R}_+$, such that the “drift with respect to $f$”
\begin{equation}
E[f(S_{n+1}) - f(S_n) \mid S_n = x]
\end{equation}
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 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 following: consider the level sets 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\(^{11}\).

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 (2.14). Actually, let us just look at Figure 2.7; 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\(^{12}\). 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

\(^{11}\) there are many examples where they are not circles/spheres; let us mention e.g. Section 4.3 of [52], which is based on [54].

\(^{12}\) observe that, similarly to the previous case, these level curves have form $\{\|x\| = a - b\}$, $\{\|x\| = a\}$, $\{\|x\| = a + b\}$ with $a = j\sqrt{2}$, $b = j\sqrt{2} - \sqrt{j^2 + (j-1)^2}$.
2.3 Lyapunov functions

Figure 2.7 The drift with respect to $f(x) = \|x\|$ is positive

norm by a larger amount than the jumps down and to the right. In fact, it is 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 outside, then the drift with respect to the function would become nonpositive, look at Figure 2.8. 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 $y \cdot z$ the scalar product of $y, z \in \mathbb{Z}^2$)

$$
\|x + e\|^\alpha - \|x\|^\alpha = \|x\|^\alpha \left( \left( \frac{\|x + e\|}{\|x\|} \right)^\alpha - 1 \right)
= \|x\|^\alpha \left( \left( \frac{(x + e) \cdot (x + e)}{\|x\|^2} \right)^{\alpha/2} - 1 \right)
$$
Let $a = g(\|x\|)$, $a \pm b = g(\|x \pm e_1\|)$; note that $\|x + e_1\| - \|x\| > \|x\| - \|x - e_1\|$.}

$$\|x\|^{\alpha} = \|x\|^{\alpha} \left( \frac{\|x\|^2 + 2x \cdot e + 1}{\|x\|^2} \right)^{\alpha/2} - 1 = \|x\|^{\alpha} \left( \left( 1 + \frac{2x \cdot e + 1}{\|x\|^2} \right)^{\alpha/2} - 1 \right).$$

Now, observe that $|x \cdot e| \leq \|x\|$, so the term $\frac{2x \cdot e + 1}{\|x\|^2}$ should be small (at most $O(\|x\|^{-1})$); let us also recall the Taylor expansion $(1 + y)^{\alpha/2} = 1 + \frac{\alpha}{2} y - \frac{\alpha}{4} \left( 1 - \frac{\alpha}{2} \right) y^2 + O(y^3)$. Using that, we continue the above calculation:

$$\|x + e\|^{\alpha} - \|x\|^{\alpha} = \|x\|^{\alpha} \left( \frac{\alpha x \cdot e}{\|x\|^2} + \frac{\alpha}{2 \|x\|^2} - \frac{\alpha}{4} \left( 1 - \frac{\alpha}{2} \right) \frac{(2x \cdot e + 1)^2}{\|x\|^4} + O(\|x\|^{-3}) \right)$$
2.3 Lyapunov functions

\[
= \alpha \|x\|^{\alpha-2} \left( x \cdot e + \frac{1}{2} - \left( 1 - \frac{\alpha}{2} \right) \frac{(x \cdot e)^2}{\|x\|^2} + O(\|x\|^{-1}) \right). \tag{2.15}
\]

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 \(x = (x_1, x_2) \in \mathbb{Z}^2\),

\[
\sum_{e \in \{\pm e_1, \pm e_2\}} x \cdot e = 0, \quad \text{and} \quad \sum_{e \in \{\pm e_1, \pm e_2\}} (x \cdot e)^2 = 2x_1^2 + 2x_2^2 = 2\|x\|^2. \tag{2.16}
\]

Using (2.15) and (2.16), we then obtain for \(f(x) = \|x\|^\alpha\), as \(\|x\| \to \infty\),

\[
\mathbb{E}[f(S_{n+1}) - f(S_n) \mid S_n = x] = \frac{1}{4} \sum_{e \in \{\pm e_1, \pm e_2\}} (\|x + e\|^\alpha - \|x\|^\alpha) \\
= \alpha \|x\|^{\alpha-2} \left( \frac{1}{2} - \left( 1 - \frac{\alpha}{2} \right) \frac{\|x\|^2}{2\|x\|^2} + O(\|x\|^{-1}) \right) \\
= \alpha \|x\|^{\alpha-2} \left( \frac{\alpha}{4} + O(|x|^{-1}) \right), \tag{2.17}
\]

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\(^{13}\). 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

\[
\ln \|x + e\| - \ln \|x\| = \ln \left( \frac{(x + e) \cdot (x + e)}{\|x\|^2} \right) \\
= \ln \left( 1 + \frac{2x \cdot e + 1}{\|x\|^2} \right)
\]

\(^{13}\) 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...
Recurrence

\[\frac{2x \cdot e}{\|x\|^2} + \frac{1}{\|x\|^2} - \frac{2(x \cdot e)^2}{\|x\|^4} + O(\|x\|^{-3}),\]

so, using (2.16) again, we obtain (as \(x \to \infty\))

\[E[f(S_{n+1}) - f(S_n) \mid S_n = x] = \frac{1}{4} \sum_{e \in \{\pm e_1, \pm e_2\}} (\ln \|x + e\| - \ln \|x\|)\]

\[= \frac{1}{\|x\|^2} - \frac{1}{4} \times \frac{4\|x\|^2}{\|x\|^4} + O(\|x\|^{-3})\]

\[= O(\|x\|^{-3}),\]

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

\[
\ln\|x + e\| - \ln\|x\| = \ln \left(1 + \frac{2x \cdot e + 1}{\|x\|^2}\right)
\]

\[
= \frac{2x \cdot e}{\|x\|^2} + \frac{1}{\|x\|^2} - \frac{2(x \cdot e)^2}{\|x\|^4} - \frac{2x \cdot e}{\|x\|^4} - \frac{1}{2\|x\|^4}
\]

\[
+ \frac{8(x \cdot e)^3}{3\|x\|^6} - \frac{4(x \cdot e)^2}{\|x\|^6} - \frac{4(x \cdot e)^4}{\|x\|^8} + O(\|x\|^{-5}).
\]

Then, using (2.16) together with the fact that

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

we obtain

\[E[f(S_{n+1}) - f(S_n) \mid S_n = x] = \frac{1}{\|x\|^2} - \frac{2\|x\|^2}{\|x\|^4} - \frac{2(x_1^4 + x_2^4)}{\|x\|^8} + O(\|x\|^{-5})\]

\[= \|x\|^{-4} \left(\frac{3}{2} - \frac{2(x_1^4 + x_2^4)}{\|x\|^4} + O(\|x\|^{-1})\right). \tag{2.19}\]

We want the right-hand side of (2.19) 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.,

\[\text{the reader is invited to check that only one extra term is not enough.}\]
2.3 Lyapunov functions

$x = (\pm a, \pm a)$, so that $\frac{2(x^i + x^j)}{\|x\|^4} = \frac{2(a^i + a^j)}{4a^2} = 1$, we obtain that the expression in the parentheses is $\frac{3}{2} - 1 + O(||x||^{-1})$, which is strictly positive for all large enough $x$.

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)

\[
\ln^\alpha \|x + e\| - \ln^\alpha \|x\| = \ln^\alpha \|x\| \left(\ln^\alpha \|x + e\| - 1\right) = \ln^\alpha \|x\| \left(\frac{\ln((\|x\|^2(1 + \frac{2x \cdot e + 1}{\|x\|^2}))^{\alpha}}{\ln \|x\|^2} - 1\right) = \ln^\alpha \|x\| \left(1 + (\ln \|x\|^2)^{-1} \ln \left(1 + \frac{2x \cdot e + 1}{\|x\|^2}\right)\right)^\alpha - 1
\]

(by (2.18))

\[
= \ln^\alpha \|x\| \left(1 + (\ln \|x\|^2)^{-1} \left(\frac{2x \cdot e}{\|x\|^2} + \frac{1}{\|x\|^2} \left(\frac{2(x \cdot e)^2}{\|x\|^4} + O(||x||^{-3})\right)\right)\right)^\alpha - 1
\]

\[
= \ln^\alpha \|x\| \left(\frac{\alpha(\ln \|x\|^2)^{-1}}{\ln \|x\|^2} \left(\frac{2x \cdot e}{\|x\|^2} + \frac{1}{\|x\|^2} - \frac{2(x \cdot e)^2}{\|x\|^4} + O(||x||^{-3})\right)\right)
\]

\[
- \frac{\alpha(1-\alpha)}{2} (\ln \|x\|^2)^{-2} \frac{4(x \cdot e)^2}{\|x\|^4} + O(||x||^{-3}(\ln \|x\||^{-2}))
\]

Then, using (2.16) we obtain

\[
E[f(S_{n+1}) - f(S_n) \mid S_n = x] = \frac{\alpha}{2} \ln^\alpha \|x\| \left(\frac{1}{\|x\|^2} - \frac{\|x\|^2}{\|x\|^3} + O(||x||^{-3})\right)
\]

\[
- \frac{(1-\alpha)}{2} (\ln \|x\|^2)^{-1} \frac{2||x||^2}{\|x||^3} + O((||x|| \ln \|x\||)^{-2})
\]
Recurrence

\[ -\frac{\alpha}{2\|x\|^2 \ln^{2-\alpha}\|x\|} \left( \frac{1 - \alpha}{2} + O(\|x\|^{-1} \ln \|x\|) \right) , \]

which is negative for all sufficiently large $x$. Thus Theorem 2.4 shows that SRW on $\mathbb{Z}^2$ is recurrent, 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 $f(x) = \ln^n \|x\|$ with $\alpha \in (0, 1)$)

\[
\mathbb{E}[f(X_{n+1}) - f(X_n) \mid X_n = x] = -\frac{\alpha}{\|x\|^2 \ln^{2-\alpha}\|x\|} \left( -\ln \|x\|^2 \mathbb{E}_x x \cdot D_x - \ln \|x\|^2 \mathbb{E}_x \|D_x\|^2 
\right.
\]

\[
+ \ln \|x\|^2 \frac{2\mathbb{E}_x (x \cdot D_x)^2}{\|x\|^2} + 2(1 - \alpha) \frac{\mathbb{E}_x (x \cdot D_x)^2}{\|x\|^2} + O((\ln \|x\|)^{-1}) \bigg) .
\]

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 2.14 for some further ideas.

2.4 Exercises

Combinatorial proofs (Section 2.1):

**Exercise 2.1** Understand the original proof of Polya [56]. (Warning: it uses generating functions, and it is in German.)

\[ ^{15} \text{finally!} \]
2.4 Exercises

**Exercise 2.2** Let $p_{2n}^{(d)}$ be the probability that $d$-dimensional SRW finds itself at its starting position after $2n$ steps. In the end of Section 2.1 we have seen that $p_{2n}^{(2)} = (p_{2n}^{(1)})^2$, because it is possible to decouple the components of a two-dimensional SRW. Can this decoupling be done for at least some $d \geq 3$ as well (so that, in particular, $p_{2n}(0,0) = (p_{2n}^{(1)}(0,0))^d$ would hold)?

**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.

**Exercise 2.4** Being $S_n$ the two-dimensional SRW, prove that $S_n \cdot (e_1 + e_2)$ and $S_n \cdot (e_1 - e_2)$ are independent one-dimensional SRWs.

**Exercise 2.5** Using the result of the previous exercise, can one derive a version of local Central Limit Theorem for the two-dimensional SRW from the de Moivre-Laplace theorem?

**Electrics networks (Section 2.2):**

**Exercise 2.6** Consider a nearest-neighbour random walk on $\mathbb{Z}$ with transition probabilities $p_x = p(x, x-1)$, $q_x = 1-p_x = p(x, x+1)$. Show that it is reversible and write its reversible measure.

**Exercise 2.7** 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.

**Exercise 2.8** Consider a reversible Markov chain on a finite state space $\Sigma$ with reversible measure $\pi$. Let us think of it as a linear operator $P$ acting on functions $f : \Sigma \to \mathbb{R}$ in the following way: $Pf(x) = \sum_{y \in \Sigma} p(x,y) f(y)$. Prove that all the eigenvalues of $P$ are real and belong to $[-1, 1]$.

**Lyapunov functions (Section 2.3):**

**Exercise 2.9** Prove the “only if” part of the transience criterion (Theorem 2.5).

**Exercise 2.10** Find a Lyapunov-function proof of transience of one-dimensional nearest-neighbor random walk with constant drift.
Exercise 2.11  Now, consider a Markov chain \((X_n)\) on \(\mathbb{Z}_+\) such that

- there exists \(K > 0\) such that \(|X_{n+1} - X_n| \leq K\) a.s. (i.e., the jumps are uniformly bounded);
- there exists \(\varepsilon > 0\) such that \(\mathbb{E}_x X_1 \geq x + \varepsilon\) for all \(x\) (i.e., the drift is uniformly positive).

Prove that it is transient.

Exercise 2.12  Using Theorem 2.5, prove that simple random walk in dimensions \(d \geq 3\) is transient. Hint: use \(f(x) = \|x\|^{-\alpha}\) for some \(\alpha > 0\).

Exercise 2.13  Show that \(f(x) = \ln \ln \|x\|\) (suitably redefined at the origin and at the sites which are at distance at most \(e\) from it) would also work for proving the recurrence of the two-dimensional simple random walk.

Exercise 2.14  Using Lyapunov functions, prove the recurrence of a two-dimensional spatially homogeneous zero-mean random walk with bounded jumps.

Exercise 2.15  Understand the proof of the “only if” part of the recurrence criterion (Theorem 2.4) — see the proof of Theorem 2.2.1 of [33]. Can you find\(^\text{16}\) a simpler proof?

Exercise 2.16  Is it always possible to find a function which is a martingale outside the origin (i.e., in Theorem 2.4 \(A\) is a singleton and (2.4) holds with equality) for a recurrent Markov chain? Also, prove that the above is possible for a recurrent Markov chain on \(\mathbb{Z}_+\) with nearest-neighbour jumps by writing down such a function explicitly.

Exercise 2.17  The following result (also known as Foster’s criterion or Foster-Lyapunov theorem) provides a criterion for the positive recurrence of an irreducible Markov chain:

An irreducible Markov chain \(X_n\) on a countable state space \(\Sigma\) 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

\[
\mathbb{E}[f(X_{n+1}) - f(X_n) \mid X_n = x] \leq -\varepsilon, \text{ for all } x \in \Sigma \setminus A, \quad (2.20)
\]

\[
\mathbb{E}[f(X_{n+1}) \mid X_n = x] < \infty, \text{ for all } x \in A. \quad (2.21)
\]

\(^{16}\) if you find it, please, let me know
2.4 Exercises

(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 [52]).

**Exercise 2.18** For the $d$-dimensional simple random walk, show that the first and the second moments of $\Delta_{x} := \|S_{1}\| - \|x\|$ under $E_{x}$ are given by

\[
E_{x}\Delta_{x} = \frac{d - 1}{2d\|x\|} + O(\|x\|^{-2}), \quad (2.22)
\]
\[
E_{x}\Delta_{x}^{2} = \frac{1}{d} + O(\|x\|^{-1}). \quad (2.23)
\]

**Exercise 2.19** Suppose now that $(X_{n}, n \geq 0)$ is a time-homogeneous Markov chain on an unbounded subset $\Sigma$ of $\mathbb{R}_{+}$. Assume that $X_{n}$ has uniformly bounded increments, so that

\[
P\{|X_{n+1} - X_{n}| \leq b\} = 1 \quad (2.24)
\]

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

\[
\mu_{k}(x) := E[(X_{n+1} - X_{n})^{k} | X_{n} = x].
\]

The first moment function, $\mu_{1}(x)$, is also called the one-step mean drift of $X_{n}$ at $x$.

Lamperti [42, 43, 44] 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 2.20** Let $d \geq 3$. Prove that for any $\varepsilon > 0$ there exists large enough $C_{d} = C_{d}(\varepsilon)$ such that $\|S_{n \wedge \tau}\|^{-(d - 2) + \varepsilon}$ is a supermartingale and $\|S_{n \wedge \tau}\|^{-(d - 2) - \varepsilon}$ is a submartingale, where $\tau = \tau_{B(C_{d})}$. What happens in the case $\varepsilon = 0$?
Some potential theory for simple random walks

Disclaimer: this chapter is by no means a systematic treatment of the subject, not even remotely so. If the reader is looking for one, the author can recommend e.g. [29, 62] or Chapters 4 and 6 of [46]. Here we rather adopt a “customer’s point of view”: we only recall a few general notions and tools that permit us to obtain estimates on what we are concerned about in this book — probabilities related to simple random walks. We also do not try to discuss the reason why it is called “potential theory” and what exactly are its relations to the classical theory of harmonic functions — this would take quite some time, and the author has to confess that he does not understand it quite well anyway.

We are going to consider the transient and the recurrent cases separately; it is true that in this book we are mainly concerned with the latter one, but it is still more convenient to begin with the transient case (i.e., SRW in dimensions \(d \geq 3\)) which is somehow conceptually simpler\(^1\). Let us begin, though, with the following result, which is dimension-independent:

**Proposition 3.1** Let \(h : \mathbb{Z}^d \to \mathbb{R}\) be a harmonic function, i.e.,

\[
h(x) = \frac{1}{2d} \sum_{y \sim x} h(y) \quad \text{for all } x \in \mathbb{Z}^d.
\]  

(3.1)

Assume also that it is bounded, i.e., there exists \(K > 0\) such that \(|h(x)| \leq K\) for all \(x \in \mathbb{Z}^d\). Then \(h\) is constant.

**Proof** Let us reason by contradiction: assume that a function \(h\) as above is not constant. Then it must be nonconstant on both the set of even\(^2\) sites and the set of odd sites. Indeed, its value

\(^1\) also, we’ll need some of these transient-case results in Chapter 6 when discussing the “classical” random interlacement model in dimensions \(d \geq 3\)

\(^2\) recall that even (odd) sites are sites with even (odd) sum of their coordinates
3.1 Transient case

on an even site is the average of its values on the neighbouring odd sites and vice versa; so, if it is equal to a constant on one of these two sets, it will have to be the same constant on the other one. Now, we can find two sites \( x \) and \( y \) of the same parity such that \( x - y = \pm 2e_k \) for some \( k \in \{1, \ldots, d\} \) and \( h(x) \neq h(y) \). Note that, if \( h \) is harmonic in the sense of (3.1), then so are its translations/rotations/translations. Therefore, without restricting generality, we can assume that \( h(e_1) \neq h(-e_1) \).

For \( x = (x_1, x_2, \ldots, x_d) \in \mathbb{Z}^d \), let us denote \( \bar{x} = (-x_1, x_2, \ldots, x_d) \) the “mirrored” site with respect to the hyperplane orthogonal to the first coordinate axis. Let \((S_n)\) be the SRW starting at \( e_1 \), and \((\bar{S}_n)\) be the corresponding “mirrored” SRW (which, clearly, starts at \((-e_1)\)). Define the stopping time

\[
\sigma = \min\{k : S_k \cdot e_1 = 0\} = \min\{k : S_k = \bar{S}_k\}
\]

to be the moment when \( S \) meets \( \bar{S} \). Note that \( \sigma < \infty \) a.s., due to the recurrence of one-dimensional SRW.

Now, the harmonicity of \( h \) implies that \( h(S_n) \) is a martingale: indeed, \( \mathbb{E}(h(S_{n+1}) | S_n = x) = (2d)^{-1} \sum_{y \sim x} h(y) = h(x) \) by (3.1); clearly, the same argument also implies that \( h(\bar{S}_n) \) is a martingale. Since \( h \) is bounded, Corollary 1.7 implies that

\[
\mathbb{E}h(S_\sigma) = \mathbb{E}h(\bar{S}_\sigma) = h(-e_1),
\]

which is the desired contradiction (recall that we just assumed that \( h(x) \neq h(y) \)).

3.1 Transient case

First, let us go to dimensions \( d \geq 3 \), where the SRW is transient. We specifically concentrate on the simple random walk case, although a similar theory can be developed for random walks with arbitrary jump distribution, or even general transient reversible Markov chains. We need first to recall some basic definitions related to simple random walks in higher dimensions.

The three main objects that we need are: the Green’s function, the capacity, and the harmonic measure. Specifically, it holds that:

- The Green’s function \( G \) is harmonic outside the origin, and so the process \( G(S_{n\wedge \tau_0}) \) is a martingale. This gives us a convenient
tool for calculating certain exit probabilities via the Optional Stopping Theorem\(^3\).

- Informally speaking, the capacity of a set measures how big is this set from the point of view of the simple random walk. This permits us to obtain some refined bounds on probabilities of hitting sets.

- The harmonic measure is a probability distribution that lives on the boundary of a set, and is the “conditional entrance measure from infinity”. When we are concerned with entrance measures starting from some fixed site, sometimes it is possible to argue that this entrance measure is not much different from the harmonic measure, thus allowing us to have some control on where the random walk enters that set.

Let us now elaborate.

Green’s function.

For \(d \geq 3\), the Green’s function \(G : (\mathbb{Z}^d)^2 \to \mathbb{R}_+\) is defined in the following way:

\[
G(x, y) = \mathbb{E}_x \left( \sum_{k=0}^{\infty} 1 \{ S_k = y \} \right) = \sum_{k=0}^{\infty} \mathbb{P}_x[S_k = y].
\] (3.2)

That is, \(G(x, y)\) is equal to the mean number of visits to the origin starting from \(x\). It is important to note that in the case \(x = y\) we do count this as one “initial” visit (so, in particular, \(G(x, x) > 1\) in all dimensions). By symmetry it holds that \(G(x, y) = G(y, x) = G(0, y - x)\); thus, we can abbreviate \(G(y) := G(0, y)\) so that \(G(x, y) = G(x - y) = G(y - x)\). Now, a very important property of \(G(\cdot)\) is that it is harmonic outside the origin, i.e.,

\[
G(x) = \frac{1}{2d} \sum_{y \sim x} G(y) \quad \text{for all } x \in \mathbb{Z}^d \setminus \{0\}. \tag{3.3}
\]

Since, as observed above, \(G(x)\) is the mean number of visits to the origin starting from \(x\), 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. An immediate consequence of (3.3) is the following

**Proposition 3.2** *The process \(G(S_{n \wedge \tau_0})\) is a martingale.*

\(^3\) stated as Theorem 1.6 in this book; in fact, Corollary 1.7 will be usually enough
3.1 Transient case

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\(^4\) that $G(x)$ must be of order $\|x\|^{-(d-2)}$, due to the following heuristic argument. Fix $x \in \mathbb{Z}^d$, $x \neq 0$. First, as is well-known (think e.g. of the Central Limit Theorem), the simple random walk is \textit{diffusive}, i.e., it 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\(^5\) 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\(^6\) of order $m^{-d/2}$; therefore, the Green’s function’s value in $x$ is roughly

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

Note also that

$$G(x) = P_x[\tau_0 < \infty]G(0); \quad (3.4)$$

indeed, starting from $x$, the mean number of visits to 0 is zero given that $\tau_0 = \infty$ and is $G(0)$ given that $\tau_0 < \infty$, so the above again comes out from the total expectation formula. This implies that the probability of ever visiting $y$ starting from $x$ (which is the same as the probability of ever visiting 0 starting from $x-y$) is also of order $\|x-y\|^{-(d-2)}$.

Now, since the SRW is “roughly spherically symmetric”, it is reasonable to expect that the Green’s function should be asymptotically “well behaved” and, in particular, depend (almost) only on $\|x\|$ as $x \to \infty$. In fact, it is possible to obtain that

$$G(x) = \frac{\gamma_d}{\|x\|^{d-2}} + O(\|x\|^{-d}), \quad (3.5)$$

with $\gamma_d = \frac{\Gamma(d/2)^d}{\pi^{d/2}(d-2)}$, see Theorem 4.3.1 of [46]. We prefer not to include the complete proof here (that is, we will take this fact for granted), but see Exercises 3.3, 3.4, and 3.5.

\(^4\) “to see” does not mean “to prove” \(^5\) well, not really (observe that the simple random walk has period two), but you understand what I mean \(^6\) recall that we used a very similar heuristic argument in the beginning of Section 2.1
We are now able to obtain a straightforward (and useful) estimate for the probability that the simple random walk escapes an annulus through its outer boundary:

**Lemma 3.3** For all \( x \in \mathbb{Z}^d, d \geq 3 \), and \( R > r > 0 \) such that \( x \in \mathbb{B}(R) \setminus \mathbb{B}(r) \) we have

\[
\mathbb{P}_x \left[ \tau_{\partial \mathbb{B}(R)}^+ \leq \tau_{\partial \mathbb{B}(r)}^+ \right] = \frac{r^{-(d-2)} - \|x\|^{-(d-2)} + O(r^{-(d-1)})}{r^{-(d-2)} - R^{-(d-2)} + O(R^{-(d-1)})},
\]

(3.6)
as \( r \to \infty \).

**Proof** This comes out of an application of the Optional Stopping Theorem to the martingale \( G(S_{\tau_{\partial \mathbb{B}(r)}}) \). Indeed, let us abbreviate by \( p \) the probability in the left-hand side of (3.5); also, let \( g_\downarrow \) and \( g_\uparrow \) be the (conditional) expected values of \( G(S) \) on first-hitting the inner and the outer boundaries of the annulus, i.e.,

\[
g_\downarrow = \mathbb{E}_x \left( G(S_{\tau_{\partial \mathbb{B}(r)}}) \mid \tau_{\partial \mathbb{B}(r)}^+ < \tau_{\partial \mathbb{B}(R)}^+ \right),
\]

\[
g_\uparrow = \mathbb{E}_x \left( G(S_{\tau_{\partial \mathbb{B}(r)}}) \mid \tau_{\partial \mathbb{B}(r)}^+ < \tau_{\partial \mathbb{B}(R)}^+ \right).
\]

Since \( S_0 = x \), the Optional Stopping Theorem (Corollary 1.7) implies that

\[
G(x) = G(S_0) = \mathbb{E}_x G(S_{\tau_{\partial \mathbb{B}(r)}}) = pg_\uparrow + (1 - p)g_\downarrow,
\]

meaning that

\[
p = \frac{g_\downarrow - G(x)}{g_\downarrow - g_\uparrow},
\]

(3.7)

Note that \( S_{\tau_{\partial \mathbb{B}(r)}} \in \partial \mathbb{B}(r) \) because \( x \notin \mathbb{B}(r) \); since for any \( y \in \partial \mathbb{B}(h) \) it holds that \( h - 1 < \|y\| \leq h \), we have by (3.5) that

\[
g_\downarrow = \frac{\gamma_d}{r^{d-2}} \left( 1 + O(r^{-1}) \right),
\]

and

\[
g_\uparrow = \frac{\gamma_d}{R^{d-2}} \left( 1 + O(R^{-1}) \right).
\]

Plugging (3.5) and the above into (3.7), we obtain (3.6). \( \square \)

Sending \( R \) to infinity in (3.6), we obtain that, for any \( x \notin \mathbb{B}(r) \),

\[
\mathbb{P}_x \left[ \tau_{\mathbb{B}(r)} = \infty \right] = 1 - \frac{\|x\|^{-(d-2)} + O(r^{-(d-1)})}{r^{-(d-2)} + O(r^{-(d-1)})}
\]

as \( r \to \infty \).
Let us now move on to the next fundamental notion of the (discrete) potential theory.

Capacity.
For finite $A \subset \mathbb{Z}^d$ and $x \in \mathbb{Z}^d$, let us denote

$$E_A(x) = \mathbb{P}_x[\tau_A^+ = \infty] 1\{x \in A\}. \quad (3.9)$$

By definition, this quantity is 0 outside $A$, and, at $x \in A$, it is the escape probability from $A$. Note that $E_A(x)$ can be positive only if $x \in \partial A$.

The capacity of a finite set $A \subset \mathbb{Z}^d$ is defined by

$$\text{cap}(A) = \sum_{x \in A} E_A(x); \quad (3.10)$$

clearly, it holds that the capacity is translationally invariant, and $\text{cap}(A) = \text{cap}(\partial A)$.

What this notion of capacity is good for? To answer this question, we need some preparatory steps. Consider a finite $A \subset \mathbb{Z}^d$.

Let us prove a relation that will be used many times later in this chapter: for any $y \in \mathbb{Z}^d$, it holds that

$$\mathbb{P}_x[\tau_A < \infty] = \sum_{y \in A} G(x,y) E_A(y) = \sum_{y \in \mathbb{Z}^d} G(x,y) E_A(y). \quad (3.11)$$

For the proof, we use an important idea called the last-visit decomposition. On the event $\{\tau_A < \infty\}$, 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, just set $\sigma$ to be 0). By transience (recall that $A$ is finite!), it is clear that $\sigma$ is a.s. finite. It is important to note that $\sigma$ is not a stopping time, which actually turns out to be good! To understand why, let us first observe that, by the strong Markov property, the walk’s trajectory after any stopping time is “free”, that is, it just behaves as simple random walk starting from the position it had at that stopping time. Now, if we know that $\sigma$ happened at a given moment, then we know something about the future, namely, we know that the walk must not return to $A$ anymore. In other
words, after $\sigma$ the walk’s law is the \textit{conditioned} (on $\tau_A^+ = \infty$) one\footnote{strictly speaking, this statement needs a rigorous proof; we leave it to the reader as an exercise}. Now, look at Figure 3.1: what is the probability that the walker visits $y \in A$ exactly $k$ times (on the picture, $k = 2$), and then escapes to infinity, being $y$ the last visited point of $A$?

This probability is the total weight of the trajectories such that, first, they visit $y$ exactly $k$ times and then escape to infinity not touching $A$ anymore. This means that for any $y \in A$ and $k \geq 1$ it holds that

$$P_x[\text{exactly } k \text{ visits to } y, S_\sigma = y] = P_x[\text{at least } k \text{ visits to } y] E_{S_\sigma}(y).$$ \hfill (3.12)

Okay, maybe, at first sight it is not clear why “exactly $k$ visits” in the left-hand side became “at least $k$ visits” in the right-hand side. To understand this, think again of the piece of the trajectory till $\sigma$. If we consider \textit{only} such trajectories, they correspond to the event \{y is visited \textit{at least} $k$ times\} – indeed, if we only observe the trajectory till the $k$th visit, we then know that this last event occurred.

Then, summing (3.12) in $k$ from 1 to $\infty$ we obtain$^8$

$$P_x[\tau_A < \infty, S_\sigma = y] = G(x, y) E_{S_\sigma}(y),$$ \hfill (3.13)

and summing the above in $y \in A$ we obtain (3.11).

Now, we are able to obtain the following useful corollary of (3.11):

$$\text{cap}(A) \min_{z \in A} G(x, z) \leq P_x[\tau_A < \infty] \leq \text{cap}(A) \max_{z \in A} G(x, z);$$ \hfill (3.14)

informally, at least in the case when $\min_{z \in A} \|x - z\|$ and $\max_{z \in A} \|x -$
3.1 Transient case

...z∥ are of the same order, the above means that the probability of ever hitting the set A is proportional to its capacity, and (by (3.5)) is inversely proportional to the distance to that set to power \(d - 2\). This justifies the (already mentioned) intuition that the capacity measures how large is the set from the point of view of the simple random walk.

Next, let us obtain the exact expressions (in terms of the Green function \(G\)) for capacities of one- and two-point sets. We are going to prove that

\[
\text{cap}({0}) = \frac{1}{G(0)}, \quad (3.15)
\]
\[
\text{cap}({0, x}) = \frac{2}{G(0) + G(x)} \quad (3.16)
\]

(by translation invariance, the above also yeild the expressions for \(\text{cap}({x})\) and \(\text{cap}({x, y})\)). Indeed, first, under \(P_0\), the number of visits to the origin (counting the “initial” one at time 0) is a Geometric random variable with success probability \(E_{S_{(0)}}(0)\). So, its mean is \(1/E_{S_{(0)}}(0)\) on one hand and \(G(0)\) on the other hand, meaning that \(E_{S_{(0)}}(0) = 1/G(0)\). Since, by definition, \(\text{cap}({0}) = E_{S_{(0)}}(0)\), we obtain (3.15). The argument for two-point sets is very similar: let \(p := E_{S_{(0, x)}}(0); by symmetry, it holds also that \(p = E_{S_{(0, x)}}(x)\). So, the total number of visits to \(\{0, x\}\) (starting from either 0 or \(x\)) has Geometric distribution with success probability \(p\), meaning that \(p^{-1} = G(0) + G(x)\). Since \(\text{cap}({0, x}) = E_{S_{(0, x)}}(0) + E_{S_{(0, x)}}(x) = 2p\), (3.16) follows.

The above argument can be also used to calculate the capacities of other symmetric sets, like e.g. the four vertices of a “spacial rectangle” or other similar things.

As for the capacity of a \(d\)-dimensional (discrete) ball, let us show that (with the constant \(\gamma_d\) from (3.5))

\[
\text{cap}(B(r)) \sim \frac{r^{d-2}}{\gamma_d} \quad \text{as } r \to \infty. \quad (3.17)
\]

To understand why \(\text{cap}(B(r))\) should be of order \(r^{d-2}\), consider \(x\) such that \(\|x\| \in [r + c, r + c + 1]\), where \(c > 0\) is a large enough constant, see Figure 3.2. Note that (3.8) yeilds

\[
\mathbb{P}_x[\tau_{B(r)}^+ = \infty] = 1 - \left(\frac{r}{\|x\|}\right)^{d-2} + O(r^{-1})
\]
42

Potential theory

\begin{equation}
\begin{aligned}
    B(r) &\xrightarrow{r+c} x \\
    r+c+1 &\xrightarrow{r+c} + c
\end{aligned}
\end{equation}

Figure 3.2 Escaping from a ball.

\[
\begin{aligned}
    &= 1 - \left(1 + \frac{\|x\| - r}{r}\right)^{-d-2} + O(r^{-1}) \\
    &= -\frac{(d-2)c}{r} + O(r^{-1})
\end{aligned}
\]

(where there is no dependence on \(c\) in the \(O\)'s). The two terms in the above expression are of the same order, but we are allowed to make \(c\) as large as we want; this will imply that, for such \(x\), \(\mathbb{P}_x[\tau_{B(r)}^+] = \infty \land r^{-1}\). This by its turn means that \(\mathbb{E}_y = r\). Since \(|\partial B(r)| \sim r^{d-1}\), we see that the capacity of \(B(r)\) is indeed of order \(r^{d-2}\).

To obtain the more precise relation (3.17), we need the following

Proposition 3.4  For any finite \(A \subset \mathbb{Z}^d\) it holds that

\[
\begin{aligned}
    \text{cap}(A) &= \lim_{x \to \infty} \frac{\mathbb{P}_x[\tau_A < \infty]}{G(x)} = \lim_{x \to \infty} \frac{\|x\|^{d-2}}{\gamma_d} = \mathbb{P}_x[\tau_A < \infty].
\end{aligned}
\]  (3.18)

Proof  This is a direct consequence of (3.14) and (3.5): as \(x \to \infty\), both \(\min_{z \in A} G(x, z)\) and \(\max_{z \in A} G(x, z)\) are asymptotically equivalent to \(\gamma^d \|x\|^{-d-2}\).

It remains to observe that the asymptotics of the capacity of a ball (relation (3.17)) follows from (3.8) and Proposition 3.4.

We are now going to present an interesting application of the technique we just developed. Let us recall the following

Definition 3.5  We say that \(A \subset \mathbb{Z}^d\) is recurrent, if \(\mathbb{P}_x[\tau_A < \infty] = 1\) for all \(x \in \mathbb{Z}^d\). Otherwise, we call the set \(A\) transient.
3.1 Transient case

Clearly, the question if a set is recurrent or transient is, in principle, not so trivial. As a start, we obtain the following result:

**Proposition 3.6** If \( A \) is recurrent, then

\[
\sum_{k=1}^{\infty} 1\{S_k \in A\} = \infty \quad \mathbb{P}_x\text{-a.s.,}
\]

for all \( x \in \mathbb{Z}^d \), that is, regardless of the starting point, \( A \) is visited infinitely many times a.s.. If \( A \) is transient, then

\[
\sum_{k=1}^{\infty} 1\{S_k \in A\} < \infty \quad \mathbb{P}_x\text{-a.s.,}
\]

for all \( x \in \mathbb{Z}^d \).

**Proof** The first part (that recurrence implies that \( A \) is visited infinitely many times a.s.) is evident; let us prove the second part. Let \( A \) be a transient set and let us define the function \( h \) by

\[
h(x) = \mathbb{P}_x[A \text{ is visited infinitely often}].
\]

An immediate observation is that \( h \) is harmonic – just use the formula of total probability, conditioning on the first step. So, since \( h \) is also obviously bounded, Proposition 3.1 implies that this function is constant, \( h(x) = p \in [0, 1] \) for all \( x \).

Now, what can be the value of \( p \)? First, it has to be strictly less than 1 by transience of \( A \) (there is at least one site \( x_0 \) such that \( \mathbb{P}_{x_0}[\tau_A < \infty] < 1 \) and, obviously, \( h(x_0) \leq \mathbb{P}_{x_0}[\tau_A < \infty] \)). Next, write (conditioning on the first entrance to \( A \), if any)

\[
p = h(x_0)
\]

\[
= \sum_{y \in A} \mathbb{P}_{x_0}[\tau_A < \infty, S_{\tau_A} = y]h(y)
\]

\[
= p \sum_{y \in A} \mathbb{P}_{x_0}[\tau_A < \infty, S_{\tau_A} = y]
\]

\[
= p \mathbb{P}_{x_0}[\tau_A < \infty],
\]

and (since, as we just assumed, \( \mathbb{P}_{x_0}[\tau_A < \infty] < 1 \)) this implies that \( p = 0 \).

As we know already, in dimensions \( d \geq 3 \) the one-point sets are transient; Proposition 3.6 then implies\(^9\) that all finite sets are

\(^9\) how exactly?
transient as well. But what can we say about infinite sets? The answer is given by the following theorem:

**Theorem 3.7 (Wiener’s criterion)** For \( d \geq 3 \), \( A \subset \mathbb{Z}^d \) is recurrent if and only if

\[
\sum_{k=1}^{\infty} \frac{\text{cap}(A_k)}{2^{(d-2)k}} = \infty,
\]

where

\[ A_k = \{ x \in A : 2^{k-1} \leq \|x\| \leq 2^k \} \]

is the intersection of \( A \) with the annulus \( B(2^k) \setminus B(2^{k-1}) \).

The proof is left to the reader (Exercises 3.11 and 3.12).

The last of the three main ingredients that make our naive view of the potential theory is harmonic measure; we start looking at it now.

**Harmonic measure.**

As before, let \( A \) be a finite subset of \( \mathbb{Z}^d, \ d \geq 3 \). The harmonic measure \( \text{hm}_A(\cdot) \) on \( A \) is defined by

\[
\text{hm}_A(x) = \frac{E_{S_A}(x)}{\text{cap}(A)}, \quad x \in A,
\]

that is, the value of \( \text{hm}_A(x) \) is proportional to the escape probability from \( x \) to infinity. Remarkably, it is also true that \( \text{hm}_A \) is the “entrance measure to \( A \) from infinity”, that is, the following result holds:

**Theorem 3.8** For all \( y \in A \), we have

\[
\text{hm}_A(y) = \lim_{x \to \infty} \mathbb{P}_x[S_{\tau_A} = y \mid \tau_A < \infty].
\]

Why should the above be valid? Let us first give an informal explanation. Consider \( y, z \in \partial A \), such that \( y \neq z \) and both \( E_{S_A}(y) \) and \( E_{S_A}(z) \) are strictly positive. Then, the ratio of the “total weights” of trajectories which escape \( A \) from, respectively, \( y \) and \( z \), equals \( E_{S_A}(y) / E_{S_A}(z) \). Now, if \( x \) is very far away from \( A \) and the walker that started somewhere at \( A \) happens to pass through \( x \), it likely does not “remember” its exact starting position. Since the
time reversal does not change the “weight” of the trajectory\textsuperscript{10}, the ratio of the chances that a trajectory passing through $x$ will end up in $y$ (respectively, in $z$) should be then $\text{Es}_A(y)/\text{Es}_A(z)$ as well.

Now, the rigorous proof of the above result may look not very intuitive at first sight, but note that it also makes use of this reversibility property.

**Proof of Theorem 3.8.** We now use a trajectory-counting argument very similar to the one in the proof of (3.11). For $x \notin A$, $y \in \partial A$, and $n \geq 1$, let us denote by $\Theta_{xy}^{(n)}$ the set of nearest-neighbour trajectories $\varphi = (z_0, \ldots, z_k)$ such that

1. $z_0 = x$, $z_k = y$, and $z_j \notin A$ for all $j \leq k - 1$, i.e., the trajectory ends on the first entrance to $A$, which takes place in $y$;
2. $\sum_{j=0}^{k} \mathbf{1}\{z_j = x\} = n$, i.e., the trajectory visits $x$ exactly $n$ times (note that we do count $z_0 = x$ as one visit);

see Figure 3.3. For such trajectory, we also write $|\varphi| = k$ to denote its length, and $P_{\varphi} = (2d)^{-|\varphi|}$ to denote its weight (a.k.a. probability). Let us also denote by

$$N_x = \sum_{j=0}^{\infty} \mathbf{1}\{S_j = x\}$$

the total number of visits to $x \notin A$, by

$$N_x^\flat = \sum_{j=0}^{\tau_x - 1} \mathbf{1}\{S_j = x\}$$

\textsuperscript{10} formally, for infinite trajectories this only means that $0 = 0$, but you understand what I wanted to say.
the number of visits to $x$ before the first return to $A$, and by

$$N_x^\dagger = \sum_{j=\tau^+_A}^{\infty} 1\{S_j = x\}$$

the number of visits to $x$ after the first return to $A$ (naturally, setting $N_x^\dagger = 0$ on $\{\tau^+_A = \infty\}$).

Now, it is clear that

$$P_x[\tau_A < \infty, S_{\tau_A} = y] = \sum_{n=1}^{\infty} \sum_{\wp \in \Theta^{(n)}(y)} P_{\wp}$$

(3.22)

(we just sum the weights of all trajectories starting at $x$ and entering $A$ at $y$). The next relation may seem a bit less clear, but it is here where we use the reversibility property:

$$P_y[N_x^\dagger \geq n] = \sum_{\wp \in \Theta^{(n)}(x)} P_{\wp}.$$  

(3.23)

Indeed (quite analogously to the proof of (3.11)), when, starting at $y$, we see a reversal of a trajectory from $\Theta^{(n)}(x,y)$, we are sure that the event $\{N_x^\dagger \geq n\}$ occurs. Therefore, we can write

$$P_x[S_{\tau_A} = y \mid \tau_A < \infty] = \frac{P_x[\tau_A < \infty, S_{\tau_A} = y]}{P_x[\tau_A < \infty]}$$

(by (3.22) and (3.23))

$$= (P_x[\tau_A < \infty])^{-1} \sum_{n=1}^{\infty} P_y[N_x^\dagger \geq n]$$

$$= (P_x[\tau_A < \infty])^{-1} E_y N_x^\dagger$$

(since, clearly, $N_x = N_x^\dagger + N_x^\sharp$)

$$= (P_x[\tau_A < \infty])^{-1} (E_y N_x - E_y N_x^\sharp)$$

(conditioning on the position of the first re-entry to $A$)

$$= (P_x[\tau_A < \infty])^{-1} \left( G(y,x) - \sum_{z \in \partial A} P_y[\tau_A^+ < \infty, S_{\tau_A^+} = z] G(z, x) \right).$$

(3.24)
Then, Proposition 3.4 together with (3.5) imply that, for any fixed $z \in \mathbb{Z}^d$,

$$\frac{G(z, x)}{\mathbb{P}_x[\tau_A < \infty]} \to \frac{1}{\text{cap}(A)} \quad \text{as } x \to \infty.$$ 

So, sending $x$ to infinity in (3.24), we obtain

$$\lim_{x \to \infty} \mathbb{P}_x[S_{\tau_A} = y \mid \tau_A < \infty] = \frac{1}{\text{cap}(A)} \left(1 - \sum_{z \in \partial A} \mathbb{P}_y[\tau_A^+ < \infty, S_{\tau_A^+} = z] \right)$$

$$= \frac{1}{\text{cap}(A)} \left(1 - \mathbb{P}_y[\tau_A^+ < \infty] \right)$$

$$= \frac{\mathbb{P}_y[\tau_A^+ = \infty]}{\text{cap}(A)}$$

$$= \text{hm}_A(y),$$

thus concluding the proof of Theorem 3.8. \hfill \Box

Many other interesting things can be said about the transient case, but we prefer to stop here and pass to the recurrent one.  

### 3.2 Potential theory in two dimensions

In this section, we try to do roughly the same as in the previous one, only in two dimensions. As we know, there is one big difference between the dimension two and higher dimensions: as shown in Chapter 2, unlike the higher-dimensional SRW, the two-dimensional SRW 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, here 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. Let us do it now.

\[^{11}\text{revenons à nos moutons} \odot\]
Potential theory

Figure 3.4 The coupling of two random walks starting at 0 and \( x = (4,2) \). Note that their first coordinate become equal at time 3 (when the first walk is at \((2,-1)\) and the second one is at \((2,3)\)), and the walks meet at time 6 at site \((1,1)\).

Potential kernel.

Namely, let us define the potential kernel \( a(\cdot) \) by

\[
a(x) = \sum_{k=0}^{\infty} \left( \mathbb{P}_0[S_k = 0] - \mathbb{P}_x[S_k = 0] \right), \quad x \in \mathbb{Z}^2.
\]  

(3.25)

By definition, it holds that \( a(0) = 0 \). To see that the limit (finite or infinite) in (3.25) actually exists is a little bit more subtle, but still quite elementary. Indeed, Exercise 3.25 (iii) implies that

- if \( x \) is an even site (i.e., the sum of its coordinates is even), then all terms in the summation (3.25) are nonnegative (more precisely, they are positive for even \( k \) and zero for odd \( k \));
- if \( x \) is an odd site, then we have a series with alternating signs in (3.25), but the sums of each two consecutive terms (i.e., \( \mathbb{P}_0[S_{2k} = 0] - \mathbb{P}_x[S_{2k+1} = 0] \)) are again strictly positive and converge to zero, which clearly implies that the sum is well-defined.

Note that the above argument also implies that \( a(x) > 0 \) for all \( x \neq 0 \).

Now, let us convince ourselves that the series converges (i.e., \( a(x) \) is finite) for all \( x \in \mathbb{Z}^2 \), and figure out how large \( a(x) \) should be. The “normal” approach would be employing the local Central
3.2 Potential theory in two dimensions

Limit Theorem\textsuperscript{12} for this, but we prefer to use another interesting and very useful tool called coupling\textsuperscript{13}. It will be a long argument (sorry for that!) since we will need to do a three-parts divide-and-conquer argument (later it will be clear what is meant by this), but it is still nice and instructive. 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: 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 of the two walks are different, and in the same direction in case they are equal, see Figure 3.4. Formally, assume that, 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 $J_n$ and $Z_n$ be independent random variables assuming values in $\{1, 2\}$ and $\{-1, 1\}$ respectively, with equal (to $\frac{1}{2}$) probabilities. Then, we set:

$$
(S_{n+1}', S_{n+1}'') = \begin{cases} 
(S_n' + Z_n e_{J_n}, S_n'' - Z_n e_{J_n}), & \text{if } S_n' \cdot e_{J_n} \neq S_n'' \cdot e_{J_n}, \\
(S_n' + Z_n e_{J_n}, S_n'' + Z_n e_{J_n}), & \text{if } S_n' \cdot e_{J_n} = S_n'' \cdot e_{J_n}.
\end{cases}
$$

Note that if the first (second) coordinates of the two walks are equal at some moment, then they will remain so forever. This means, in particular, that, when the two walks meet, they stay together. Let us assume, for definiteness, that $x$ belongs to the first quadrant, that is, $x = (2b_1, 2b_2)$ for $b_1, b_2 \geq 0$. Let

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

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

$$
T_j = \min\{n \geq 0 : S_n' \cdot e_j = b_j\} = \min\{n \geq 0 : S_n'' \cdot e_j = b_j\} \tag{3.26}
$$

(clearly, they have to meet exactly in the middle). Let also $T = T_1 \lor T_2$ be the coupling time, i.e., the moment when the two walks meet and stay together.

\textsuperscript{12} analogous to the De Moivre-Laplace one, only in two dimensions; see e.g. Theorem 2.1.1 of [46]

\textsuperscript{13} we used it already (without calling it so) in the proof of Proposition 3.1; here, we take it further
Now, we go back to (3.25) and use the strategy usually called “divide and conquer”: write

\[
a(x) = \sum_{k < \|x\|} (\mathbb{P}_0[S_k = 0] - \mathbb{P}_x[S_k = 0]) \\
+ \sum_{k \in [\|x\|, \|x\|^3]} (\mathbb{P}_0[S_k = 0] - \mathbb{P}_x[S_k = 0]) \\
+ \sum_{k > \|x\|^3} (\mathbb{P}_0[S_k = 0] - \mathbb{P}_x[S_k = 0])
\]

\[=: M_1 + M_2 + M_3,
\]

and then let us deal with the three terms separately.

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

\[\mathbb{P}_0[S_{2k} = 0] \asymp \frac{1}{k}. \tag{3.27}\]

To deal with the term \(M_1\), just observe that \(\mathbb{P}_x[S_k = 0] = 0\) for \(k < \|x\|\) — there is simply not enough time for the walker to go from \(x\) to 0. The relation (3.27) then implies that

\[M_1 \asymp \ln \|x\|. \tag{3.28}\]

For the second term, we have already observed that all summands there are nonnegative, so (3.27) implies that

\[0 \leq M_2 \lesssim \ln \|x\|. \tag{3.29}\]

That is, \(M_1\) is of order \(\ln \|x\|\), and \(M_2\) is nonnegative and at most of order \(\ln \|x\|\); this clearly implies that the sum of them is also of order \(\ln \|x\|\).

It remains to deal with the term \(M_3\). It is here that we use the coupling idea: let us write

\[
\sum_{k > \|x\|^3} (\mathbb{P}_0[S_k = 0] - \mathbb{P}_x[S_k = 0])
\]

\[= \mathbb{E} \sum_{k > \|x\|^3} (1\{S'_k = 0\} - 1\{S''_k = 0\})
\]

\[(\text{write } 1 = 1\{T \leq k\} + 1\{T > k\}, \text{ and note that the } k\text{th term is 0 on } \{T \leq k\})
\]

\[= \mathbb{E} \sum_{k > \|x\|^3} (1\{S'_k = 0\} - 1\{S''_k = 0\}) 1\{T > k\}\]
(if \( T > k \) then \( S_k' \) cannot be at the origin, recall (3.26))
\[
= E \sum_{k > \|x\|^3} 1\{S_k' = 0\}1\{T > k\}
\]
\[
= \sum_{k > \|x\|^3} P[S_k' = 0, T > k]
\]
(since \( \{T > k\} = \{T_1 > k\} \cup \{T_2 > k\} \))
\[
\leq \sum_{k > \|x\|^3} P[S_k' = 0, T_1 > k] + \sum_{k > \|x\|^3} P[S_k' = 0, T_2 > k]
\]
(by symmetry)
\[
= 2 \sum_{k > \|x\|^3} P[S_k' = 0, T_1 > k].
\] (3.30)

We are now going to prove that the terms in the above sum are at most of order \( k^{-4/3} \). For this, we first prove that, for all \( m \geq b_1^3 \),
\[
P_0[X_{2m} = 0, \hat{T}(b_1) > 2m] \leq m^{-5/6}, \quad (3.31)
\]
where \( X \) is a one-dimensional simple random walk, and \( \hat{T}(s) = \min\{\ell > 0 : X_\ell = s\} \). To show (3.31), we use the following well-known fact:

**Proposition 3.9 (The Reflection Principle)** Let us consider oriented paths\(^{14} \) in \( \mathbb{Z}^2 \), such that from \( x \) the path can go to either to \( x + e_1 - e_2 \) or to \( x + e_1 + e_2 \) (that is, it can go to north-east and south-east directions). 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 \( x = (x_1, -x_2) \) to \( y \).

**Proof** Just look at Figure 3.5 (on the right). In a way, it is the same coupling as before, only in dimension 1. \( \square \)

Now, we write
\[
P_0[X_{2m} = 0, \hat{T}(b_1) \leq 2m] = P_{b_1}[X_{2m} = b_1, \hat{T}(0) \leq 2m]
\]
(by the Reflection Principle)
\[
= P_{-b_1}[X_{2m} = b_1]
\]

\(^{14}\) these are the space-time paths of one-dimensional simple random walk, the horizontal axis represents time and the vertical axis represents space
Potential theory

Figure 3.5 Two famous visual proofs

\[= 2^{-2m} \left( \frac{2m}{m - b_1} \right).\]

So, we have for \(m \geq b_1^3\),

\[
\mathbb{P}_0[X_{2m} = 0, \hat{T}(b_1) > 2m] = 2^{-2m} \left( \frac{2m}{m} - \left( \frac{2m}{m-b_1} \right) \right) \\
= 2^{-2m} \left( \frac{2m}{m} \right) \left( 1 - \frac{(m - b_1 + 1) \cdots (m - 1)m}{(m + 1) \cdots (m + b_1)} \right) \\
= 2^{-2m} \left( \frac{2m}{m} \right) \left( 1 - \left( 1 - \frac{b_1}{m + 1} \right) \cdots \left( 1 - \frac{b_1}{m + b_1} \right) \right) \\
\leq 2^{-2m} \left( \frac{2m}{m} \right) \left( 1 - \left( 1 - \frac{b_1}{m} \right)^{b_1} \right)
\]

(using the simple inequality \((1-x)^a \geq 1 - ax\) for \(x \in [0,1]\) and \(a \geq 1\))

\[\leq 2^{-2m} \left( \frac{2m}{m} \right) \times \frac{b_1^2}{m}\]
3.2 Potential theory in two dimensions

(recall the calculation (2.2) from Section 2.1 and use that $b_1^2 \leq m^{2/3}$)

\[
\frac{1}{m^{1/2}} \times \frac{m^{2/3}}{m} = \frac{1}{m^{5/6}},
\]

thus proving (3.31). Note also that, if $B_k$ 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\footnote{use e.g. the Chernoff’s bound}

\[
P\left[\frac{2}{3}k \leq B_k \leq \frac{4}{3}k \right] = P\left[|B_k - \mathbb{E}B_k| \leq \frac{1}{3}k\right] \geq 1 - e^{-ck}.
\] (3.32)

So, we have

\[
P[S_{2k}^r = 0, T_1 > 2k]
= \sum_{m=0}^{k} P[B_k = m] P_0[X_{2m} = 0, \tilde{T}(b_1) > 2m] P_0[X_{2(k-m)} = 0]
\]

(due to (3.32), only the “middle terms” matter, and they are all of the same order)

\[
\lesssim \frac{1}{k^{5/6}} \times \frac{1}{k^{1/2}} = k^{-4/3};
\]

going back to (3.30) we find that the term $M_3$ is bounded above
by a constant (in fact, even polynomially small in $\|x\|$), and this
finally shows that, for $x \in \mathbb{Z}^2$ with both coordinates even, $a(x)$
exists and is of order $\ln \|x\|$. We are finally done with the long
proof of existence, but the reader is advised to see Exercise 3.26.

The above argument can be tweaked to treat all $x \in \mathbb{Z}^2$, but, as
we will see now, this is unnecessary. Let us show that the function $a$
is harmonic outside the origin, i.e.,

\[
a(x) = \frac{1}{4} \sum_{y \sim x} a(y) \quad \text{for all } x \neq 0.
\] (3.33)

Then, assuming that (3.33) holds, it is clear that the fact that
$a(x) < \infty$ for some $x \neq 0$ implies that $a(x) < \infty$ for all $x \in \mathbb{Z}^2$
(indeed, if $a(x) < \infty$ then $a(y)$ should also be finite for all $y \sim x$, etc.). Also, if $x, y \neq 0$ and $y \sim x$, then (3.33) implies that $\frac{1}{4}a(x) \leq$
Potential theory

\[ a(y) \leq 4a(x), \] and this shows that \( a(x) \) should be of order \( \ln \|x\| \) as \( x \to \infty \) indeed.

Now, we prove (3.33); this is again a consequence of the total expectation formula, we only need to take some more care this time because of the limits involved. Let

\[ N_z^{(k)} = \sum_{j=0}^{k} 1\{S_k = z\} \quad (3.34) \]

be the number of visits to \( z \) up to time \( k \); we thus have \( a(x) = \lim_{k \to \infty} (E_0 N_0^{(k)} - E_x N_0^{(k)}) \). The total expectation formula (conditional on the first step) gives us that, for \( x \neq 0 \),

\[ E_x N_0^{(k)} = \frac{1}{4} \sum_{y \sim x} E_y N_0^{(k-1)}, \]

so

\[
E_0 N_0^{(k)} - E_x N_0^{(k)} = E_0 N_0^{(k)} - \frac{1}{4} \sum_{y \sim x} E_y N_0^{(k-1)} \\
= E_0 N_0^{(k-1)} + P_0[S_k = 0] - \frac{1}{4} \sum_{y \sim x} E_y N_0^{(k-1)} \\
= P_0[S_k = 0] + \frac{1}{4} \sum_{y \sim x} (E_0 N_0^{(k-1)} - E_y N_0^{(k-1)}).
\]

Sending \( k \) to infinity in the above, we obtain (3.33).

Let \( \mathcal{N} = \{ \pm e_i, i = 1, 2 \} \) be the set of the four neighbours of the origin. Another useful fact is that

\[ a(x) = 1 \text{ for all } x \in \mathcal{N}. \quad (3.35) \]

To see this, first, observe that, by symmetry, \( a(\cdot) \) must have the same value on the sites of \( \mathcal{N} \). Then, again by the total expectation formula,

\[ E_0 N_0^{(k)} = 1 + \frac{1}{4} \sum_{y \in \mathcal{N}} E_y N_0^{(k-1)} = 1 + E_{e_1} N_0^{(k-1)} \]

(note that the time-zero visit counts, and then use symmetry).

The above implies that

\[
E_0 N_0^{(k)} - E_{e_1} N_0^{(k-1)} = (1 + E_{e_1} N_0^{(k-1)}) - (E_{e_1} N_0^{(k-1)} + P_{e_1}[S_k = 0]) \\
= 1 - P_{e_1}[S_k = 0],
\]
and, sending \( k \) to infinity, we obtain (3.35).

Again, as with the Green’s function in the previous section (expression (3.5)), we argue that the common wisdom suggests that the potential kernel should be roughly spherically symmetric, and “well-behaved” in general. Indeed, it is possible to prove that, as \( x \to \infty \),

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

(3.36)

where, being \( \gamma = 0.5772156\ldots \) the Euler-Mascheroni constant\(^{16} \),

\[
\gamma' = \frac{2\gamma + \ln 8}{\pi} = 1.0293737\ldots,
\]

(3.37)

cf. Theorem 4.4.4 of [46], and see Exercises 3.28 and 3.29. We also note that it is possible to obtain exact values of \( a(\cdot) \) in other sites (close to the origin); for example, it holds that \( a(e_1 + e_2) = \frac{4}{\pi} \), \( a(2e_1) = 4 - \frac{8}{\pi} \), and so on, see Section III.15 of [66].

Observe that the harmonicity of \( a \) outside the origin (established in (3.33)) immediately implies that the following result holds:

**Proposition 3.10** The process \( a(S_{k \wedge \tau_0}) \) is a martingale.

We will repeatedly use this fact in the sequel. What we will also repeatedly use, is that, due to (3.36),

\[
a(x + y) - a(x) = O(\|y\|) \quad (3.38)
\]

for all \( x, y \in \mathbb{Z}^2 \) such that (say) \( \|x\| > 2\|y\| \).

With some (slight) 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 \( a(\|x\|) \), although they are of course quite close for large \( x \). The advantage of using this notation is e.g. that, due to (3.36) and (3.38), we may write (for fixed \( x \) or at least \( x \) such that \( 2\|x\| \leq r \))

\[
\sum_{y \in \partial \mathcal{B}(x,r)} \nu(y) a(y) = a(r) + O\left(\frac{\|x\|}{r}\right) \quad \text{as } r \to \infty
\]

(3.39)

for any probability measure \( \nu \) on \( \partial \mathcal{B}(x,r) \).

\(^{16} \gamma = \lim_{n \to \infty} (1 + \frac{1}{2} + \cdots + \frac{1}{n} - \ln n)\)
As in the higher-dimensional case, we need an asymptotic expression for the probability that a random walk in an annulus leaves it through its outer boundary. Quite analogously to the proof of Lemma 3.3, we can obtain the following result from (3.36), Proposition 3.10, and the Optional Stopping Theorem:

**Lemma 3.11** For all \( x \in \mathbb{Z}^2 \) and \( R > r > 0 \) such that \( x \in B(y, R) \setminus B(r) \) we have

\[
P_x \left[ \tau_{\partial B(y, R)} < \tau_{B(r)} \right] = \frac{a(x) - a(r) + O(r^{-1})}{a(R) - a(r) + O(r^{-1} + \|y\| + \frac{1}{r})},
\]

as \( r, R \to \infty \). The above also holds with \( \tau_{\partial B(y, R)} \) on the place of \( \tau_{\partial B(y, r)} \).

We do not write its proof here (since it is really analogous); but, in case the reader prefers to see a similar proof again, we prove the following result for the probability of escaping the origin:

**Lemma 3.12** Assume that \( x \in B(y, r) \) and \( x \neq 0 \). Then

\[
P_x \left[ \tau_{\partial B(y, r)} < \tau_{0}^+ \right] = \frac{a(x)}{a(r) + O\left(\frac{\|y\|+1}{r}\right)},
\]

as \( r \to \infty \). As before, the above also holds with \( \tau_{\partial B(y, r)} \) on the place of \( \tau_{\partial B(y, r)} \).

**Proof** Indeed, use Proposition 3.10, and the Optional Stopping Theorem to write (recall that \( a(0) = 0 \))

\[a(x) = P_x \left[ \tau_{\partial B(y, r)} < \tau_{0}^+ \right] E_x \left( a(S_{\tau_{\partial B(y, r)}}) \mid \tau_{\partial B(y, r)} < \tau_{0}^+ \right),\]

and then use (3.39).

Note that Lemma 3.12 implies that (since, from the origin, on the next step the walk will go to a site of \( \mathcal{N} \) where the potential kernel equals 1)

\[
P_0[\tau_{\partial B(r)} < \tau_{0}^+] = \frac{1}{a(r) + O(r^{-1})} = \left(\frac{2}{\pi} \ln r + \gamma + O(r^{-1})\right)^{-1}.
\]

(3.42)

The reader may have recalled that this was the formula that was used in the introduction to calculate the probability of going to the edge of our galaxy before returning to the initial point.
Green’s function.

Wait, but didn’t we agree in the beginning of this section that there is no Green’s function in two dimensions? Well, this applies to the Green’s function in the whole space, but we also can define a “restricted Green’s function”, which can be still quite useful. Let \( \Lambda \) be a (typically, finite) subset of \( \mathbb{Z}^2 \). For \( x, y \in \Lambda \) let us define

\[
G_{\Lambda}(x, y) = \mathbb{E}_x \sum_{k=0}^{\tau_{\partial\Lambda}-1} \mathbb{1}\{S_k = y\}
\]

(3.43)

to be the mean number of visits to \( y \) starting from \( x \) before stepping out of \( \Lambda \). Notice that this definition formally makes sense also in the case when at least one of the arguments is outside of \( \Lambda \), in which case \( G_{\Lambda}(x, y) = 0 \).

This notion is, of course, less convenient than that of the Green’s function in the whole space, since we (clearly) lose the translation invariance, and also (apparently) lose the symmetry. It is quite remarkable, however, that, in fact, the symmetry is not lost! Indeed, let us prove that

\[
G_{\Lambda}(x, y) = G_{\Lambda}(y, x)
\]

for any \( x, y \in \Lambda \). (3.44)

To prove the above, we use the usual trick of getting rid of random sums (such as the one in (3.43)):

\[
\mathbb{E}_x \sum_{k=0}^{\tau_{\partial\Lambda}-1} \mathbb{1}\{S_k = y\} = \mathbb{E}_x \sum_{k=0}^{\infty} \mathbb{1}\{S_k = y, \tau_{\partial\Lambda} > k\}
\]

\[
= \sum_{k=0}^{\infty} \mathbb{P}_x[S_k = y, \tau_{\partial\Lambda} > k]
\]

and so, to prove (3.44), it is enough to show that

\[
\mathbb{P}_x[S_k = y, \tau_{\partial\Lambda} > k] = \mathbb{P}_y[S_k = x, \tau_{\partial\Lambda} > k]
\]

for any \( k \). But this is quite evident: indeed, the number of \( k \)-step trajectories that lie fully inside \( \Lambda \), start at \( x \) and end at \( y \), is obviously the same as the number of such trajectories that start at \( y \) and end at \( x \).

Next, there is a useful relation connecting the restricted Green’s function to the potential kernel:

\[\text{with the usual convention that } \sum_{k=0}^{\infty} = 0\]
Theorem 3.13  Assume that \( \Lambda \) is finite. Then, it holds that
\[
G_\Lambda(x, y) = \mathbb{E}_x a(S_{\tau_{\partial \Lambda}} - y) - a(x - y). \tag{3.45}
\]

Proof  First, Proposition 3.10 together with (3.35) imply that the process \( a(S_n - y) \) is a submartingale: indeed, when the walk is not at \( y \), its expected drift equals zero, while, when it is at \( y \), its value is 0 and will become 1 on the next step (so the expected drift equals 1). With a moment of thought, one can see\(^{18}\) that the process (recall the notation from (3.34))
\[
Y_n = a(S_n - y) - N_y^{(n-1)} \tag{3.46}
\]
is a martingale – the drift of \( a(S_n - y) \) which is present when the walker is at \( y \) is “compensated” by the increase of the value of \( N_y \). Since \( G_\Lambda(x, y) = \mathbb{E}_x N_y^{(\tau_{\partial \Lambda}-1)} \), it is enough to apply the Optional Stopping Theorem\(^{19}\) to the martingale \( Y_n \) and the stopping time \( \tau_{\partial \Lambda} \):
\[
a(x - y) = \mathbb{E}_x Y_0 = \mathbb{E}_x Y_{\tau_{\partial \Lambda}} = \mathbb{E}_x a(S_{\tau_{\partial \Lambda}} - y) - G_\Lambda(x, y),
\]
thus proving (3.45).

We use the above theorem to obtain a convenient expression for the restricted Green’s function in the case \( \Lambda = B(R) \), with a large \( R \). Let \( x, y \in B(R) \). Note that, for any \( z \in \partial_e B(R) \), it holds that
\[
a(z - y) = a(R) + O\left(\frac{\|z\| + 1}{R}\right)
\]
(indeed, analogously to (3.38) just observe that \( ||z - y|| - R|| \leq ||y|| + 1 \) and use that \( \ln ||z - y|| = \ln R + \ln \left(1 + \frac{||z - y|| - R}{R}\right)\)). So, Theorem 3.13 implies that, for \( x, y \in B(R) \)
\[
G_{B(R)}(x, y) = a(R) - a(x - y) + O\left(\frac{1 + ||x|| ||y||}{R}\right) \tag{3.47}
\]
(as \( ||x - y|| \to \infty \))
\[
= \frac{2}{\pi} \ln \frac{R}{||x - y||} + O\left(\frac{1 + ||x|| ||y||}{R} + \frac{1}{||x - y||}\right) \tag{3.48}
\]
\(^{18}\) exercise: prove it formally
\(^{19}\) indeed, item (iii) of Corollary 1.7 works fine here, due to the finiteness of \( \Lambda \) and Exercise 3.27
3.2 Potential theory in two dimensions

(note that, by the symmetry property (3.44), we can assume without loss of generality that $\|y\| \leq \|x\|$, so we can choose “the better term” in the above $O$’s).

Harmonic measure.

Here (I mean, in two dimensions) we take a different route: we first define and discuss the notion of harmonic measure, and only then pass to that of capacity. For a finite $A \subset \mathbb{Z}^2$ and $x \in \mathbb{Z}^2$ let us define

$$q_A(x) = a(x - y_0) - \mathbb{E}_x a(S_{\tau_A} - y_0), \quad (3.49)$$

where $y_0$ is some site of $A$ (later we will prove that the value of $q_A(x)$ does not depend on the choice of this $y_0$). Note that $q_A(x) = 0$ for all $x \in A$ (since $\tau_A = 0$ when the walk starts at $x \in A$); also, the above definition is invariant under translations, i.e., $q_{A+z}(x + z) = q_A(x)$ for any $z \in \mathbb{Z}^2$. The importance of this quantity is underlined by the following fact:

**Proposition 3.14** For any finite $A \subset \mathbb{Z}^2$ and any $x, y \in \mathbb{Z}^2$ it holds that

$$q_A(x) = \lim_{R \to \infty} a(R)\mathbb{P}_x[\tau_{B(y,R)^c} < \tau_A] = \frac{2}{\pi} \lim_{R \to \infty} \mathbb{P}_x[\tau_{B(y,R)^c} < \tau_A] \ln R. \quad (3.50)$$

**Proof** We use a martingale argument similar to the proofs of Lemmas 3.3 and 3.11–3.12; we only need to take a bit more care in order to “separate” the term $\mathbb{E}_x a(S_{\tau_A})$ so that it remains “free from conditioning”. Assume without restricting generality that $y_0 = 0$ (otherwise we can just “shift” the origin there) and let us apply the Optional Stopping Theorem\(^{20}\) to the martingale $a(S_{n \wedge \tau_A - y_0})$ with the stopping time $\tau_A \wedge \tau_{B(y,R)^c}$:

$$a(x) = \mathbb{E}_x a(S_{\tau_A \wedge \tau_{B(y,R)^c}})$$

$$= \mathbb{E}_x a(S_{\tau_{B(y,R)^c}}) 1\{\tau_{B(y,R)^c} < \tau_A\} + \mathbb{E}_x a(S_{\tau_A}) 1\{\tau_A < \tau_{B(y,R)^c}\}$$

writing $1\{\tau_A < \tau_{B(y,R)^c}\} = 1 - 1\{\tau_{B(y,R)^c} < \tau_A\}$ in the second term

$$= \mathbb{P}_x[\tau_{B(y,R)^c} < \tau_A] \mathbb{P}_x \left( a(S_{\tau_{B(y,R)^c}}) \mid \tau_{B(y,R)^c} < \tau_A \right)$$

$$+ \mathbb{E}_x \left( a(S_{\tau_A}) - a(S_{\tau_A}) 1\{\tau_{B(y,R)^c} < \tau_A\} \right)$$

$$= \mathbb{P}_x[\tau_{B(y,R)^c} < \tau_A] \mathbb{P}_x \left( a(S_{\tau_{B(y,R)^c}}) \mid \tau_{B(y,R)^c} < \tau_A \right)$$

\(^{20}\) formally, to justify its use, Corollary 1.7 (ii) is enough
so, abbreviating \( b = 1 + \max_{x \in A} \|x\| \), we obtain from (3.39) that

\[
\mathbb{P}_x[\tau_{B(y,R)^c} < \tau_A] = \frac{a(x) - \mathbb{E}_x a(S_{\tau_A})}{\mathbb{E}_x(a(S_{\tau_B(y,R)^c}) - a(S_{\tau_A}) \mid \tau_{B(y,R)^c} < \tau_A)}
\]

(3.51)

and this implies (3.50).

Since the limit in (3.50) does not depend on \( y \), this means that Proposition 3.14 indeed shows that the definition (3.49) of \( q_A \) does not depend on the choice of \( y_0 \in A \) (since “moving \( y_0 \) within \( A \)” effectively amounts to changing \( y \) in (3.50)).

Now, we are ready to define the notion of the harmonic measure in two dimensions:

**Definition 3.15** For a finite set \( A \subset \mathbb{Z}^2 \), the harmonic measure \( (h\text{m}_A(y), y \in A) \) is defined as follows:

\[
\text{hm}_A(y) = \frac{1}{4} \sum_{z \sim y} q_A(z) = \frac{1}{4} \sum_{z \not\in A, z \sim y} (a(z) - \mathbb{E}_z a(S_{\tau_A})).
\]

(3.53)

Admittedly, at this point it may be not completely clear why \( \text{hm}_A(\cdot) \) should be even nonnegative (it is because of Proposition 3.14); what is definitely not clear, is why it sums to 1 on \( \partial A \). Things start to make sense, though, when we observe that the above definition is very similar to (3.20) (i.e., the many-dimensional one): the harmonic measure is proportional to the escape probability. To see this, observe that Proposition 3.14 implies that

\[
\text{hm}_A(y) = \lim_{R \to \infty} a(R)\mathbb{P}_y[\tau_A^+ > \tau_{B(R)^c}] = \frac{2}{\pi} \lim_{R \to \infty} \mathbb{P}_y[\tau_A^+ > \tau_{B(R)^c}] \ln R;
\]

(3.54)

indeed, just write, conditioning on the first step,

\[
\mathbb{P}_y[\tau_A^+ > \tau_{B(R)^c}] = \frac{1}{4} \sum_{z \not\in A, z \sim y} \mathbb{P}_z[\tau_{B(R)^c} < \tau_A],
\]

then multiply both sides by \( a(R) \) and pass to the limit.
To complete the analogy with the many-dimensional case, we have to prove that the harmonic measure is the entrance law “starting at infinity” (compare to Theorem 3.8; since the walk is recurrent in two dimensions, we do not need to condition on \( \{\tau_A < \infty\} \):

**Theorem 3.16** For all finite \( A \subset \mathbb{Z}^2 \) and all \( y \in A \) we have

\[
\lim_{x \to \infty} \mathbb{P}_x[S_{\tau_A} = y] = \lim_{x \to \infty} \mathbb{P}_x[S_{\tau_A} = y].
\] (3.55)

**Proof** First, it is clear that, without restricting generality, we can assume that \( 0 \in A \). Recall the notation \( N^b_x = \sum_{k=0}^{\tau_A - 1} \mathbb{1}\{S_k = x\} \) from the proof of Theorem 3.8; also, let us define

\[
N^b_{x,R} = \sum_{k=0}^{(\tau_A - 1) \wedge \tau_{B(R)^c}} \mathbb{1}\{S_k = x\}, \quad N^2_{x,R} = \sum_{k=\tau_A}^{\tau_{B(R)^c}} \mathbb{1}\{S_k = x\},
\]

and \( N_{x,R} := N^b_{x,R} + N^2_{x,R} = \sum_{k=0}^{\tau_{B(R)^c}} \mathbb{1}\{S_k = x\} \) be the corresponding (i.e., before the first re-entry, after the first re-entry, and total) visit counts “restricted” on \( B(R) \). Quite similarly to the proof of Theorem 3.8 one can write

\[
\mathbb{P}_x[S_{\tau_A} = y] = \mathbb{E}_y N^b_x
\]

(by the Monotone Convergence Theorem)

\[
= \lim_{R \to \infty} \mathbb{E}_y N^b_{x,R}
\]

\[
= \lim_{R \to \infty} \left( \mathbb{E}_y N_{x,R} - \mathbb{E}_y N^2_{x,R} \right)
\]

\[
= \lim_{R \to \infty} \left( G_{B(R)}(y,x) - \sum_{z \in A} \mathbb{P}_y[\tau_A^+ < \tau_{B(R)^c}, S_{\tau_A^+} = z] G_{B(R)}(z,x) \right)
\]

(using (3.47))

\[
= \lim_{R \to \infty} \left( a(R) - a(y - x) + O\left(\frac{\|y\| + 1}{R}\right) \right)

- \sum_{z \in A} \mathbb{P}_y[\tau_A^+ < \tau_{B(R)^c}, S_{\tau_A^+} = z] \left( a(R) - a(z - x) + O\left(\frac{\|z\| + 1}{R}\right) \right)
\]
(note that $P_{y}[\tau_{A}^{+} < \tau_{B(R)}^{+}, S_{r_{A}} = z] \to P_{y}[\tau_{A}^{+} = z]$ as $R \to \infty$)

$$= \lim_{R \to \infty} a(R) \left( 1 - \sum_{z \in A} P_{y}[\tau_{A}^{+} < \tau_{B(R)}^{+}, S_{r_{A}} = z] \right)$$

$$- a(y - x) + \sum_{z \in A} P_{y}[S_{r_{A}} = z]a(z - x)$$

(observe that, in the first parentheses, we have $P_{y}[\tau_{A}^{+} > \tau_{B(R)}^{+}]$, then use (3.54))

$$= hm_{A}(y) - a(y - x) + \sum_{z \in A} P_{y}[S_{r_{A}} = z]a(z - x). \quad (3.56)$$

From (3.38) it is straightforward to obtain that

$$a(y - x) - \sum_{z \in A} P_{y}[S_{r_{A}} = z]a(z - x)$$

$$= \sum_{z \in A} P_{y}[S_{r_{A}} = z](a(y - x) - a(z - x))$$

$$= O\left( \frac{\text{diam}(A)}{\|x\|_{3}} \right)$$

which converges to 0 as $x \to \infty$, and so the proof of (3.55) is concluded.

Then, Theorem 3.16 implies that $hm_{A}(\cdot)$ is indeed a probability measure (because, due to the recurrence, so is the entrance measure to $A$ for any fixed starting point).

Now, we are happy with Theorem 3.16, but not completely so. This is because it is a qualitative result, which does not say how fast the convergence occurs. Imagine, for example, that the set $A$ is “large” (say, a disk of radius $r^{2}$), and the distance from $x$ to $A$ is even larger (say, of order $r^{3}$). How the entrance measure from $x$ to $A$ compares to the harmonic measure then? Well, in the end of the proof of Theorem 3.16 we obtained some estimate, but it is not quite sharp -- in the example we just considered, the term $O\left( \frac{\text{diam}(A)}{\|x\|_{3}} \right)$ would be of order $r^{-1}$, while $hm_{A}(y)$ itself would be of order $r^{-2}$ (since there are $O(r^{2})$ sites on the boundary of the disk), which, to put it mildly, is not quite satisfactory. The following theorem gives a much better estimate:

**Theorem 3.17** Let $A$ be a finite subset of $\mathbb{Z}^{2}$ and assume that $\text{dist}(x, A) \geq 3 \text{diam}(A) + 1$. Then it holds that

$$P_{x}[S_{r_{A}} = y] = hm_{A}(y)\left( 1 + O\left( \frac{\text{diam}(A)}{\text{dist}(x, A)} \right) \right). \quad (3.57)$$
Proof. Again, without restricting generality, we assume that $0 \in A$ and $|A| \geq 2$, so $\text{diam}(A) \geq 1$. Recall that in the calculation (3.56) we obtained

$$
P_x[S_{\tau_A} = y] = -a(y - x) + \sum_{z \in A} P_y[S_{\tau_A}^+ = z]a(z - x).$$

(3.58)

The idea is to estimate the right-hand side of the above expression using a martingale argument similar to that in the proof of Proposition 3.14. We need some preparations, though. From the asymptotic expression (3.36) for $a$, it is straightforward to obtain that there exist constants $\theta_1, \theta_2 > 0$ such that whenever $\|x\| > \theta_1$ and $2\|y\| \leq \|x\|$ it holds that $a(x) - a(y) > \theta_2$ (in fact, it is even clear that for any $\theta_2 < \frac{2}{\pi} \ln 2$ it is possible to choose a large enough $\theta_1$ such that the above holds, but we do not need to be so precise).

Let us now abbreviate $V = \partial eB((2 \text{diam}(A)) \lor \theta_1)$, so that, by the above,

$$a(v) - a(z) \geq \theta_2 \text{ for all } v \in V \text{ and } z \in A. \quad (3.59)$$

We assume additionally that $\|x\| > \theta_1$; the reader is invited to check (or to just accept) that this assumption does not restrict generality.

Let us apply the Optional Stopping Theorem to the martingale $a(S_{n \wedge \tau} - x)$ and the stopping time $\tau^+_A \wedge \tau_V$ (see Figure 3.6; observe that $\tau_V < \tau_x$ for the walk that starts at $y \in \partial A$)

$$a(y - x) = E_y a(S_{\tau^+_A \wedge \tau_V} - x)$$

$$= E_y (a(S_{\tau^+_A} - x)1\{\tau^+_A < \tau_V\}) + E_y (a(S_{\tau_V} - x)1\{\tau_V < \tau^+_A\})$$

(writing $1\{\tau^+_A < \tau_V\} = 1 - 1\{\tau_V < \tau^+_A\}$ in the first term)

$$= E_y a(S_{\tau^+_A} - x)$$

$$+ E_y ((a(S_{\tau_V} - x) - a(S_{\tau^+_A} - x))1\{\tau_V < \tau^+_A\})$$

$$= \sum_{z \in A} P_y[S_{\tau_A}^+ = z]a(z - x)$$

$$+ E_y (a(S_{\tau_V} - x) - a(S_{\tau_A}^+ - x) \mid \tau_V < \tau_A^+)P_y[\tau_V < \tau_A^+].$$
So, the above together with (3.58) imply that
\[
\begin{align*}
\operatorname{hm}_A(y) - P_x[S_{\tau_A} = y] &= \mathbb{E}_y(a(S_{\tau_V} - x) - a(S_{\tau_A}^+ - x) \mid \tau_V < \tau_A^+) \\
&\quad \times P_y[\tau_V < \tau_A^+]. 
\end{align*}
\]

(3.60)

Let us now recall the expression (3.51) from the proof of Proposition 3.14 together with the definition (3.53) of the harmonic measure. For the second factor in the right-hand side of (3.60), we have
\[
\mathbb{P}_y[\tau_V < \tau_A^+] = \frac{1}{4} \sum_{z \sim y} \mathbb{P}_u[\tau_V < \tau_A^+] = \frac{1}{4} \sum_{z \sim y} q_A(z) \mathbb{E}_z(a(S_{\tau_V}) - a(S_{\tau_A}) \mid \tau_V < \tau_A) \\
\quad \leq \frac{\operatorname{hm}_A(y)}{\theta_2}.
\]

(3.61)

We also observe that, for any \( v \in V \) and \( z \in A \), from (3.38) we obtain that \( a(x - v) - a(x - z) = O\left(\frac{\text{diam}(A)}{\text{dist}(x, A)}\right) \) (look again at Figure 3.6), so the first factor in the right-hand side of (3.60) is \( O\left(\frac{\text{diam}(A)}{\text{dist}(x, A)}\right) \) as well. This shows that the right-hand side of (3.60)
3.2 Potential theory in two dimensions

is indeed $O\left(\frac{\text{diam}(A)}{\text{dist}(x,A)}\right) \times \text{hm}_A(y)$ and therefore concludes the proof of Theorem 3.17.

Capacity.

When one learns something new, it is a good idea to review one’s old notes and see if there was something interesting that went unnoticed at that time. Specifically, let us revisit the calculations in and around Proposition 3.14. Recall that, for a finite $A \subset \mathbb{Z}^2$ and $y_0 \in A$, we defined in (3.49) the quantities

$$q_A(x) = a(x - y_0) - \mathbb{E}_x a(S_{\tau_A} - y_0), \quad x \in \mathbb{Z}^2,$$

and proved that $q_A(x)$ does not depend on the choice of $y_0$ (as long as $y_0 \in A$). Let us re-examine the second term in that definition in the light of Theorem 3.17. If $x$ is far away from $A$, the entrance measure to $A$ is “almost harmonic” and, quantitatively,

$$\mathbb{E}_x a(S_{\tau_A} - y_0) = \sum_{z \in A} \mathbb{P}_x[S_{\tau_A} = z] a(z - y_0) = \sum_{z \in A} \text{hm}_A(z) a(z - y_0)(1 + O\left(\frac{\text{diam}(A)}{\text{dist}(x,A)}\right)), \quad (3.61)$$

The main term in (3.61) does not depend on $x$ and therefore looks as something important. It is so indeed:

**Definition 3.18** For a finite set $A$ with $y_0 \in A$, we define its capacity by

$$\text{cap}(A) = \sum_{x \in A} a(x - y_0) \text{hm}_A(x). \quad (3.62)$$

First, we need to show that the above definition does not depend on the choice of $y_0 \in A$. Basically, this is a consequence of the fact that $q_A(x)$ does not depend on the choice of $y_0 \in A$. Indeed, if $y_1 \in A$, we also have $q_A(x) = a(x - y_1) - \mathbb{E}_x a(S_{\tau_A} - y_1)$ and so, by (3.61),

$$a(x-y_0) - a(x-y_1) = \left(\text{cap}(A) - \sum_{z \in A} \text{hm}_A(z) a(z-y_1)\right)(1 + O\left(\frac{\text{diam}(A)}{\text{dist}(x,A)}\right)).$$

Since the left-hand side of the above clearly converges to 0 as $x \to \infty$, the expression in the first parentheses in the right-hand side must be equal to 0 (since it does not depend on $x$).
Now, assume that \( y_0 \in A \subset B(r_A) \). Then, (3.61) implies that
\[
q_A(x) = a(x - y_0) - \text{cap}(A) + O\left(\frac{r_A \ln r_A}{\|x\|}\right),
\]
and then, recalling the calculation (3.51), we can write
\[
P_x[\tau_{B(R)}^c < \tau_A] = \frac{a(x - y_0) - \text{cap}(A) + O\left(\frac{r_A \ln r_A}{\|x\|}\right)}{a(R) - \text{cap}(A) + O\left(\frac{r_A \ln r_A}{R}\right)}. \tag{3.63}
\]
That is, if we know the capacity of \( A \), we are then able to compute the escape probabilities as above with higher precision. Notice that, the larger \( \text{cap}(A) \) is, the smaller is the probability in (3.63). This again justifies the intuition “the capacity measures how big is the set from the point of view of SRW”; only in two dimension it enters to the second-order term (and not to the principal one, as in the higher-dimensional case).

Now, let us discuss the simplest cases where the two-dimensional capacities can be calculated. To start with, what can we say about capacities of one- and two-point sets? The first question is easy: since \( a(0) = 0 \), it holds that \( \text{cap}\left(\{x\}\right) = 0 \) for any \( x \in \mathbb{Z}^2 \). As for two-point sets, observe that, by symmetry, the harmonic measure of any two-point set is uniform, so \( \text{cap}\left(\{x, y\}\right) = \frac{1}{2}a(y - x) \) for any \( x, y \in \mathbb{Z}^2, x \neq y \). As for the capacity of a disk, (3.39) implies that
\[
\text{cap}\left(\mathbb{B}(r)\right) = a(r) + O(r^{-1}). \tag{3.64}
\]
It is remarkable to observe that the capacities of a two-point set \( \{0, x\} \) with \( \|x\| = r \) and the whole disk \( \mathbb{B}(r) \) only differ by a factor of 2 (this is asymptotically, as \( r \to \infty \)). Dimension two sometimes brings surprises.

Next, it is not difficult to obtain from (3.63) that
\[
\text{cap}(A) = \lim_{t \to \infty} \left( a(x - y_0) - \lim_{R \to \infty} \left( a(R)P_x[\tau_{B(R)}^c < \tau_A]\right) \right) \tag{3.65}
\]
(first, multiply it by \( a(R) \) and let \( R \to \infty \) to get rid of the denominator, then “separate” the term \( \text{cap}(A) \) from the numerator). An easy corollary of this fact is the following

**Proposition 3.19** (i) Assume that \( A \subset B \) are finite subsets of \( \mathbb{Z}^2 \). Then \( \text{cap}(A) \leq \text{cap}(B) \).

---

\(^{21}\) observe that \( E_x(S_{\tau_A} | \tau_{B(R)}^c < \tau_A) \) is also quite close to \( \text{cap}(A) \), as can be seen by conditioning on the location of \( S_{\tau_{B(R)}^c} \).
(ii) Let $A, B$ be finite subsets of $\mathbb{Z}^2$ such that $A \cap B$ is nonempty. Then $\text{cap}(A \cup B) \leq \text{cap}(A) + \text{cap}(B) - \text{cap}(A \cap B)$.

**Proof** Note that, in both cases, we may choose $y_0$ which belongs to both $A$ and $B$. The part (i) is easy: just observe that $\mathbb{P}_x[\tau_{B(x)} < \tau_A] \geq \mathbb{P}_x[\tau_{B(x)} < \tau_B]$ for $A \subset B$ (indeed, it is easier to avoid a smaller set), and use (3.65). The proof of part (ii) is only a bit more complicated: write

$$
\mathbb{P}_x[\tau_{B(x)} < \tau_{A \cup B}] \\
= 1 - \mathbb{P}_x[\{\tau_A < \tau_{B(x)}\} \cup \{\tau_B < \tau_{B(x)}\}] \\
= 1 - \mathbb{P}_x[\tau_A < \tau_{B(x)}] - \mathbb{P}_x[\tau_B < \tau_{B(x)}] + \mathbb{P}_x[\tau_A \lor \tau_B < \tau_{B(x)}]
$$

(since $\tau_{A \cap B} \geq \tau_A \lor \tau_B$)

$$
\geq 1 - \mathbb{P}_x[\tau_A < \tau_{B(x)}] - \mathbb{P}_x[\tau_B < \tau_{B(x)}] + \mathbb{P}_x[\tau_A \cap \tau_B < \tau_{B(x)}] \\
= \mathbb{P}_x[\tau_{B(x)} < \tau_A] + \mathbb{P}_x[\tau_{B(x)} < \tau_B] - \mathbb{P}_x[\tau_{B(x)} < \tau_{A \cap B}],
$$

and then use (3.65) again.

Observe that is it essential to require in (ii) that the two sets have nonempty intersection – otherwise $A = \{x\}$ and $B = \{y\}$ with $x \neq y$ would, sort of, provide a counterexample. A possible way to get around it (so that (ii) would be valid for all $A, B$) would be to declare the capacity of an empty set to be equal to $(-\infty)$ – this is formally in agreement with (3.65), by the way.

### 3.3 Exercises

**Exercise 3.1** Prove that in Proposition 3.1 it is enough to assume that $h$ is bounded only from one side (i.e., either $h(x) \leq K$ for all $x$, or $h(x) \geq K$ for all $x$).

**Exercise 3.2** Prove that $G(0) - G(e_1) = 1$.

**Exercise 3.3** Prove the following exact expression for the Green’s function:

$$
G(x) = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \frac{e^{i(\theta,x)}}{1 - \Phi(\theta)} \, d\theta, \quad (3.66)
$$
where $\Phi(\theta) = d^{-1} \sum_{k=1}^{d} \cos \theta_k$. In particular, it holds that
\[
G(0) = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \frac{1}{1 - \Phi(\theta)} d\theta. \tag{3.67}
\]

**Exercise 3.4** Follow the proof of Theorem 4.3.1 of [46] to see how to obtain (3.5) from the local CLT.

**Exercise 3.5** Assume that we only know that, for some $c > 0$, 
$G(x) \sim c \|x\|^{-(d-2)}$ as $x \to \infty$. Prove that $\frac{d-2}{2c}$ equals the volume of the unit ball in $\mathbb{R}^d$.

**Exercise 3.6** Obtain a direct proof (i.e., a proof that does not use (3.5)) that $\text{Es}_{B(r)}(y) \asymp r^{-1}$ for any $y \in B(R)^c$ (and therefore also that $\text{cap}(B(r)) \asymp r^{d-2}$). Suggestion: use Lyapunov functions.

**Exercise 3.7** For $A \subset \mathbb{Z}^d$ and $x \in \mathbb{Z}^d$, denote $G(x, A) = \sum_{y \in A} G(x, y)$ to be the mean number of visits to $A$ starting from $x$. Prove that, for finite $A$,
\[
\frac{|A|}{\max_{y \in A} G(y, A)} \leq \text{cap}(A) \leq \frac{|A|}{\min_{y \in A} G(y, A)}. \tag{3.68}
\]

We have to mention that the above provides another useful way of estimating capacities.

**Exercise 3.8** For a finite $A \subset \mathbb{Z}^d$, let $\mathcal{K}_A$ be a class of nonnegative functions, defined in the following way:
\[
\mathcal{K}_A = \left\{ h : \mathbb{Z}^d \to \mathbb{R}_+ \text{ such that } h(x) = 0 \text{ for all } x \notin A \right. \\
\left. \text{and } \sum_{x \in \mathbb{Z}^d} G(y, x) h(x) \leq 1 \text{ for all } y \in \mathbb{Z}^d \right\}.
\]
Prove that 
\[
\text{cap}(A) = \sup_{h \in \mathcal{K}_A} \sum_{x \in A} h(x).
\]

**Exercise 3.9** Prove that the capacity (defined as in (3.10)) of any infinite transient set is infinite.

**Exercise 3.10** Estimate (in the sense of “$\asymp$”) the capacities of various sets (and in various dimensions), such as
- a line segment (i.e., a sequence of neighboring sites lying on a straight line);
3.3 Exercises

- a “plaquette” (i.e., a discrete two-dimensional square immersed in $\mathbb{Z}^d$);
- a cylinder (product of a line segment with a $(d-1)$-dimensional plaquette), with height/width ratio varying arbitrarily;
- whatever else you can imagine.

**Exercise 3.11** Prove the “only if” part of Theorem 3.7.

**Exercise 3.12** Prove the “if” part of Theorem 3.7.

**Exercise 3.13** Using the Wiener’s criterion, prove that, in three dimensions, the “ray” $\{(0,0,k), k \geq 0\} \subset \mathbb{Z}^3$ is a recurrent set.

**Exercise 3.14** Give an example of a transient set such that the expected number of visits there is infinite.

**Exercise 3.15** Let $f : \mathbb{R}_+ \to (0,1]$ be a monotonously decreasing function with $\lim_{s \to \infty} f(s) = 0$. Let us construct a random set $A_f \subset \mathbb{Z}^d$ in the following way: $x \in \mathbb{Z}^d$ is included to $A_f$ with probability $f(||x||)$, independently.

(i) Is it true that, for any such $f$, $P[A_f \text{ is transient}]$ must be either 0 or 1?

(ii) Give (nontrivial) examples of functions for which the resulting random set is recurrent/transient. What should be the “critical rate of decay” of $f$ which separates recurrence from transience?

**Exercise 3.16** Give an example of a transient set $A \subset \mathbb{Z}^d$ such that $P_x[\tau_A < \infty] = 1$ for infinitely many $x \in \mathbb{Z}^d$.

**Exercise 3.17** For any transient set $A$ and any $\varepsilon > 0$ prove that $P_x[\tau_A < \infty] < \varepsilon$ for infinitely many $x \in \mathbb{Z}^d$.

**Exercise 3.18** Prove that, for finite $A$,

$$\sum_{y \in A} h_m_A(y) G(y,A) = \frac{|A|}{\text{cap}(A)}. \quad (3.69)$$

**Exercise 3.19** For $r \geq 1$, prove that $h_m_{B(r)}(x) \asymp r^{-(d-1)}$ for all $x \in \partial B(r)$.

**Exercise 3.20** Show that, for finite $A \subset B \subset \mathbb{Z}^d$ ($d \geq 3$) it holds

$$P_{h_m_A}[\tau_A < \infty] = \frac{\text{cap}(A)}{\text{cap}(B)}$$
Potential theory

(here $P_{hm_B}$ means the probability for the walk with the initial position chosen according to the measure $hm_B$)

Exercise 3.21 Prove that the harmonic measure is consistent, in the sense that, for any finite $A \subset B$

$$P_{hm_B}(S_{\tau_A} = y \mid \tau_A < \infty) = hm_A(y)$$

for all $y \in A$.

Exercise 3.22 Can you obtain an analogue of Theorem 3.17 in the many-dimensional case (i.e., $d \geq 3$)? That is, prove that

$$P_x[S_{\tau_A} = y \mid \tau_A < \infty] = hm_A(y)(1 + O(\frac{diam(A)}{dist(x,A)})).$$

(3.70)

Exercise 3.23 Let $A \subset \mathbb{Z}^d$ be finite and $x \notin A$, $y \in \partial A$.

(i) Prove that

$$P_x[\tau_A < \infty, S_{\tau_A} = y] = P_y[\tau_x < \tau_A^+] |(G(0) - O(\frac{(diam(A))^{d-2}}{dist(x,A)}))),$$

and that $P_x[\tau_A < \infty, S_{\tau_A} = y] > G(0)P_y[\tau_x < \tau_A^+]$ (that is, the $O(\cdot)$ in (3.71) is always strictly positive).

(ii) Prove that

$$P_y[\tau_x < \infty, \tau_A^+ = \infty] = P_y[\tau_x < \infty] Es_A(y)(1 + O(\frac{diam(A)}{dist(x,A)}))$$

(3.72)

that is, the events \{eventually hit $x$\} and \{escape from $A$\} are approximately independent under $P_y$ when $\frac{diam(A)}{dist(x,A)} \to 0$.

Exercise 3.24 Consider a transient and reversible Markov chain with the reversible measure $\mu$. Prove that (for the Green’s function $G$ defined in the usual way)

$$\mu(x)G(x, y) = \mu(y)G(y, x)$$

for any $x, y$. Prove that the above also holds for the restricted Green’s function $G_A$.

Recurrent case (Section 3.2):

Exercise 3.25 Prove, preferably without any calculations, that

(i) $P_0[S_{2n} = 0] > P_0[S_{2n} = e_1 + e_2]$ for all $n$. 

3.3 Exercises

(ii) \( \mathbb{P}_0[S_k = x] \geq \mathbb{P}_0[S_k = x+2e_1] \), for all \( k \) and \( x \) such that \( x \cdot e_1 \geq 0 \), and the inequality is strict in the case when the first probability is strictly positive.

(iii) Use the above to conclude that, for any \( x \neq 0 \) and any \( n \)

\[
\mathbb{P}_0[S_{2n} = 0] > \mathbb{P}_0[S_{2n} = x] + \mathbb{P}_0[S_{2n+1} = x].
\]

Exercise 3.26 Can you simplify the proof of existence of the potential kernel \( a \) using the coordinate decoupling idea from the end of Section 2.1?

Exercise 3.27 Give a rigorous proof that for any \( d \) and any finite \( \Lambda \subset \mathbb{Z}^d \) it holds that \( \mathbb{E}_x \tau_{\partial \Lambda} < \infty \) for all \( x \).

Exercise 3.28 Prove that

\[
a(x) = \frac{1}{2\pi} \int_{[-\pi,\pi]^2} \frac{1 - \cos(\theta_1 x_1 + \theta_2 x_2)}{1 - \frac{1}{2}(\cos \theta_1 + \cos \theta_2)} d\theta.
\]

Exercise 3.29 Derive (3.36) either by analysing (3.74) (cf. Section III.12 of [66]) or by analysing (3.25) directly, via the Local Central Limit Theorem (cf. Section 4.4 of [46]).

Exercise 3.30 In the proof of Theorem 3.13, show that

\[ a(x - y) = \mathbb{E}_x Y_{\tau_{\partial \Lambda}} \]

without invoking the Optional Stopping Theorem.

Exercise 3.31 Let \( x \neq 0 \) and define

\[ \eta_x = \sum_{k=0}^{\tau_x^+ - 1} 1\{S_k = x\} \]

to be the number of visits to \( x \) before hitting the origin. Prove that \( \mathbb{E}_x \eta_x = 2a(x) \) and \( \mathbb{E}_0 \eta_x = 1 \).

Exercise 3.32 Prove that the harmonic measure is consistent in two dimensions as well (recall Exercise 3.21): for any finite \( A \subset B \)

\[
\mathbb{P}_{hm_A}[S_{\tau_A} = y] = hm_A(y)
\]

for all \( y \in A \).
Exercise 3.33  Prove that
\[
\text{cap}(A) = \left( \sup_{y \in A} f(y) \right)^{-1},
\]
where the supremum is over all nonnegative functions \( f \) on \( A \) such that \( \sum_{y \in A} a(x - y) f(y) \leq 1 \) for all \( x \in A \).

Exercise 3.34  Let \( x_1, x_2, x_3 \in \mathbb{Z}^2 \) be three distinct sites, and abbreviate \( v_1 = x_2 - x_1, v_2 = x_3 - x_2, v_3 = x_1 - x_3 \). Prove that the capacity of the set \( A = \{x_1, x_2, x_3\} \) is given by the formula
\[
a(v_1)a(v_2)a(v_3)
a(v_1)a(v_3) + a(v_1)a(v_3) + a(v_2)a(v_3) - \frac{1}{2}(a^2(v_1) + a^2(v_2) + a^2(v_3))
\]
\[(3.76)\]

Exercise 3.35  What if we try to develop the same theory as in Section 3.1, only on a finite set \( \Lambda \subset \mathbb{Z}^2 \) instead of the whole space \( \mathbb{Z}^d, d \geq 3 \)? The analogy is clear: we regard the outside of \( \Lambda \) as “infinity”. Let \( A \subset \Lambda \). Analogously to (3.9), let us define for any \( x \in \Lambda \)
\[
\text{Es}_{A,\Lambda}(x) = \mathbb{P}_x[\tau^+_A > \tau_{\partial,\Lambda}] \mathbf{1}\{x \in A\};
\]
on \( A \), this equals the probability of “escaping from \( A \) within \( \Lambda \)”. Also, let
\[
\text{u}_{A,\Lambda}(x) = \mathbb{P}_x[\tau_A < \tau_{\partial,\Lambda}]
\]
be the probability of “hitting \( A \) within \( \Lambda \)”; note that \( u_{A,\Lambda} = 1 \) for all \( x \in A \).

(i) Prove a relation analogous to (3.11):
\[
u_{A,\Lambda}(y) = \sum_{x \in \Lambda} G_{\Lambda}(y, x) \text{Es}_{A,\Lambda}(x),
\]
or, in the matrix form, \( u_{A,\Lambda} = G_{\Lambda} \text{Es}_{A,\Lambda} \).

(ii) Let us define
\[
\text{cap}_{\Lambda}(A) = \sum_{x \in \Lambda} \text{Es}_{A,\Lambda}(x).
\]
\[(3.78)\]

What general properties of this notion of capacity can one obtain?
(iii) Now, consider the case $\Lambda = B(R)$, and let $R$ grow to infinity. What happens then to the objects we just considered? How does the “canonical” two-dimensional capacity (as defined in (3.18)) then relates to the one defined in (3.78)?

**Exercise 3.36** Lemma 3.11 permits us to obtain good approximations for probabilities of exiting annuli at inner/outer boundaries, in the case when the two circumferences are (almost) concentric. But what can be said about the situation depicted on Figure 3.7 (of course, in the regime when the radii of the circumferences are large)? Can you propose a method of obtaining a good approximation for that probability?

**Exercise 3.37** Can you prove the existence of the potential kernel (defined as in (3.25)) for the one-dimensional SRW? Can you actually calculate it?

**Exercise 3.38** Obtain an explicit formula (that is, more explicit than (3.45)) for the one-dimensional Green’s function (defined as in (3.43)) restricted on an interval $\Lambda$. 

**Figure 3.7** The walk hits the outer boundary
Simple random walk conditioned on not hitting the origin

4.1 Doob’s $h$-transforms

Let us start with a one-dimensional example. Let $(S_n, n \geq 0)$ be the simple random walk in dimension 1. It is well known that for any $0 < x < R$

$$P_x[\tau_R < \tau_0] = \frac{x}{R}$$  (4.1)

— this is the solution of the Gambler’s Ruin Problem (found in elementary probability books, most of them) for players of equal strength. For the purposes of this chapter, however, it is also important to notice that the above fact follows, in a quite straightforward way, from the Optional Stopping Theorem applied to the martingale $(S_n)$ and the stopping time $\tau_0 \wedge \tau_R$. Now, how will the walk behave if we condition it to reach $R$ before reaching the origin? Using (4.1), we write

$$P_x[S_1 = x + 1 \mid \tau_R < \tau_0]$$

$$= \frac{P_x[S_1 = x + 1, \tau_R < \tau_0]}{P_x[\tau_R < \tau_0]}$$

$$= \frac{P_x[S_1 = x + 1]P_x[\tau_R < \tau_0 \mid S_1 = x + 1]}{P_x[\tau_R < \tau_0]}$$

$$= \frac{\frac{1}{2}P_{x+1}[\tau_R < \tau_0]}{P_x[\tau_R < \tau_0]}$$

$$= \frac{\frac{1}{2} \times \frac{x + 1}{R}}{\frac{x}{R}}$$

$$= \frac{1}{2} \times \frac{x + 1}{x},$$

which also implies that $P_x[S_1 = x - 1 \mid \tau_R < \tau_0] = \frac{1}{2} \times \frac{x - 1}{x}$. Notice that, by the way, the drift at $x$ of the conditioned walk is of
4.1 Doob’s h-transforms

order $\frac{1}{2}$ – this is the so-called Lamperti’s process which we have already seen in this book: recall Exercises 2.18 and 2.19. The above calculation does not yet formally show that the conditioned walk is a Markov process (strictly speaking, we would have needed to condition on the whole history, to begin with), but let us forget about that for now, and examine the new transition probabilities we just obtained, $\hat{p}(x, x - 1) = \frac{1}{2} \times \frac{|x - 1|}{x}$ and $\hat{p}(x, x + 1) = \frac{1}{2} \times \frac{|x + 1|}{x}$.

First, it is remarkable that they do not depend on $R$, which suggests that we can just send $R$ to infinity and obtain “the random walk conditioned on never returning to the origin”. Secondly, just look at the arguments of $\hat{p}$’s and the second fraction in the right-hand sides: these new transition probabilities are related to the old ones (which are $p(x, y) = \frac{1}{2}$ for $x \sim y$) in a special way:

\[ \hat{p}(x, y) = p(x, y) \times \frac{h(y)}{h(x)} \]  

(4.2)

with $h(x) = |x|$ (soon it will be clear why do we prefer to keep the function nonnegative). What is special about this function $h$ is that it is harmonic outside the origin, so that $h(S_{n, \tau_0})$ is a martingale. It is precisely this fact that permitted us to obtain (4.1) with the help of the Optional Stopping Theorem.

Keeping the above in mind, let us spend some time with generalities. Consider a countable Markov chain on a state space $\Sigma$, and let $A \subset \Sigma$ be finite. Let $h : \Sigma \to \mathbb{R}_+$ be a nonnegative function which is zero on $A$ and strictly positive and harmonic outside $A$, i.e., $h(x) = \sum_y p(x, y)h(y)$ for all $x \notin A$. We assume also that $h(x) \to \infty$ as $x \to \infty$; this clearly implies that the Markov chain is recurrent, recall Theorem 2.4.

**Definition 4.1** The new Markov chain with transition probabilities defined as in (4.2) is called the Doob’s $h$-transform of the original Markov chain with respect to $h$.

Observe that the harmonicity of $h$ implies that $\hat{p}$’s are transition probabilities indeed:

\[ \sum_y \hat{p}(x, y) = \frac{1}{h(x)} \sum_y p(x, y)h(y) = \frac{1}{h(x)} \times h(x) = 1. \]

To the best of the author’s knowledge, this kind of object first appeared in [28], in the continuous-space-and-time context. In this book, we again (as in Chapter 3) do not try build a comprehensive
Conditioned SRW

genereal theory of Doob’s $h$-transforms, but rather only look at the simplified setup (only simple random walks) and take our profits. Further information can be found e.g. in [12, 47, 77], and the book [29] provides a systematic treatment of the subject in full generality.

Note the following simple calculation: for any $x \notin A \cup \partial_e A$, we have (note that $h(y) \neq 0$ for all $y \sim x$)

$$
\mathbb{E}_x \frac{1}{h(X_1)} = \sum_{y \sim x} \hat{p}(x, y) \frac{1}{h(y)}
= \sum_{y \sim x} p(x, y) \frac{h(y)}{h(x) h(y)}
= \frac{1}{h(x)} \sum_{y \sim x} p(x, y)
= \frac{1}{h(x)},
$$

which implies the following

**Proposition 4.2** The process $1/h(\hat{X}_{n \wedge \tau_A \cup \partial_e A})$ is a martingale and the Markov chain $\hat{X}$ is transient.

(The last statement follows from Theorem 2.5 since $h(x) \to \infty$ as $x \to \infty$.)

Now, let us try to get an idea about what the $h$-transformed chain really does. For technical reasons, let us make another assumption\(^1\): there exists $c > 0$ such that $|h(x) - h(y)| \leq c$ for all $x \sim y$ (for general Markov chains, $x \sim y$ means $p(x, y) + p(y, x) > 0$).

For $R > 0$, let us define

$$
\Lambda_R = \{ x \in \Sigma : h(x) \leq R \};
$$

under the above assumptions, $\Lambda_R$ is finite for any $R$. Note that the Optional Stopping Theorem implies that, for $x_0 \in \Lambda_R \setminus A$

$$
h(x_0) = \mathbb{P}_{x_0}[\tau_{\partial_e \Lambda_R} < \tau_A] \mathbb{E}_{x_0}(h(X_{\tau_{\partial_e \Lambda_R}}) \mid \tau_{\partial_e \Lambda_R} < \tau_A),
$$

(recall that $\mathbb{E}_{x_0}(h(X_{\tau_A}) \mid \tau_A < \tau_{\partial_e \Lambda_R}) = 0$ because $h$ is zero on $A$)

\(^1\) one can live without this assumption e.g. in the one-dimensional nearest-neighbour case, see Exercises 4.3 and 4.4
4.1 Doob’s $h$-transforms

Figure 4.1 Comparing the weights of the path

and, since the second factor in the above display is in $[R, R + c]$, we have

$$\mathbb{P}_{x_0} [\tau_{h, \Lambda_R} < \tau_A] = \frac{h(x_0)}{R} (1 + O(R^{-1})). \quad (4.3)$$

Then, we consider another countable Markov chain $\tilde{X}$ on the state space $\Sigma \setminus A$ with transition probabilities $\tilde{p}(\cdot, \cdot)$ defined as in (4.2) for $x \notin A$. Now, consider a path $\varphi = (x_0, \ldots, x_{n-1}, x_n)$ where $x_0, \ldots, x_{n-1} \in \Lambda_R \setminus A$ and $x_n \in \Sigma \setminus \Lambda_R$ (here, “path” is simply a sequence of neighbouring sites; in particular, it need not be self-avoiding). The original weight of that path (i.e., the probability that the Markov chain $X$ follows it starting from $x_0$) is

$$P_\varphi = p(x_0, x_1)p(x_1, x_2) \ldots p(x_{n-1}, x_n),$$

and the weight of the path for the new Markov chain $\tilde{X}$ will be

$$\tilde{P}_\varphi = p(x_0, x_1) \frac{h(x_1)}{h(x_0)} p(x_1, x_2) \frac{h(x_2)}{h(x_1)} \ldots p(x_{n-1}, x_n) \frac{h(x_n)}{h(x_{n-1})}$$

$$= p(x_0, x_1)p(x_1, x_2) \ldots p(x_{n-1}, x_n) \frac{h(x_n)}{h(x_0)}$$

$$= \tilde{P}_\varphi \frac{h(x_n)}{h(x_0)}. \quad (4.4)$$

Here comes the key observation: the last term in (4.4) actually equals $\frac{R}{h(x_0)} (1 + O(R^{-1}))$, that is, it is almost inverse of the ex-
Conditioned SRW

pression in the right-hand side of (4.3). So, we have

\[ \hat{P}_\nu = \frac{P_\nu}{\mathbb{P}_{x_0}[\tau_{\partial\Lambda_R} < \tau_A]} (1 + O(R^{-1})) , \]

that is, the probability that the \( \hat{X} \) chain follows a path is almost the conditional probability that that the original chain \( X \) follows that path, under the condition that it goes out of \( \Lambda_R \) before reaching \( A \) (and the relative error goes to 0 as \( R \to \infty \)). Now, the (decreasing) sequence of events \( \{\tau_{\partial\Lambda_R} < \tau_A\} \) converges to \( \{\tau_A = \infty\} \) as \( R \to \infty \). Therefore, we can rightfully call \( \hat{X} \) the Markov chain conditioned on never reaching \( A \), even though the probability of the latter event equals zero.

We end this section with an unexpected\(^2\) remark. Recall the conditioned one dimensional SRW we just constructed: denoting \( \Delta_x = S_1 - x \), we calculated that \( E_x \Delta_x = \frac{1}{2} \); and also, obviously, it holds that \( E_x \Delta_x^2 = 1 \). So,

\[ x E_x \Delta_x = E_x \Delta_x^2 , \]

notice, by the way, that this relation remains unaffected if one rescales the space by a constant factor. Now, recall Exercise 2.18: the above equality will (asymptotically) hold in three dimensions (and only in three dimensions) for the norm of the SRW (i.e., its distance to the origin). This suggests that there may be some “hidden relationship” between the one-dimensional conditioned SRW, and the norm of the three-dimensional SRW. Since in the continuous limit such relationships often reveal themselves better, one may wonder if the (suitably defined) one-dimensional conditioned (on not hitting the origin) Brownian motion is somehow related to the norm of the three-dimensional Brownian motion. The reader is invited to look at this question more closely.

4.2 Conditioned simple random walk in two dimensions:

basic properties

As you probably expected, we now turn our attention to the two-dimensional SRW. By (3.33) the potential kernel \( a \) is ready to play the role of the \( h \), so let us define another random walk \( (\hat{S}_n, n \geq 0) \)

\(^2\) well, maybe, quite the contrary
on \( \mathbb{Z}^2 \setminus \{0\} \) in the following way: its transition probability matrix equals
\[
\hat{p}(x, y) = \begin{cases} 
\frac{a(y)}{4a(x)}, & \text{if } x \sim y, x \neq 0, \\
0, & \text{otherwise.}
\end{cases}
\]
(4.5)

The discussion of the previous section then means that the random walk \( \hat{S} \) is the Doob h-transform of the simple random walk, under the condition of not hitting the origin. Let \( \hat{\tau} \) and \( \hat{\tau}^+ \) be the entrance and the hitting times for \( \hat{S} \); they are defined as in (1.1)–(1.2), only with \( \hat{S} \). We summarize the basic properties of the random walk \( \hat{S} \) in the following

**Proposition 4.3**  The following statements hold:

(i) The walk \( \hat{S} \) is reversible, with the reversible measure \( \mu(x) = a^2(x) \).

(ii) In fact, it can be represented as a random walk on the two-dimensional lattice with the set of conductances \( (a(x)a(y), x, y \in \mathbb{Z}^2, x \sim y) \).

(iii) The process \( 1/a(\hat{S}_{n \wedge \hat{\tau}_\nu}) \) is a martingale\(^3\).

(iv) The walk \( \hat{S} \) is transient.

**Proof**   Indeed, for (i) and (ii) note that
\[
a^2(x)\hat{p}(x, y) = \frac{a(x)a(y)}{4} = a^2(y)\hat{p}(y, x)
\]
for all adjacent \( x, y \in \mathbb{Z}^2 \setminus \{0\} \), and, since \( a \) is harmonic outside the origin,
\[
\frac{a(x)a(y)}{\sum_{z \sim x} a(x)a(z)} = \frac{a(y)}{4 \sum_{z \sim x} \frac{1}{4} a(z)} = \frac{a(y)}{4a(x)} = \hat{p}(x, y).
\]

Items (iii)–(iv) are Proposition 4.2. \( \square \)

Next, we relate the probabilities of certain events for the walks \( S \) and \( \hat{S} \). For \( D \subset \mathbb{Z}^2 \), let \( \Gamma_D^{(x)} \) be the set of all nearest-neighbour finite trajectories that start at \( x \in D \setminus \{0\} \) and end when entering \( \partial D \) for the first time; denote also \( \Gamma_y^{(x)} = \Gamma_{B(y,R)}^{(x)} \). For \( \mathcal{H} \subset \Gamma_D^{(x)} \) write \( S \in \mathcal{H} \) (respectively, \( \hat{S} \in \mathcal{H} \)) if there exists \( k \) such that \( (S_0, \ldots, S_k) \in \mathcal{H} \)

\(^3\) recall that \( \mathcal{N} \) is the set of the four neighbours of the origin
Conditioned SRW

(respectively, \((\hat{S}_0, \ldots, \hat{S}_k) \in \mathcal{H}\)). In the next result we show that \(\mathbb{P}_x[S \in \cdot \mid \tau_0 > \tau_{\partial B(R)}]\) and \(\mathbb{P}_x[\hat{S} \in \cdot] \) are almost indistinguishable on \(\Gamma_{0,R}^{(x)}\) (that is, the conditional law of \(S\) almost coincides with the unconditional law of \(\hat{S}\)).

**Lemma 4.4** Let \(x \in B(R) \setminus \{0\}\), and assume \(\mathcal{H} \subset \Gamma_{0,R}^{(x)}\). We have

\[
\mathbb{P}_x[ S \in \mathcal{H} \mid \tau_0 > \tau_{\partial B(R)} ] = \mathbb{P}_x[ \hat{S} \in \mathcal{H} ] (1 + O((R \ln R)^{-1})).
\]  

(4.6)

On sets which are distant from the origin, however, we show that \(S\) and \(\hat{S}\) have almost the same behaviour (without any conditioning):

**Lemma 4.5** Suppose that \(0 \notin D\), and assume that \(\mathcal{H} \subset \Gamma_D^{(x)}\); also, denote \(s = \text{dist}(0, D)\), \(r = \text{diam}(D)\). Then, for \(x \in D\),

\[
\mathbb{P}_x[S \in \mathcal{H}] = \mathbb{P}_x[\hat{S} \in \mathcal{H}] \left(1 + O\left(\frac{r}{s \ln s}\right)\right).
\]  

(4.7)

**Proof of Lemmas 4.4 and 4.5** Let us prove (4.6). Assume without loss of generality that no trajectory from \(\mathcal{H}\) passes through the origin. Note that for any path \(\varrho = (x_0, x_1, \ldots, x_n) \in \mathbb{Z}^2 \setminus \{0\}\) we have (as in (4.4))

\[
\mathbb{P}_{x_0}[\hat{S}_1 = x_1, \ldots, \hat{S}_n = x_n] = \frac{a(x_1)}{4a(x_0)} \times \frac{a(x_2)}{4a(x_1)} \times \cdots \times \frac{a(x_n)}{4a(x_{n-1})}
\]

\[
= \frac{a(x_n)}{a(x_0)} \left(\frac{1}{4}\right)^n,
\]

(4.8)

and therefore it holds that

\[
\mathbb{P}_x[\hat{S} \in \mathcal{H}] = \sum_{\varrho \in \mathcal{H}} \frac{a(\varrho_{\text{end}})}{a(x)} \left(\frac{1}{4}\right)^{|\varrho|},
\]

where \(|\varrho|\) is the length of \(\varrho\) and \(\varrho_{\text{end}}\) is the last site of \(\varrho\). On the other hand, by Lemma 3.12

\[
\mathbb{P}_x[S \in \mathcal{H} \mid \tau_0 > \tau_{\partial B(R)}] = \frac{a(R) + O(R^{-1})}{a(x)} \sum_{\varrho \in \mathcal{H}} \left(\frac{1}{4}\right)^{|\varrho|}.
\]

Since \(\varrho_{\text{end}} \in \partial B(R)\), we have \(a(\varrho_{\text{end}}) = a(R) + O(R^{-1})\) by (3.38), and so (4.6) follows.

The proof of (4.7) is essentially the same, only this time the factor \(\frac{a(\varrho_{\text{end}})}{a(x)}\) will be equal to \(1 + O\left(\frac{r}{s \ln s}\right)\), again due to (3.38). \(\square\)
4.2 Conditioned SRW: basic properties

Next, we need an analogue of Lemma 3.11 for the conditioned random walk, i.e., we would like to estimate the probability that the $\hat{S}$-walk escapes an annulus through its outer boundary:

**Lemma 4.6** For all $x \in \mathbb{Z}^2$ and $R > r > 0$ such that $x \in B(R) \setminus B(r)$ we have

$$P_x[\hat{\tau}_{\partial B(R)} < \hat{\tau}_{B(r)}] = 1 - \frac{(a(x))^{-1} - (a(R) + O(R^{-1}))^{-1}}{(a(r) + O(r^{-1}))^{-1} - (a(R) + O(R^{-1}))^{-1}},$$

as $r, R \to \infty$.

**Proof** This is an argument of the type we have seen already many times in this book: use the Optional Stopping Theorem for the martingale $1/(\hat{S}_k \wedge \hat{\tau}_N)$ with the stopping time $\hat{\tau}_{\partial B(R)} \wedge \hat{\tau}_{B(r)}$. We let the reader fill the details. \[\square\]

Letting $R \to \infty$ in (4.9) yields

**Corollary 4.7** Assume $r \geq 1$ and $\|x\| \geq r + 1$. We have

$$P_x[\hat{\tau}_{B(r)} = \infty] = 1 - \frac{a(r) + O(r^{-1})}{a(x)}.$$  \hspace{1cm} (4.10)

One typically needs relations such as (4.9)–(4.10) when dealing with *excursions* of the conditioned random walk. If $A \subset A'$ are (finite) subsets of $\mathbb{Z}^2$, then the *excursions* between $\partial A$ and $\partial A'$ are pieces of nearest-neighbour trajectories that begin on $\partial A$ and end on $\partial A'$, see Figure 4.2, which is, hopefully, self-explanatory. We refer to Section 3.4 of [18] for formal definitions.

For example, let $A = B(r)$ and $A' = B(2r)$. Then, for each $x \in \partial A'$, (4.10) and (3.36) imply that

$$P_x[\hat{\tau}_{B(r)} = \infty] = 1 - \frac{\frac{\pi}{2} \ln r + \gamma' + O(r^{-1})}{\frac{\pi}{2} \ln 2r + \gamma' + O(r^{-1})}$$

$$= \frac{\ln 2 + O(r^{-1})}{\ln r + (\frac{\pi}{2} \gamma' + \ln 2) + O(r^{-1})},$$

that is, the distribution of the total number of such excursions is very close to Geometric with success probability $\frac{\ln 2}{\ln r + (\frac{\pi}{2} \gamma' + \ln 2)}$.

Arguments of this sort will appear several times in the subsequent exposition.
What we do next, is develop some potential theory for the conditioned walk $\hat{S}$.

### 4.3 Green’s function and capacity

The Green’s function of the conditioned walk is defined in the following way (completely analogous to (3.2)): for $x, y \in \mathbb{Z}^2 \setminus \{0\}$

$$
\hat{G}(x, y) = \mathbb{E}_x \sum_{k=0}^{\infty} 1\{\hat{S}_k = y\}.
$$

(4.11)

It is remarkable that one is actually able to calculate this function in terms of the potential kernel $a$:

**Theorem 4.8** For all $x, y \in \mathbb{Z}^2 \setminus \{0\}$ it holds that

$$
\hat{G}(x, y) = \frac{a(y)}{a(x)} (a(x) + a(y) - a(x - y)).
$$

(4.12)

**Proof** First, we need a very simple general fact about hitting times of recurrent Markov chains:

**Lemma 4.9** Let $(X_n)$ be a recurrent Markov chain on a state space $\Sigma$, and $x \in \Sigma$, $A, B \subset \Sigma$ are such that $A \cap B = \emptyset$ and
Theorem 4.8. Fix (a large) $R > 0$, abbreviate $\Lambda_R = B(R) \setminus \{0\}$, and let us denote for $y \in \Lambda_R$

$$N^*_y,R = \sum_{k=0}^{\tau_{\Lambda_R}^y} 1\{S_k = y\},$$  

$$\hat{N}^*_y,R = \sum_{k=0}^{\hat{\tau}_{\Lambda_R}^y} 1\{\hat{S}_k = y\},$$

to be the numbers of visits to $y$ before hitting 0 or $\partial_e B(R)$, for the SRW and the conditioned SRW. Let us also denote $\hat{G}_R(x, y) = \mathbb{E}_x \hat{N}^*_y,R$. Before looking at the next argument, it is a good idea to recall how “exactly $k$ visits” became “at least $k$ visits” in (3.12). Now, let $x \in \Lambda_R$ and observe that, on one hand,

$$\mathbb{P}_x[N^*_y,R = n, \tau_0 < \tau_{L,R}(\Lambda_R^y)] = \mathbb{P}_x[N^*_y,R \geq n][\mathbb{P}_y[\tau_0 < \tau_{L,R}(\Lambda_R^y)] \mathbb{P}_y[\tau_{\partial_e B(R)} < \tau_{L,R}(\Lambda_R^y)]],$$

$$\mathbb{P}_x[N^*_y,R \geq n][\mathbb{P}_y[\tau_{\partial_e B(R)} < \tau_{L,R}(\Lambda_R^y)] \mathbb{P}_y[\tau_{L,R}(\Lambda_R^y)]],$$

(4.13)
Conditioned SRW

(by Lemma 4.9)
$$\Pr_x[N_{y,R}^* \geq n] \Pr_y[\tau_{\Lambda^c_R}^+ > \tau_{0,B(R)}] \Pr_y[\tau_{0,B(R)} < \tau_0]$$

$$= \Pr_x[N_{y,R}^* = n] \Pr_y[\tau_{0,B(R)} < \tau_0]$$

(by Lemma 3.12)
$$= \Pr_x[N_{y,R}^* = n] \frac{a(y)}{a(R) + O(R^{-1})}, \quad (4.14)$$

and, on the other hand, the same expression can be also treated in the following way:

$$\Pr_x[N_{y,R}^* = n, \tau_{0,B(R)} < \tau_0]$$

$$= \Pr_x[N_{y,R}^* = n] \Pr_x[\tau_{0,B(R)} < \tau_0]$$

(by Lemma 4.4)
$$= \Pr_x[\hat{N}_{y,R}^* = n](1 + O((R \ln R)^{-1})) \Pr_x[\tau_{0,B(R)} < \tau_0]$$

(again, by Lemma 3.12)
$$= \Pr_x[\hat{N}_{y,R}^* = n](1 + O((R \ln R)^{-1})) \frac{a(x)}{a(R) + O(R^{-1})}. \quad (4.15)$$

Note also that $a(R) + O(R^{-1}) = a(R)(1 + O((R \ln R)^{-1}))$. So, since (4.14) and (4.15) must be equal, we have

$$a(x) \Pr_x[\hat{N}_{y,R}^* = n] = a(y) \Pr_x[N_{y,R}^* = n](1 + O((R \ln R)^{-1}))$$

multiplying by $n$ and summing in $n \geq 1$, we obtain

$$a(x)G_R(x,y) = a(y)G_{\Lambda_R}(x,y)(1 + O((R \ln R)^{-1})). \quad (4.16)$$

Note that $G_R(x,y) \to G(x,y)$ as $R \to \infty$, due to the Monotone Convergence Theorem. Next, we are actually able to say something about $G_{\Lambda_R}(x,y)$: by Theorem 3.13, it holds that

$$G_{\Lambda_R}(x,y) = \mathbb{E}_x a(S_{\tau_{\Lambda_R}} - y) - a(x - y)$$

(once again, by Lemma 3.41)
$$= \frac{a(x)}{a(R) + O(R^{-1})}(a(R) + O(\|y\|^{-1}R))$$

$$+ \left(1 - \frac{a(x)}{a(R) + O(R^{-1})}\right)a(y) - a(x - y)$$

$^4$ by the way, recall the solution of Exercise 3.31
\[ = a(x) + a(y) - a(x - y) + O\left(\frac{\|y\| + 1}{R} + \frac{a(x)a(y)}{a(R)}\right). \]

Inserting this back to (4.16) and sending \( R \) to infinity, we finally obtain (4.12).

At this point, let us recall that the function \( 1/a(\cdot) \) is harmonic\(^5\) on \( \mathbb{Z}^2 \setminus (\mathcal{N} \cup \{0\}) \), and observe that the Green’s function \( \hat{G}(\cdot, y) \) is harmonic on \( \mathbb{Z}^2 \setminus \{0, y\} \) (as before, this is an immediate consequence of the total expectation formula, recall (3.3)). It turns out that this “small” difference will be quite important: indeed, the latter fact will be operational in some places below, for applying the Optional Stopping Theorem in some particular settings. For future reference, we formulate the above fact in the equivalent form:

**Proposition 4.10** For any \( y \in \mathbb{Z}^2 \setminus \{0\} \) it holds that the process \( (\hat{G}(\hat{S}_n \wedge \hat{\tau}_y, y), n \geq 0) \) is a martingale. Moreover, let us define

\[
\hat{\ell}(x, y) = 1 + \frac{a(y) - a(x - y)}{a(x)} = \frac{\hat{G}(x, y)}{a(y)}. \tag{4.17}
\]

Then the process \( (\hat{\ell}(\hat{S}_n \wedge \hat{\tau}_y), n \geq 0) \) is a martingale.

By the way, notice that

\[
\lim_{x \to \infty} \hat{\ell}(x, y) = 0 \tag{4.18}
\]

for any fixed \( y \), so the last process is a “martingale vanishing at infinity”, which makes it more convenient for applications via the Optional Stopping Theorem (so this is why we kept “1+” in (4.17)).

Now, as we already know from Section 3.1, it is possible to obtain exact expressions (in terms of the Green’s function) for one-site escape probabilities, and probabilities of (not) hitting a given site. Indeed, since, under \( \mathbb{P}_x \), the number of visits\(^6\) to \( x \) is Geometric with success probability \( \mathbb{P}_x[\hat{\tau}_x = \infty] \), using (4.13), we obtain

\[
\mathbb{P}_x[\hat{\tau}_x^+ < \infty] = 1 - \frac{1}{\hat{G}(x, x)} = 1 - \frac{1}{2a(x)} \tag{4.19}
\]

\(^5\) with respect to the conditioned walk
\(^6\) counting the one at time 0
for \( x \neq 0 \). Also, quite analogously to (3.4), since

\[
\hat{G}(x, y) = \mathbb{P}_x[\hat{\tau}_y^+ < \infty] \hat{G}(y, y) \quad \text{for } x \neq y, \ x, y \neq 0
\]

(one needs to go to \( y \) first, to start counting visits there), we have

\[
\mathbb{P}_x[\hat{\tau}_y < \infty] = \frac{\hat{G}(x, y)}{\hat{G}(y, y)} = \frac{a(x) + a(y) - a(x - y)}{2a(x)}.
\]

(4.20)

Let us also observe that (4.20) implies the following surprising fact: for any \( x \neq 0 \),

\[
\lim_{y \to \infty} \mathbb{P}_x[\hat{\tau}_y < \infty] = \frac{1}{2}.
\]

(4.21)

It is interesting to note that this fact permits us to obtain a criterion for recurrence of a set with respect to the conditioned walk. Quite analogously to Definition 3.5, we say that a set is recurrent with respect to a (transient) Markov chain, if it is visited infinitely many times almost surely; a set is called transient, if it is visited only finitely many times almost surely. Recall that, for SRW in dimension \( d \geq 3 \), the characterization is provided by the Wiener’s criterion (Theorem 3.7) formulated in terms of capacities of intersections of the set with exponentially growing annuli. Although this result does provide a complete classification, it may be difficult to apply it in practice, because it is not always trivial to calculate (even to estimate) capacities. Now, it turns out that for the conditioned two-dimensional walk \( \hat{S} \) the characterization of recurrent and transient sets is particularly simple:

**Theorem 4.11**  A set \( A \subset \mathbb{Z}^2 \) is recurrent with respect to \( \hat{S} \) if and only if \( A \) is infinite.

**Proof of Theorem 4.11**  Clearly, we only need to prove that every infinite subset of \( \mathbb{Z}^2 \) is recurrent for \( \hat{S} \). As mentioned before, this is basically a consequence of (4.21). Indeed, let \( \hat{S}_0 = x_0 \); since \( A \) is infinite, by (4.21) one can find \( y_0 \in A \) and \( R_0 \) such that \( \{x_0, y_0\} \subset B(R_0) \) and

\[
\mathbb{P}_{x_0}[\hat{\tau}_{y_0} < \hat{\tau}_{\partial B(R_0)}] \geq \frac{1}{3}.
\]

Then, for any \( x_1 \in \partial B(R_0) \), we can find \( y_1 \in A \) and \( R_1 > R_0 \) such
4.3 Green’s function and capacity

that \( y_1 \in B(R_1) \setminus B(R_0) \) and

\[
\mathbb{P}_{x_1} \left[ \hat{\tau}_{y_1} < \hat{\tau}_{\partial B(R_1)} \right] \geq \frac{1}{3}.
\]

Continuing in this way, we can construct a sequence \( R_0 < R_1 < R_2 < \ldots \) (depending on the set \( A \)) such that, for each \( k \geq 0 \), the walk \( \hat{S} \) hits \( A \) on its way from \( \partial B(R_k) \) to \( \partial B(R_{k+1}) \) with probability at least \( \frac{1}{3} \), regardless of the past. This clearly implies that \( A \) is a recurrent set.

Next, following the script of Section 3.1, let us discuss the notion of capacity for the conditioned walk. It is tempting to just repeat the previous definitions by reusing (3.10), but, at this point, some care has to be taken. Namely, recall that in Section 3.1 we used path reversals a few times — for SRW, the probability that it follows a path is the same as the probability that it follows the reversed path. This is no longer true for the conditioned walk; it is still reversible, but the reversible measure is not constant (recall Proposition 4.3).

As a consequence of the above, we have that the Green’s function (4.12) for the conditioned walk is no longer symmetric; instead, it holds that

\[
a^2(x)\hat{G}(x,y) = a^2(y)\hat{G}(y,x) \tag{4.22}
\]

for all \( x, y \in \mathbb{Z}^2 \setminus \{0\} \). This, of course, follows directly from (4.12), but can be also obtained independently, analogously to the proof of (3.44) (recall also Exercise 3.24).

For finite \( A \subset \mathbb{Z}^d \) and \( x \in \mathbb{Z}^d \), quite analogously to (3.9), let us denote by

\[
\hat{\text{ES}}_A(x) = \mathbb{P}_{x}[\hat{\tau}_A^+ = \infty]1\{x \in A\} \tag{4.23}
\]

the escape probability from \( x \in A \) with respect to the conditioned walk. Now, the crucial observation (that the reader is strongly invited to check) is that one can show that, for any \( A \subset \mathbb{Z}^d \setminus \{0\} \),

\[
\mathbb{P}_{x}[\hat{\tau}_A < \infty] = \sum_{y \in A} \hat{G}(x,y)\hat{\text{ES}}_A(y) = \sum_{y \in \mathbb{Z}^d} \hat{G}(x,y)\hat{\text{ES}}_A(y) \tag{4.24}
\]

exactly in the same way as (3.11) was proved!

However, maybe somewhat surprisingly, it is not a good idea to define the capacity for the conditioned walk as in (3.10). To explain this, consider first a toy model. Let \( X \) be the random
walk on the three-dimensional integer lattice with conductances on all the horizontal planes defined as on Figure 4.4, and the conductancies of all vertical edges being equal to 1. Clearly, $X$ is transient and reversible, with the reversible measure

$$
\mu(x) = \begin{cases} 
6, & \text{if } x \cdot e_1 \text{ is even}, \\
7, & \text{if } x \cdot e_1 \text{ is odd}.
\end{cases}
$$

This informally means that the odds that the process is in a site with an odd abscissa is $\frac{6}{6+7} = \frac{6}{13}$. Now, consider the $n$-step transition probability $p^{(n)}(x, y)$, where $n$ is of the same parity as $x - y$. Intuitively, it should not depend so much on $x$ (since the walk normally “forgets” its initial point anyway), but there should be a substantial dependence on $y$: by the above discussion, if, for example, $y \cdot e_1$ is even, the ratio $\frac{p^{(n)}(x, y)}{p^{(n)}(x, y')}$ should be close to $\frac{6}{7}$ in case when $y' \cdot e_1$ is odd and $y'$ has the same parity as $y$ and is “not far away” from it. So, if we divide $p^{(n)}(x, y)$ by $\mu(y)$, this will
4.3 Green’s function and capacity

(almost) remove the dependence on the second argument; since the Green’s function is the sum of $p^{(n)}$’s, $G(x, y)/\mu(y)$ looks like the “right” object to consider (it “should” depend on the distance between $x$ and $y$, but not so much on $x$ and $y$ themselves). Then, an analogue of (3.11) would be

$$\mathbb{P}_x[\tau_A < \infty] = \sum_{y \in A} G(x, y) \mathbb{E}_A(y) = \sum_{y \in A} \frac{G(x, y)}{\mu(y)} \times \mu(y) \mathbb{E}_A(y),$$

so, at least if $A$ is finite and the starting point $x$ is far away from $A$, the probability of eventually hitting $A$ would be a product of a factor which (almost) only depends on the distance with $\sum_{y \in A} \mu(y) \mathbb{E}_A(y)$. This indicates that the last quantity might be the correct definition of the capacity.

It is important to notice, however, that the above correct definition is, in principle, non-unique: as we remember from Section 2.2, the reversible measure is not unique (one can always multiply it by a positive constant). The same is also true with respect to the conductancies: if we multiply all them by the same positive constant, the corresponding random walk will remain the same. Still, we will see below that there is a canonical way to define the capacity for conditioned random walks (i.e., we can choose that multiplicative constant in a natural way).

Let us go back to the conditioned walk $\hat{S}$. First, note that, by (3.38),

$$\hat{G}(x, y) = \frac{a(y)}{a(x)}(a(x) + a(y) - a(x - y)) = \frac{a(y)}{a(x)}(a(y) + O(\frac{||y||}{||x||}))$$

as $x \to \infty$ and $\frac{||y||}{||x||} \to 0$. By (4.12), we have

$$\frac{\hat{G}(x, y)}{a^2(y)} = \frac{\hat{G}(y, x)}{a^2(x)} = \frac{a(x) + a(y) - a(x - y)}{a(x)a(y)},$$

and so it is natural to introduce new notation $\hat{g}(x, y) = \frac{\hat{G}(x, y)}{a^2(y)} = \hat{g}(y, x)$ for the “symmetrized” Green’s function. Then, (4.25) implies that

$$\hat{g}(x, y) = \frac{\hat{G}(x, y)}{a^2(y)} = \frac{1}{a(x)}(1 + O(\frac{||y||}{||x|| \ln(||y|| + 1)}))$$

as $x \to \infty$, which indeed essentially depends on $||x||$. Therefore, if
we rewrite (4.24) in the following way,

$$
P_x[\hat{\tau}_A < \infty] = \sum_{y \in A} g(x, y) \times a^2(y) \hat{E}s_A(y),$$

(4.27)

we see that the first terms in the above summation are “almost” the same for large $x$. According to the above discussion, it is then reasonable to adopt the following definition for the capacity $\hat{\text{cap}}(\cdot)$ with respect to the conditioned walk:

$$\hat{\text{cap}}(A) = \sum_{y \in A} a^2(y) \hat{E}s_A(y).$$

(4.28)

Let us go back to the recent observation that, in principle, the capacity is defined up to a multiplicative factor. Why $a$ is better than, say, $3a$ for the role of the function $h$ of Section 4.1? How to choose one of them canonically? A reasonable way to do it is the following: $h$ should be such that $E_0 h(\hat{S}_1) = 1$; as we know (recall (3.35)) $h \equiv a$ is then the right choice indeed.

Now, we have two notions of two-dimensional capacity: one for the original recurrent walk, and another one for the transient conditioned walk. What is the relationship between them? Remarkably, it is very simple:

**Theorem 4.12** For all $A \subset \mathbb{Z}^2 \setminus \{0\}$, we have

$$\hat{\text{cap}}(A) = \text{cap} (A \cup \{0\}).$$

(4.29)

**Proof** Indeed, let us write for $x \in A$

$$\hat{E}s_A(x) = P_x[\hat{\tau}_A^+ = \infty]
= \lim_{R \to \infty} P_x[\hat{\tau}_A^+ > \tau_{\partial B(R)}]
= \lim_{R \to \infty} P_x[\tau_A^+ > \tau_{\partial B(R)} \mid \tau_{\partial B(R)} < \tau_0]
= \lim_{R \to \infty} P_x[\tau_A^+ > \tau_{\partial B(R)}] \frac{a(R) + O(R^{-1})}{a(x)}
= \frac{\text{hm}_A(x)}{a(x)}.$$

(4.30)
To conclude the proof, just insert (4.30) to (4.28) and then compare to the “usual” definition (3.62) of the two-dimensional capacity (with \( y_0 = 0 \)).

Let us also obtain an expression for the probability of avoiding any finite set. The following estimate, however, only works well when the starting point is much farther from the origin than the “target” set; nevertheless, once again it validates the intuition that “capacity measures how big is the set from the walker’s point of view”. Assume that \( A \subseteq B(r) \), and \(|x| \geq r + 1\). Then, with (4.28) in mind, just look at (4.26) and (4.27) to obtain that

\[
P_x[\hat{\tau}_A < \infty] = \frac{\hat{\text{cap}}(A)}{a(x)} \left(1 + O\left(\frac{r}{|x| \ln(r+1)}\right)\right).
\]

(4.31)

The remainder of this section will be a bit too technical for the author’s taste, but, unfortunately, the lack of translational invariance of the conditioned walk claims its price. The reader certainly remembers that the estimate (3.38) on the two-point differences of the potential kernel was quite instrumental in several arguments of this book; likewise, it will be important to have difference estimates for the function \( \hat{g} \) as well:

**Lemma 4.13** Assume that \( x, y, z \in \mathbb{Z}^2 \setminus \{0\} \) are distinct and such that \(|x - y| \wedge |x - z| \geq 5|y - z|\). Then

\[
|\hat{g}(x, y) - \hat{g}(x, z)| \leq O\left(\frac{|y - z|}{|x - y| \ln(1+|x| |y||y||z||z|) \ln(1+|y||y||z||z|)}\right).
\]

(4.32)

**Proof** First, let us write

\[
\hat{g}(x, y) - \hat{g}(x, z) = \frac{a(x) + a(y) - a(x - y)}{a(x)a(y)} - \frac{a(x) + a(z) - a(x - z)}{a(x)a(z)}
\]

\[
= \frac{a(x)a(z) - a(x - y)a(z) - a(x)a(y) + a(x - z)a(y)}{a(x)a(y)a(z)}
\]

(put \( \pm a(x - z)a(z) \) to the numerator, then group accordingly)

\[
= \frac{a(x)(a(z) - a(y)) - a(x - z)(a(z) - a(y)) + a(z)(a(x - z) - a(x - y))}{a(x)a(y)a(z)}.
\]

(4.33)

Throughout this proof, let us assume without loss of generality that \(|y| \geq |z|\). Since the walk \( \hat{S} \) is not spatially homogeneous
Figure 4.5 On the proof of Lemma 4.13, the three cases to consider (from left to right): (1) \(\|x\|, \|y\|\) are of the same logarithmic order and \(\|x\|\) is not much larger than \(\|y\|\), (2) \(\|x\|\) is much smaller than \(\|y\|\), (3) \(\|x\|\) is significantly larger than \(\|y\|\).

(and, therefore, \(\hat{g}\) is not translationally invariant), we need to take into account the relative positions of the three sites with respect to the origin. Specifically, we will consider the following three different cases (see Figure 4.5).

Case 1: \(\|y\|^1/2 \leq \|x\| \leq 2\|y\|\).

In this case, the first thing to note is that

\[
\|y - z\| \leq \frac{\|x - y\|}{5} \leq \frac{\|x\| + \|y\|}{5} \leq \frac{2\|y\| + \|y\|}{5} = \frac{3}{5}\|y\|,
\]

so \(\|z\| \geq \frac{2}{5}\|y\|\), meaning that \(\|y\|\) and \(\|z\|\) must be of the same order; this then implies that \(a(x), a(y), a(z)\) are all of the same order too. Then, we use (3.38) on the three parentheses in the numerator of (4.33), to obtain after some elementary calculations that the expression there is at most of order \(\|y - z\|\|x - y\|\ln\|y\|\), while the denominator is of order \(\ln^3\|y\|\). This proves (4.32) in case 1.

Case 2: \(\|x\| < \|y\|^{1/2}\).

Here, it is again easy to see that \(\|y\|\) and \(\|z\|\) must be of the same order. Now, we note that, by (3.38), \(a(x - z) = a(z) + O(\|x\|/\|y\|)\), so, inserting this to (4.33) (and also using that \(a(y) - a(z) = O(\|y - z\|/\|y\|)\)), we find that it is equal to

\[
a(x)(a(z) - a(y)) + a(z)(a(y) - a(z) + a(x - z) - a(x - y)) + O(\|x\|/\|y\| \cdot \|y - z\|/\|y\|) \\
= a(x)a(y)a(z)
\]

\[
\frac{a(x)(a(z) - a(y)) + a(z)(a(y) - a(z) + a(x - z) - a(x - y)) + O(\|x\|/\|y\| \cdot \|y - z\|/\|y\|)}{a(x)a(y)a(z)}
\]
\[4.3 \text{ Green’s function and capacity}\]

\[
d = \frac{a(z) - a(y)}{a(y)a(z)} + \frac{a(y) - a(z) + a(x - z) - a(x - y)}{a(x)a(y)} + O\left(\frac{\|x - z\|}{\|y\|^2 \ln^2 \|y\|}\right).
\]

(4.34)

Now, by (3.38) the first term is \(O\left(\frac{\|y - z\|}{\|y\|^2 \ln^2 \|y\|}\right)\) (that is, exactly what we need, since \(\|y\|\) and \(\|y - x\|\) are of the same order), and the third term is clearly of smaller order. As for the second term, note that, by (3.36) and using the fact that \(\|x - y\| \cdot \|z\| - \|y\| \cdot \|x - z\| \leq O(\|x\| \cdot \|y - z\|)\) by the Ptolemy’s inequality, we obtain

\[
a(y) - a(z) + a(x - z) - a(x - y)
\]

\[= \frac{2}{\pi} \ln \frac{\|y\| \cdot \|x - z\|}{\|x - y\| \cdot \|z\|} + O(\|y\|^{-2})\]

\[= \frac{2}{\pi} \ln \left(1 - \frac{\|x - y\| \cdot \|z\| - \|y\| \cdot \|x - z\|}{\|x - y\| \cdot \|z\|}\right) + O(\|y\|^{-2})\]

\[= O\left(\frac{\|y - z\|}{\|x - y\| \cdot \|z\|} + \|y\|^{-2}\right) = O\left(\frac{\|y - z\|}{\|x - y\| \cdot \|z\|}\right),\]

so it is again of smaller order than the first term. This shows (4.32) in case 2.

Case 3: \(\|x\| > 2\|y\|\).

Notice that, in this case, \(\|z\|\) need not be of the same order as \(\|y\|\), it may happen to be significantly smaller. Here (by also grouping the first two terms in the numerator) we rewrite (4.33) as

\[
\frac{(a(x) - a(x - z))(a(z) - a(y))}{a(x)a(y)a(z)} + \frac{a(x - z) - a(x - y)}{a(x)a(y)}.
\]

(4.35)

By (3.38), the second term is \(O\left(\frac{\|y - z\|}{\|y\| \ln(1 + \|x\|) \ln(1 + \|y\|)}\right)\) (that is, exactly what we need). Next, observe that (recall that we assumed that \(\|y\| \geq \|z\|\))

\[
\ln \|y\| - \ln \|z\| = \ln \frac{\|y\|}{\|z\|} \leq \|y\| - \|z\| - 1 = \frac{\|y\| - \|z\|}{\|z\|} \leq \frac{\|y - z\|}{\|z\|}.
\]

Therefore (also using (3.38) on the first factor), the numerator of the first term is \(O\left(\frac{\|y\|}{\|z\|} \times \frac{\|y - z\|}{\|y\|} \right)\) = \(O\left(\frac{\|y - z\|}{\|y\|}\right)\), and so (since the denominator is not less than \(a(x)a(y)\)) the first term in (4.35) is at most of the same order as the second one. This concludes the proof of Lemma 4.13. \(\square\)
Lemma 4.13 permits us to obtain the following useful expression for the probability of ever hitting $A$ from a distant site (which can be seen as a generalization of (4.31)). Consider a finite $A \subset \mathbb{Z}^2 \setminus \{0\}$ with $y_0 \in A$ and note first that $r_0 := \|y_0\| + \text{diam}(A)$ will be of the same order regardless of the choice of $y_0$. Assume now that $\text{dist}(x, A) > \text{diam}(A)$; then, it also holds that $\|x\| \vee \|y\| \vee |z|$ is of the same order as $\|x\| \vee (\|y_0\| + \text{diam}(A))$ for any choice of $y, z \in A$. Indeed, trivially $\|x\| \vee \|y\| \vee |z| \leq \|x\| \vee (\|y_0\| + \text{diam}(A))$; on the other hand, $\|y\| + \text{diam}(A) < \|y\| + \text{dist}(x, A) \leq \|y\| + \|x - y\| \leq \|x\| + 2\|y\| \leq 3(\|x\| \vee \|y\| \vee |z|)$. Then, note that $1 + \|y\| + \|y - z\| \leq 1 + 2\|y\| + \|z\| \leq 3(1 + \|y\| \vee |z|) \leq (1 + \|y\| \vee |z|)^2$ (since the expression in parentheses is at least 2), so

$$\frac{|y - z|}{\ln(1 + \|y\| \vee |z|)} \leq 3 \frac{|y - z|}{\ln(1 + \|y\| + \|y - z\|)} \leq 3 \frac{\text{diam}(A)}{\ln(1 + \|y\| + \text{diam}(A))}.$$ 

where the second inequality is due to the fact that the function $f(x) = \frac{x}{\ln(a + x)}$ is increasing on $(0, +\infty)$ for any $a \geq 1$. So, recalling (4.28) and (4.24), we see that Lemma 4.13 gives us (in the case $\text{dist}(x, A) > \text{diam}(A)$) that

$$\mathbb{P}_x[\hat{\tau}_A < \infty] = \hat{c}\hat{\text{ap}}(A)\left(\hat{g}(x, y_0) + O\left(\frac{\text{diam}(A)}{\text{dist}(x, A) \ln(1 + \|x\| \vee |y_0|) \ln(1 + |y_0|)}\right)\right).$$

(4.36)

So, following Section 3.1, we have now delivered the Green’s function and the capacity to the conditioned walk; we pass now to the third important ingredient.

### 4.4 Harmonic measure

As before, we would like to define the harmonic measure as “the entrance measure from infinity”. To figure out how it should be defined, recall the informal explanation of the harmonic measure’s definition (for the transient case $d \geq 3$) on page 44, just after Theorem 3.8. There is a sentence there which reads “Since the time reversal does not change the “weight” of the trajectory, . . .”; now, this is no longer true! Since (recall Proposition 4.3 (i)) the conditioned walk is reversible with the reversible measure $(a^2(x), x \in \mathbb{Z}^2 \setminus \{0\})$, the “weight” of the reversed trajectory which starts at $y \in \partial A$ should now be multiplied by $a^2(y)$. So, that heuristic argument suggests that, very similarly to what we have seen in
Section 3.1, the harmonic measure should be proportional to \( y \)'s contribution in the summation (4.28).

Indeed, it turns out to be the case. Let us define the harmonic measure with respect to the conditioned walk in the following way:

\[
\hat{\text{hm}}_A(y) = \frac{a^2(y) \hat{\text{Es}}_A(y)}{\text{cap}(A)} = \frac{a(y) \text{hm}_A(y)}{\text{cap}(A)} \quad (4.37)
\]

(the last equality holds by (4.30)). We can interpret the above in the following way: \( \hat{\text{hm}} \) is \( \text{hm} \) biased by \( a \).

Now, the goal will be to prove that, analogously to the harmonic measures considered in Chapter 3, \( \hat{\text{hm}}_A \) is indeed the “entrance measure from infinity”. We will aim directly to prove a quantitative result, i.e., one that contains explicit estimates on the Radon-Nikodym derivative of the entrance measure from a distant starting point with respect to the harmonic measure. We will see that these estimates match those that we obtained for SRW. However, before we will be able to prove the above, we will need a couple of technical facts. These are, honestly, a bit tedious to prove, and, again, they are so because of the lack of the translational invariance of the conditioned walk.

The first technical fact that we need is that the conditioned walk can go out of an annulus with uniformly positive probability:

**Lemma 4.14** Let \( b, C \) be positive constants such that \( 1 + b < C \), and assume that \( r \geq 1, x_0, y_0 \in \mathbb{Z}^2 \setminus \{0\} \) are such that \( x_0 \in \mathcal{B}(y_0, Cr) \) and \( \|x_0 - y_0\| > (1 + b)r \). Then, there exists a constant \( c' > 0 \) (depending only on \( b \) and \( C \)) such that

\[
P_{x_0}[\hat{\tau}_{\partial\mathcal{B}(y_0, Cr)} < \hat{\tau}_{\mathcal{B}(y_0, r)}] \geq c'. \quad (4.38)
\]

*Proof* Note that we can assume that \( r \) is large enough, otherwise the uniform ellipticity of the conditioned walk\(^7\) will imply the claim. First of all, it is clear (and can be proved in several possible ways; see Exercise 4.11 below) that the analogue of (4.38) holds for SRW, i.e., for all \( r, x_0, y_0 \) as above it holds that

\[
P_{x_0}[\tau_{\partial\mathcal{B}(y_0, 2Cr)} < \tau_{\mathcal{B}(y_0, r)}] \geq c'' \quad (4.39)
\]

for some \( c'' > 0 \) which depends only on \( b \) and \( C \). Now, the idea is to derive (4.38) from (4.39). Recall that (as in (4.8)) the weight \( \hat{P}_\rho \)

\(^7\) the fact that \( \hat{S} \) is uniformly elliptic is almost evident, but see Exercise 4.10
with respect to $\hat{S}$ of a finite path $\varrho$ which starts at $x_0$ and does not pass through the origin equals $a(\varrho_{\text{end}}) P_{\varrho}$, where $P_{\varrho}$ is the weight of the same path with respect to the SRW. We can then write for any $R > 0$ such that $x_0 \in B(y_0, R)$

\[
P_{x_0} \left[ \tau_{\partial B(y_0, R)} < \tau_{B(y_0, r)} \right] \geq \min_{z \in \partial B(y_0, R)} \left( \frac{a(z)}{a(x_0)} \right) \times \left( P_{x_0} \left[ \tau_{\partial B(y_0, R)} < \tau_{B(y_0, r)} \right] - P_{x_0} \left[ \tau_0 < \tau_{\partial B(y_0, R)} \right] \right).
\]

Now, a key observation is that, for all large enough $r$, the property

\[
\frac{a(z)}{a(x_0)} \geq \frac{1}{2} \quad \text{for all } z \in \partial B(y_0, R)
\]

(4.41) holds for either $R = Cr$ or $R = 2Cr$ (or both). Indeed, roughly speaking, for (4.41) to hold it would be enough that $\|z\|$ is of order $r + \|y_0\|$ for all $z \in \partial B(y_0, R)$; this can be seen to be so in at least one of the above cases (look at the left side of Figure 4.6: if $\partial B(y_0, Cr)$ is “too close” to the origin, then $\partial B(y_0, 2Cr)$ is not).

For definiteness, assume now that (4.41) holds for $R = 2Cr$. By (4.40), we have then

\[
P_{x_0} \left[ \tau_{\partial B(y_0, Cr)} < \tau_{B(y_0, r)} \right] \geq P_{x_0} \left[ \tau_{\partial B(y_0, 2Cr)} < \tau_{B(y_0, r)} \right]
\]
4.4 Harmonic measure

\[
\begin{align*}
\geq & \frac{1}{2} \left( \mathbb{P}_{x_0} [\tau_{B(y_0, 2Cr)} < \tau_{B(y_0, r)}] - \mathbb{P}_{x_0} [\tau_0 < \tau_{B(y_0, 2Cr)}] \right) \\
\geq & \frac{1}{2} c'' - \mathbb{P}_{x_0} [\tau_0 < \tau_{B(y_0, 2Cr)}] \\
\geq & \frac{c''}{4},
\end{align*}
\]

provided that

\[
\mathbb{P}_{x_0} [\tau_0 < \tau_{B(y_0, 2Cr)}] \leq \frac{c''}{2}.
\]

(4.42)

Now, if \(0 \notin B(y_0, Cr)\), then (4.42) trivially holds; so, let us assume that \(0 \in B(y_0, Cr)\). We then consider two cases: \(\|x_0\| \geq \frac{b}{4} r\), and \(\|x_0\| < \frac{b}{4} r\). In the first case, Lemma 3.12 implies that \(\mathbb{P}_{x_0} [\tau_0 < \tau_{B(y_0, 2Cr)}] \approx \frac{1}{\ln r}\), so (4.42) holds for large enough \(r\). In the second case, note first that

\[
\mathbb{P}_{x_0} [\tau_{B(y_0, Cr)} < \tau_{B(y_0, r)}] \geq \min_{z \in \partial B(y_0, (1 + \frac{b}{2})r)} \mathbb{P}_z [\tau_{B(y_0, Cr)} < \tau_{B(y_0, r)}]
\]

and, for all \(z \in \partial B(y_0, (1 + \frac{b}{2})r)\), it holds that \(\|z\| \geq \frac{b}{4} r\) (see Figure 4.6 on the right). We may then repeat the above reasoning with an arbitrary \(z \in \partial B(y_0, (1 + \frac{b}{2})r)\) on the place of \(x_0\) to finally obtain the claim.

Next technical fact we need is the following lower bound on the probability that the conditioned walk never hits a disk:

**Lemma 4.15**  Fix \(b > 0\) and assume that \(x_0, y_0 \in \mathbb{Z}^2 \setminus \{0\}\) and \(r \geq 1\) are such that \(\|x_0 - y_0\| \geq (1 + b)r\). Then there exists \(c = c(b)\) such that for all \(x_0, y_0, r\) as above it holds that

\[
\mathbb{P}_{x_0} [\tau_{B(y_0, r)} = \infty] \geq \frac{c}{\ln(\|y_0\| + r)}.
\]

(4.43)

**Proof**  We need to consider two cases: \(B(y_0, r)\) is (relatively to its size) close to/far from the origin. First, let us assume that \(\|y_0\| < 12r\) (so that the disk is relatively close to the origin). In this case, we can assume additionally that \(\|x_0 - y_0\| \geq 51r\) (indeed, Lemma 4.14 implies that, for any starting position \(x_0\) such that \(\|x_0 - y_0\| \geq (1 + b)r\), with at least a constant probability the walk reaches \(\partial B(y_0, 51r)\) before hitting \(B(y_0, r))\). Then, it holds that \(B(y_0, r) \subset B(13r)\), and \(B(26r) \subset B(y_0, 51r)\). Now, Corollary 4.7

\(^8\) make a picture!
Conditioned SRW easily implies that, if \( r \geq 1 \) and \( \|x\| \geq 2r \)
\[ \mathbb{P}_x[\hat{\tau}_B(r) = \infty] \geq \frac{c'}{\ln r} \]
(because (4.10) will work for large enough \( \|x\| \), and one can use the uniform ellipticity of \( \hat{S} \) otherwise); this proves (4.43) in the first case.

Now, suppose that \( \|y_0\| \geq 12r \) (that is, \( r \leq \frac{1}{12} \|y_0\| \)). Analogously to the previous case, Lemma 4.14 permits us to assume without loss of generality that \( x_0 \in \partial B(y_0, 3r) \). We now use the martingale (recall Proposition 4.10)
\[ \hat{\ell}(\hat{S}_{n \wedge \tau_{y_0}}; y_0) = 1 + \frac{a(y_0) - a(\hat{S}_{n \wedge \tau_{y_0}} - y_0)}{a(\hat{S}_{n \wedge \tau_{y_0}})} \]

In exactly\(^9\) the same way as in the proof of Theorem 2.5 (page 22), we obtain from the Optional Stopping Theorem that
\[ \hat{\ell}(x_0, y_0) = \sum_{z \in \partial B(y_0, r)} \mathbb{P}_{x_0}[\hat{\tau}_{B(y_0, r)} < \infty, \hat{S}_{\hat{\tau}_{B(y_0, r)}} = z] \hat{\ell}(z, y_0) \]
\[ \geq \mathbb{P}_{x_0}[\hat{\tau}_{B(y_0, r)} < \infty] \min_{z \in \partial B(y_0, r)} \hat{\ell}(z, y_0), \]

so
\[ \mathbb{P}_{x_0}[\hat{\tau}_{B(y_0, r)} < \infty] \leq \frac{\hat{\ell}(x_0, y_0)}{\min_{z \in \partial B(y_0, r)} \hat{\ell}(z, y_0)}. \]

Assume \( z \in \partial B(y_0, r) \) and write, using (3.36) and with \( \gamma'' := \pi \gamma' / 2 \),
\[ \hat{\ell}(z, y_0) = \frac{a(z) + a(y_0) - a(y_0 - z)}{a(z)} \]
\[ = \frac{\ln \|z\| + \ln \|y_0\| - \ln r + \gamma'' + O(\|y_0\|^{-2} + r^{-1})}{\ln \|z\| + \gamma'' + O(\|z\|^{-2})} \]
\[ \geq \frac{\ln(\|y_0\| - r) + \ln \|y_0\| - \ln r + \gamma'' + O(\|y_0\|^{-2} + r^{-1})}{\ln(\|y_0\| + r) + \gamma'' + O(\|y_0\|^{-2} + r^{-1})} \]
\[ = \frac{2 \ln \|y_0\| + \ln (1 - \frac{r}{\|y_0\|}) - \ln r + \gamma'' + O(r^{-1})}{\ln \|y_0\| + \ln (1 + \frac{r}{\|y_0\|}) + \gamma'' + O(r^{-1})} \]
\[ := \frac{T_1}{T_2}. \]

\(^9\) recall also (4.18)
and, denoting $R := \|x_0 - y_0\|$ (so that $R = 3r + O(1)$),

$$
\hat{\ell}(x_0, y_0) = \frac{a(x_0) + a(y_0) - a(y_0 - x_0)}{a(x_0)}
= \frac{\ln \|x_0\| + \ln \|y_0\| - \ln R + \gamma'' + O(\|y_0\|^{-2} + R^{-2})}{\ln \|x_0\| + \gamma'' + O(\|x_0\|^{-2})}
\leq \frac{\ln(\|y_0\| + R) + \ln \|y_0\| - \ln R + \gamma'' + O(\|y_0\|^{-2} + R^{-2})}{\ln(\|y_0\| - R) + \gamma'' + O(\|y_0\|^{-2} + R^{-2})}
= \frac{2 \ln \|y_0\| + \ln (1 + \frac{R}{\|y_0\|}) - \ln R + \gamma'' + O(R^{-2})}{\ln \|y_0\| + \ln (1 - \frac{R}{\|y_0\|}) + \gamma'' + O(R^{-2})}
= \frac{T_3}{T_4}.
$$

(4.46)

Now, a straightforward calculation yields

$$
\frac{T_2}{T_4} = 1 + \frac{\ln \left(1 + \frac{r}{\|y_0\|} + O(r^{-1})\right)}{\ln \|y_0\| + \ln (1 - \frac{R}{\|y_0\|}) + \gamma'' + O(R^{-2})},
$$

and

$$
\frac{T_3}{T_1} = 1 - \frac{\ln \left(1 - \frac{R}{\|y_0\|} + O(r^{-1})\right)}{2 \ln \|y_0\| + \gamma''}.
$$

Therefore, by (4.44) we have (after some more calculations, sorry)

$$
P_{x_0}[\hat{\tau}_{B(y_0, r)} < \infty]
\leq \frac{T_2}{T_4} \times \frac{T_3}{T_1}
\leq 1 - \frac{\ln \left(1 - \frac{R}{\|y_0\|} + O(r^{-1})\right)}{\ln \|y_0\|(1 + O(\ln \|y_0\|))}.
$$

(4.47)

It remains only to observe that, if $r$ is large enough, the numerator in (4.47) is bounded from below by a positive constant: indeed, observe that $\frac{R}{r}$ is (asymptotically) 3, $\frac{r}{\|y_0\|}$ and $\frac{R}{\|y_0\|}$ are at most $\frac{1}{12}$ and $\frac{1}{4}$ respectively, and

$$
\sqrt{3 \times \frac{1 - \frac{1}{12}}{1 + \frac{1}{4}} \times \frac{1 - \frac{1}{3}}{1 + \frac{1}{12}}} = \sqrt{\frac{891}{845}} > 1.
$$
This concludes the proof of Lemma 4.15 in the case when $r$ is large enough; the case of smaller values of $r$, though, can be easily reduced to the former one by using the uniform ellipticity of the $\widehat{S}$-walk.

Now, we profit from the previous technical lemmas to obtain a quantitative result for the entrance measure to a finite set; it is quite analogous to Theorem 3.17, only this time for the conditioned walk:

**Theorem 4.16** Assume that $A \subset \mathbb{Z}^2 \setminus \{0\}$ is finite and $x \notin A$ is such that $\text{dist}(x, A) \geq 12(\text{diam}(A) + 1)$. For all $y \in A$, we have

$$P_x[\widehat{S}_{\widehat{\tau}_A} = y \mid \widehat{\tau}_A < \infty] = \widehat{\text{hm}}_A(y)(1 + O(\frac{\text{diam}(A)}{\text{dist}(x, A)})).$$

(4.48)

**Proof** We will begin as in the proof of Theorem 3.8 (page 44), and the reader is also advised to recall the solution of Exercise 3.22, since the next proof will have quite some similarities to it. Let us assume without restricting generality that $A$ contains at least two sites, so that $\text{diam}(A) \geq 1$. First, we proceed very similarly to the proof of Theorem 3.8. We keep the notation $\Theta^{(n)}_{x y}$ from that proof; also, recall that $P_{\nu}$ (respectively, $\widehat{P}_{\nu}$) is the weight of the trajectory $\nu$ with respect to the SRW (respectively, to the $\widehat{S}$-walk). We also keep\(^{10}\) the notations $\widehat{N}_x$, $\widehat{N}^y_x$, $\widehat{N}^z_x$ for, respectively, the total number of visits to $x \notin A$, the number of visits to $x$ before the first return to $A$, and the number of visits to $x$ after the first return to $A$ (again, setting $\widehat{N}^z_x = 0$ on $\{\widehat{\tau}_A^x = \infty\}$).

As in (3.22), it is clear that

$$P_x[\widehat{\tau}_A < \infty, \widehat{S}_{\widehat{\tau}_A} = y] = \sum_{n=1}^{\infty} \sum_{\nu \in \Theta^{(n)}_{x y}} \widehat{P}_{\nu},$$

but the relation corresponding to (3.23) will now be (due to Proposition 4.3 (i))

$$P_y[\widehat{N}^z_x \geq n] \frac{a^2(y)}{a^2(x)} = \sum_{\nu \in \Theta^{(n)}_{y x}} \widehat{P}_{\nu}.$$\(^{10}\) now with hats, as we are dealing with the $\widehat{S}$-walk
4.4 Harmonic measure

Then, analogously to (3.24) we write

\[ P_x[\hat{S}_{\tau_A} = y \mid \hat{\tau}_A < \infty] = \frac{P_x[\hat{\tau}_A < \infty, \hat{S}_{\tau_A} = y]}{P_x[\hat{\tau}_A < \infty]} \]

\[ = \frac{1}{P_x[\hat{\tau}_A < \infty]} \sum_{n=1}^{\infty} \frac{a^2(y)}{a^2(x)} P_y[\hat{N}_x^y \geq n] \]

\[ = \frac{a^2(y)}{a^2(x) P_x[\hat{\tau}_A < \infty]} \mathbb{E}_y \hat{N}_x^y \]

\[ = \frac{a^2(y)}{a^2(x) P_x[\hat{\tau}_A < \infty]} (\mathbb{E}_y \hat{N}_x^y - \mathbb{E}_y \hat{N}_x^y) \]

\[ = \frac{a^2(y)}{P_x[\hat{\tau}_A < \infty]} \left( \frac{\hat{G}(y, x)}{a^2(x)} - \sum_{z \in \partial A} P_y[\hat{\tau}_A^z \hat{S}_{\tau_A} = z] \frac{\hat{G}(z, x)}{a^2(x)} \right) \]

\[ = \frac{a^2(y)}{P_x[\hat{\tau}_A < \infty]} \left( \hat{g}(y, x) - \sum_{z \in \partial A} P_y[\hat{\tau}_A^z \hat{S}_{\tau_A} = z] \hat{g}(z, x) \right) \]

\[ = \frac{a^2(y)}{P_x[\hat{\tau}_A < \infty]} \left( \hat{g}(y, x) \left( \hat{E}_{S_A}(y) + \sum_{z \in \partial A} P_y[\hat{\tau}_A^z \hat{S}_{\tau_A} = z] \hat{g}(z, x) \right) \right. \]

\[ - \left. \sum_{z \in \partial A} P_y[\hat{\tau}_A^z < \infty, \hat{S}_{\tau_A} = z] \hat{g}(z, x) \right) \]

\[ = \frac{a^2(y) \hat{g}(y, x) \hat{E}_{S_A}(y)}{P_x[\hat{\tau}_A < \infty]} + \frac{a^2(y)}{P_x[\hat{\tau}_A < \infty]} \sum_{z \in \partial A} P_y[\hat{\tau}_A^z < \infty, \hat{S}_{\tau_A} = z] (\hat{g}(y, x) - \hat{g}(z, x)) \]  
(4.49)

Next, it is straightforward to obtain that

\[ \frac{1}{\hat{g}(x, y)} = O\left( \ln(1 + \|x\| \vee \|y\|) \right) \]  
(4.50)

(see Exercise 4.12), So, by (4.36) it holds that

\[ \frac{a^2(y) \hat{g}(y, x) \hat{E}_{S_A}(y)}{P_x[\hat{\tau}_A < \infty]} = \hat{h}_A(y) \left( 1 + O\left( \frac{\text{diam}(A)}{\text{dist}(x, A)} \right) \right). \]
(4.51)

So, it remains to show that the second term in (4.49) is

\[ O\left( \frac{\text{diam}(A)}{\text{dist}(x, A)} \right). \]
Similarly to the proof of Theorem 3.8 and the solution of Exercise 3.22, we are going to use the Optional Stopping Theorem with the martingale \( \tilde{M}_{n \wedge \tau} \), where

\[
\tilde{M}_n = \hat{g}(y, x) - \hat{g}(\tilde{S}_{n \wedge \tau}, x)
\]

to estimate the second term in (4.49). Recall that \( y \in \partial A \), and let us define

\[
V = \partial B(y, 2 \text{diam}(A))
\]

(the reader is advised to look again at Figure 3.6 on page 64). Let \( \tau = \hat{\tau}_A^+ \wedge \hat{\tau}_V \). We have (note that \( \tau < \hat{\tau}_x \))

\[
0 = E_y \tilde{M}_0 = E_y \tilde{M}_\tau
= E_y (\tilde{M}_{\hat{\tau}_A} 1\{\hat{\tau}_A < \hat{\tau}_V\}) + E_y (\tilde{M}_{\hat{\tau}_V} 1\{\hat{\tau}_V < \hat{\tau}_A^+\})
\]

(since \( 1\{\hat{\tau}_A < \infty\} = 1\{\hat{\tau}_A < \hat{\tau}_V\} + 1\{\hat{\tau}_V < \hat{\tau}_A^+ < \infty\})

\[
= E_y (\tilde{M}_{\hat{\tau}_A} 1\{\hat{\tau}_A < \infty\}) - E_y (\tilde{M}_{\hat{\tau}_A} 1\{\hat{\tau}_V < \hat{\tau}_A^+ < \infty\})
+ E_y (\tilde{M}_{\hat{\tau}_V} 1\{\hat{\tau}_V < \hat{\tau}_A^+\}).
\]

Note that for any \( z \in V \cup \partial A \) it holds that \( \|y - z\| \leq 2 \text{diam}(A) \) and \( \|x - z\| \geq 10(\text{diam}(A) + 1) \), so in the following we will be able to apply Lemma 4.13. Since

\[
E_y (\tilde{M}_{\hat{\tau}_A} 1\{\hat{\tau}_A < \infty\}) = \sum_{z \in \partial A} P_y [\hat{\tau}_A < \infty, \tilde{S}_{\hat{\tau}_A} = z] (\hat{g}(y, x) - \hat{g}(z, x)),
\]

we obtain that

\[
\sum_{z \in \partial A} P_y [\hat{\tau}_A < \infty, \tilde{S}_{\hat{\tau}_A} = z] (\hat{g}(y, x) - \hat{g}(z, x))
= E_y (\tilde{M}_{\hat{\tau}_A} 1\{\hat{\tau}_V < \hat{\tau}_A^+ < \infty\}) - E_y (\tilde{M}_{\hat{\tau}_V} 1\{\hat{\tau}_V < \hat{\tau}_A^+\})
= P_y [\hat{\tau}_V < \hat{\tau}_A^+] \left( E_y (\tilde{M}_{\hat{\tau}_A} 1\{\hat{\tau}_V < \hat{\tau}_A^+ < \infty\} \mid \hat{\tau}_V < \hat{\tau}_A^+) - E_y (\tilde{M}_{\hat{\tau}_V} \mid \hat{\tau}_V < \hat{\tau}_A^+) \right)
\]

(recall Lemma 4.13 and (4.36); note that \( \tilde{M}_\tau = \hat{g}(y, x) - \hat{g}(z, x) \) on \( \{\tilde{S}_\tau = z\} \)

\[
\leq P_y [\hat{\tau}_V < \hat{\tau}_A^+] \times O\left( \frac{1}{\text{diam}(A)} \ln(1 + \|y\| + \text{diam}(A)) \frac{1}{\ln(1 + \|x\| + \|y\| + \text{diam}(A))} \right).
\]

Next, we can write

\[
\hat{E} s_A (y) = P_y [\hat{\tau}_A = \infty]
\]
4.5 Range of the conditioned SRW

\[ \sum_{v \in V} P_y[\tau_V < \tau_A^+, S_{\tau_v} = v]P_x[\tau_A = \infty] \]

(by Lemma 4.15)

\[ \geq \frac{c}{\ln(\|y\| + \text{diam}(A))} P_y[\tau_V < \tau_A^+], \]

which means that

\[ P_y[\tau_V < \tau_A^+] \leq O(\tilde{E}_{\text{SA}}(y) \ln(\|y\| + \text{diam}(A))). \] (4.53)

Also, (4.51) implies that

\[ \frac{a^2(y)}{P_x[\tau_A < \infty]} = O(\frac{\text{hm}_A(y)}{\text{Es}_A(y)g(x,y)}). \] (4.54)

It only remains to combine (4.52)–(4.54) with (4.50) to see that the second term in (4.49) is indeed \( O(\text{hm}_A(y)\text{diam}(A)) \), thus concluding the proof of Theorem 4.16.

In the remaining part of this chapter, we are going to discuss some even more surprising properties of the conditioned walk \( \hat{S} \).

4.5 Range of the conditioned SRW

For a set \( T \subset \mathbb{Z}_+ \) (thought of as a set of time moments) let

\[ \hat{S}_T = \bigcup_{m \in T} \{\hat{S}_m\} \]

be the range of the walk \( \hat{S} \) with respect to that set, i.e., it is made of sites that are visited by \( \hat{S} \) over \( T \). For simplicity, we assume in the following that the walk \( \hat{S} \) starts at a fixed neighbour \( x_0 \) of the origin, and we write \( P \) for \( P_{x_0} \). For a nonempty and finite set \( A \subset \mathbb{Z}^2 \), let us consider the random variable

\[ \mathcal{R}(A) = \frac{|A \cap \hat{S}_{[0,\infty]}|}{|A|}; \]

that is, \( \mathcal{R}(A) \) is the proportion of visited sites of \( A \) by the walk \( \hat{S} \) (and, therefore, \( 1 - \mathcal{R}(A) \) is the proportion of unvisited sites of \( A \)). In this section we are interested in the following question: how \( \mathcal{R}(A) \) should behave for “large” sets? By (4.21), in average approximately half of \( A \) should be covered, i.e., \( \mathbb{E}\mathcal{R}(A) \) should be
close to 1/2. Now, keeping in mind how difficult is to find something really new in these modern times, let us think what are the usual examples of random variables which are concentrated on $[0, 1]$ and have expected value $1/2$. Well, three examples come to mind: Uniform$[0, 1]$, Bernoulli$(1/2)$, and $1/2$ itself. Which one of them shall the walk $\hat{S}$ choose? It turns out that it is the first one which is the most relevant for the conditioned walk\textsuperscript{11}. Indeed, the main result of this section is the following surprising fact: the proportion is visited sites of a “typical” large set (e.g., a disk, a rectangle, a segment) is a random variable which is close in distribution to the Uniform$[0, 1]$ law. The paper [35] contains the corresponding results in greater generality, but here we content ourselves in proving the result for a particular case of a large disk which does not “touch” the origin:

**Theorem 4.17** Let $D \subset \mathbb{R}^2$ be a closed disk such that $0 \notin D$, and denote $D_n = nD \cap \mathbb{Z}^2$. Then, for all $s \in [0, 1]$, we have, with positive constant $c_1$ depending only on $D$,

$$|\mathbb{P}[\mathcal{R}(D_n) \leq s] - s| \leq c_1 \left(\frac{\ln \ln n}{\ln n}\right)^{1/3}. \tag{4.55}$$

Also, we prove that the range of $\hat{S}$ contains many “big holes”.

**Theorem 4.18** Let $D$ and $D_n$ be as in Theorem 4.17. Then,

$$\mathbb{P}[D_n \cap \hat{S}_{[0, \infty)} = \emptyset \text{ for infinitely many } n] = 1. \tag{4.56}$$

Theorem 4.18 invites the following

**Remark 4.19** A natural question to ask is whether there are also large disks that are completely filled, that is, if a.s. there are infinitely many $n$ such that $D_n \subset \hat{S}_{[0, \infty)}$. It is not difficult to see that the answer to this question is “no”. We, however, postpone explaining that to Chapter 6 (see Exercise 6.1).

We now obtain some refined bounds on the hitting probabilities for excursions of the conditioned walk.

**Lemma 4.20** Assume that $x, y \in \mathbb{Z}^2 \setminus \{0\}$ with $x \neq y$, and $R > \max\{\|x\|, \|y\|\} + 1$. Then, we have

$$\mathbb{P}_x[\hat{\tau}_y < \hat{\tau}_{\partial B(R)}]$$

\textsuperscript{11} although there are examples when the second and the third options also reveal themselves; see Exercises 4.16 and 4.17
Figure 4.7 Excursions that may visit $D_n$

\[
\frac{a(R)(a(x) + a(y) - a(x - y)) - a(x)a(y)\left(1 + O\left(\frac{\|y\|}{R\ln(\|y\|+1)}\right)\right)}{a(x)(2a(R) - a(y)\left(1 + O\left(\frac{\|y\|}{R\ln(\|y\|+1)}\right)\right))}.
\]

(4.57)

Proof Once again in this book, we use the Optional Stopping Theorem, but this time with a new martingale, namely, $M_n = \hat{\ell}(\hat{S}_n \wedge \hat{\tau}_y, y)$ (recall Proposition 4.10). We have (recall (4.25))

\[
\hat{\ell}(z, y) = \frac{1}{a(z)}(a(y) + a(z) - a(z - y))
\]

\[
= \frac{1}{a(R)}\left(1 + O\left(\frac{1}{R\ln R}\right)\right) \times (a(y) + O\left(\frac{\|y\|}{R}\right))
\]

\[
= \frac{a(y)}{a(R)}\left(1 + O\left(\frac{\|y\|}{R\ln(\|y\|+1)}\right)\right)
\]

(4.58)

for any $z \in \partial B(R)$. Let us then abbreviate $p = \mathbb{P}_x[\hat{\tau}_y < \hat{\tau}_{\partial B(R)}]$ and write, using the above-mentioned theorem with stopping time $\hat{\tau}_y \wedge \hat{\tau}_{\partial B(R)}$

\[
E_x M_0 = \frac{1}{a(x)}(a(x) + a(y) - a(x - y))
\]

(note that $\hat{\ell}(y, y) = 2$)

\[
= 2p + (1 - p)E_x(\hat{\ell}(\hat{S}_{\hat{\tau}_{\partial B(R)}}, y) \mid \hat{\tau}_{\partial B(R)} < \hat{\tau}_y)
\]
Conditioned SRW

(by (4.58))

\[ = 2p + (1 - p) \frac{a(y)}{a(R)} (1 + O\left(\frac{\|y\|}{R \ln(\|y\|+1)}\right)). \]

Solving the above for \( p \), we obtain (4.57).

Let us assume that \( x \in B(n \ln n), y \in B(n) \), and abbreviate \( R = n \ln^2 n \). A quick calculation using Lemma 4.20 then shows that

\[
\mathbb{P}_x[y < \hat{\tau}_{\partial B(R)}] = \frac{a(R)(a(x) + a(y) - a(x - y)) - a(x)a(y)(1 + O(\ln^{-3} n))}{a(x)(2a(R) - a(y)(1 + O(\ln^{-3} n)))}.
\]

(4.59)

We proceed with

Proof of Theorem 4.17  First, we describe informally the idea of the proof. We consider the visits to the set \( D_n \) during excursions of the walk from \( \partial B(n \ln n) \) to \( \partial B(n \ln^2 n) \), see Figure 4.7. The crucial argument is the following: the randomness of \( \mathcal{R}(D_n) \) comes from the number \( Q \) of those excursions and not from the excursions themselves. If the number of excursions is approximately \( c \times \ln n \ln \ln n \), then it is possible to show (using a standard weak-LLN argument) that the proportion of covered sites in \( D_n \) is concentrated around \( 1 - e^{-c} \). On the other hand, that number of excursions can be modeled roughly as \( Y \times \ln n \ln \ln n \), where \( Y \) is an Exponential random variable with rate 1. Then,

\[ \mathbb{P}[\mathcal{R}(D_n) \geq 1 - s] \approx \mathbb{P}[Y \geq \ln s^{-1}] = s, \]

as required.

In the following, we will assume, for concreteness, that \( B(1/2) \subset D \subset B(1) \) so that \( D_n \subset B(n) \setminus B(n/2) \) for all \( n \); the extension of the proof to the general case is straightforward.

We now give a rigorous argument. Let \( \hat{H} \) be the conditional entrance measure for the (conditioned) walk \( \hat{S} \), i.e.,

\[
\hat{H}_{D_n}(x, y) = \mathbb{P}_x[\hat{S}_{\hat{\tau}_{D_n}} = y \mid \hat{\tau}_{D_n} < \infty].
\]

(4.60)

Let us denote the initial piece of the trajectory by \( \mathcal{E}_0 = \hat{S}_{[0, \hat{\tau}_{\partial B(n \ln n)}]} \). Then, we consider a Markov chain \( \mathcal{E}_k, k \geq 1 \) of excursions between \( \partial B(n \ln n) \) and \( \partial B(n \ln^2 n) \), defined in the following way:
for $k \geq 2$ the initial site of $\mathcal{E}_k$ is chosen according to the measure $\tilde{H}_{\partial B(n \ln n)}(\cdot)$, where $z_{k-1} \in \partial B(n \ln n)$ is the last site of the excursion $\mathcal{E}_{k-1}$; also, the initial site of $\mathcal{E}_1$ is the last site of $\mathcal{E}_0$; the weights of trajectories are chosen according to (4.5) (i.e., each excursion is an $\hat{S}$-walk trajectory). It is important to observe that one may couple $(\mathcal{E}_k, k \geq 1)$ with the “true” excursions of the walk $\hat{S}$ in an obvious way: one just picks the excursions subsequently, each time tossing a coin to decide if the walk returns to $B(n \ln n)$.

Let

$$\psi_n = \min_{x \in \partial B(n \ln^2 n)} P_x[\tilde{\tau}_{\partial B(n \ln n)} = \infty],$$

$$\psi^*_n = \max_{x \in \partial B(n \ln^2 n)} P_x[\tilde{\tau}_{\partial B(n \ln n)} = \infty]$$

be the minimal and maximal probabilities to avoid $B(n \ln n)$, starting at sites of $\partial B(n \ln^2 n)$. Using (4.10) together with (3.36), we obtain

$$P_x[\tilde{\tau}_{\partial B(n \ln n)} = \infty] = 1 - \frac{a(n \ln n) + O((n \ln n)^{-1})}{a(x)}$$

$$= \frac{a(n \ln^2 n) - a(n \ln n) + O((n \ln n)^{-1})}{a(n \ln^2 n) + O((n \ln^2 n)^{-1})}$$

(by (3.36)–(3.37), with $\gamma'' = \frac{\pi^2}{2} = \gamma + \frac{3}{2} \ln 2$)

$$= \frac{\ln \ln n}{\ln n + 2 \ln \ln n + \gamma'' (1 + o(n^{-1}))}$$

(4.61)

for any $x \in \partial B(n \ln^2 n)$, and so it also holds that

$$\psi_n = \frac{\ln \ln n}{\ln n + 2 \ln \ln n + \gamma'' (1 + o(n^{-1}))}.$$  

(4.62)

Let us consider a sequence of i.i.d. random variables $(\eta_k, k \geq 0)$ such that $P[\eta_k = 1] = 1 - P[\eta_k = 0] = \psi_n$. Let $\hat{Q} = \min\{k : \eta_k = 1\}$, so that $\hat{Q}$ is a Geometric random variable with mean $\psi_n^{-1}$.

**Lemma 4.21** The random variable $\hat{Q}$ can be coupled with the actual number of excursions $Q$ in such a way that $Q \leq \hat{Q}$ a.s. and

$$P[Q \neq \hat{Q}] \leq o(n^{-1}).$$

(4.63)
Moreover, this coupling can be constructed in such a way that \( \hat{Q} \) is independent from the excursion sequence \((\mathcal{E}_k, k \geq 1)\) itself.

**Proof** Let \( U_1, U_2, U_3, \ldots \) be a sequence of i.i.d. random variables with Uniform\([0,1]\) distribution; we will now explain how to construct all random variables of interest using this sequence. First, we set

\[
\eta_k = 1\{U_k \leq \psi_n\}.
\]

Next, let \( \tilde{\eta}_k = 1\{Q < k\} \) be the indicator of the event that the walk \( \hat{S} \) does not make its \( k \)th excursion. Given that \( \tilde{\eta}_{k-1} = 0 \) and \( z_{k-1} \) is the last site of \( \mathcal{E}_{k-1} \), set

\[
\tilde{\eta}_k = 1\{U_k \leq \hat{\tau}_{\partial B(n\ln n)} = \infty\}.
\]

It is clear that, with this construction, \( Q \leq \hat{Q} \) and that \( \hat{Q} \) is independent of the excursions themselves. Then, we write

\[
\{Q > \hat{Q}\} = \{\text{there exists } k \leq \hat{Q} - 1 \text{ such that } \psi_n < U_k \leq \psi^*_n\}
\]

so

\[
P[Q > \hat{Q}] \leq P[\text{there exists } k \leq \hat{Q} - 1 \text{ such that } \psi_n < U_k \leq \psi^*_n].
\]

(4.64)

In the following, assume without loss of generality that \( n \) is large enough to that \( \psi_n < 1/2 \). By (4.64), we have

\[
P[Q < k | \hat{Q} = k] \leq P[\text{there exists } j \leq k - 1 \text{ such that } \psi_n < U_k \leq \psi^*_n \text{ for } j = 1, \ldots, k - 1, U_k \leq \varphi_n] \\
\leq 2(\psi^*_n - \psi_n)k.
\]

Now, (4.61) implies that \( \psi^*_n - \psi_n \leq o\left(\frac{\ln \ln n}{n \ln n}\right) \) for any \( x \in \partial B(n \ln^2 n) \), and so (recall that, by (4.62), \( 1/\psi_n = O\left(\frac{\ln n}{n \ln n}\right) \))

\[
P[Q < \hat{Q}] \leq \sum_{k=1}^{\infty} (1 - \varphi_n)^{k-1} \psi_n \times 2(\psi^*_n - \psi_n)k \\
= \frac{2(\psi^*_n - \psi_n)}{\psi_n} = o(n),
\]

which concludes the proof. \( \square \)
4.5 Range of the conditioned SRW

We continue proving Theorem 4.17. Define
\[
R^{(k)} = \frac{|D_n \cap (E_0 \cup E_1 \cup \ldots \cup E_k)|}{|D_n|},
\]
to be the proportion of visited sites in \(D_n\) with respect to the first \(k\) excursions together with the initial piece \(E_0\). According to the programme outlined in the beginning of the proof, we now aim to verify that \(R^{(k)}\) is concentrated around its mean value; we do this using a typical “weak LNN” argument.

Now, it is straightforward to check that (4.59) implies that, for any \(x \in \partial B(n \ln n)\) and \(y \in D_n\)
\[
\mathbb{P}_x \left[ \hat{\tau}_y < \hat{\tau}_{\partial B(n \ln^2 n)} \right] = \ln \ln n \ln n \left(1 + O\left(\frac{\ln \ln n}{\ln n}\right)\right),
\]
and, for \(y, z \in B(n) \setminus B(n/2)\) such that \(\|y - z\| = n/b\) with \(b \leq 2 \ln n\)
\[
\mathbb{P}_z \left[ \hat{\tau}_y < \hat{\tau}_{\partial B(n \ln^3 n)} \right] = \frac{2\ln n + \ln b}{\ln n} \left(1 + O\left(\frac{\ln \ln n}{\ln n}\right)\right),
\]
(we leave checking this to the reader, as Exercise 4.15).

Next, for \(y \in D_n\) and a fixed \(k \geq 1\) consider the random variable
\[
\xi^{(k)}_y = 1\{y \notin E_0 \cup E_1 \cup \ldots \cup E_k\},
\]
so that \(<R^{(k)} = |D_n|^{-1} \sum_{y \in D_n} \xi^{(k)}_y\). Now (4.65) implies that, for all \(j \geq 1\),
\[
\mathbb{P}[y \notin E_j] = 1 - \frac{\ln \ln n}{\ln n} \left(1 + O\left(\frac{\ln \ln n}{\ln n}\right)\right),
\]
and (4.66) implies that
\[
\mathbb{P}[y \notin E_0] = 1 - O\left(\frac{\ln \ln n}{\ln n}\right)
\]
for any \(y \in D_n\). Let \(\mu^{(k)}_y = \mathbb{E} \xi^{(k)}_y\). Then we have
\[
\mu^{(k)}_y = \mathbb{P}[y \notin E_0 \cup E_1 \cup \ldots \cup E_k]
\]
\[
= \left(1 - O\left(\frac{\ln \ln n}{\ln n}\right)\right) \times \left(1 - \frac{\ln \ln n}{\ln n} \left(1 + O\left(\frac{\ln \ln n}{\ln n}\right)\right)\right)^k
\]
\[
= \exp\left(-k \frac{\ln \ln n}{\ln n} \left(1 + O\left(k^{-1} + \frac{\ln \ln n}{\ln n}\right)\right)\right).
\]

\(^{12}\) the calculation is a bit long, though; consistently with the approach of this book it is omitted
Next, we need to estimate the covariance\(^{13}\) of \(\xi^{(k)}_y\) and \(\xi^{(k)}_z\) in case \(\|y - z\| \geq \frac{\ln n}{\ln n}\). First note that, for any \(x \in \partial B(n \ln n)\)

\[
P_x \{y, z\} \cap \mathcal{E}_1 = \emptyset = 1 - P_x[y \in \mathcal{E}_1] - P_x[z \in \mathcal{E}_1] + P_x \{y, z\} \subset \mathcal{E}_1
\]

\[
= 1 - 2 \left( \frac{\ln \ln n}{\ln n} \right) \left( 1 + O\left( \frac{\ln \ln n}{\ln n} \right) \right) + P_x \{y, z\} \subset \mathcal{E}_1
\]

by (4.65); also, since

\[
\{ \tau_y < \tau_z < \tau_{\partial B(n \ln^2 n)} \} \subset \{ \tau_y < \tau_{\partial B(n \ln^2 n)} \},
\]

\(\hat{S}_k = z\) for some \(\hat{y} < k < \hat{\tau}_{\partial B(n \ln^2 n)}\),

from (4.65)–(4.66) we obtain

\[
P_x \{y, z\} \subset \mathcal{E}_1 = P_x \left[\max \{\tau_y, \tau_z\} < \tau_{\partial B(n \ln^2 n)}\right]
\]

\[
\leq P_x \left[\tau_y < \tau_{\partial B(n \ln^2 n)}\right] P_y \left[\tau_z < \tau_{\partial B(n \ln^2 n)}\right]
\]

\[
+ P_x \left[\tau_z < \tau_{\partial B(n \ln^2 n)}\right] P_z \left[\tau_y < \tau_{\partial B(n \ln^2 n)}\right]
\]

\[
\leq 2 \left( \frac{\ln \ln n}{\ln n} \right) \times 3 \ln \ln n \left( 1 + O\left( \frac{\ln \ln n}{\ln n} \right) \right)
\]

\[
= O\left( \left( \frac{\ln \ln n}{\ln n} \right)^2 \right).
\]

Therefore, similarly to (4.67) we obtain

\[
E(\xi^{(k)}_y \xi^{(k)}_z) = \exp \left( -2k \frac{\ln \ln n}{\ln n} \left( 1 + O\left( \frac{\ln \ln n}{\ln n} \right) \right) \right),
\]

which, together with (4.67), implies after some elementary calculations that, for all \(y, z \in D_n\) such that \(\|y - z\| \geq \frac{\ln n}{\ln n}\)

\[
\text{cov}(\xi^{(k)}_y, \xi^{(k)}_z) = O\left( \frac{\ln \ln n}{\ln n} \right)
\]

(4.68)

uniformly in \(k\), since

\[
\left( \frac{\ln \ln n}{\ln n} + k \left( \frac{\ln \ln n}{\ln n} \right)^2 \right) \exp \left( -2k \frac{\ln \ln n}{\ln n} \right) = O\left( \frac{\ln \ln n}{\ln n} \right)
\]

uniformly in \(k\). At this point, we are ready to write that weak-LLN argument. Using Chebyshev’s inequality, we write

\[
P\left[ \left| D_n \right|^{-1} \sum_{y \in D_n} (\xi^{(k)}_y - \mu^{(k)}_y) \right] > \varepsilon
\]

\(^{13}\) recall that we want a weak-LLN-type argument, that is, we need to estimate

\[
\text{Var} \sum_{y \in D_n} \xi^{(k)}_y ; \text{so, we will need these covariances indeed}
\]
\[ 4.5 \text{ Range of the conditioned SRW} \]

\[ \leq (\varepsilon|D_n|)^{-2} \text{Var} \left( \sum_{y \in D_n} \xi_y^{(k)} \right) \]

\[ = (\varepsilon|D_n|)^{-2} \sum_{y,z \in D_n} \text{cov}(\xi_y^{(k)}, \xi_z^{(k)}) \]

\[ = (\varepsilon|D_n|)^{-2} \left( \sum_{y \in D_n, \|y-z\| \leq \frac{n}{\ln n}} \text{cov}(\xi_y^{(k)}, \xi_z^{(k)}) + \sum_{y \in D_n, \|y-z\| \geq \frac{n}{\ln n}} \text{cov}(\xi_y^{(k)}, \xi_z^{(k)}) \right) \]

\[ \leq (\varepsilon|D_n|)^{-2} \left( \sum_{y \in D_n} |D_n \cap B(y, \frac{n}{\ln n})| + |D_n|^2 O(\frac{\ln \ln n}{\ln n}) \right) \]

\[ \leq \varepsilon^{-2} O(\ln^{-2} n + \frac{\ln \ln n}{\ln n}) = \varepsilon^{-2} O\left(\frac{\ln \ln n}{\ln n}\right). \] (4.69)

Now, having established that \( R^{(k)} \) is “almost deterministic”, we recall that the number of excursions is “almost geometrically distributed”, and work with this fact. Let

\[ \Phi^{(s)} = \min \{ k : R^{(k)} \geq 1 - s \} \]

be the number of excursions necessary to make the unvisited proportion of \( D_n \) at most \( s \). We have

\[ \mathbb{P}[R(D_n) \geq 1 - s] \]

\[ = \mathbb{P}[\Phi^{(s)} \leq Q] \]

\[ = \mathbb{P}[\Phi^{(s)} \leq Q, Q = \tilde{Q}] + \mathbb{P}[\Phi^{(s)} \leq Q, Q \neq \tilde{Q}] \]

\[ = \mathbb{P}[\Phi^{(s)} \leq \tilde{Q}] + \mathbb{P}[\Phi^{(s)} \leq Q, Q \neq \tilde{Q}] - \mathbb{P}[\Phi^{(s)} \leq \tilde{Q}, Q \neq \tilde{Q}], \]

so, recalling (4.63),

\[ |\mathbb{P}[R(D_n) \geq 1 - s] - \mathbb{P}[\Phi^{(s)} \leq \tilde{Q}]| \leq \mathbb{P}[Q \neq \tilde{Q}] \leq O(n^{-1}). \] (4.70)

Next, by the “independence” part of Lemma 4.21, we write

\[ \mathbb{P}[\Phi^{(s)} \leq \tilde{Q}] = \mathbb{E}(\mathbb{P}[\tilde{Q} \geq \Phi^{(s)} | \Phi^{(s)}]) \]

\[ = \mathbb{E}(1 - \psi_n)^{\Phi^{(s)}}, \] (4.71)

and focus on obtaining lower and upper bounds on the expectation in the right-hand side of (4.71). The idea is that, since \( R^{(k)} \) is concentrated, \( \Phi^{(s)} \) has to be concentrated as well; however, since it is in the exponent, some extra care needs to be taken. Let us assume that \( s \in (0, 1) \) is fixed and abbreviate

\[ \delta_n = \left(\frac{\ln \ln n}{\ln n}\right)^{1/3} \]
\[ k_n^- = \left\lfloor (1 - \delta_n) \ln s^{-1} \frac{\ln n}{\ln \ln n} \right\rfloor, \]
\[ k_n^+ = \left\lceil (1 + \delta_n) \ln s^{-1} \frac{\ln n}{\ln \ln n} \right\rceil; \]
we also assume that \( n \) is sufficiently large so that \( \delta_n \in (0, \frac{1}{2}) \) and \( 1 < k_n^- < k_n^+ \). Now, according to (4.67),
\[
\mu_y^{(k_n^\pm)} = \exp \left( - (1 \pm \delta_n) \ln s^{-1} \left( 1 + O((k_n^\pm)^{-1} + \frac{\ln \ln n}{\ln n}) \right) \right)
\]
\[
= s \exp \left( - \ln s^{-1} \left( \pm \delta_n + O((k_n^\pm)^{-1} + \frac{\ln \ln n}{\ln n}) \right) \right)
\]
\[
= s \left( 1 + O(\delta_n \ln s^{-1} + \frac{\ln \ln n}{\ln n}(1 + \ln s^{-1})) \right),
\]
so in both cases it holds that (observe that \( s \ln s^{-1} \leq 1/e \) for all \( s \in [0,1] \))
\[
\mu_y^{(k_n^\pm)} = s + O(\delta_n + \frac{\ln \ln n}{\ln n}) = s + O(\delta_n). \quad (4.72)
\]

With a similar calculation, one can also obtain that
\[
(1 - \psi_n)^{(k_n^\pm)} = s + O(\delta_n). \quad (4.73)
\]

We then write, using (4.72)
\[
P[\Phi(s) > k_n^+] = P[\mathcal{R}^{(k_n^+)} < 1 - s]
\]
\[
= P\left[ D_n^{-1} \sum_{y \in D_n} (\xi_y^{(k_n^+)} > s) \right]
\]
\[
= P\left[ D_n^{-1} \sum_{y \in D_n} (\xi_y^{(k_n^+)} - \mu_y^{(k_n^+)} > s - |D_n|^{-1} \sum_{y \in D_n} \mu_y^{(k_n^+)} ) \right]
\]
\[
= P\left[ |D_n|^{-1} \sum_{y \in D_n} (\xi_y^{(k_n^+)} - \mu_y^{(k_n^+)} ) > O(\delta_n) \right]. \quad (4.74)
\]

Then, (4.69) implies that
\[
P[\Phi(s) > k_n^+] \leq O\left( \left( \frac{\ln \ln n}{\ln n} \right)^{1/3} \right). \quad (4.75)
\]

Quite analogously, one can also obtain that
\[
P[\Phi(s) < k_n^-] \leq O\left( \left( \frac{\ln \ln n}{\ln n} \right)^{1/3} \right). \quad (4.76)
\]

Using (4.73) and (4.75), we then write
\[
\mathbb{E}(1 - \psi_n)^{\Phi(s)} \geq \mathbb{E}\left( (1 - \psi_n)^{\Phi(s)} I\{\Phi(s) \leq k_n^+\} \right)
\]
4.5 Range of the conditioned SRW

\[
\begin{align*}
\geq (1 - \psi_n) k_n^+ \mathbb{P}[\Phi(s) \leq k_n^+] \\
\geq (s - O((\frac{\ln \ln n}{\ln n})^{1/3})) (1 - O((\frac{\ln \ln n}{\ln n})^{1/3})), \quad (4.77)
\end{align*}
\]

and, using (4.73) and (4.76),

\[
\begin{align*}
\mathbb{E}(1 - \psi_n) \Phi(s) &= \mathbb{E}((1 - \psi_n) \Phi(s) \mathbb{1}\{\Phi(s) \geq k_n^+\}) \\
&\quad + \mathbb{E}((1 - \psi_n) \Phi(s) \mathbb{1}\{\Phi(s) < k_n^+\}) \\
&\leq (1 - \psi_n) k_n^+ \mathbb{P}[\Phi(s) < k_n^+] \\
&\leq \left(s + O\left((\frac{\ln \ln n}{\ln n})^{1/3}\right)\right) + O\left((\frac{\ln \ln n}{\ln n})^{1/3}\right). \quad (4.78)
\end{align*}
\]

Therefore, recalling (4.70)–(4.71), we obtain (4.55); this concludes the proof of Theorem 4.17.

Now, we prove that there are “big holes” in the range of \(\hat{S}\):

**Proof of Theorem 4.18** For the sake of simplicity, let us still assume that \(D \subset B(1) \setminus B(1/2)\); the general case can be treated in a completely analogous way.

Consider two sequences of events

\[
\begin{align*}
E_n &= \{\tau_{D,2n-1} > \tau_{\partial B(2^n)}, \|\hat{S}_j\| > 2^{2n-1} \text{ for all } j \geq \tau_{\partial B(2^n)}\}, \\
E'_n &= \{\|\hat{S}_j\| > 2^{2n-1} \text{ for all } j \geq \tau_{\partial B(2^n)}\}
\end{align*}
\]

and note that \(E_n \subset E'_n\) and \(2^{2n-1} D \cap \hat{S}_{[0,\infty)} = \emptyset\) on \(E_n\). Our goal is to show that, almost surely, an infinite number of events \((E_n, n \geq 1)\) occurs. Observe, however, that the events in each of the above two sequences are not independent, so the “basic” second Borel-Cantelli lemma will not work.

In the following, we use a generalization of the second Borel-Cantelli lemma, known as the Kochen-Stone theorem [41]: it holds that

\[
\mathbb{P}\left[\sum_{k=1}^{\infty} \mathbb{1}\{E_k\} = \infty\right] \geq \limsup_{k \to \infty} \frac{\left(\sum_{i=1}^{k} \mathbb{P}[E_i]\right)^2}{\sum_{i,j=1}^{k} \mathbb{P}[E_i \cap E_j]} \quad (4.79)
\]

We will now prove that there exists a positive constant \(c_4\) such that

\[
\mathbb{P}[E'_n] \geq \frac{c_4}{n} \quad \text{for all } n \geq 1. \quad (4.80)
\]
It is elementary to obtain\textsuperscript{14} that, for some \( c_5 > 0 \),
\[
\mathbb{P}_x \left[ \tilde{\tau}_{D;2^{3n-1}} > \tilde{\tau}_{\partial B(2^{3n})}, \tau_{\partial B(2^{3n})}^+ > \tau_{\partial B(2^{3n})} \right] > c_5
\]
for all \( x \in \partial B(2^{3(n-1)}) \). Lemma 4.4 then implies that, for some \( c_6 > 0 \),
\[
\mathbb{P}_x \left[ \tilde{\tau}_{D;2^{3n-1}} > \tilde{\tau}_{\partial B(2^{3n})} \right] = (1 + o(2^{-3n})) \mathbb{P}_x \left[ \tau_{D;2^{3n-1}} > \tau_{\partial B(2^{3n})} \mid \tau_{\partial B(2^{3n})}^+ > \tau_{\partial B(2^{3n})} \right]
\]
\[
= (1 + o(2^{-3n})) \mathbb{P}_x \left[ \tau_{D;2^{3n-1}} > \tau_{\partial B(2^{3n})}, \tau_{\partial B(2^{3n})}^+ > \tau_{\partial B(2^{3n})} \right] > c_6
\]
for all \( x \in \partial B(2^{3(n-1)}) \). Let us denote, recalling (3.36), \( \gamma^* = \frac{\pi}{2} \times \frac{1}{\ln 2} \times \gamma' = \frac{2n+3\ln 2}{2n} \). Using (4.10), we then obtain
\[
\mathbb{P}_z \left[ ||S_j|| > 2^{3n-1} \right. \text{ for all } j \geq 0] = 1 - \frac{a(2^{3n-1}) + O(2^{-3n})}{a(2^{3n}) + O(2^{-3n})}
\]
\[
= 1 - \frac{\frac{\pi}{2}(3n - 1) \ln 2 + \gamma' + O(2^{-3n})}{\frac{\pi}{2}(3n) \ln 2 + \gamma' + O(2^{-3n})}
\]
\[
= \frac{1}{3n + \gamma^*} \left( 1 + o(2^{-3n}) \right)
\]
(4.82)
for any \( z \in \partial B(2^{3n}) \). The inequality (4.80) follows from (4.81) and (4.82).

Now, we need an upper bound for \( \mathbb{P}[E_m \cap E_n], m \leq n \). Clearly, \( E'_m \cap E'_n \subset E'_m \cap E'_n \), and note that the event \( E'_m \cap E'_n \) means that the particle hits \( \partial B(2^{3n}) \) before \( \partial B(2^{3m-1}) \) starting from a site on \( \partial B(2^{3m}) \), and then never hits \( \partial B(2^{3n-1}) \) starting from a site on \( \partial B(2^{3n}) \). So, again using (4.10) and Lemma 4.4, we write analogously to (4.82) (and also omitting a couple of lines of elementary calculations)
\[
\mathbb{P}[E_m \cap E_n] \leq \mathbb{P}[E'_m \cap E'_n]
\]
\[
= \frac{(a(2^{3m-1}))^{-1} - (a(2^{3m}))^{-1} + O(2^{-3m})}{(a(2^{3m-1}))^{-1} - (a(2^{3m}))^{-1} + O(2^{-3m})}
\times \left( 1 - \frac{a(2^{3n-1}) + O(2^{-3n})}{a(2^{3n}) + O(2^{-3n})} \right)
\]
\[
= \frac{1}{(3(n - m) + 1)(3m + \gamma^*)} \left( 1 + o(2^{-3m}) \right)
\]
\textsuperscript{14} e.g., by comparison/coupling to Brownian Motion, or directly, in the spirit of Exercise 4.11.
Therefore, we can write
\[
\sum_{m,n=1}^{k} \mathbb{P}[E_i \cap E_j] \leq 2 \sum_{1 \leq m \leq n \leq k} \mathbb{P}[E_i \cap E_j] \\
\leq 2c_9 \sum_{1 \leq m \leq n \leq k} \frac{1}{m(n - m + 1)} \\
= 2c_9 \sum_{m=1}^{k} \frac{1}{m} \sum_{n=m}^{k} \frac{1}{n - m + 1} \\
\leq 2c_9 \sum_{m=1}^{k} \frac{1}{m} \sum_{n=m}^{k-m+1} \frac{1}{n - m + 1} \\
= 2c_9 \left( \sum_{m=1}^{k} \frac{1}{m} \right)^2,
\]
and so \( \sum_{i,j=1}^{k} \mathbb{P}[E_i \cap E_j] \leq c_{10} \ln^2 k \). Now, (4.80) implies that \( \sum_{i=1}^{k} \mathbb{P}[E_i] \geq c_{11} \ln k \), so, using (4.79), we obtain that
\[
\mathbb{P} \left[ \sum_{k=1}^{\infty} 1\{E_k\} = \infty \right] \geq c_{12} > 0.
\]
To obtain that the probability in the above display must be equal to 1, we need a suitable 0-1 law. Conveniently enough, it is provided by Proposition 3.8 in Chapter 2 of [62]: if every set is either recurrent or transient with respect to \( \hat{S} \), then every tail event must have probability 0 or 1. Now, note that Theorem 4.11 implies that every set must be recurrent or transient indeed. This concludes the proof of Theorem 4.18.

### 4.6 Exercises

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

**Exercise 4.2** Calculate the Green’s function of the conditioned one-dimensional SRW.
Exercise 4.3  Consider a nearest neighbour random walk on $\mathbb{Z}_+$ with drift towards the origin, $p(n, n+1) = 1 - p(n, n-1) = p < \frac{1}{2}$. What will be its conditioned (on never hitting the origin) version?

Exercise 4.4  Now, consider a one-dimensional nearest neighbour Lamperti random walk (recall Exercise 2.19): $p(n, n+1) = 1 - p(n, n-1) = \frac{1}{2} + \frac{c}{n}$ with $c < \frac{1}{4}$, so that the random walk is recurrent. What can you say about its conditioned version?

Exercise 4.5  Assume that the original Markov chain is reversible; must its $h$-transform be reversible as well?

Exercise 4.6  For any finite $A \subset \mathbb{Z}^2$, find a (nontrivial) nonnegative function which is zero on $A$ and harmonic outside $A$.

Exercise 4.7  Prove that
\[ P_x[\hat{\tau}_N < \infty] = \frac{1}{a(x)} \] (4.84)
for any $x \in \mathbb{Z}^2 \setminus \{0\}$.

Exercise 4.8  Can you find an expression for the Green’s function of the $h$-transform in the general case (as in Section 4.1), i.e., an analogue of (4.12)? What else should we assume for that?

Exercise 4.9  Do we actually need recurrence in Lemma 4.9? prosledit chtob byl odin i tot zhe nomer

Exercise 4.10  Prove that the conditioned walk $\hat{S}$ is uniformly elliptic, i.e., there exists $c > 0$ such that $P_x[S_1 = y] > c$ for all $x, y \in \mathbb{Z}^2 \setminus \{0\}$ such that $x \sim y$.

Exercise 4.11  Give at least a couple of different proofs of (4.39) which do not use Lemma 3.11.

Exercise 4.12  Prove that there exist two positive constants $c_1, c_2$ such that, for all $x, y \in \mathbb{Z}^2 \setminus \{0\}$,
\[ \frac{c_1}{\ln(1 + ||x|| \vee ||y||)} \leq \hat{g}(x, y) \leq \frac{c_2}{\ln(1 + ||x|| \vee ||y||)}. \] (4.85)

Exercise 4.13  Prove that the process $(\hat{\ell}(\hat{S}_{n \wedge \hat{\tau}_y}, y), n \geq 0)$ (recall (4.17)) is a martingale by a direct computation (i.e., not using the general fact that Green’s functions give rise to martingales).

Exercise 4.14  In Lemma 4.13, can one substitute the constant 5 by any fixed positive number?
Exercise 4.15 Check (4.65)–(4.66).

Exercise 4.16 Give examples of $A \subset \mathbb{Z}^2$ with $|A_n| \to \infty$ such that

(i) $\mathcal{R}(A_n)$ converges in distribution to Bernoulli($1/2$);
(ii) $\mathcal{R}(A_n)$ converges in probability to $1/2$.

Exercise 4.17 Can we somehow obtain mixtures of the three main limit distributions of $\mathcal{R}(\cdot)$, and how shall we recognise their “domains of attraction”? For example, try to analyse the following case: let $A_n = B(r_n e_1, n)$, where $r_n \to \infty$ in some way.
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 $\sigma$, 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 $\sigma$ in these cases?

After some thought, one will come with the following solution, simple and straightforward: just generate the i.i.d. sequence

\(^1\) 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 $\mathbb{P}[Z \in F] = 1$ (therefore, in particular, the support is a closed set).

\(^2\) even more, the permutation $\sigma$ 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.
Y_1, Y_2, Y_3, \ldots \text{ independently, then there is a permutation that sends } X\text{-sequence to the } Y\text{-sequence. Indeed (this argument works for any pair of discrete random variables with the same support), almost surely any possible value of } X \text{ (and } Y) \text{ occurs infinitely many times both in the } X\text{-sequence and the } Y\text{-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 Σ × R_+.

Naturally, we assume that the reader knows what is a Poisson point process in R^d with (not necessarily constant) rate λ. If one needs to consider a Poisson process on, say, Z × 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 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 δλ(x) (up to terms of smaller order), and the probability that there are at least two points in that small domain is o(δ). Here, the tradition dictates that the author cites a comprehensive book on the subject, so [61].

Coming back to the soft local times method, we mention that, in full generality, it was introduced in [58]; see also [17, 18] which contain short surveys of this method applied to constructions of excursion processes^3. The idea of using projections of Poisson pro-

\[ ^3 \text{ but this will be treated in a detailed way in Section 6.3.1 below.} \]
cesses for constructions of other (point) processes is not new, see e.g. [48, 36]. The key tool of this method (Lemma 5.1) appears in [75] 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 $\Sigma$, which has enough structure\footnote{for example, the following is enough: let $\Sigma$ be a locally compact and Polish metric space, and $\mu$ is a Radon measure (i.e., every compact set has finite $\mu$-measure) on the measurable space $(\Sigma, \mathcal{B})$, where $\mathcal{B}$ is the Borel $\sigma$-algebra on $\Sigma$.} that permits us to construct a Poisson point process on $\Sigma$ of rate $\mu$, where $\mu$ is a measure on $\Sigma$.

Now, the 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, $\Xi$ 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

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

be the set\footnote{endowed with sigma-algebra $\mathcal{D}$ generated by the evaluation maps $\eta \mapsto \eta(A)$, $A \in \mathcal{B} \otimes \mathcal{B}(\mathbb{R})$.} of point configurations of this process. It is a general fact that one can canonically construct a Poisson point process $\eta$ as above; see e.g. Proposition 3.6 on p.130 of [61] for details of this construction.

The result below is our “magic lemma”: it provides us with a way to simulate a random element of $\Sigma$ with law absolutely continuous with respect to $\mu$, using the Poisson point process $\eta$. We first write it formally, and then explain, what does it mean.

**Lemma 5.1** Let $g : \Sigma \rightarrow \mathbb{R}_+$ 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

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

Then, under the law $\mathbb{Q}$ of the Poisson point process $\eta$,
5.1 Soft local times

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

1. there exists a.s. a unique $\hat{\vartheta} \in \Xi$ such that $\xi g(z_{\hat{\vartheta}}) = v_{\hat{\vartheta}}$.
2. $(z_{\hat{\vartheta}}, \xi)$ is distributed as $g(z)\mu(dz) \otimes \text{Exp}(1)$.
3. $\eta' := \sum_{\vartheta \neq \hat{\vartheta}} \delta_{(z_{\vartheta}, v_{\vartheta} - \xi g(z_{\vartheta}))}$ has the same law as $\eta$ and is independent of $(\xi, \vartheta)$.

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

- In (5.2) we define $\xi$ as the smallest positive number such that there is exactly one point $(z_{\hat{\vartheta}}, v_{\hat{\vartheta}})$ 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 $\mu$). Also, $\xi$ 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 $\eta$ down by amount $\xi g(z)$. Call this new configuration $\eta'$. Then, $\eta'$ is also a Poisson point process on $\Sigma \times \mathbb{R}_+$ with rate $\mu \otimes dw$, and it is independent of $\xi$ and $Z$.

Sketch of the proof of Lemma 5.1. The formal proof can be found in [58] (Lemma 5.1 is Proposition 4.1 of [58]), 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 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 is valid 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 $\mu$, and therefore the area\(^6\) below $\varepsilon g$ equals $\varepsilon$. So, the number of points below $\varepsilon g$ is Poisson with rate $\varepsilon$, 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 $N_\varepsilon$ 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

\(^6\) with respect to $\mu \otimes dv$
5.1 Soft local times

with parameter 1 as \( \varepsilon \to 0 \) (note that \( 1 - e^{-\varepsilon} = \varepsilon + o(\varepsilon) \) as \( \varepsilon \to 0 \)). Therefore, \( \xi \) should indeed be Exponential(1).

The above fact could have been established in a direct way (note that \( Q[\xi > t] \) 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 \( 1 - \exp \left( -\varepsilon \int_R g(z) \mu(dz) \right) = \varepsilon \int_R g(z) \mu(dz) + o(\varepsilon) \).

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 \( \xi \) 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 \( (\xi, Z) \). 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 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 \( F_{k-1} \) is a.s. absolutely continuous with respect to the Lebesgue measure on \([0, 1] \), where \( F_{k-1} \) is the sigma-algebra generated by \( X_1, \ldots, X_{k-1} \).

This idea of using the method of soft local times to couple (possi-

\(^7\) that is, we effectively condition on \( \xi \), and show that the conditional law of \( Z \) does not depend on the value of \( \xi \).
bly 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.

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 coordinates, and then taking $Y_1, Y_2, Y_3, \ldots$ to be the first coordinates 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 $\xi_1$ 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 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 $(\xi_1, X_1)$.
- Thus, we are ready to use Lemma 5.1 again in order to construct $X_2$.
- So, consider the conditional density $g(\cdot | F_1)$ of $X_2$ given $F_1$ and find the smallest positive number $\xi_2$ in such a way that exactly one point lies on the graph of $\xi_2 g(\cdot | F_1) + \xi_1 g(\cdot)$ 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 $g(\cdot | F_2)$, 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 [58]): here it is only a bit more general since we formulate it for general adapted processes.

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

$$\xi_1 = \inf \left\{ t \geq 0 : \text{there exists } g \in \Xi \text{ such that } tg(z_\varrho) \geq v_\varrho \right\},$$

$$G_1(z) = \xi_1 g(z | \mathcal{F}_0), \text{ for } z \in \Sigma,$$

\footnote{this point has coordinates $(X_1, \xi_1 g(X_1))$}
5.1 Soft local times

\[ 5.1 \quad \text{Soft local times} \]

\[ \xi_1 g(\cdot) + \xi_2 g(\cdot | X_1) \]

\[ \xi_1 g(\cdot) \]

\[ \Sigma \]

\[ X_2 \quad X_4 = Y_3 \quad X_5 \quad X_3 = Y_4 \quad X_1 \]

\[ \mathbb{R}_+ \]

**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.

where \( g(\cdot | F_0) \) is the density of \( Z_1 \) given \( F_0 \), and

\[ (z_1, v_1) \]

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

Denote also \( R_1 = \{(z_1, v_1)\} \). Then, for \( n \geq 2 \) we proceed inductively,

\[ \xi_n = \inf \{ t \geq 0 : \text{there exists } (z_\varrho, v_\varrho) \notin R_{n-1} \text{ such that } G_{n-1}(z_\varrho) + t g(z_\varrho | F_{n-1}) \geq v_\varrho \} \]

\[ G_n(z) = G_{n-1}(z) + \xi_n g(z | F_{n-1}) \]

(5.3)

and

\[ (z_n, v_n) \]

is the unique pair \( (z_\varrho, v_\varrho) \notin R_{n-1} \) 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) \( (z_1, \ldots, z_n) \overset{\text{law}}{=} (Z_1, \ldots, Z_n) \) and they are independent from \( \xi_1, \ldots, \xi_n \),
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.

(ii) the point process

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

is distributed as $\eta$ and independent of the above, 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 $\mu$. 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

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

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

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

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 [58]), but, to formulate it properly, one would need to define the so-called expected local time density first (cf. (4.16) of [58]), which we prefer not to do here.

**Proof of Proposition 5.3.** It is an easy calculation that uses conditioning and induction. First, observe that 

\[ g(z | \mathcal{F}_{n-1}) = \mathbb{P}[X_n = z | \mathcal{F}_{n-1}] \],

so we have

\[ \mathbb{E}G_1(z) = \mathbb{E}(\mathbb{P}[X_1 = z | \mathcal{F}_0]) = \mathbb{P}[X_1 = z] = EL_1(z). \]

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

\[
\begin{align*}
\mathbb{E}G_n(z) &= \mathbb{E}\left(\mathbb{E}(G_n(z) | \mathcal{F}_{n-1})\right) \\
&= \mathbb{E}G_{n-1}(z) + \mathbb{E}\left(\mathbb{E}(\xi_ng(z | \mathcal{F}_{n-1}) | \mathcal{F}_{n-1})\right) \\
&= \mathbb{E}G_{n-1}(z) + \mathbb{E}(g(z | \mathcal{F}_{n-1})\mathbb{E}(\xi_n | \mathcal{F}_{n-1})) \\
&= \mathbb{E}G_{n-1}(z) + \mathbb{E}(\mathbb{P}[X_n = z | \mathcal{F}_{n-1}]) \\
&= \mathbb{E}G_{n-1}(z) + \mathbb{P}[X_n = z].
\end{align*}
\]

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

\[
\mathbb{P}\left[\{X_1, \ldots, X_m\} \subset \{Y_1, \ldots, Y_n\}\right] \geq \mathbb{P}\left[G_m(z) \leq G'_n(z)\right] \text{ for all } z \in \Sigma,
\]

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 \( (\xi_1, \ldots, \xi_m) \) 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

\[
\begin{align*}
\mathbb{P}\left[G_m(z) \leq G'_n(z)\right] &\geq \mathbb{P}\left[G_m(z) \leq \varphi(z)\right] \text{ for all } z \in \Sigma
\end{align*}
\]
Note that, in the right-hand side of the above relation we do not have this “conflict of $\xi$’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 $\Sigma$ is finite, this difficulty can be usually circumvented by considering the values of the functions separately in each point of $\Sigma$ 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; see e.g. Section 6 of [22] and Lemma 2.9 of [16] for examples of how large deviations for soft local times may be treated.

To underline the importance of finding couplings as above (and also the exact couplings of Section 5.1.1 below), observe that there are many quantities of interest that can be expressed in terms of local times (as defined in (5.4)) only (that is, they do not depend on the order). Such quantities are, for example,

- hitting time of a site $x$: $\tau(x) = \min\{n : L_n(x) > 0\}$;
- cover time: $\min\{n : L_n(x) > 0 \text{ for all } x \in \Sigma\}$, where $\Sigma$ is the space where the process lives;
- blanket time [27]: $\min\{n \geq 1 : L_n(x) \geq \delta n \pi(x)\}$, where $\pi$ is the stationary measure of the process and $\delta \in (0, 1)$ is a parameter;
- disconnection time [24, 68]: loosely speaking, it is the time $n$ when the set $\{x : L_n(x) > 0\}$ becomes “big enough” to “disconnect” the space $\Sigma$ in some precise sense;
5.1 Soft local times

- the set of favorite (most visited) sites (e.g. [38, 74]): \{x : L_n(x) \geq L_n(y) \text{ for all } y \in \Sigma\};
- and so on.

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 \(\sigma\) such that \(X_{\sigma(1)}, X_{\sigma(2)}, X_{\sigma(3)}, \ldots\) is also an i.i.d. sequence, however, sampled from another distribution (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|)1\{y \in [-1, 1]\}.

The first step is to reconstruct a Poisson process in \([-1, 1] \times \mathbb{R}_+\), using \(X\)’s and \(\xi\)’s. This can be done in the following way (see Figure 5.6): for all \(n \geq 1\), put a point to \((X_n, \frac{1}{2}(\xi_1 + \cdots + \xi_n))\). 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 an exercise for the reader to check that, this time, essentially the same solution works in the general case.

### 5.1.1 Exact matchings with soft local times

Above, we have seen how the method of soft local times is useful for dominating the range of one process with the range of another one (recall (5.7)). This is indeed the main application of this method that one can find in the literature. Still, a natural question that one may ask is: assume that the processes \(X\) and \(Y\) are “close” in some sense. Is it possible to match their ranges exactly at finite time (i.e., not as in the beginning of this chapter), i.e., is it possible to couple these processes in such a way that \((X_1, \ldots, X_n)\) and \((Y_1, \ldots, Y_n)\) are just permutations of each other with high probability?\(^9\)

The problem, however, is that the “vanilla” soft local time method will not work for finding exact matchings as above. The reason is that, while the soft local times \(G_n^X\) and \(G_n^Y\) are relatively

\(^9\) as a motivating example, see Exercise 5.5
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 $\sigma(1) = 1$, $\sigma(2) = 3$, $\sigma(3) = 2$, $\sigma(4) = 6$, $\sigma(5) = 4$, $\sigma(6) = 10$, $\sigma(7) = 5$ (i.e., $Y_7 = X_5$ etc.).

close, with high probability they will still be locally different, so that there will be points of the Poisson process which are not included to the range of $X$ but included to the range of $Y$, and vice-versa. However, it turns out to be possible to modify this method to make it work for this purpose. In the following, we explain the ideas of [22, 23]. We will formulate the results in greater generality, since we do not intend to prove them here anyway.

Let $(\Sigma, \rho)$ be a compact metric space, with $\mathcal{B}(\Sigma)$ representing its Borel $\sigma$-algebra. We say that $(\Sigma, \rho)$ is of polynomial class, when there exist some $\beta \geq 0$ and $\varphi \geq 1$ such that, for all $r \in (0, 1]$, the number of open balls of radius at most $r$ needed to cover $\Sigma$ is smaller than or equal to $\varphi r^{-\beta}$.

As an example of metric space of polynomial class, consider first a finite space $\Sigma$, endowed with the discrete metric

$$\rho(x, y) = 1\{x \neq y\}, \text{ for } x, y \in \Sigma.$$
In this case, we can choose $\beta = 0$ and $\varphi = |\Sigma|$. As a second example, let $\Sigma$ be a compact $k$-dimensional Lipschitz submanifold of $\mathbb{R}^m$ with metric induced by the Euclidean norm of $\mathbb{R}^m$. In this case we can take $\beta = k$, but $\varphi$ will in general depend on the precise structure of $\Sigma$. It is important to observe that, for a finite $\Sigma$, it may not be the best idea to use the above discrete metric; one may be better off with another one, e.g., the metric inherited from the Euclidean space where $\Sigma$ is immersed (see e.g. the proof of Lemma 2.9 of [16]).

Here, we consider a Markov chain $X = (X_i)_{i \geq 1}$ on a general (not necessarily countable) state space $\Sigma$ as above. Such a process is characterized by its transition kernel $\mathcal{P}(x, dy)$ and initial distribution $v$, on $(\Sigma, \mathcal{B}(\Sigma))$: for any $B \in \mathcal{B}(\Sigma)$ and $n \geq 0$,

$$
P[X_{n+1} \in B \mid X_n = x, X_{n-1}, \ldots, X_0] = \mathcal{P}(x, B),$$

and

$$
P[X_0 \in B] = v(B).$$

Let us suppose that the chain has a unique invariant probability measure $\Pi$, that is,

$$
\int_\Sigma \mathcal{P}(x, B) \Pi(dx) = \Pi(B)
$$

for all $B \in \mathcal{B}(\Sigma)$. Moreover, we assume that the initial distribution and the transition kernel are absolutely continuous with respect to $\Pi$. Let us denote respectively by $\nu(\cdot)$ and $\tilde{p}(x, \cdot)$ the Radon-Nikodym derivatives (i.e., densities) of $v(\cdot)$ and $\mathcal{P}(x, \cdot)$: for all $A \in \mathcal{B}(\Sigma)$

$$
v(A) = \int_A \nu(y) \Pi(dy),$$

$$
\mathcal{P}(x, A) = \int_A \tilde{p}(x, y) \Pi(dy), \text{ for } x \in \Sigma.
$$

Let us also use assume that the density $\tilde{p}(x, \cdot)$ is uniformly Hölder continuous, that is, there exist constants $\kappa > 0$ and $\gamma \in (0, 1]$ such that for all $x, z, z' \in \Sigma$,

$$
|\tilde{p}(x, z) - \tilde{p}(x, z')| \leq \kappa(\rho(z, z'))^\gamma.
$$

(5.8)
Next, let us suppose that there exists $\varepsilon \in (0, \frac{1}{2}]$ such that
\[
\sup_{x, y \in \Sigma} |\tilde{p}(x, y) - 1| \leq \varepsilon, \tag{5.9}
\]
and
\[
\sup_{x \in \Sigma} |\nu(x) - 1| \leq \varepsilon. \tag{5.10}
\]

Observe that (5.10) is not very restrictive because, due to (5.9), the chain will anyway come quite close to stationarity already on step 2.

Additionally, let us denote by $Y = (Y_i)_{i \geq 1}$ a sequence of i.i.d. random variables with law $\Pi$.

Before stating the main result of this subsection, we recall the definition of the total variation distance between two probability measures $\bar{\mu}$ and $\hat{\mu}$ on some measurable space $(\Omega, T)$:
\[
\text{d}_{TV}(\bar{\mu}, \hat{\mu}) = \sup_{A \in T} |\bar{\mu}(A) - \hat{\mu}(A)|.
\]

When dealing with random elements $U$ and $V$, we will write (with a slight abuse of notation) $\text{d}_{TV}(U, V)$ to denote the total variation distance between the laws of $U$ and $V$. Denoting by $L_n^Z := (L_n^Z(x))_{x \in \Sigma}$ the local time field of the process $Z = X$ or $Y$ at time $n$, we are now ready to state

**Theorem 5.4** Suppose that $(\Sigma, \rho)$ is of polynomial class with parameters $\beta, \varphi$, that the transition density is uniformly Hölder continuous in the sense of (5.8), and also that both the transition density and the initial distribution are close to stationarity in the sense of (5.9)–(5.10). Then there exists a universal positive constant $K$ such that, for all $n \geq 1$, it holds that
\[
\text{d}_{TV}(L_n^X, L_n^Y) \leq K\varepsilon \sqrt{1 + \ln(\varphi 2^\beta) + \frac{2\beta}{\gamma} \ln \left( \frac{\kappa \vee (2\varepsilon)}{\varepsilon} \right)}, \tag{5.11}
\]

Notice that, in fact, the estimate (5.11) is only useful for small enough $\varepsilon$ (recall that the total variation distance never exceeds 1). Although not every Markov chain comes close to the stationary distribution in just one step, that can be sometimes circumvented by considering the process at times $k, 2k, 3k, \ldots$ with a large $k$. Nevertheless, it is also relevant to check if we can obtain a uniform control in $n$ of $\text{d}_{TV}(L_n^X, L_n^Y)$ away from the “almost stationarity” regime. We state the following
5.1 Soft local times

Theorem 5.5. As before, assume that $(\Sigma, \rho)$ is of polynomial class and that the transition density is uniformly Hölder continuous. Assume also that

$$\theta := \left( \sup_{x,y \in \Sigma} |\tilde{p}(x,y) - 1| \right) \vee \left( \sup_{x \in \Sigma} |\nu(x) - 1| \right) < 1. \tag{5.12}$$

Then, there exists $K' = K'(\beta, \varphi, \kappa, \gamma, \theta) > 0$, decreasing in $\theta$, such that

$$d_{TV}(L^X_n, L^Y_n) \leq 1 - K'$$

for all $n \geq 1$.

Such a result can be useful e.g. in the following context: if we are able to prove that, for the i.i.d. sequence, something happens with probability close to 1, then the same happens for the field of local times of the Markov chain with at least uniformly positive probability. Observe that it is not unusual that the fact that the probability of something is uniformly positive implies that it should be close to 1 then (because one frequently has general results stating that this something should converge to 0 or 1).

Now, we briefly describe the main idea of the proof of the above results (and suggest the interested reader to look at [22, 23] for a complete argument); this idea is also depicted on Figure 5.7 (we refer to [22] for some notations used there; in that paper, $G$ was used to denote soft local times, but here this notation is reserved for the Green’s function).

1. As noted in the beginning of this subsection, the “direct” soft local time coupling will likely not work (look at the two crosses in the top right corners of the first part of the picture).
2. Since the Markov chain $X$ is already close to the equilibrium, it will “regenerate” quite often, and it is therefore possible to extract an “independent part” from $(X_1, \ldots, X_n)$ (the elements of this independent part will be i.i.d. with law $\Pi$).
3. Then, to generate the local times up to time $n$, we may do it first for the “dependent part”, and then for the “independent” one.
4. Note that, to obtain the i.i.d. sequence $(Y_1, \ldots, Y_n)$, their soft local time must be “flat”, as pictured on the right part of Figure 5.7. Now, here comes the key idea: let us erase the points above the soft local time of the dependent part, and then try
Intermezzo

Figure 5.7 Resampling of the “independent parts”

to “resample” the two Poisson processes in such a way that the projections (although not necessarily the points themselves) coincide with high probability. This will amount to constructing...
5.2 Poisson processes of objects

Couplings of Binomial processes (local times of i.i.d. random elements on some common space) with slightly different laws; Proposition 5.1 of [22] shows that this is indeed possible. If that coupling of Binomial processes is successful, then the local times of \((X_1, \ldots, X_n)\) and \((Y_1, \ldots, Y_n)\) will coincide.

The author understands that the above does not even remotely qualify to be a proof; he has a positive experience, though, of showing the above to some colleagues and hear “ah yes, sure!” as a response.

5.2 Poisson processes of objects

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 objects, 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 objects 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 objects to the points independently (as e.g. on Figure 5.8). A broad example of this is the Poisson Boolean model, cf. e.g. [50]; there, these objects are balls/disks with random radii.

\(^{10}\) translations, rotations, reflections, and combinations of them.
However, the situation becomes much more complicated if we need to build a Poisson process of infinite objects. For example, what about a two-dimensional Poisson process of lines, which should look like as shown on Figure 5.9?

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 5.10: 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 5.11), and then consider a one-dimensional Poisson point process on this line. Then, through each of these points, draw a
5.2 Poisson processes of objects

Figure 5.11 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)

line with uniformly distributed direction (that is, \( \alpha \) on Figure 5.11 is uniform in \([-\pi/2, \pi/2]\); for definiteness, think that the positive values of \( \alpha \) 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 \( \mathbb{R}^2 \).

The above 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 the origin (as on Figure 5.11). For all \( n \in \mathbb{Z} \), consider the events

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

The events \( (H_n, n \in \mathbb{Z}) \) are independent by construction, and it not difficult to see that \( P[H_n] \approx \varepsilon n^{-1} \) (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 \( \alpha \) which are close to \( \pm \pi/2 \). Therefore, choos-

\(^{11}\) this again! We meet the harmonic series quite frequently in two dimensions...
Figure 5.12 Calculating the expected number of lines that intersect the small ball. It holds that $\alpha = \arccos \frac{h}{\sqrt{x^2 + h^2}}$.

If we choose 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 $\varphi$? Consider a small disk of diameter $\varepsilon$ situated at distance $h$ above the origin, as on Figure 5.12. 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 $[\alpha - \delta, \alpha + \delta]$, where $\alpha = \arccos \frac{h}{\sqrt{x^2 + h^2}}$ and (up to terms of smaller order) $\delta = \varepsilon \frac{\varepsilon}{\sqrt{x^2 + h^2}}$.

So, if $\lambda$ 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 ball, we have

$$E N(h, \varepsilon) = \lambda \varepsilon \int_{-\infty}^{+\infty} \varphi \left( \arccos \frac{h}{\sqrt{x^2 + h^2}} \right) \frac{dx}{\sqrt{x^2 + h^2}} + o(\varepsilon). \quad (5.13)$$

This does not look very nice, but notice that, if we just erase “$\varphi$” and “arccos” from (5.13)$^{12}$, the integral would become something more familiar (recall the Cauchy density)

$$\int_{-\infty}^{+\infty} \frac{h}{x^2 + h^2} \, dx = \int_{-\infty}^{+\infty} \frac{1}{(\varepsilon/h)^2 + 1} \, d\left( \frac{x}{h} \right) = \int_{-\infty}^{+\infty} \frac{du}{u^2 + 1} = \pi,$$

so the parameter $h$ disappears. And, actually, it is easy to get rid of $\varphi$ and arccos at once: just choose $\varphi(\alpha) = \frac{1}{2} \cos \alpha$. So, we obtain from (5.13) that $E N(h, \varepsilon) = \frac{1}{2} \pi \lambda \varepsilon + o(\varepsilon)$, which is a good sign

$^{12}$ and the parentheses as well, although it not strictly necessary.
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\(^\text{13}\), but, instead of doing this now, let us consider another (more general\(^\text{14}\)) approach to defining Poisson processes of objects. The idea is to represent these objects as points in the parameter space; i.e., each possible object 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 is a process of objects 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 $\theta$ 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 $\mathbb{R}_+ \times [0, 2\pi)$, and translate it to a set of lines in $\mathbb{R}^2$, as shown on Figure 5.13.

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 5.14) we have $\theta = \alpha$ and $d\varphi = \cos \alpha \, dx$, 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.

\(^{13}\) please, try to do it!

\(^{14}\) in fact, it is the approach.
At this point we prefer to end this discussion and recommend the beautiful book [40] 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 objects, but the construction may be quite nontrivial. As for the Poisson line process itself, it serves as a supporting example for the previous sentence and as a “get-some-intuition” example for the next chapter\textsuperscript{15}, but it is not directly

\textsuperscript{15} in particular, the reader is invited to pay special attention to Exercises 5.12 and 5.13 below.
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 of 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, by construction, it should obviously obey the cosine law. And here comes the paradox: almost surely 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 $\Sigma$, with transition probabilities $p(x, x')$, initial distribution $\pi_0$, and stationary measure $\pi$. Let $A$ be a subset of $\Sigma$. Prove that for any $n \geq 1$ and $\lambda > 0$ it holds that

$$
P_{\pi_0}[\tau_A \leq n] \geq P_{\pi_0} \left[ \xi_0 \pi_0(x) + \sum_{j=1}^{n-1} \xi_j p(X_j, x) \geq \lambda \pi(x), \forall x \in \Sigma \right] - e^{-\lambda \pi(A)},$$

(5.14)

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

Exercise 5.3 Find a nontrivial application of (5.14).

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

Exercise 5.5 Let $(X_j)_{j \geq 1}$ be a Markov chain on the state space $\Sigma = \{0, 1\}$, with the following transition probabilities:

$$p(k, k) = 1 - p(k, 1 - k) = \frac{1}{2} + \varepsilon$$
for $k = 0, 1$, where $\varepsilon \in (0, \frac{1}{2})$ is small. Clearly, by symmetry, $(\frac{1}{2}, \frac{1}{2})$ is the stationary distribution of this Markov chain. Next, let $(Y_i)_{i \geq 1}$ be a sequence of i.i.d. Bernoulli random variables with success probability $\frac{1}{2}$. Prove that the distance in total variation between the laws of $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$ is of order $\varepsilon$, uniformly in $n$.

Poisson processes of objects (Section 5.2):

**Exercise 5.6** 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 5.15 (from left to right):

1. choose two points uniformly and independently, and draw a chord between them;
2. first choose a radius 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;

I do not ask you to prove that the probability of the above event will be $\frac{1}{3}$, $\frac{1}{2}$, and $\frac{1}{3}$ 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

---

\[16\] i.e., a straight line segment linking the center to a boundary point
process (since it is, in a way, the canonical random collection of lines on the plane) 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 5.7** 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 (defined in the correct way in the previous exercise) of the bigger circle intersects the smaller one, thus generating a chord in it as well. Verify that this induced random chord has the correct distribution.

**Exercise 5.8** 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 non-connected) domains?

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

**Exercise 5.9** Explain the paradox in the end of Section 5.2.

**Exercise 5.10** 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 5.11** (Random billiards) A particle moves with constant speed inside some (connected, but not necessarily simply connected) domain $D$. 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 $X_t \in D$ be the location of the process at time $t$, and $V_t \in [0, 2\pi)$ be the corresponding direction; $\xi_n \in \partial D$, $n = 0, 1, 2, \ldots$ are the points where the process hits the boundary, as shown on Figure 5.16.

Prove that

\[^{17}\] recall the construction of the Poisson line process via a reference line
• the stationary measure of the random walk $\xi_n$ is uniform on $\partial D$;
• the stationary measure of the process $(X_t, V_t)$ is the product of uniform measures on $D$ and $[0, 2\pi)$.

Observe that this result holds for any (reasonable) domain $D$!

The $d$-dimensional version of this process appeared in [63] under the name of “running shake-and-bake algorithm”, and was subsequently studied in [13, 14, 15]. For some physical motivation for the cosine reflection law see e.g. [19] and references therein.

**Exercise 5.12** 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 $\partial A$, 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 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 *correct* law (i.e., the same as if we took a Poisson point process on $\partial B$ with the same intensity, and traced rays from each}

---

\(^{18}\) I mean, it is not *easier*
Figure 5.17 Poisson line process seen in a finite convex set

Figure 5.18 Poisson line process seen in two disjoint sets; if a straight line intersects $A_1$ the way shown on the picture, it must intersect $A_2$.

of its points, traced independent rays with the cosine law), see Figure 5.17.

Exercise 5.13 Now, let us consider two nonintersecting domains $A_1, A_2 \subset \mathbb{R}^2$, and abbreviate $r = \max\{\text{diam}(A_1), \text{diam}(A_2)\}$, $s = \text{dist}(A_1, A_2)$. Consider a two-dimensional Poisson line process with rate $\lambda$. It is quite clear that the restrictions of this process on $A_1$ and $A_2$ are not independent, just look at Figure 5.18. However, in the case $s \gg r$ one still can decouple them. Let $H_1$ and $H_2$ be two events supported on $A_1$ and $A_2$. This means that, informally speaking, the occurrence of the event $H_k$ is determined by the configuration seen on $A_k$, for $k = 1, 2$. Prove that, for some positive
constant $C$ we have

$$|\mathbb{P}[H_1 \cap H_2] - \mathbb{P}[H_1]\mathbb{P}[H_2]| \leq \frac{C\lambda r}{s}. \quad (5.15)$$

**Exercise 5.14** How would you define a Poisson line process in higher dimensions? Also, what about Poisson plane process in $\mathbb{R}^d$, $d \geq 3$, etc.?

**Exercise 5.15** Find the expected value of the orthogonal projection of the unit cube on a randomly oriented plane.
Random interlacements

The main subject of this chapter is the random interlacement model in two dimensions. We, however, begin with a short introduction to higher-dimensional (that is, $d \geq 3$) random interlacements because it a more “natural” object.

6.1 Introduction: random interlacements in higher dimensions

As a warm-up, let us solve Exercises 3.20 and 3.21: we want to prove that for any finite $A \subset B$

$$
P_{hm_B}[	au_A < \infty] = \frac{\text{cap}(A)}{\text{cap}(B)}, \quad (6.1)$$

and for all $y \in \partial A$

$$
P_{hm_B}[S_{\tau_A} = y \mid \tau_A < \infty] = \text{hm}_A(y). \quad (6.2)$$

To do this, write for $x \notin B$

$$
P_x[\tau_A < \infty, S_{\tau_A} = y]$$

(by the strong Markov property)

$$
= \sum_{z \in \partial B} P_x[\tau_A < \infty, S_{\tau_A} = y, S_{\tau_B} = z]
$$

$$
= \sum_{z \in \partial B} P_x[\tau_A < \infty, S_{\tau_A} = y, \tau_B < \infty, S_{\tau_B} = z]
$$

$$
= \sum_{z \in \partial B} P_x[\tau_B < \infty, S_{\tau_B} = z]P_z[\tau_A < \infty, S_{\tau_A} = y].
$$
Now, let us divide both sides by $G(x)$ and send $x$ to infinity. For the left-hand side, we have by Proposition 3.4 and Theorem 3.8
\[
\frac{\mathbb{P}_x[\tau_A < \infty, S_{\tau_A} = y]}{G(x)} = \frac{\mathbb{P}_x[S_{\tau_A} = y \mid \tau_A < \infty]}{G(x)} \mathbb{P}_x[\tau_A < \infty]
\]
\[\rightarrow \text{cap}(A) \mathbb{h}m_A(y) \quad \text{as} \quad x \rightarrow \infty,
\]
and, likewise, the right-hand side will become
\[
\sum_{z \in \partial B} \frac{\mathbb{P}_x[\tau_B < \infty, S_{\tau_B} = z]}{G(x)} \mathbb{P}_z[\tau_A < \infty, S_{\tau_A} = y]
\]
\[\rightarrow \sum_{z \in \partial B} \text{cap}(B) \mathbb{h}m_B(z) \mathbb{P}_x[\tau_A < \infty, S_{\tau_A} = y]
\]
\[= \text{cap}(B) \mathbb{P}_{\mathbb{h}m_B}[\tau_A < \infty, S_{\tau_A} = y]
\]
\[= \text{cap}(B) \mathbb{P}_{\mathbb{h}m_B}[\tau_A < \infty] \mathbb{P}_{\mathbb{h}m_B}[S_{\tau_A} = y \mid \tau_A < \infty],
\]
which means that
\[
\mathbb{h}m_A(y) = \frac{\text{cap}(B) \mathbb{P}_{\mathbb{h}m_B}[\tau_A < \infty]}{\text{cap}(A)} \times \mathbb{P}_{\mathbb{h}m_B}[S_{\tau_A} = y \mid \tau_A < \infty]. \quad (6.3)
\]
Note the following elementary fact: if $\mu$ and $\nu$ are probability measures such that $\mu = c\nu$, then $c = 1$ (and therefore $\mu = \nu$). So, since (6.3) holds for all $y \in \partial A$, this gives us (6.1) and (6.2) at once.

Let us now go back to the main subject of this section. Random interlacements were introduced by Sznitman in [68], motivated by the problem of disconnection of the discrete torus $\mathbb{Z}_d^n := \mathbb{Z}^d/n\mathbb{Z}^d$ by the trace of simple random walk, in dimension 3 or higher. Detailed accounts can be found in the survey [11] and in the recent books [31, 71]. 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 integer lattice.

Let us make a stop here, and try to understand the previous sentence. We want to figure out how a Poisson process of (bi-infinite) SRW’s trajectories should look like. It may be not clear at this point how to approach this question in the whole space at once, but let us try to fix a finite $B \subset \mathbb{Z}^d$ and think about how the part of that process which “touches” $B$ should be.
First, as mentioned in the previous chapter, that Poisson process of trajectories should be thought of as a Poisson point process in another space (a “space of trajectories”). Since the set of trajectories that intersect $B$ is a subset of that space, the number of them (i.e., the number of points in that subset) should have Poisson distribution with some parameter $\lambda_B$. Since these trajectories “come from the infinity”, they should choose the entrance location to $B$ according to $\text{hm}_B$; also, as usually happens in the Poissonian context, they should be independent between each other.

So, for now, we have a Poisson($\lambda_B$) number of independent SRW’s trajectories that enter $B$ at random sites chosen according to $\text{hm}_B$ and then continue walking; but how this parameter $\lambda_B$ should depend on $B$? To understand this, let us consider a situation depicted on Figure 6.1: take some $A$ such that $A \subset B$, so the trajectories that intersect $A$ constitute a subset of those that intersect $B$. It is here that (6.1) and (6.2) come in handy: by (6.2), the trajectories that enter $A$ do so via $\text{hm}_A$ (as they should), and, by (6.1), their number has Poisson distribution with parameter $\lambda_B \times \frac{\text{cap}(A)}{\text{cap}(B)}$. Since, on the other hand, that parameter should not depend on $B$ at all (that is, $\lambda_B$ must be equal to $u \cdot \text{cap}(B)$ for
some constant $u > 0$), we arrive to the following description of the random interlacement process with rate $u$ on $A \subset \mathbb{Z}^d$:

- take Poisson($u \text{cap}(A)$) particles and place them on $\partial A$ independently, according to $h_{\text{mA}}$;
- let them walk.

By the author’s experience, the above description is the one which is effectively used in most cases when reasoning about random interlacements. Still, it is also important to know that it is possible to define the random interlacement process in the whole $\mathbb{Z}^d$ at once, as a Poisson point process in the space of doubly-infinite SRW trajectories. We refer to Theorem 1.1 of [68] and also the aforementioned books for the precise construction. The author opted for not reproducing it here since, by his experience, it is at this point that a few students generally disappear from the course and never come back. (Nevertheless, we will briefly summarize a more general construction of random interlacements on transient weighted graphs in the next Section 6.2.) What is important to have in mind, though, is that it is possible to construct the random interlacement process for all $u \geq 0$ simultaneously. The idea is exactly the same as in the basic case one-dimensional Poisson processes: all such 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 6.2. Here, the idea is the same: instead of just considering a Poisson point process on a (properly defined) trajectory space $W^*$ for a fixed value of $u > 0$, we rather consider a Poisson point process on $W^* \times \mathbb{R}_+$. This way, the points of the latter process will be of the form $(w, s)$, where $w$ is a trajectory and $s$ is a positive number (the “level” of the trajectory).

To obtain the random interlacements on level $u$, simply take all points $(w, s)$ from that process with $s \leq u$, and then “forget” their second coordinates. From the point of view of a finite set $A \subset \mathbb{Z}^d$, as $u$ increases, new trajectories will “pop up” at rate $\text{cap}(A)$ and get included to the picture.

We denote by $I^u$ the set of all sites that belong to at least one trajectory (on level $u$); the sites of $\mathbb{Z}^d$ that are not touched by the trajectories constitute the vacant set $V^u = \mathbb{Z}^d \setminus I^u$. Note that

\footnote{in fact, even with variable rate}
the above simultaneous (for all $u > 0$) construction implies that $\mathcal{I}^{u_1} \subset \mathcal{I}^{u_2}$ and $\mathcal{V}^{u_2} \subset \mathcal{V}^{u_1}$ for $u_1 < u_2$.

Now, the probability that all sites of a set $A$ are vacant is the same as the probability that a Poisson($u \cap(A)$) random variable equals 0, that is,

$$\mathbb{P}[A \subset \mathcal{V}^u] = \exp\left(-u \cap(A)\right) \quad \text{for all finite } A \subset \mathbb{Z}^d.$$  (6.4)

It is interesting to note that the law of the vacant set $\mathcal{V}^u$ is uniquely characterized by the set of identities (6.4), see Theorem 9.2.XII of [20].

As previously mentioned, one of the motivations for studying random interlacements is their relationship with SRW on the torus $\mathbb{Z}^d_n := \mathbb{Z}^d / n\mathbb{Z}^d$. First, it was shown in [5] that the local picture left by the SRW in a small subset of $\mathbb{Z}^d_n$ up to time $un^d$ is close in distribution to that of rate-$u$ random interlacements. Then, in [76] it was established that the same is true for a macroscopic box in $\mathbb{Z}^d_n$ (did this settle the conjecture about the critical $u$ for the disconnection of the torus? – ask Augusto). Comment about disconnection time here.

From now on, we assume that the reader is familiar at least

\footnote{it is clear that it is a very reasonable assumption to make, but let us still mention e.g. the books [7, 37]}
with the definition and the basic facts about the Bernoulli (site) percolation model on $\mathbb{Z}^d$.

We now recall some known properties of the interlacement and the vacant sets; the list below is of course incomplete. For the interlacement set $\mathcal{I}^u$, it holds that:

- $\mathbb{P}[\mathcal{I}^u \text{ is connected for all } u > 0] = 1$, i.e., the interlacement set is connected simultaneously for all $u > 0$ (Theorem 1.5 of [10]).
- If we consider a graph whose vertices are the trajectories and there is an edge between two vertices iff the corresponding trajectories intersect, then the diameter of this graph equals $\lceil d/2 \rceil$ a.s. (that is, from each site of $\mathcal{I}^u$, it is possible to walk to any other site of $\mathcal{I}^u$, “changing” the underlying trajectories at most $\lceil d/2 \rceil$ times), as shown in [60].
- The graph distance (a.k.a. chemical distance) on $\mathcal{I}^u$ is essentially equivalent to Euclidean distance (except for some “local holes”), see Theorem 1.3 of [10].
- SRW on $\mathcal{I}^u$ essentially behaves as the usual SRW [59, 64].
- On the other hand, a random walk with drift on $\mathcal{I}^u$ in three dimensions always has zero (even subpolynomial) speed (Theorem 1.3 of [34]); this is in sharp contrast with the random walk with drift on the infinite cluster of Bernoulli percolation (somewhat surprisingly, this random walk is ballistic for small values of the drift, and has zero speed for large values of the drift, see [6]). It is conjectured, though, that a random walk with drift on $\mathcal{I}^u$ in dimensions $d \geq 4$ has similar behaviour to that of the drifted random walk on Bernoulli percolation cluster.

- **was noch? ask Augusto**

While the interlacement set is always connected, this is not so for the vacant set: for example, it is not difficult to convince oneself that it should contain infinitely many “isolated islands”. The most important questions, though, are related to the percolative properties of the vacant set: does it contain an infinite cluster, what is the probability of existence of a long path in this set, and so on. Some of the known results are:

- There is a critical value $u^* \in (0, \infty)$ (which depends also on dimension) such that, for all $u < u^*$, $\mathcal{V}^u$ contains an infinite connected subset a.s., and, for all $u > u^*$, all connected components
of $\mathcal{V}^u$ are a.s. finite (see [65]). It is not known (and constitutes a very difficult open problem) what happens for $u = u^*$. 

- There is $u^{**} \in [u^*, \infty)$ such that, for any $u > u^{**}$, the probability that the origin is connected to $x$ within $\mathcal{V}^u$ decays exponentially (with a logarithmic correction for $d = 3$) in $\|x\|$, see [58], Theorem 3.1. It is conjectured that $u^* = u^{**}$, but, again, so far it is unclear how to prove this. Observe that this would be an analogue of Menshikov’s theorem [51] in Bernoulli percolation.

- what else?.. ask A.

Compared to the Bernoulli percolation model, the main difficulty in studying the vacant set of random interlacements is that strong (polynomial) correlations are present there. Indeed, note that, by (3.15)–(3.16) and (6.4), for all $x, y \in \mathbb{Z}^d$ it holds that

\[
P[x \in \mathcal{V}^u] = \exp \left( - \frac{u}{G(0)} \right),
\]

\[
P[\{x, y\} \subset \mathcal{V}^u] = \exp \left( - \frac{2u}{G(0) + G(x - y)} \right).
\]

Let us now calculate the covariance and the correlation of $1\{x \in \mathcal{V}^u\}$ and $1\{y \in \mathcal{V}^u\}$:

\[
\text{cov}(1\{x \in \mathcal{V}^u\}, 1\{y \in \mathcal{V}^u\}) = \exp \left( - \frac{2u}{G(0) + G(x - y)} \right) - \exp \left( - \frac{2u}{G(0)} \right).
\]

(by the way, notice that $\text{cov}(1\{x \in \mathcal{V}^u\}, 1\{y \in \mathcal{V}^u\}) > 0$) for all $x \neq y$

\[
= \exp \left( - \frac{2u}{G(0)} \right) \left( \exp \left( \frac{2u G(x - y)}{G(0) + G(x - y)} \right) - 1 \right)
\]

(as $x - y \to \infty$)

\[
= \frac{2u}{G^2(0)} \exp \left( - \frac{2u}{G(0)} \right) \times \frac{\gamma_d}{\|x - y\|^{d-2}} (1 + O(\|x - y\|^{-2(d-2)})).
\]

Therefore,

\[
\text{corr}(1\{x \in \mathcal{V}^u\}, 1\{y \in \mathcal{V}^u\}) = \frac{2ue^{-u/G(0)}}{G^2(0)(1 - e^{-u/G(0)})} \times \frac{\gamma_d}{\|x - y\|^{d-2}} (1 + O(\|x - y\|^{-2(d-2)})).
\]
That is, when \( u \) is fixed, both the covariance and the correlation of \( 1\{x \in V^u\} \) and \( 1\{y \in V^u\} \) have the same order of polynomial decay (with respect to the distance between \( x \) and \( y \)) as the Green’s function. It is also interesting to notice that

\[
\frac{ue^{-u/G(0)}}{(1 - e^{-u/G(0)})} \sim \begin{cases} 
ue^{-u/G(0)}, & \text{as } u \to \infty, \\
G(0), & \text{as } u \to 0.
\end{cases}
\]

That is, the correlation decreases in \( u \) as \( u \to \infty \) because, intuitively, there are many trajectories that may contribute. The correlation “stabilizes” as \( u \to 0 \) because the dependence is “embedded” to just one trajectory that may commute between \( x \) and \( y \).

This was for the one-point correlations, but when one considers two nonintersecting (and, usually, distant) sets \( A_1, A_2 \subset \mathbb{Z}^d \), things start to complicate even more. Assume that the diameters of these sets are smaller or equal to \( r \) and that they are at distance \( s \) from each other. Suppose also that we are given two functions \( f_1 : \{0, 1\}^{A_1} \to [0, 1] \) and \( f_2 : \{0, 1\}^{A_2} \to [0, 1] \) that depend only on the configuration of the random interlacements inside the sets \( A_1 \) and \( A_2 \) respectively. In (2.15) of [68] it was established that

\[
\text{Cov}(f_1, f_2) \leq c_d u \frac{\text{cap}(A_1) \text{cap}(A_2)}{s^{d-2}} \leq c'_d u \left( \frac{r^2}{s} \right)^{d-2},
\]

(6.9)

see also Lemma 2.1 of [4]. For this estimate to make sense, we need that \( r^2 \ll s \), that is, the sets should be quite far away from each other relative to their sizes. It is remarkable that, even with these restrictions, the above inequality can be used to obtain highly non-trivial results. The estimate in (6.9) is not optimal and can be improved to some extent, as was shown in [21]. Still, it is an open problem to find the correct maximal order of decay of the covariance in (6.9). Since the random interlacement model is actually closely related to the so-called Gaussian Free Field (see \([69, 70]\)) for which that maximal order of decay is known to be \( O\left( \frac{\sqrt{\text{cap}(A_1) \text{cap}(A_2)}}{s^{d-2}} \right) \), as shown in [57], it would be natural to conjecture that the same should hold for random interlacements. Still, the decay of correlations would remain polynomial and the estimate only can be useful when \( r \ll s \).

As we just mentioned, even these weaker estimates can be very useful; one still naturally looks for more, though. It turns out that
one can obtain estimates with much smaller error term if one is allowed to change the level parameter $u$ (this is called sprinkling).

To explain what it means, we first need to introduce the notion of monotone (increasing or decreasing) event/function. Let $A \subseteq \mathbb{Z}^d$; for configurations of zeros and ones $\{\eta : A \to \{0, 1\} = \{0, 1\}^A\}$ there is a natural partial order: $\eta \leq \zeta$ whenever $\eta(x) \leq \zeta(x)$ for all $x \in A$. We then say that a function $f : \{0, 1\}^A \to \mathbb{R}$ is increasing if $f(\eta) \geq f(\zeta)$ for all $\eta \geq \zeta$; a function $g$ is decreasing if $(-g)$ is increasing. We say that an event determined by a configuration is increasing/decreasing if so is its indicator function. If we denote by $P^u$ and $E^u$ the probability and the expectation with respect to the level-$u$ random interlacements (using the convention that the state of $x$ is 0 if it is vacant and 1 if it is occupied), from the “simultaneous” construction of random interlacements it then immediately follows that $E^u f$ and $P^u[H]$ are increasing (respectively, decreasing) functions of $u$ for any increasing (respectively, decreasing) function $f$ or event $H$.

It turns out that most interesting events (functions) are monotone. For example:

- the proportion of occupied (vacant) sites in $A$;
- existence of path of occupied (vacant) sites in $A$ that connects a fixed site $x_0 \in A$ to $\partial A$;
- existence of a crossing (in some suitable sense) by occupied (vacant) sites in $A$;
- and so on.

An important fact about monotone events is that, in the case of random interlacements the so-called Harris-FKG inequality is still true (as it is for the Bernoulli percolation model). To formulate it, note first that a monotone function depending on $A \subset \mathbb{Z}^d$ can be also seen as a monotone function depending on $A' \subset \mathbb{Z}^d$ where $A \subset A'$ (i.e., it would not depend on the configuration on $A' \setminus A$, but we can still formally regard it as a function $\{0, 1\}^{A'} \to \mathbb{R}$).

In [73] it was proved that, given two functions $f_{1,2} : \{0, 1\}^A \to \mathbb{R}$ which are both increasing or both decreasing, it holds that

$$E^u f_1 f_2 \geq E^u f_1 E^u f_2,$$  \hspace{1cm} (6.10)

that is, the increasing functions (events) are positively correlated. It then follows that (6.10) also holds when the increasing functions $f_{1,2}$ depend on disjoint sets $A_{1,2}$, since both functions can be
Random interlacements

seen as functions on \( \{0, 1\}^{A_1 \cup A_2} \); so \( \text{Cov}(f_1, f_2) \geq 0 \) for monotone (in the same direction) functions \( f_1, f_2 \). The Harris-FKG inequality can be often useful, however, it works only in one direction; in the other direction we can only aim for inequalities of the form \( \mathbb{E}^u f_1 f_2 \leq \mathbb{E}^u f_1 \mathbb{E}^u f_2 + \text{error term} \), which follows from (6.9) and alikes. Now, here comes the key idea (known as sprinkling): if, for some (small) \( \delta \), we substitute \( \mathbb{E}^u f_1 \mathbb{E}^u f_2 \) with \( \mathbb{E}^u f_1 \mathbb{E}^u f_2 + \mathbb{E}^{u+\delta} f_1 \mathbb{E}^{u+\delta} f_2 \) in the right-hand side, the error term can be made considerably smaller! Indeed, the following result was proved in [58]:

**Theorem 6.1** Let \( A_1, A_2 \) be two nonintersecting subsets of \( \mathbb{Z}^d \), with at least one of them being finite. Let \( s \) be the distance between \( A_1 \) and \( A_2 \), and \( r \) be the minimum of their diameters. Then, for all \( u > 0 \) and \( \varepsilon \in (0, 1) \) we have

\[
(i) \quad \text{for any increasing functions } f_1 : \{0, 1\}^{A_1} \to [0, 1] \text{ and } f_2 : \{0, 1\}^{A_2} \to [0, 1], \\
\mathbb{E}^u f_1 f_2 \leq \mathbb{E}^{(1+\varepsilon)u} f_1 \mathbb{E}^{(1+\varepsilon)u} f_2 + c'(r+s)^d \exp(-c'' \varepsilon^2 us^{d-2});
\]

(6.11)

\[
(ii) \quad \text{for any decreasing functions } f_1 : \{0, 1\}^{A_1} \to [0, 1] \text{ and } f_2 : \{0, 1\}^{A_2} \to [0, 1], \\
\mathbb{E}^u f_1 f_2 \leq \mathbb{E}^{(1-\varepsilon)u} f_1 \mathbb{E}^{(1-\varepsilon)u} f_2 + c'(r+s)^d \exp(-c'' \varepsilon^2 us^{d-2}).
\]

(6.12)

Observe that these inequalities are operational even when the two sets are close (relative to the size of the smallest of them). They are consequences of the next result:

**Theorem 6.2** Let \( A_1, A_2 \) be two nonintersecting subsets of \( \mathbb{Z}^d \), with at least one of them being finite. Abbreviate \( s = \text{dist}(A_1, A_2) \) and \( r = \min\{\text{diam}(A_1), \text{diam}(A_2)\} \). Then, there are positive constants \( c' \) and \( c'' \) (depending only on the dimension \( d \)) such that for all \( u > 0 \) and \( \varepsilon \in (0, 1) \) there exists a coupling \( \mathcal{Q} \) between \( \mathcal{I}^u \) and two independent random interlacements processes, \( (\mathcal{I}^u_1)_{u \geq 0} \) and \( (\mathcal{I}^u_2)_{u \geq 0} \) such that

\[
\mathcal{Q}[\mathcal{I}^u_1 \cap A_k \subseteq \mathcal{I}^u \cap A_k \subseteq \mathcal{I}^{u(1+\varepsilon)}_k, k = 1, 2] \geq 1 - c'(r+s)^d \exp(-c'' \varepsilon^2 us^{d-2}).
\]

(6.13)
6.2 The two-dimensional case

formulate for trajectories instead of interlacement sets?

In words, it shows us that there is a way to decouple the intersection of the interlacement set $I^u$ with two disjoint subsets $A_1$ and $A_2$ of $Z^d$. Namely, we couple the original interlacement process $I^u$ with two independent interlacements processes $I^u_1$ and $I^u_2$ in such a way that $I^u$ restricted on $A_k$ is “close” to $I^u_k$, for $k = 1, 2$, with probability rapidly going to 1 as the distance between the sets increases.

Sketch the proof with soft local times. (put modified versions of Figures 4 and 5 of Popov-Teixeira)

Now, ... implies that for $x \in \partial A_k$ we have $\varphi(x) := E^P \rho_{A_k}^{A_k}(x) = E^P \rho_{A_1 \cup A_2}^{A_1 \cup A_2}(x)$. The conclusion of ... is now a simple consequence of the above display and the fact that the expectation of $G^u$ is linear in $u$ according to (??) (see Figure 6.4).

Cite also the conditional decoupling of [1].

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 $Z^2$, so the vacant set would be always empty. Nevertheless, we know that there is also a natural notion of capacity in two dimensions, so one may wonder if there is a way to construct a decreasing family $(V^\alpha, \alpha > 0)$ of random subsets of $Z^2$ in such a way that a formula analogous to (6.4) 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

$$\mathbb{P}[A \subset V^\alpha] = \exp \left(- \pi \alpha \text{cap}(A) \right)$$

holds for all sets containing the origin (the factor $\pi$ in the exponent is just for convenience, as explained below). 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 6.4 below. In addition, we will see that (similarly to the $d \geq 3$ case)
Figure 6.3 On the construction of random interlacements on the set $A$; the points of $\Sigma_A$ are substituted by points in $\partial A \times \mathbb{R}_+$ with marks representing the corresponding trajectories.
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

$$\tilde{e}_A(x) = 1\{x \in A\} \mathbb{P}[\tilde{\tau}(A) = \infty] \mu_x,$$

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

$$\tilde{\text{cap}}(A) = \sum_{x \in A} \tilde{e}_A(x).$$

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

Now, we use the general construction of random interlacements on a transient weighted graph introduced in [73]. In the following few lines we briefly summarize this construction. Let $W$ be the space of all doubly infinite nearest-neighbour transient trajectories in $\mathbb{Z}^2$,

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

We say that $\varrho$ and $\varrho'$ 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
Random interlacements

time shift, and define $\chi^*$ 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 $\chi^*$. 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 $\chi^*$. One then constructs the random interlacements as Poisson point process on $W^* \times \mathbb{R}^+$ with the intensity measure $\nu \otimes du$, where $\nu$ is described in the following way. It is the unique sigma-finite measure on $W^*$ such that for every finite $A$

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

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

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

The existence and uniqueness of $\nu$ was shown in Theorem 2.1 of [73].

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

$$\sum_{\lambda : u_\lambda \leq \pi \alpha} \delta(w^*_\lambda).$$

Observe that this definition is somewhat unconventional (we used $\pi \alpha$ instead of just $\alpha$, 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 $\alpha$ “observed” on a finite set $A \subset \mathbb{Z}^2$. Namely,

- take a Poisson($\pi \alpha \hat{\text{cap}}(A)$) number of particles;
- place these particles on the boundary of $A$ independently, with distribution $\bar{c}_A = ((\hat{\text{cap}} A)^{-1} \hat{e}_A(x), x \in A)$;
- let the particles perform independent $\hat{S}$-random walks (since $\hat{S}$ 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 [73].
6.2 The two-dimensional case

The vacant set at level \( \alpha \),

\[
\mathcal{V}_\alpha = \mathbb{Z}^2 \setminus \bigcup_{\lambda : u_\lambda \leq \pi \alpha} \omega_\lambda^*(\mathbb{Z})
\]
is the set of lattice points not covered by the random interlacement. It contains the origin by definition. In Figure 6.5 we present a simulation\(^3\) 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 (6.14) is verified. For this, we need to prove the following fact:

**Proposition 6.3** For any finite set \( A \subset \mathbb{Z}^2 \) such that \( 0 \in A \) it holds that \( \mathrm{cap}(A) = \hat{\mathrm{cap}}(A) \).

**Proof** Indeed, consider an arbitrary \( x \in \partial A, x \neq 0 \), and (large) \( r \) such that \( A \subset B(r - 2) \). Write using (3.36)

\[
\mathbb{P}_x[\tilde{\tau}_1(A) > \tilde{\tau}_1(\partial B(r))] = \sum_{\varrho} \frac{a(\varrho_{\text{end}})}{a(x)} \left(\frac{1}{4}\right)^{|\varrho|} = (1 + o(1)) \frac{2 \ln r}{a(x)} \sum_{\varrho} \left(\frac{1}{4}\right)^{|\varrho|} = (1 + o(1)) \frac{2 \ln r}{a(x)} P_x[\tau^+_1 > \tau_1(\partial B(r))],
\]

where the sums are taken over all trajectories \( \varrho \) 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 (??) to obtain that, if \( 0 \in A \),

\[
a(x) \mathbb{P}_x[\tilde{\tau}_1(A) = \infty] = \text{hm}_A(x). \tag{6.17}
\]

Multiplying by \( a(x) \) and summing over \( x \in A \) (recall that \( \mu_x = a^2(x) \)) we obtain the expressions in (3.18) and (6.16) and thus conclude the proof. \( \square \)

Together with formula (1.1) of [73], Proposition 6.3 shows the fundamental relation (6.14) announced in introduction: for all finite subsets \( A \) of \( \mathbb{Z}^2 \) containing the origin,

\[
\mathbb{P}[A \subset \mathcal{V}_\alpha^c] = \exp\left(-\pi \alpha \text{cap}(A)\right).
\]

\(^3\) many thanks to Darcy Camargo!
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...
6.2 The two-dimensional case

Let us describe some other basic properties of two-dimensional random interlacements:

**Theorem 6.4** (i) For any $\alpha > 0$, $x \in \mathbb{Z}^2$, $A \subset \mathbb{Z}^2$, it holds that

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

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

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

(ii) For $\gamma'$ from (3.36) we have

$$\mathbb{P}[x \in \mathcal{V}^\alpha] = \exp\left(-\pi \alpha \frac{a(x)}{2}\right) = e^{-\pi \gamma' - \alpha \gamma' \frac{1}{2} \ln \|x\|} (1 + O(\|x\|^{-2})).$$

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

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

(iv) For $x, y \neq 0, x \neq y$, we have

$$\mathbb{P}[\{x, y\} \subset \mathcal{V}^\alpha] = \exp\left(-\pi \alpha \Psi\right),$$

where

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

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

$$\mathbb{P}[\{x, y\} \subset \mathcal{V}^\alpha] = s^{-\frac{2\beta}{4} + o(1)}.$$ 

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

$$\mathbb{P}[B(x, r) \subset \mathcal{V}^\alpha] = s^{-\frac{2\beta}{4} + o(1)}.$$ 

The relation (4.21) leads to the following heuristic explanation for Theorem 6.4 (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.

These results invite a few comments.

Remark 6.5

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 \( \text{RI}(\alpha) \) given that a distant site \( x \) is vacant, is similar – near the origin – to the unconditional law of \( \text{RI}(\alpha/4) \). Combined with (i), the similarity holds near \( x \) as well. Moreover, one can also estimate the “local rate” away from the origin, see Figure 6.6. More specifically, observe from Lemma 6.9 (ii) that \( \text{cap}(A_2) \ll \ln s \) with \( s = \text{dist}(0, A_2) \) large implies \( \text{cap}\left(\{0\} \cup A_2\right) = \frac{a(s)}{2}(1 + o(1)) \).

If \( x \) is at a much larger distance from the origin than \( A_2 \), say \( \ln \|x\| \sim \ln(s^2) \), then (6.21) reveals a “local rate” equal to \( \frac{7}{8}\alpha \), that is, \( P[A_2 \subset \mathcal{V}^\alpha | x \in \mathcal{V}^\alpha] = \exp \left( -\frac{7}{8}\pi \alpha \text{cap}\left(\{0\} \cup A_2\right)(1 + o(1)) \right) \); indeed, the expression in the denominator in (6.21) equals approximately \( 1 - \frac{\text{cap}(\{0\} \cup A_2)}{2n(x)} \approx 1 - \frac{a(s)/2}{2n(x)} \approx \frac{7}{8} \).

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 \( \text{cap}\left(\{0, x\} \cup A\right) = \text{cap}\left(\{0, x\} \cup (-A + x)\right) \) by symmetry. For the second statement in (i), note that, for \( A' = \{0, x\} \cup A \), it holds that \( \text{cap}(A') = \text{cap}(MA') = \text{cap}(\{0, x\} \cup \)
6.2 The two-dimensional case

MA). Item (ii) follows from the above mentioned fact that \( \text{cap} \{0, x\} = \frac{1}{2} a(x) \) together with (3.36).

We postpone the proof of other parts of Theorem 6.4, 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 6.6** (i) We have

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

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

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

\[
\mathbb{P}[V^\alpha \cap (B(r) \setminus B(r/2)) = \emptyset] \leq r^{-2(1-\sqrt{\alpha})^2 + o(1)}.
\]  

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 6.7 below.

**Proof of (i) and (ii) (incomplete, in the latter case)** Part (i) immediately follows from Theorem 6.4 (ii).

The proof of the part (ii) is easy in the case \( \alpha > 2 \). Indeed, observe first that \( \mathbb{E}|V^\alpha| < \infty \) implies that \( V^\alpha \) itself is a.s. finite. Also, Theorem 6.4 (ii) actually implies that \( \mathbb{E}|V^\alpha \setminus \{0\}| \to 0 \) as \( \alpha \to \infty \), so \( \mathbb{P}[V^\alpha = \{0\}] \to 1 \) by the Chebyshev inequality.

Now, let us prove that, in general, \( \mathbb{P}[|V^\alpha| < \infty] = 1 \) implies that \( \mathbb{P}[V^\alpha = \{0\}] > 0 \). Indeed, if \( V^\alpha \) is a.s. finite, then one can find a sufficiently large \( R \) such that \( \mathbb{P}[|V^\alpha \cap (\mathbb{Z}^2 \setminus B(R))| = 0] > 0 \). Since \( \mathbb{P}[x \notin V^\alpha] > 0 \) for any \( x \neq 0 \), the claim that \( \mathbb{P}[V^\alpha = \{0\}] > 0 \) follows from the FKG inequality applied to events \( \{x \notin V^\alpha\} \), \( x \in B(R) \) together with \( \{|V^\alpha \cap (\mathbb{Z}^2 \setminus B(R))| = 0\} \).

As before, we postpone the proof of part (iii) and the rest of part (ii) of Theorem 6.6. Let us remark that we believe that the right-hand side of (6.24) 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 \( V^1 \) 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 $\hat{S}$-walks that hit $B(r)$ has Poisson law with rate of order $\ln r$ (recall (??) (2-dim capacity of a disk)). 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 $V^\alpha \cap B(r)$ 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 than the expected number of uncovered sites there.

Denote by $\Upsilon_n : \mathbb{Z}^2 \to \mathbb{Z}_n^2$, $\Upsilon_n(x, y) = (x \mod n, y \mod n)$, the natural projection modulo $n$. Then, if $S_0$ were chosen uniformly at random on any fixed $n \times n$ square, we can write $X_k = \Upsilon_n(S_k)$. Similarly, $B(y, r) \subset \mathbb{Z}_n^2$ is defined by $B(y, r) = \Upsilon_nB(z, r)$, where $z \in \mathbb{Z}^2$ is such that $\Upsilon_nz = y$. Let also

$$t_\alpha := \frac{4\alpha}{\pi} n^2 \ln^2 n$$

(recall that, as mentioned in Section ??, $\alpha = 1$ corresponds to the leading-order term of the expected cover time of the torus). In the following theorem, we prove that, given that 0 is uncovered, the law of the uncovered set around 0 at time $t_\alpha$ is close to that of $\text{RI}(\alpha)$:

**Theorem 6.7** Let $\alpha > 0$ and $A$ be a finite subset of $\mathbb{Z}^2$ such that $0 \in A$. Then, we have

$$\lim_{n \to \infty} \mathbb{P}[\Upsilon_nA \subset U_{t_\alpha}^{(n)} \mid 0 \in U_{t_\alpha}^{(n)}] = \exp \left( -\pi \alpha \text{cap}(A) \right).$$

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 $\mathbb{Z}_n^2 = \mathbb{Z}^2/n\mathbb{Z}^2$. 


Here and in the sequel we denote by \((Z^{(i)}, i \geq 1)\) the (complete) excursions of the walk \(X\) between \(\partial A\) and \(\partial A'\), and by \((\hat{Z}^{(i)}, i \geq 1)\) the RI’s excursions between \(\partial A\) and \(\partial A'\) (dependence 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 \(\text{RI}(\alpha)\), say, between \(\partial B(y_0, n)\) and \(\partial B(y_0, cn)\) for some \(c > 0\) and \(y_0 \in \mathbb{Z}^2\). Also, (let us identify the torus \(\mathbb{Z}_n^2\) 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 \(\mathbb{Z}_n^2\) between \(\partial B(y_0, n)\) and \(\partial B(y_0, cn)\), 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 [58] for the general theory (see also Figure 1 of [58] which gives some quick insight on what is going on), and also to Section 2 of [17]. 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 \(\mathbb{Z}_n^2\), between \(\partial A\) 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 \(x \not\in A\) and \(y \in \partial A\) let us denote \(\varphi(x, y) = \mathbb{P}_x[X_{\tau_A} = y]\). 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 \((\iota(Z^{(j)}), \ell(Z^{(j)})), j \geq 1)\) is a Markov chain with transition probabilities

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

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 \(\partial A\)) and run until hitting \(\partial A'\). Then (see Figure 6.7; observe that \(A\) and \(A'\) need not be necessarily connected, as shown on the picture)

- let \(\xi_1\) 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;
Random interlacements

- the mark of the chosen point is the first excursion (call it \( Z^{(1)} \)) that we obtain;
- then, let \( \xi_2 \) 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 \( k \)th excursion is indeed distributed according to \( \varphi(\ell(Z^{(k-1)}), \cdot) \)); moreover, the \( \xi \)'s are i.i.d. random variables with Exponential(1) distribution.

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

\[
L_k(y) = \xi_1 \varphi(x_0, y) + \sum_{j=2}^{k} \xi_j \varphi(\ell(Z^{(j-1)}), y). \tag{6.26}
\]

For the random interlacements, the soft local times are defined analogously. Recall that \( \hat{\text{hm}}_A \) defines the (normalized) harmonic measure on \( A \) with respect to the \( \hat{S} \)-walk. For \( x \not\in A \) and \( y \in \partial A \) let

\[
\hat{\varphi}(x, y) = P_x[\hat{S}_{\hat{\tau}_1(A)} = y, \hat{\tau}_1(A) < \infty] + P_x[\hat{\tau}_1(A) = \infty] \hat{\text{hm}}_A(y). \tag{6.27}
\]

Analogously, for the random interlacements, it holds that the sequence \(((\ell(Z^{(j)}), \ell(Z^{(j)})), j \geq 1)\) is also a Markov chain, with transition probabilities

\[
\hat{P}_{(y,z), (y',z')} = \varphi(z, y') P_y[S_{\hat{\tau}_1(\partial A')} = z'].
\]

The process of picking the excursions for the random interlacements is quite analogous: if the last excursion was \( \hat{Z} \), we use the probability distribution \( \varphi(\ell(\hat{Z}), \cdot) \) to choose the starting point of the next excursion. Clearly, the last term in (6.27) is needed for \( \hat{\varphi} \) to have total mass 1; informally, if the \( \hat{S} \)-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 [76]). 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 param-

**Figure 6.7** Construction of the first three excursions between $\partial A$ and $\partial A'$ on the torus $\mathbb{Z}_2^2$ using the soft local times (here, $A = A_1 \cup A_2$ and $A' = A'_1 \cup A'_2$)
eter 1. Then, define the soft local time of random interlacement of \( k \)th excursion as

\[
\hat{L}_k(y) = \hat{\xi}_1 \hat{\varphi}(x_0, y) + \sum_{j=2}^{k} \hat{\xi}_j \hat{\varphi}(\ell(\hat{Z}^{(j-1)}), y).
\] (6.28)

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 6.8** Assume that \( x \notin B(y, r) \) and \( \|y\| > 2r \geq 1 \). Abbreiviate also \( \Psi_1 = \|y\|^{-1}r, \Psi_2 = r \ln \frac{\ln \|x\|}{\|y\|}, \Psi_3 = r \ln r \left( \ln \frac{\|x-y\|}{\|y\|} + \ln \frac{\|x\|}{\|y\|} \right) \).

(i) We have

\[
\mathbb{P}_x[\tilde{\tau}_1(B(y, r)) < \infty] = \frac{(a(y) + O(\Psi_1))(a(y) + a(x) - a(x - y) + O(r^{-1}))}{a(x)(2a(y) - a(r) + O(r^{-1} + \Psi_1))}.
\] (6.29)

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

\[
\mathbb{P}_x[\tilde{\tau}_1(A) < \infty] = \frac{(a(y) + O(\Psi_1))(a(y) + a(x) - a(x - y) + O(r^{-1} + \Psi_3))}{a(x)(2a(y) - \text{cap}(A) + O(\Psi_2))}.
\] (6.30)

Observe that (6.29) is not a particular case of (6.30); this is because (??) typically provides a more precise estimate than (??). (do we need the above)

### 6.3.2 Harmonic measure and capacities

Before proceeding, let us notice the following immediate consequence of Lemma 6.8: for any finite \( A \subset \mathbb{Z}^2 \) such that \( 0 \in A \), we have

\[
\text{cap}(A) = \lim_{\|x\| \to \infty} a(x)\hat{P}_x[\tilde{\tau}_1(A) < \infty].
\] (6.31)

Next, we need estimates for the \( \tilde{S} \)-capacity of a “distant” set, and, in particular of a ball which does not contain the origin. Recall the notations \( \Psi_{1,2,3} \) from Lemma 6.8.

**Lemma 6.9** Assume that \( \|y\| > 2r \geq 1 \).
(i) We have
\[
\text{cap } (\{0\} \cup B(y, r)) = \frac{(a(y) + O(\Psi_1))(a(y) + O(r^{-1}))}{2a(y) - a(r) + O(r^{-1} + \Psi_1)}. \tag{6.32}
\]

(ii) Suppose that \(A \subset B(y, r)\). Then
\[
\text{cap } (\{0\} \cup A) = \frac{(a(y) + O(\Psi_1))(a(y) + O(r^{-1} + \Psi_2))}{2a(y) - \text{cap}(A) + O(\Psi_2)}. \tag{6.33}
\]

Proof. This immediately follows from (6.31) and Lemma 6.8 (observe that \(a(x) - a(x-y) \to 0\) as \(x \to \infty\) and \(\Psi_3\) becomes \(\Psi_2\)).

We also need to compare the harmonic measure on a set (distant from the origin) to the entrance measure of the \(\hat{S}\)-walk started far away from that set.

**Lemma 6.10** Assume that \(A\) is a finite subset of \(\mathbb{Z}^2\), \(0 \notin A, x \neq 0\), and also that \(2\, \text{diam}(A) < \dist(x, A) < \frac{1}{2} \cdot \text{dist}(0, A)\). Abbreviate \(u = \text{diam}(A)\), \(s = \text{dist}(x, A)\). Assume also that \(A' \subset \mathbb{Z}^2\) (finite, infinite, or even possibly empty) is such that \(\text{dist}(A, A') \geq s + 1\) (for definiteness, we adopt the convention \(\text{dist}(A, \emptyset) = \infty\) for any \(A\)). Then, for \(y \in A\), it holds that
\[
\hat{P}_z[\hat{S}_{\hat{\tau}_1(A)} = y \mid \hat{\tau}_1(A) < \infty, \hat{\tau}_1(A) < \hat{\tau}_1(A')] = \text{hm}_A(y) \left( 1 + O \left( \frac{u \ln s}{s} \right) \right). \tag{6.34}
\]

Proof. Let \(z_0 \in A\) be such that \(\|z_0 - x\| = s\), and observe that \(A' \cap B(z_0, s) = \emptyset\). Define the discrete circle \(L = \partial B(z_0, s)\); observe that \(x \in L\) and \(\text{dist}(z', A) \geq s/2\) for all \(z' \in L\). Let
\[
\sigma = \sup \{0 \leq k \leq \hat{\tau}_1(A) : \hat{S}_k \in L\}
\]
be the last time before \(\hat{\tau}_1(A)\) when the trajectory passes through \(L\). Note also that for all \(z \in L\) (recall (??))
\[
P_z[\hat{S}_{\hat{\tau}_1(A)} = y \mid \tau_1(A) < \hat{\tau}_1(L)] = \text{hm}_A(y) \left( 1 + O \left( \frac{u \ln s}{s} \right) \right). \tag{6.35}
\]

Using the Markov property of \(\hat{S}\), we write
\[
\hat{P}_z[\hat{\tau}_1(A) < \infty, \hat{\tau}_1(A) < \hat{\tau}_1(A'), \hat{S}_{\hat{\tau}_1(A)} = y] = \sum_{k \geq 0, z \in L} \hat{P}_z[\hat{\tau}_1(A) < \infty, \hat{\tau}_1(A) < \hat{\tau}_1(A'), \sigma = k, \hat{S}_\sigma = z, \hat{S}_{\hat{\tau}_1(A)} = y]
\]
Random interlacements

\[
= \sum_{k \geq 0, z \in L} \hat{P}_z[\hat{S}_k = z, \hat{S}_\ell \notin A \cup A' \text{ for all } \ell \leq k]
\times \hat{P}_z[\hat{\tau}_1(A) < \infty, \hat{\tau}_1(A) < \hat{\tau}_1(A'), \hat{S}_{\hat{\tau}_1(A)} = y, \hat{S}_\ell \notin L \text{ for all } \ell \leq \hat{\tau}_1(A)].
\]

(6.36)

Abbreviate \( r = \text{dist}(0, A) \). Now, observe that the last term in (6.36) only involves trajectories that lie in \( B(z_0, s) \), and we have \( \text{dist}(0, B(z_0, s)) \geq r/2 \). So, we can use Lemma 4.4 (ii) together with (6.35) to write

\[
\hat{P}_z[\hat{\tau}_1(A) < \infty, \hat{\tau}_1(A) < \hat{\tau}_1(A'), \hat{S}_{\hat{\tau}_1(A)} = y, \hat{S}_\ell \notin L \text{ for all } \ell \leq \hat{\tau}_1(A)]
= P_z[S_{\tau_1(A)} = y, S_\ell \notin L \text{ for all } \ell \leq \tau_1(A)] \left(1 + O\left(\frac{u}{r \ln r}\right)\right)
= P_z[S_{\tau_1(A)} = y \mid \tau_1(A) < \tau_1(L)] P_z[\tau_1(A) < \tau_1(L)] \left(1 + O\left(\frac{u}{r \ln r}\right)\right)
= \text{hm}_A(y) \hat{P}_z[\hat{\tau}_1(A) < \hat{\tau}_1(L)] \left(1 + O\left(\frac{u}{r \ln r} + \frac{u \ln s}{s}\right)\right).
\]

Inserting this back to (6.36), we obtain

\[
\hat{P}_x[\hat{\tau}_1(A) < \infty, \hat{\tau}_1(A) < \hat{\tau}_1(A'), \hat{S}_{\hat{\tau}_1(A)} = y]
= \text{hm}_A(y) \left(1 + O\left(\frac{u}{r \ln r} + \frac{u \ln u}{s}\right)\right)
\times \sum_{k \geq 0, z \in L} \hat{P}_z[\hat{S}_k = z, \hat{S}_\ell \notin A \cup A' \text{ for all } \ell \leq k] \hat{P}_z[\hat{\tau}_1(A) < \hat{\tau}_1(L)]
= \text{hm}_A(y) \hat{P}_z[\hat{\tau}_1(A) < \infty, \hat{\tau}_1(A) < \hat{\tau}_1(A')] \left(1 + O\left(\frac{u}{r \ln r} + \frac{u \ln s}{s}\right)\right),
\]

and this concludes the proof of Lemma 6.10 (observe that the first term in \( O(\cdot) \) is of smaller order than the second one). \( \square \)

6.3.3 Proofs for random interlacements

First of all, we finish the proof of Theorem 6.4.

Proof of Theorem 6.4, parts (iii)–(v). Recall the fundamental formula (6.14) for the random interlacement and the relation (3.36). Then, the statement (iv) follows from (3.76) and from (6.20), while (v) is a consequence of Lemma 6.9 (i).

Finally, observe that, by symmetry, Theorem 6.4 (ii), and Lemma 6.9 (ii) we have

\[
P[A \subset V^x \mid x \in V^x]
\]
= \exp \left( -\pi \alpha \left( \operatorname{cap}(A \cup \{x\}) - \operatorname{cap}(\{0, x\}) \right) \right)

= \exp \left( -\pi \alpha \left( \frac{a(x) + O\left(\frac{\ln \ln \|x\|}{\|x\|}\right)^2}{2a(x) - \operatorname{cap}(A) + O\left(\frac{\ln \ln \|x\|}{\|x\|}\right)} - \frac{a(x)}{2} \right) \right)

= \exp \left( -\frac{\pi \alpha}{4} \operatorname{cap}(A) \frac{1 + O\left(\frac{\ln \ln \|x\|}{\|x\|}\right)}{1 - \frac{\operatorname{cap}(A)}{2a(x)} + O\left(\frac{\ln \ln \|x\|}{\|x\|}\right)} \right),

thus proving the part (iii).

Proof of Theorem 6.6 (iii) We start by observing that the first part of (iii) follows from the bound (6.24) and Borel-Cantelli. So, let us concentrate on proving (6.24). Recall the following elementary fact: let $N$ be a Poisson random variable with parameter $\lambda$, and $Y_1, Y_2, Y_3, \ldots$ be independent (also of $N$) random variables with exponential distribution $\mathcal{E}(p)$ with parameter $p$. Let $\Theta = \sum_{j=1}^{N} Y_j$ be the corresponding compound Poisson random variable. Its Cramér transform $b \mapsto \lambda (\sqrt{b} - 1)^2$ is easily computed, and Chernov’s bound gives, for all $b > 1$,

$$P[\Theta \geq b\lambda p^{-1}] \leq \exp \left( -\lambda (\sqrt{b} - 1)^2 \right). \quad (6.37)$$

Now, assume that $\alpha < 1$. Fix $\beta \in (0, 1)$, which will be later taken close to 1, and fix some set of non-intersecting disks $B'_1 = B(x_1, r^3), \ldots, B'_{k_r} = B(x_{k_r}, r^3) \subset B(r) \setminus B(r/2)$, with cardinality $k_r = \frac{1}{4} r^{2(1-\beta)}$. Denote also $B_j := B(x_j, \frac{r^3}{\ln r^3})$, $j = 1, \ldots, k_r$.

Before going to the heart of the matter we briefly sketch the strategy of proof (also, one may find it helpful to look at Figure ??). We start to show that at least a half of these balls $B_j$ will receive at most $b \frac{2\ln \ln r}{\ln \ln r^3}$ excursions from $\partial B_j$ to $\partial B'_j$ up to time $t_\alpha$, where $b > 1$ is a parameter (the above number of excursions is larger than the typical number of excursions by factor $b$). Moreover, using the method of soft local times (recall Section 5.1), we couple such excursions from $RI(\alpha)$ with a slightly larger number of independent excursions from the $\hat{S}$-walk: with overwhelming probability, the trace on $\bigcup_j B_j$ of the latter excursion process contains the trace of the former, so the vacant set $\mathcal{V}^\alpha$ restricted to balls $B_j$ is smaller than the set of unvisited points by the independent process. Now, by independence, it will be possible to estimate the probability for leaving that many balls partially uncovered, and this will conclude the proof.
Figure 6.8 On the proof of Theorem 6.6 (iii). With high probability, at least a positive proportion of the inner circles is not completely covered.

Let us observe that the number of $\hat{S}$-walks in $RI(\alpha)$ intersecting a given disk $B_j$ has Poisson law with parameter $\lambda = (1 + o(1))\frac{\alpha^2}{2\beta} \ln r$. Indeed, the law is Poisson by construction, the parameter $\pi \alpha \text{cap}(B_j \cup \{0\})$ is found in (6.14) and then estimated using Lemma 6.9 (i).

Next, by Lemma 6.8 (i), the probability that the walk $\hat{S}$ started from any $y \in \partial B'_j$ does not hit $B_j$ is $(1 + o(1))\frac{3 \ln \ln r}{(2 - \beta) \ln r}$. This depends on the starting point, however, after the first visit to $\partial B_j$, each $\hat{S}$-walk generates a number of excursions between $\partial B_j$ and $\partial B'_j$. 

Random interlacements
6.3 Proofs for random interlacements

which is dominated by a geometric law \( G(p') \) (supported on \( \{1, 2, 3, \ldots\} \)) with success parameter \( p' = (1 + o(1)) \frac{\ln r}{(2 - \beta) \ln r} \). Recall also that the integer part of \( \mathcal{E}(u) \) is geometric \( G(1 - e^{-u}) \). So, with \( p = -\ln(1 - p') \), the total number \( \hat{N}_\alpha^{(j)} \) of excursions between \( \partial B_j \) and \( \partial B'_j \) in \( \text{RI}(\alpha) \) can be dominated by a compound Poisson law with \( \mathcal{E}(p) \) terms in the sum with expectation

\[
\lambda p^{-1} = (1 + o(1)) \frac{2\alpha \ln^2 r}{3 \ln \ln r^\beta}.
\]

Then, using (6.37), we obtain for \( b > 1 \)

\[
\mathbb{P}\left[ \frac{\hat{N}_\alpha^{(j)}}{\hat{Q}^2} \geq b \frac{2\alpha \ln^2 r}{3 \ln \ln r^\beta} \right] \leq \exp\left( - (1 + o(1)) \left( \sqrt{b} - 1 \right)^2 \frac{2\alpha \ln r}{2 - \beta} \right)
= r^{-(1 + o(1)) (\sqrt{b} - 1)^2} \frac{2\alpha}{2 - \beta}.
\]

(6.38)

Now, let \( W_b \) be the set

\[
W_b = \left\{ j \leq k_r : \frac{\hat{N}_\alpha^{(j)}}{\hat{Q}^2} < b \frac{2\alpha \ln^2 r}{3 \ln \ln r^\beta} \right\}.
\]

Combining (6.38) with Markov inequality, we obtain

\[
\frac{k_r}{2} \mathbb{P}\left[ \left\{ 1, \ldots, k_r \right\} \setminus W_b \right] > k_r/2 \leq \mathbb{E}\left[ \left\{ 1, \ldots, k_r \right\} \setminus W_b \right]
\leq k_r r^{-1 + o(1)} (\sqrt{b} - 1)^2 \frac{2\alpha}{2 - \beta},
\]

so

\[
\mathbb{P}\left[ \left\{ 1, \ldots, k_r \right\} \setminus W_b \right] > k_r/2 \geq 1 - 2r^{-1 + o(1)} (\sqrt{b} - 1)^2 \frac{2\alpha}{2 - \beta}.
\]

(6.39)

Assume that \( 1 < b < \alpha^{-1} \) and \( \beta \in (0, 1) \) is close enough to 1, so that \( \frac{b}{2 - \beta} < 1 \). As in Section ??, we denote by \( \hat{Z}^{(1)}(j), \ldots, \hat{Z}^{(N^{(j)})}(j) \) the excursions of \( \text{RI}(\alpha) \) between \( \partial B_j \) and \( \partial B'_j \). Also, let \( \tilde{Z}^{(1)}(j), \tilde{Z}^{(2)}(j), \tilde{Z}^{(3)}(j), \ldots \) be a sequence of i.i.d. \( \hat{S} \)-excursions between \( \partial B_j \) and \( \partial B'_j \), started with the law \( \hat{h}_B \); these sequences themselves are also assumed to be independent.

Abbreviate \( m = b \frac{2\alpha \ln^2 r}{3 \ln \ln r^\beta} \). Next, for \( j = 1, \ldots, k_r \) we consider the events

\[
D_j = \left\{ \tilde{Z}^{(1)}(j), \ldots, \tilde{Z}^{(N^{(j)})}(j) \right\} \subset \left\{ \tilde{Z}^{(1)}(j), \ldots, \tilde{Z}^{(1 + m)}(j) \right\}.
\]

Lemma 6.11  It is possible to construct the excursions \( \hat{Z}^{(k)}(j), k =
Random interlacements

1, \ldots, \tilde{N}^{(j)}, (\tilde{Z}^{(k),j}, k = 1, 2, 3, \ldots), j = 1, \ldots, k_r, on a same probability space in such a way that for a fixed $C' > 0$

$$
\mathbb{P}[D^c] \leq \exp \left( -C' \frac{\ln^2 r}{(\ln \ln r)^2} \right),
$$

(6.40)

for all $j = 1, \ldots, k_r$.

Proof This coupling can be built using the method of soft local times described in Section 5.1; in this specific situation, the exposition of [17] is better suited. Roughly speaking, this method consists of using a marked Poisson point process on $(\bigcup \partial B_j) \times \mathbb{R}_+$, where the “marks” are corresponding excursions; see Section 2 of [17] for details. Then, the $\tilde{Z}$-excursions are simply the marks of the points of the Poisson processes on $\partial B_j \times \mathbb{R}_+$ ordered according to the second coordinate, so they are independent by construction. The $\hat{Z}$-excursions are also the marks of the points of these Poisson processes, but generally taken in a different order using a special procedure; Figure 1 of [17] is, hopefully, self-explanatory and may provide some quick insight.

The inequality (6.40) then follows from Lemma 2.1 of [17] (observe that, by Lemma 6.10, the parameter $v$ in Lemma 2.1 of [17] can be anything exceeding $O(1/\ln^2 r)$, so we choose e.g. $v = (\ln \ln r)^{-1}$; also, that lemma is clearly valid for $\hat{S}$-excursions as well).

We continue the proof of part (iii) of Theorem 6.6. Define

$$
D = \bigcap_{j \leq k_r} D_j;
$$

using (6.40), we obtain by the union bound the subpolynomial estimate

$$
\mathbb{P}[D^c] \leq \frac{1}{4} r^{2(1-\beta)} \exp \left( -C' \frac{\ln^2 r}{(\ln \ln r)^2} \right).
$$

(6.41)

Let $\delta > 0$ be such that $(1 + \delta) \frac{\ln r}{\ln \ln r} < 1$. Define the events

$$
\tilde{G}_j = \{ B'_j \text{ is completely covered by } \tilde{Z}^{(1),j} \cup \cdots \cup \tilde{Z}^{(1+\delta)m),j} \}.
$$

Then, for all $j \leq k_r$ it holds that

$$
\mathbb{P}[\tilde{G}_j] \leq \frac{1}{5}
$$

(6.42)
6.3 Proofs for random interlacements

for all large enough $r$. Indeed, if the $\tilde{Z}$’s were independent SRW-excursions, the above inequality (with any fixed constant in the right-hand side) could be obtained as in the proof of Lemma 3.2 of [17] (take $\tilde{H} = \ln n$, and $n = 3r^\beta + 1$ there). On the other hand, Lemma 4.4 (ii) implies that the first $(1 + \delta)m$ $\tilde{S}$-excursions can be coupled with SRW-excursions with high probability, so (6.42) holds for $\tilde{S}$-excursions as well.

Next, define the set

$$\tilde{W} = \{ j \leq k_r : \tilde{G}_j \text{ occurs} \}.$$ 

Since the events $(\tilde{G}_j, j \leq k_r)$ are independent, by (6.42) we have (recall that $k_r = \frac{1}{4} r^{2(1-\beta)}$)

$$\mathbb{P}\left[ |\tilde{W}| \geq \frac{3}{5} k_r \right] \geq 1 - \exp \left( - C r^{2(1-\beta)} \right)$$

for all $r$ large enough.

Observe that, by construction, on the event $D$ we have $\mathcal{V}^\alpha \cap B'_j \neq \emptyset$ for all $j \in \tilde{W} \cap W_b$. So, using (6.39), (6.43), and (6.41), we obtain

$$\mathbb{P}\left[ |\mathcal{V}^\alpha \cap (B(r) \setminus B(r/2))| = 0 \right] \leq \mathbb{P}\left[ |W_b| < \frac{k_r}{2} \right] + \mathbb{P}[D^c] + \mathbb{P}\left[ |\tilde{W}| < \frac{3k_r}{5} \right]$$

$$\leq 2r^{-(1+o(1))(\sqrt{5}-1)^2} + \mathbb{P}\left[ |\tilde{W}| < \frac{3k_r}{5} \right] + \exp \left( - C r^{2(1-\beta)} \right) + \frac{1}{4} r^{2(1-\beta)} \exp \left( - C' \frac{\ln^2 r}{(\ln \ln r)^2} \right).$$

Since $b \in (1, \alpha^{-1})$ can be arbitrarily close to $\alpha^{-1}$ and $\beta \in (0, 1)$ can be arbitrarily close to 1, this concludes the proof of (6.24).

**Proof of Theorem 6.6 (ii)** To complete the proofs in Section 6.2, it remains to prove that $|\mathcal{V}^\alpha| < \infty$ a.s. for $\alpha > 1$. First, we establish the following elementary fact. For $x \in \partial B(2r)$ and $y \in B(r) \setminus B(r/2)$, it holds

$$\tilde{P}_x \left[ \tilde{\tau}_i(B(y)) < \tilde{\tau}_i(B(r \ln r)) \right] = \frac{\ln \ln r}{\ln r} (1 + o(1)).$$

(6.44)

Indeed, define the events

$$G_0 = \{ \tau_i(B(0)) < \tau_i(B(r \ln r)) \},$$

$$G_1 = \{ \tau_i(B(y)) < \tau_i(B(r \ln r)) \};$$
then, Lemma 3.2 implies that
\[ P_x[G_0] = \frac{\ln \ln r}{\ln r} (1 + o(1)) = P_x[G_1](1 + o(1)). \]

Observe that
\[ P_x[G_0 \cap G_1] = P_x[G_0 \cup G_1]P_x[G_0 \cap G_1 | G_0 \cup G_1] \leq (P_x[G_0] + P_x[G_1])(P_y[G_1] + P_y[G_0]) \leq 4\left(\frac{\ln \ln r}{\ln r}\right)^2 (1 + o(1)). \]

So,
\[ P_x[G_1 | G_0] = \frac{P_x[G_1] - P_x[G_0 \cap G_1]}{1 - P_x[G_0]} = \frac{\ln \ln r}{\ln r}(1 + o(1)), \]

and we use Lemma 4.4 to conclude the proof of (6.44).

Now, the goal is to prove that, for \( \alpha > 1 \)
\[ \mathbb{P}[\text{there exists } y \in B(r) \setminus B(r/2) \text{ such that } y \in V^\alpha] \leq r^{-\frac{\alpha}{2}(1-\alpha^{-1})^2(1+o(1))}. \] (6.45)

This would clearly imply that the set \( V^\alpha \) is a.s. finite, since
\[ \{V^\alpha \text{ is infinite}\} = \{V^\alpha \cap (B(2^n) \setminus B(2^{n-1})) \neq \emptyset \text{ for infinitely many } n\}, \] (6.46)

and the Borel-Cantelli lemma together with (6.45) imply that the probability of the latter event equals 0.

Let \( N_{\alpha,r} \) be the number of \( \tilde{S} \)-excursions of \( \mathcal{RI}(\alpha) \) between \( \partial B(r) \) and \( \partial B(r \ln r) \). Analogously to (6.38) (using Lemma 3.2 in place of Lemma 6.8 (i)), it is straightforward to show that, for \( b < 1 \),
\[ \mathbb{P}\left[N_{\alpha,r} \leq b \frac{2 \alpha \ln^2 r}{\ln \ln r}\right] \leq r^{-2\alpha(1-\sqrt{b})^2(1+o(1))}. \] (6.47)

Now, (6.44) implies that for \( y \in B(r) \setminus B(r/2) \)
\[ \mathbb{P}\left[y \text{ is uncovered by first } b \frac{2 \alpha \ln^2 r}{\ln \ln r} \text{ excursions}\right] \leq (1 - \frac{\ln \ln r}{\ln r}(1 + o(1)))^{b \frac{2 \alpha \ln^2 r}{\ln \ln r}} = r^{-2\alpha(1+o(1))}, \] (6.48)
so, using the union bound,
\[
P\left[ \exists y \in B(r) \setminus B(r/2) : y \in \mathcal{V}^{\alpha}, N_{\alpha,r} > b \frac{2\alpha \ln^2 r}{\ln r} \right] \leq r^{-2(b\alpha - 1)(1+o(1))}.
\]
(6.49)

Using (6.47) and (6.49) with \( b = \frac{1}{4} \left( 1 + \frac{1}{\alpha} \right)^2 \) we conclude the proof of (6.45) and of Theorem 6.6 (ii).

\[ \square \]

**6.3.4 Proof of Theorem 6.7**

Let us first give a more detailed heuristic argument for (6.25). As usual, we consider the *excursions* of the random walk \( X \) between \( \partial B(\frac{n}{3} \ln n) \) and \( \partial B(n/3) \) up to time \( t_\alpha \). Recall that \( N_\alpha \) denotes the number of these excursions. Lemma ?? shows that this (random) number is concentrated around \( 2^{\alpha \ln^2 n} \), with deviation probabilities of subpolynomial order. This is for unconditional probabilities, but, since the probability of the event \( \{ 0 \in U_{t_\alpha}^{(n)} \} \) is only polynomially small (actually, it is \( n^{-2\alpha + o(1)} \)), the same holds for the deviation probabilities conditioned on this event. So, let us just assume for now that the number of the excursions is *exactly* \( 2^{\alpha \ln^2 n} \) a.s., and see where will it lead us.

Assume without restriction of generality that \( 0 \in A \). Then, Lemmas ?? and ?? imply that

- the probability that an excursion hits the origin is roughly \( \ln^2 n \ln(n/3) \);
- provided that \( \text{cap}(A) \ll \ln n \), the probability that an excursion hits the set \( A \) is roughly \( \frac{\ln n}{\ln(n/3)} \left( 1 + \frac{\pi \text{cap}(A)}{2 \ln(n/3)} \right) \).

So, the *conditional* probability \( p_* \) that an excursion does not hit \( A \) given that it does not hit the origin is
\[
p_* \approx 1 - \frac{\ln n}{\ln(n/3)} \left( 1 + \frac{\pi \text{cap}(A)}{2 \ln(n/3)} \right) \approx 1 - \frac{\pi \ln n}{2 \ln^2 n} \text{cap}(A),
\]
and then we obtain
\[
P[\Upsilon_{t_\alpha} \cap U_{t_\alpha}^{(n)} \mid 0 \in U_{t_\alpha}^{(n)}] \approx p_*^{N_{\alpha}} \approx \left( 1 - \frac{\pi \ln n}{2 \ln^2 n} \text{cap}(A) \right)^{2n \ln^2 n} \approx \exp \left( - \pi \alpha \text{cap}(A) \right),
\]
which agrees with the statement of Theorem 6.7.

However, turning the above heuristics to a rigorous proof is not an easy task. The reason for this is that, although $N_\alpha$ is indeed concentrated around $\frac{2\alpha \ln^2 n}{\ln n}$, it is not concentrated enough: the probability that 0 is not hit during $k$ excursions, where $k$ varies over the “typical” values of $N_\alpha$, changes too much. Therefore, in the proof of Theorem 6.7 we take a different route by considering the suitable $h$-transform of the walk, as explained below.

Define for $x \in \mathbb{Z}^2_n$ and any $t$

$$h(t, x) = P_x[T_n(0) > t]$$

(so that $h(t, x) = 1$ for $t < 0$). To simplify the notations, let us also assume that $t_\alpha$ is integer. We will represent the conditioned random walk as a time-dependent Markov chain, using the Doob’s $h$-transform. Indeed, it is well known and easily checked that the simple random walk on $\mathbb{Z}^2_n$ conditioned on the event $\{0 \in U_t(n)\}$ is a time-dependent Markov chain $\tilde{X}$ with transition probabilities given by

$$P[\tilde{X}_{s+1} = y \mid \tilde{X}_s = x] = \frac{h(t_\alpha - s - 1, y)}{h(t_\alpha - s, x)} \times \frac{1}{4},$$

(6.50)

if $x$ and $y$ are neighbours, and equal to 0 otherwise. For simpler notations, we do not indicate the dependence on $t_\alpha$ in the notation $\tilde{X}$.

### 6.4 Exercises

**Exercise 6.1** Prove the claim of Remark 4.19 by proving that (analogously to ???), with any fixed $\delta > 0$,

$$\mathbb{P}[nG \cap \mathbb{Z}^d \subset \widehat{S}_{[0, \infty]}] \leq n^{-2+\delta}$$

for all large enough $n$; the claim then would follow from the (first) Borel-Cantelli lemma.
Hints and solutions to selected exercises

Exercise 2.2.
*Unfortunately*, not. If $S_n$ is a $d$-dimensional SRW, then $P_0[S_2 = 0] = (\frac{1}{2d} \times \frac{1}{2d}) \times (2d) = \frac{1}{2d}$. If we want the claim in the exercise to be true, this would then mean that $\frac{1}{2d} = (\frac{1}{2})^d$ or $d = 2^{d-1}$, which holds only for $d = 1, 2$.

Exercise 2.3.
You may find it useful to read [26].

Exercise 2.6.
Set $\pi(0) = 1$, and then prove by induction that
$$\pi(x) = \frac{q_0 \cdots q_{x-1}}{p_1 \cdots p_x}$$
for $x > 0$ (and analogously for $x < 0$).

Exercise 2.7.
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 2.8.
Hint: define a scalar product $\langle f, g \rangle = \sum_{x \in \Sigma} \pi(x)f(x)g(x)$ and show that $P$ is self-adjoint.

Exercise 2.9.
Fix an arbitrary $x_0 \in \Sigma$, set $A = \{x_0\}$, and
$$f(x) = P_x[\tau_{x_0} < \infty] \text{ for } x \in \Sigma$$
(so, in particular, $f(x_0) = 1$). Then (2.12) 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 2.10.
Let $p = p(n, n+1)$ (for all $n$), and assume for definiteness that $p > \frac{1}{2}$. Consider the function $f(x) = \left(\frac{1-x}{p}\right)^x$ and the set $A = (-\infty, 0]$; then use Theorem 2.5.

Note also that, for proving that this random walk is transient, one may also use Theorem 2.5.15 of [52] (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 2.11.
Hint: use the Lyapunov function $f(x) = (1-\delta)^x$ for small enough $\delta > 0$.

Exercise 2.12.
Quite analogously to (2.15)–(2.17), it is elementary to obtain for $f(x) = \|x\|^{-\alpha}$

$$E[f(X_{n+1})-f(X_n) \mid X_n = x] = -\alpha\|x\|^{-\alpha-2} \left(\frac{1}{2} - \left(1 + \frac{\alpha}{2}\right) \frac{1}{d} + O(\|x\|^{-1})\right).$$

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

By the way, in your opinion, is it surprising that the “critical” value for $\alpha$ is equal to $d - 2$? (To answer this question, it is maybe a good idea to read Section 3.1.)

Exercise 2.14.
Hint: being $X_n$ the two-dimensional walk, define first its covariance matrix by $M := E_0((X_1)\top X_1)$. Find a suitable linear transformation$^4$ 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 2.17 (a).
Fix an arbitrary $x_0 \in \Sigma$, and set $A = \{x_0\}$. Observe that, for $x \neq x_0$,

$$E_x\tau_{x_0} = \sum_{y \in \Sigma} p(x, y) E_y(1 + \tau_{x_0}),$$

$^4$ why does it exist?
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 (2.20)–(2.21) with \( \varepsilon = 1 \).

Exercise 2.18.

Note that the calculation (2.15) is dimension-independent, and (2.16) remains valid as well, with obvious changes. Then obtaining (2.22) is straightforward (use (2.15) with \( \alpha = 1 \) and observe that the factor \( \frac{1}{4} \) in the next display after (2.15) will become \( \frac{1}{2d} \) in the general case). As for (2.23), show first that
\[ \mathbb{E}_x (\|S_1\|^2 - \|x\|^2) = 1 \text{ for all } x \in \mathbb{Z}^d, \]
and then use (2.22) together with the identity \( (b - a)^2 = b^2 - a^2 - 2a(b - a) \) with \( a = \|x\|, b = \|S_1\| \).

Exercise 2.19.

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 \( \alpha \) will depend on \( \varepsilon \) in (b) and (c)!

Exercise 2.20.

Hint: use Exercise 2.18.

Exercise 3.3.

Hint: first, prove that the (multivariate) characteristic function of \( S_n \) (starting at the origin) equals \( \Phi^n(\theta) \). Then, think how to extract \( \mathbb{P}_0[S_n = x] \) from there.

Exercise 3.5.

See Exercise 6.18 of [46].

Exercise 3.6.

Hint: use Exercise 2.20.
Exercise 3.7.

For \( x \notin A \), using the definition of \( G \) and the strong Markov property, let us write,

\[
G(x, A) = \sum_{y \in A} \mathbb{P}_x[\tau_A < \infty, S_{\tau_A} = y] G(y, A)
\]

\[
\geq \left( \sum_{y \in A} \mathbb{P}_x[\tau_A < \infty, S_{\tau_A} = y] \right) \times \min_{y \in A} G(y, A)
\]

\[
= \mathbb{P}_x[\tau_A < \infty] \min_{y \in A} G(y, A),
\]

and, analogously,

\[
G(x, A) \leq \mathbb{P}_x[\tau_A < \infty] \max_{y \in A} G(y, A),
\]

so

\[
\frac{G(x, A)}{\max_{y \in A} G(y, A)} \leq \mathbb{P}_x[\tau_A < \infty] \leq \frac{G(x, A)}{\min_{y \in A} G(y, A)}.
\]

(6.51)

From the asymptotic expression (3.5) for the Green’s function it is straightforward to obtain that

\[
\lim_{x \to \infty} \frac{\| x \|_\gamma^d}{\gamma^d} G(x, A) = |A|,
\]

so multiplying (6.51) by \( \frac{\| x \|_\gamma^d}{\gamma^d} \) and using Proposition 3.4 we obtain the desired result.

Exercise 3.8.

First of all, note that \( \text{Es}_A \in \mathcal{K}_A \) by (3.11), so it is enough to prove that \( \sum_{x \in A} h(x) \leq \text{cap}(A) \) for any \( h \in \mathcal{K}_A \). Let us abbreviate \( v_A(x) = \mathbb{P}_x[\tau_A < \infty] \), so that (in matrix notation) (3.11) becomes

\[
v_A = G \text{Es}_A.
\]

Now, for two (nonnegative) functions \( f, g : \mathbb{Z}_d \to \mathbb{R} \), define the usual scalar product by \( (f, g) = \sum_{x \in \mathbb{Z}_d} f(x) g(x) \). We argue that, by symmetry, \( G \) is a self-adjoint linear operator in the sense that \( (Gf, g) = (f, Gg) \). Then, for any \( h \in \mathcal{K}_A \), write (being \( 1 \) the function with value 1 at all \( x \))

\[
\sum_{x \in A} h(x) = (h, v_A) = (h, G \text{Es}_A) = (Gh, \text{Es}_A) \leq (1, \text{Es}_A) = \text{cap}(A),
\]

and we are done.
Exercise 3.11.
Use the second inequality in (3.14) to obtain that, when the series in (3.19) converges, it holds that
\[ \sum_{k=1}^{\infty} P_0[\tau_{A_k} < \infty] < \infty; \]
then use Borel-Cantelli.

Exercise 3.12.
Analogously to the above, it is straightforward to obtain that, in this case,
\[ \sum_{k=1}^{\infty} P_0[\tau_{A_k} < \infty] = \infty, \]
but obtaining the recurrence of \( A \) from this is not immediate (note that the events \( \{\tau_{A_k} < \infty\}, k \geq 1 \), need not be independent). One possible workaround is the following:
(i) Divide the sum in (3.19) into four sums
\[ \sum_{j=1}^{\infty} \frac{\text{cap}(A_{4j+i})}{2^{(d-2)(4j+i)}}, \quad i = 0, 1, 2, 3; \]
note that at least one of these sums must diverge. For definiteness, assume that it is the first one (i.e., with \( i = 0 \)).
(ii) For \( j \geq 2 \), consider any \( y \in \partial B(2^{4j-2}) \), and, using the first inequality in (3.14) together with (3.5), show that \( \mathbb{P}_y[\tau_{A_{4j}}] \geq c_1 2^{-(d-2)(4j)} \text{cap}(A_{4j}) \).
(iii) Next, for any \( z \in \partial B(2^{4j+2}) \) show that
\[ \mathbb{P}_y[\tau_{A_{4j}}] \leq c_2 \frac{\text{cap}(A_{4j})}{2^{4j(d-2)}}, \]
where \( c_2 < c_1 \). Note that \( A_{4j} \) is much closer to \( \partial B(2^{4j-2}) \) than to \( \partial B(2^{4j+2}) \) – consider drawing a picture to see it better!
(iv) Using (ii)–(iii), obtain that, for any \( y \in \partial B(2^{4j-2}) \),
\[ \mathbb{P}_y[\tau_{A_{4j}} < \tau_{\partial B(2^{4j+2})}] \geq c_3 \frac{\text{cap}(A_{4j})}{2^{4j(d-2)}}, \]
i.e., on its way from \( \partial B(2^{4j-2}) \) to \( \partial B(2^{4j+2}) \), the walker will hit \( A_{4j} \) with probability at least of order \( 2^{-4j(d-2)} \text{cap}(A_{4j}) \).
(v) Argue that the “if” part of Theorem 3.7 follows from (iv) in a straightforward way.

Of course, it is probably possible to construct many such examples; one possibility (in three dimensions) is to consider a set of the form \( \{ b_k e_1, k \geq 1 \} \), where \( (b_k) \) is a strictly increasing sequence of positive integer numbers. Then, it is clear that the expected number of visits to this set is infinite if and only if \( \sum_{k=1}^{\infty} \frac{1}{b_k} = \infty \). As shown in the paper [9], it is possible to construct a transient set for which the above series sums to infinity.

Exercise 3.18.  
Proof 1: note that (3.69) is equivalent to
\[
\text{cap}(A) = \frac{|A|}{\sum_{y \in A} \text{hm}_A(y)G(y, A)};
\]
proceed as in the solution of Exercise 3.7\(^5\) and then use basic properties of the harmonic measure.  
Proof 2: note that (3.69) is equivalent to
\[
\sum_{y \in A} \text{Es}_A(y)G(y, A) = |A|,
\]
and then obtain it directly from (3.11) by summing in \( y \in A \).

Exercise 3.19.  
Hint: recall the calculation after (3.17).

Exercises 3.20 and 3.21.  
See the beginning of Section 6.1.

Exercise 3.22.  
First, recall (3.24) — we now need to do a finer analysis of it. Let us rewrite (3.24) in the following way:
\[
\mathbb{P}_x[S_{\tau_A} = y \mid \tau_A < \infty]
\]
\(^5\) I mean, only as in the first line of the first display there
Hints and solutions to selected exercises

\[
= \frac{G(y, x) \text{ES}_A(y)}{\mathbb{P}_x[\tau_A < \infty]} + \sum_{z \in \partial A} \mathbb{P}_y[\tau^+_A < \infty, S_{\tau^+_A} = z] (G(y, x) - G(z, x)) \frac{\mathbb{P}_x[\tau_A < \infty]}{\mathbb{P}_x[\tau_A < \infty]}.
\]

(6.52)

It is not difficult to check that, for distinct \(x, y, z\) (think about the situation when \(x, y, z \in \mathbb{Z}^d\) are such that \(z\) is much closer to \(y\) than \(x\) is)

\[
G(x, y) = G(x, z) \left(1 + O\left(\frac{\|y - z\|}{\|x - y\|}\right)\right).
\]

(6.53)

So, (3.14) implies that

\[
\mathbb{P}_x[\tau_A < \infty] = \text{cap}(A) G(y, x) \left(1 + O\left(\frac{\text{diam}(A)}{\text{dist}(x, A)}\right)\right),
\]

(6.54)

which means that the first term in the right-hand side of (6.52) equals \(\text{lim}_A(y) \left(1 + O\left(\frac{\text{diam}(A)}{\text{dist}(x, A)}\right)\right)\).

To deal with the second term, denote \(V = \partial \mathcal{B}(y, 2 \text{diam}(A) + 1)\); by e.g. (3.8), there exists \(c > 0\) such that \(\mathbb{P}_z[\tau_A = \infty] \geq c\) for all \(z \in V\), and this permits us to obtain\(^6\) that

\[
\mathbb{P}_y[\tau_V < \tau^+_A] \leq c^{-1} \text{ES}_A(y).
\]

(6.55)

We then use the Optional Stopping Theorem with the martingale

\[M_n = G(y, x) - G(S_n \wedge \tau_V, x)\]

and the stopping time \(\tau = \tau^+_A \wedge \tau_V\) to obtain that\(^7\)

\[0 = \mathbb{E}_y M_0 = \mathbb{E}_y M_\tau\]

(note that \(\mathbb{P}_y[\tau^+_A < \tau_V] = \mathbb{P}_y[\tau^+_A < \infty] - \mathbb{P}_y[\tau^+_A > \tau_V, \tau^+_A < \infty]\))

\[
= \sum_{z \in \partial A} \mathbb{P}_y[\tau^+_A < \infty, S_{\tau^+_A} = z] (G(y, x) - G(z, x)) - \mathbb{P}_y[\tau^+_A > \tau_V, \tau^+_A < \infty] (G(y, x) - \mathbb{E}_y(G(S_{\tau_V}) | \tau^+_A > \tau_V, \tau^+_A < \infty)) + \mathbb{P}_y[\tau_V < \tau^+_A] (G(y, x) - \mathbb{E}_y(G(S_{\tau^+_A}) | \tau_V < \tau^+_A)).
\]

(6.56)

The first term in the right-hand side of (6.56) is what we need to estimate for (6.52), and, by (6.53) and (6.55), the second and third terms are both \(O\left(\text{ES}_A(y) G(y, x) \frac{\text{diam}(A)}{\text{dist}(x, A)}\right)\). Gathering the pieces, we obtain (3.70).

\(^6\) please, elaborate

\(^7\) of course, we also assume that \(x\) is outside \(\mathcal{B}(y, 2 \text{diam}(A) + 1)\)
Hints and solutions to selected exercises

Exercise 3.23.

Start with almost (3.24):

\[
\mathbb{P}_x[\tau_A < \infty, S_{\tau_A} = y] = G(y, x) - \sum_{z \in \partial A} \mathbb{P}_y[\tau_A^+ < \infty, S_{\tau_A^+} = z]G(z, x).
\]

Then, consider the martingale \(G(S_{\tau \wedge \tau_A}, x)\) and use the Optional Stopping Theorem with the (possibly infinite\(^8\)) stopping time \(\tau \wedge \tau_A^+\) to obtain that

\[
G(y, x) = \mathbb{E}_y G(S_{\tau \wedge \tau_A^+}, x)
= \sum_{z \in \partial A} \mathbb{P}_y[\tau_A^+ < \tau_x, S_{\tau_A^+} = z]G(z, x) + \mathbb{P}_y[\tau_x < \tau_A^+]G(0).
\]

So, using that

\[
\mathbb{P}_y[\tau_A^+ < \tau_x, S_{\tau_A^+} = z]
= \mathbb{P}_y[\tau_A^+ < \infty, S_{\tau_A^+} = z] - \mathbb{P}_y[\tau_x < \tau_A^+ < \infty, S_{\tau_A^+} = z]
= \mathbb{P}_y[\tau_A^+ < \infty, S_{\tau_A^+} = z] - \mathbb{P}_y[\tau_x < \tau_A^+]\mathbb{P}_x[\tau_A^+ < \infty, S_{\tau_A^+} = z],
\]

we obtain

\[
\mathbb{P}_x[\tau_A < \infty, S_{\tau_A} = y]
= \mathbb{P}_y[\tau_x < \tau_A^+]\left(G(0) - \sum_{z \in \partial A} \mathbb{P}_x[\tau_A^+ < \infty, S_{\tau_A^+} = z]G(z, x)\right).
\]

The second term in the above parentheses is clearly negative, and one can use (3.5) and (3.8) to obtain its magnitude and therefore show (3.71). To prove the second part, write

\[
\mathbb{P}_y[\tau_x < \infty, \tau_A^+ = \infty] = \mathbb{P}_y[\tau_x < \tau_A^+]\mathbb{P}_x[\tau_A = \infty],
\]

and then use (3.70) together with (3.4) and (3.8).

Exercise 3.25.

(i)–(ii) See Figure 6.9 (the left part for (i) and the right part for (ii)). Specifically, consider first a path starting from the origin and ending on the first hitting of the diagonal (respectively, vertical) line that separates the origin from \(e_1 + e_2\) (respectively, \(x_1\) from \(x_2\)). From the site \(y\) (respectively, \(z\)) where it ends, it is

\(^8\) we leave as an exercise checking that, in this case, one can still apply the Optional Stopping Theorem; note that \(G(S_\infty, x) = 0\) by transience.
equally probable to go to 0 and $x_1$ (respectively, to $x_1$ and $x_2$). There are also paths that go from 0 to 0 (respectively, from 0 to $x_1$) which do not cross that line at all; this implies that the inequalities are strict.

(iii) Assume without restricting generality that $x$ belongs to the first quadrant. Then, for an even $x$, the fact that $\mathbb{P}_0[S_{2n} = 0] > \mathbb{P}_0[S_{2n} = x]$ follows from (i)–(ii) by induction (there is a chain of sites that goes either from 0 or to $e_1 + e_2$ to $x$ with “steps” $2e_1$ and $2e_2$). As for the case of an odd $x$ (where we need to prove that $\mathbb{P}_0[S_{2n} = 0] > \mathbb{P}_0[S_{2n+1} = x]$), use the fact that

$$\mathbb{P}_0[S_{2n+1} = x] = \frac{1}{4} \sum_{y \sim x} \mathbb{P}_0[S_{2n} = y]$$

and note that all $y$’s in the above sum are even.

Exercise 3.27.

Hint: prove that $\tau_{\partial \Lambda}$ has exponentially small tails using an argument of the sort “from any $y \in \Lambda$ the walker can go out of $\Lambda$ in at most $n_0$ steps with probability at least $p_0 > 0$”, with some explicit $n_0$ and $p_0$.

Exercise 3.30.

Indeed,

$$Y_{n \wedge \tau_{\partial \Lambda}} = a(S_{n \wedge \tau_{\partial \Lambda}} - y) - N_y^{(n \wedge \tau_{\partial \Lambda} - 1)}$$

is still a martingale, so $a(x - y) = \mathbb{E}_x Y_{n \wedge \tau_{\partial \Lambda}}$ for all $n$. Now, $\mathbb{E}_x a(S_{n \wedge \tau_{\partial \Lambda}} - y) \to \mathbb{E}_x a(S_{\tau_{\partial \Lambda}} - y)$ by the Dominated Conver-
gence Theorem, and \( \mathbb{E}_x N_y^{(n \wedge \tau_{\mathbb{B}(\lambda)} - 1)} \to \mathbb{E}_x N_y^{(\tau_{\mathbb{B}(\lambda)} - 1)} \) by the Monotone Convergence Theorem.

Exercise 3.31.

The idea is to use Theorem 3.13 with \( \Lambda = \mathbb{B}(R) \setminus \{0\} \), and then send \( R \) to infinity. Now, instead of just writing the arguments in a detailed way, let me show how one proves such statements in practice. Lemma 3.12 implies that

\[
\mathbb{E}_x \eta_x \approx G_\Lambda(x, x) \approx a(x) \times \left(1 - \frac{a(x)}{a(R)}\right) + a(R) \times \frac{a(x)}{a(R)} \approx 2a(x).
\]

For the second statement, write

\[
\mathbb{E}_0 \eta_x = \frac{1}{4} \sum_{y \in \{\pm e_1, \pm e_2\}} \mathbb{E}_y \eta_x
\]

\[
\approx \frac{1}{4} \sum_{y \in \{\pm e_1, \pm e_2\}} G_\Lambda(y, x)
\]

\[
\approx \frac{1}{4} \sum_{y \in \{\pm e_1, \pm e_2\}} \left(a(x) \times \left(1 - \frac{1}{a(R)}\right) + a(R) \times \frac{1}{a(R)} - a(x - y)\right)
\]

\[
\approx a(x) + 1 - \frac{1}{4} \sum_{z \sim x} a(z)
\]

(since \( a \) is harmonic outside the origin)

\[= 1.\]

Inserting suitable \( O \)'s and formally passing to the limits as \( R \to \infty \) is really left as an exercise.

Exercise 3.32.

Indeed, conditioning on the location of the first entrance to \( \mathbb{B} \), we have

\[
\mathbb{P}_x[S_{\tau_A} = y] = \sum_{z \in \mathbb{B}} \mathbb{P}_x[S_{\tau_B} = z] \mathbb{P}_z[S_{\tau_A} = y].
\]

Theorem 3.16 then implies that, as \( x \to \infty \), the left-hand side converges to \( hm_A(y) \), and the right-hand side converges to

\[
\sum_{z \in \mathbb{B}} hm_B(z) \mathbb{P}_z[S_{\tau_A} = y] = \mathbb{P}_{hm_B}[S_{\tau_A} = y],
\]
as required.

Exercise 3.36.
One method that would probably work is to first approximate
the SRW with the two-dimensional Brownian Motion (possibly
using a KMT-like strong approximation theorem), and then find a
suitable M"obius transform that sends the domain on Figure 3.7 to
an annulus formed by concentric circumferences (and then use the
conformal invariance of Brownian trajectories). The author has to
confess that he did not do any concrete calculations because he
never needed such a result, but, nevertheless, it seems that this
program looks reasonably fine.

Exercise 4.6.
Hint: recall (3.49), the definition of $q_A$.

Exercise 4.10.
This is an easy consequence of the fact that the transition proba-
bilities from $x$ for the conditioned walk converge to those for the
SRW as $x \to \infty$.

Exercise 4.11.
Hint: look at Figure 6.10. Use the fact that the projection of the
SRW on any fixed direction is a martingale, and then use the
Optional Stopping Theorem.
Alternatively: just notice that $\|S_{n \wedge \tau_{y_0}} - y_0\|$ is a submartingale.

Exercise 4.12.
Assume without restricting generality that $\|x\| \geq \|y\|$ (recall that
$\hat{g}(x, y) = \hat{g}(y, x)$), and consider the following two cases.
Case 1: $\|y\| > \|x\|^{1/2}$. In this case $a(x)$ and $a(y)$ are of the same
order, and, since $\|x - y\| \leq 2\|x\|$, due to (3.36), $a(x - y) - a(x)$
is bounded above by a positive constant; therefore, the expres-
sion $a(x) + a(y) - a(x - y)$ will be of order $\ln \|x\|$. This implies
that $\hat{g}(x, y)$ will be of order $\frac{1}{\ln \|x\|}$ indeed.
Case 2: $\|y\| \leq \|x\|^{1/2}$. Here, (3.38) implies that $a(x) - a(x - y) =
O\left(\frac{\|x\|^{1/2}}{\|x\|} \right) = O(\|x\|^{-1/2})$, so
$$
\hat{g}(x, y) = \frac{a(y) + O(\|x\|^{-1/2})}{a(x)a(y)} = \frac{1}{a(x)} \left(1 + O\left(\frac{1}{\|x\|^{1/2} \ln(1 + \|y\|)}\right)\right),
$$
and this again implies (4.85).

Exercise 4.16.

(i) Think about $A_n = B(r_ne_1, n)$, where $r_n \to \infty$ very quickly. Then, typically, if the conditioned walk ever visits $r_ne_1$, it will cover the whole $A_n$ by “local recurrence” (recall that, when far away from the origin, the walk $\hat{S}$ resembles very much the SRW by Lemma 4.5).

(ii) Think about

$$A_n = \{r_1e_1, r_2e_1, \ldots, r_ne_1\},$$

where $(r_n)$ is as above\(^9\), and apply the reasoning similar to that used in the proof of Theorem 4.11.

Exercise 4.17.

This is open, I think (though should not be very difficult).

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

\(^9\) the author thanks Hubert Lacoin for suggesting this example
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 $Q$ between the Markov chain $(X_i)$ and an i.i.d. collection $(Y_i)$ (with law $\pi$), in such a way that for any $\lambda > 0$ and $t \geq 0$,

$$Q[\{Y_1, \ldots, Y_R\} \subset \{X_1, \ldots, X_t\}] \geq P_{\pi_0} \left[ \xi_0 \pi_0(x) + \sum_{j=1}^{t-1} \xi_j p(X_j, x) \geq \lambda \pi(x), \text{ for all } x \in \Sigma \right] (6.58)$$

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

Exercise 5.3.
If you find one, please, let me know.

Exercise 5.6.
To the second one. See also [39] for a more complete discussion of this.

Exercise 5.10.
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 5.15.
Answer: $3/2$. It is the last problem of the famous Mathematical Trivium [2] 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 three-dimensional convex set, not only for the cube. One may reason in the following way:
• imagine the surface of a convex set 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 set times a constant (that is, it does not depend on the surface’s shape itself!);
• to obtain this constant, consider a certain special convex set whose projections are the same in all directions (finding such a set is left as an exercise).
References


References


References


### Index

<table>
<thead>
<tr>
<th>Term</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>adapted process</td>
<td>4</td>
</tr>
<tr>
<td>Bessel process</td>
<td>78</td>
</tr>
<tr>
<td>Binomial process</td>
<td>135</td>
</tr>
<tr>
<td>capacity</td>
<td></td>
</tr>
<tr>
<td>for SRW in $d \geq 3$</td>
<td>39</td>
</tr>
<tr>
<td>for the conditioned SRW</td>
<td>90</td>
</tr>
<tr>
<td>for two-dimensional SRW</td>
<td>65</td>
</tr>
<tr>
<td>of $d$-dimensional ball</td>
<td>41</td>
</tr>
<tr>
<td>of a disk</td>
<td>66</td>
</tr>
<tr>
<td>of one- and two-point sets</td>
<td>41, 66</td>
</tr>
<tr>
<td>of three-point sets</td>
<td>72</td>
</tr>
<tr>
<td>relationship of two-dimensional capacities</td>
<td>90</td>
</tr>
<tr>
<td>Chapman-Kolmogorov equation</td>
<td>7</td>
</tr>
<tr>
<td>conductance</td>
<td>16</td>
</tr>
<tr>
<td>cosine reflection law</td>
<td>138, 143</td>
</tr>
<tr>
<td>coupling</td>
<td>35, 48-51, 108, 123, 124, 127, 128, 133, 135, 156, 166, 176, 193</td>
</tr>
<tr>
<td>cover time</td>
<td>128</td>
</tr>
<tr>
<td>decoupling</td>
<td>13, 31, 71, 145, 156</td>
</tr>
<tr>
<td>Doob’s $h$-transform</td>
<td>75, 76</td>
</tr>
<tr>
<td>drift (with respect to a Lyapunov function)</td>
<td>23</td>
</tr>
<tr>
<td>effective resistance</td>
<td>16</td>
</tr>
<tr>
<td>electric network</td>
<td>16</td>
</tr>
<tr>
<td>empirical processes</td>
<td>128</td>
</tr>
<tr>
<td>escape probability</td>
<td></td>
</tr>
<tr>
<td>for SRW in $d \geq 3$</td>
<td>39</td>
</tr>
<tr>
<td>for the conditioned SRW</td>
<td>87</td>
</tr>
<tr>
<td>filtration</td>
<td>4</td>
</tr>
<tr>
<td>Foster’s criterion</td>
<td>32</td>
</tr>
<tr>
<td>Foster-Lyapunov theorem</td>
<td>32</td>
</tr>
<tr>
<td>Gambler’s Ruin Problem</td>
<td>74</td>
</tr>
<tr>
<td>Green’s function</td>
<td></td>
</tr>
<tr>
<td>asymptotic behaviour</td>
<td>37, 59, 68</td>
</tr>
<tr>
<td>exact expression</td>
<td>68</td>
</tr>
<tr>
<td>for SRW in $d \geq 3$</td>
<td>36</td>
</tr>
<tr>
<td>for the conditioned SRW</td>
<td>82, 89, 115</td>
</tr>
<tr>
<td>relation to the potential kernel</td>
<td>58</td>
</tr>
<tr>
<td>restricted on a set</td>
<td>57, 58</td>
</tr>
<tr>
<td>Hölder continuity</td>
<td>131</td>
</tr>
<tr>
<td>harmonic function</td>
<td>34</td>
</tr>
<tr>
<td>harmonic measure</td>
<td></td>
</tr>
<tr>
<td>as entrance measure from infinity</td>
<td>44, 61, 63, 100</td>
</tr>
<tr>
<td>consistency</td>
<td>70, 71, 147, 148</td>
</tr>
<tr>
<td>for SRW in $d \geq 3$</td>
<td>44</td>
</tr>
<tr>
<td>for the conditioned SRW</td>
<td>95, 100</td>
</tr>
<tr>
<td>for two-dimensional SRW</td>
<td>60, 61, 63</td>
</tr>
<tr>
<td>harmonic series</td>
<td>11, 13, 19, 137</td>
</tr>
<tr>
<td>hitting time</td>
<td>5, 8</td>
</tr>
<tr>
<td>irreducibility</td>
<td>8</td>
</tr>
<tr>
<td>Kochen-Stone theorem</td>
<td>113</td>
</tr>
<tr>
<td>Lamperti problem</td>
<td>33</td>
</tr>
<tr>
<td>last-visit decomposition</td>
<td>39, 40, 87</td>
</tr>
<tr>
<td>local time</td>
<td>126, 128, 129, 132, 133</td>
</tr>
<tr>
<td>Lyapunov function</td>
<td>20, 21, 24</td>
</tr>
<tr>
<td>Lyapunov function criterion</td>
<td></td>
</tr>
<tr>
<td>for positive recurrence</td>
<td>32</td>
</tr>
<tr>
<td>for recurrence</td>
<td>20</td>
</tr>
<tr>
<td>for transience</td>
<td>22</td>
</tr>
<tr>
<td>magic lemma</td>
<td>120</td>
</tr>
<tr>
<td>Markov chain</td>
<td></td>
</tr>
<tr>
<td>aperiodic</td>
<td>8</td>
</tr>
<tr>
<td>criterion for positive recurrence</td>
<td>32</td>
</tr>
<tr>
<td>criterion for recurrence</td>
<td>20</td>
</tr>
<tr>
<td>criterion for reversibility</td>
<td>14</td>
</tr>
<tr>
<td>criterion for transience</td>
<td>22</td>
</tr>
<tr>
<td>discrete-time</td>
<td>7</td>
</tr>
<tr>
<td>invariant measure</td>
<td>8</td>
</tr>
<tr>
<td>irreducible</td>
<td>8</td>
</tr>
</tbody>
</table>
null recurrent, 8  
period, 8  
positive recurrent, 8, 32  
recurrent, 8, 9, 20  
reversible, 14  
time-homogeneous, 7  
transient, 8, 9, 22, 76  
transition probabilities, 7  
Markov property, 7  
strong, 8  
martingale, 5, 35, 36, 38, 55, 58, 59, 63, 74–76, 79, 81, 85, 98, 102, 105, 116, 187–189, 191  
convergence theorem  
for submartingales, 6  
for supermartingales, 6  
optional stopping theorem, see  
optional stopping theorem  
submartingale, 5, 58, 191  
supermartingale, 5, 21, 23, 33  
mean drift, 33  
metric space, 130  
natural filtration, 5  
optional stopping theorem  
for submartingales, 6, 7  
passage time, see hitting time  
Poisson point process, 119, 121  
Polya’s theorem, 1, 30  
polynomial class, 130  
potential kernel, 48  
asymptotic behaviour, 55  
extact expression, 71  
relation to the restricted Green’s function, 58  
Radon measure, 120  
random billiard, 143, 144  
recurrence, 9, 20, 33, 42, 44  
Reflection Principle, 51  
resampling, 134  
resistance, 16  
reversibility, 14, 79  
soft local time, 124, 126, 127, 129, 135, 157  
sprinkling, 155, 156  
Stirling’s approximation, 12  
stochastic process, 4  
stopping time, 5  
total variation distance, 132, 133  
transience, 9, 23, 33, 42, 44, 76  
transition kernel, 131  
weight (of a trajectory), 45, 46, 77, 96, 100, 107