Two-Dimensional Random Walk

The main subject of this introductory book is simple random walk on the integer lattice, with special attention to the two-dimensional case. This fascinating mathematical object is the point of departure for an intuitive and richly illustrated tour of related topics at the active edge of research. The book starts with three different proofs of the recurrence of the two-dimensional walk, via direct combinatorial arguments, electrical networks, and Lyapunov functions. Then, after reviewing some relevant potential-theoretic tools, the reader is guided towards the relatively new topic of random interlacements – which can be viewed as a "canonical soup" of nearest-neighbour loops through infinity – again with emphasis on two dimensions. On the way, readers will visit conditioned simple random walks – which are the "noodles" in the soup – and also discover how Poisson processes of infinite objects are constructed and review the recently introduced method of soft local times. Each chapter ends with many exercises, making the book suitable for courses and for independent study.

Serguei Popov works on questions related to random walks (also in random environments), random interlacements and others. He also wrote (together with Mikhail Menshikov and Andrew Wade) a book on the Lyapunov functions method for Markov chains. Recent works of the author on random interlacements (including the two-dimensional case) attracted considerable interest in the probabilistic community. Perhaps his most important recent contribution is the soft local times method for constructing couplings of stochastic processes, developed in a joint work with Augusto Teixeira. This method not only permitted strong advances in the field of random interlacements but also proved its usefulness in other topics.

INSTITUTE OF MATHEMATICAL STATISTICS
TEXTBOOKS

IMS Textbooks give introductory accounts of topics of current concern suitable for advanced courses at master's level, for doctoral students and for individual study. They are typically shorter than a fully developed textbook, often arising from material created for a topical course. Lengths of 100–290 pages are envisaged. The books typically contain exercises.

In collaboration with the International Society for Bayesian Analysis (ISBA), selected volumes in the IMS Textbooks series carry the "with ISBA" designation at the recommendation of the ISBA editorial representative.

Other Books in the Series (*with ISBA)

1. *Probability on Graphs*, by Geoffrey Grimmett
2. *Stochastic Networks*, by Frank Kelly and Elena Yudovina
3. *Bayesian Filtering and Smoothing*, by Simo Särkkä
4. *The Surprising Mathematics of Longest Increasing Subsequences*, by Dan Romik
5. *Noise Sensitivity of Boolean Functions and Percolation*, by Christophe Garban and Jeffrey E. Steif
6. *Core Statistics*, by Simon N. Wood
7. *Lectures on the Poisson Process*, by Günter Last and Mathew Penrose
8. *Probability on Graphs (Second Edition)*, by Geoffrey Grimmett
9. *Introduction to Malliavin Calculus*, by David Nualart and Eulàlia Nualart
10. *Applied Stochastic Differential Equations*, by Simo Särkkä and Arno Solin
11.* *Computational Bayesian Statistics*, by M. Antónia Amaral Turkman, Carlos Daniel Paulino, and Peter Müller
12. *Statistical Modeling by Exponential Families*, by Rolf Sundberg

Two-Dimensional Random Walk

From Path Counting to Random Interlacements

SERGUEI POPOV
University of Porto

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314-321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

79 Anson Road, #06-04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781108472456
DOI: 10.1017/9781108680134

First published 2021

A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data
Names: Popov, Serguei, 1972– author.
Title: Two-dimensional random walk : from path counting to random interlacements / Serguei Popov, University of Porto.
Description: Cambridge, United Kingdom ; New York, NY : Cambridge University Press, 2020. |
Series: Institute of mathematical statistics textbooks |
Includes bibliographical references and index.
Identifiers: LCCN 2020020947 (print) | LCCN 2020020948 (ebook) |
ISBN 9781108472456 (hardback) | ISBN 9781108459693 (paperback) |
ISBN 9781108680134 (epub)
Subjects: LCSH: Random walks (Mathematics)
Classification: LCC QA274.73 .P66 2020 (print) | LCC QA274.73 (ebook) |
DDC 519.2/82–dc23
LC record available at https://lccn.loc.gov/2020020947
LC ebook record available at https://lccn.loc.gov/2020020948

ISBN 978-1-108-47245-6 Hardback
ISBN 978-1-108-45969-3 Paperback

Contents

Cambridge University Press
978-1-108-47245-6 — Two-Dimensional Random Walk
Serguei Popov
Frontmatter
More Information

Cambridge University Press
978-1-108-47245-6 — Two-Dimensional Random Walk
Serguei Popov
Frontmatter
More Information

Preface

What does it look like when a mathematician explains something to a fellow mathematician? Everyone knows: lots of writing on the blackboard, lots of intuition flying around, and so on. It is not surprising that mathematicians often prefer a conversation with a colleague to "simply" reading a book. So, in view of this, my initial goal was to write a book as if I were 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 if a person (hmm... well, a mathematician) could just read it in a bus (bed, park, sofa, etc.) and still learn some ideas from contemporary mathematics? I have to confess that, unfortunately, as attested by many early readers, I have not always been successful in creating a text with the aforementioned properties. Still, I hope that at least some pieces will be up to the mission.

To explain my motivation, consider the following well-known fact: frequently, the proof of a mathematical result is difficult, containing lots of technicalities which are hard to follow. It is not uncommon that people struggle to understand such proofs without first getting a "general idea" about what is going on. Also, one forgets technicalities[1] but general ideas remain (and if the ideas are retained, the technical details can usually be reconstructed with some work). So, in this book the following approach is used. I will *always* prefer to explain the intuition first. If the proof is instructive and not too long, it will be included. Otherwise, I will let the interested reader look up the details in other books and/or papers.

The approach can be characterized as striving to understand all things through direct probabilistic intuition. Yes, I am aware that this is not always possible. Nonetheless, when facing a complex task, it is frequently easier to

[1] Even of one's own proofs, as the author has learned on quite a few occasions.

tackle it using familiar tools[2] (even in a non-optimal way) as much as possible than to employ other (possibly more adequate) tools one is unfamiliar with. Also, advanced approaches applied to basic tasks have a tendency to "hide" what is *really* going on (one becomes enchanted with the magic, while still not being able to perform it oneself).

This book revolves around two-dimensional simple random walk, which is not actually so simple, but in fact a rich and fascinating mathematical object. Our purpose here is not to provide a complete treatment of that object, but rather to make an interesting tour around it. In the end, we will come to the relatively new topic of random interlacements (which can be viewed as "canonical" nearest-neighbour loops through infinity). Also, on the way there will be several "side-quests": we will take our time to digress to some related topics which are somewhat underrepresented in the literature, such as Lyapunov functions and Doob's h-transforms for Markov chains.

Intended audience

I expect this book to be of interest to research students and postdocs working with random walks, and to mathematicians in neighbouring fields. Given the approach I take, it is better suited to 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. I am aware that not everybody likes this approach, but I hope that the book will find its audience. Although this book is designed primarily for self-study, it can also be used for a one-semester course on additional topics in Markov chains.

The technical prerequisites are rather mild. The material in the book will be at a level accessible to readers familiar with the basic concepts of probability theory, including convergence of random variables and uniform integrability, with also some background in martingales and Markov chains (at the level of [44], for example). The book is meant to be mostly self-contained (and we recall all necessary definitions and results in Chapter 1).

Many topics in the book are treated at length in the literature, e.g. [41, 63, 71, 91]; on the other hand, we also discuss some recent advances (namely, soft local times and two-dimensional random interlacements) that have not been 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.

[2] In Russia, the ability to build a log house using *only* an axe was considered proof of a carpenter's craftsmanship.

Overview of content

The content of the book is described here. Each chapter (except for the introduction) ends with a list of exercises, and a section with hints and solutions to selected exercises appears at the end of the book. 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*. We recall here some basic definitions and facts for Markov chains and martingales, mainly for reference purposes.
2. *Recurrence of two-dimensional simple random walk.* First, we recall two well-known proofs of recurrence of simple random walk in two dimensions: the classical combinatorial proof and the proof with electrical networks. We then observe that the first proof relies heavily on specific combinatorics and so is very sensitive to small changes in the model's parameters, and the second proof applies only to reversible Markov chains. Then, we present a very short introduction to the Lyapunov function method – which neither requires reversibility nor is sensitive to small perturbations of transition probabilities.
3. *Some potential theory for simple random walks.* This chapter gives 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 to recall and discuss the basic concepts (such as Green's function, potential kernel, harmonic measure) needed in the rest of the book; this chapter is not intended to provide a profound treatment of the subject.
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 transition probabilities. We consider a few classical examples and discuss some properties of this construction. Then, we work with Doob's transform of 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 in its own right: just to cite one of its properties, the probability that a site y is ever visited by a 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 subjects, apparently unrelated to simple random walk.

One is called soft local times; generally speaking, the method of soft local times is a way to construct an adapted stochastic process on a general space Σ using an auxiliary Poisson point process on $\Sigma \times \mathbb{R}_+$. In Chapter 6, this method will be an important tool for dealing with *excursion processes*. Another topic we discuss is "Poisson processes of infinite objects", using as an introductory example the Poisson line process. While this example per se is not formally necessary for the book, it helps us to build some intuition about what will happen in the next chapter.

6. *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 [93] in dimension $d \geq 3$. Then we discuss the two-dimensional case recently introduced in [26]; it is here that various plot lines of this book finally meet. This model will be built of the trajectories of simple random walk conditioned on not hitting the origin, studied in Chapter 4. Using the estimates of two-dimensional capacities and hitting probabilities obtained with the technique of Chapters 3 and 4, we then prove several properties of the model, and the soft local times of Chapter 5 will enter as an important tool in some of these proofs. As stated by Sznitman in [97], "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".

Acknowledgements

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, Darcy Camargo, Caio Alves, Gustavo Henrique Tasca, Victor Freguglia Souza, Matheus Gorito de Paula, and Thainá Soares Silva, who read the manuscript at different stages and made many useful comments and suggestions.

Most of this work was done when the author held a professor position at the University of Campinas (UNICAMP), Brazil. The author takes this occasion to express his profound gratitude to all his Brazilian friends and colleagues.

Notation

Here we list the notation recurrently used in this book.

- We write $X := \dots$ to indicate the definition of X, and will also occasionally use $\dots =: X$.
- $a \wedge b := \min\{a, b\}$, $a \vee b := \max\{a, b\}$.
- For a real number x, $\lfloor x \rfloor$ is the largest integer not exceeding x, and $\lceil x \rceil$ is the smallest integer no less than x.

Sets

- $|A|$ is the cardinality of a finite set A.
- $\mathbb{R}$ is the set of real numbers, and $\mathbb{R}_+ = [0, +\infty)$ is the set of real nonnegative numbers.
- $\mathbb{Z}$ is the set of integer numbers, $\mathbb{N} = \{1, 2, 3, \dots\}$ is the set of natural numbers, $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ is the set of integer nonnegative numbers, $\overline{\mathbb{Z}}_+ = \mathbb{Z}_+ \cup \{+\infty\}$.
- $\mathbb{R}^d$ is the d-dimensional Euclidean space and $\mathbb{Z}^d$ is the d-dimensional integer lattice (with the usual graph structure).
- $\mathbb{Z}^d_n = \mathbb{Z}^d / n\mathbb{Z}^d$ is the d-dimensional torus of (linear) size n (with the graph structure inherited from $\mathbb{Z}^d$).
- For $A \subset \mathbb{Z}^d$, $A^{\complement} = \mathbb{Z}^d \setminus A$ is the complement of A, $\partial A = \{x \in A :$ there exists $y \in A^{\complement}$ such that $x \sim y\}$ is the boundary of A, and $\partial_e A = \partial(A^{\complement})$ is the external boundary of A.
- $\mathcal{N} = \partial_e\{0\} = \{\pm e_{1,2}\} \subset \mathbb{Z}^2$ is the set of the four neighbours of the origin (in two dimensions).
- $\mathsf{B}(x, r) = \{y : \|y - x\| \le r\}$ is the ball (disk) in $\mathbb{R}^d$ or $\mathbb{Z}^d$; $\mathsf{B}(r)$ stands for $\mathsf{B}(0, r)$.

Asymptotics of functions

- $f(x) \asymp g(x)$ 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.
- $f(x) = O(g(x))$ as $x \to a$ means that $\limsup_{x\to a} \left|\frac{f(x)}{g(x)}\right| < \infty$, where $a \in \mathbb{R} \cup \{\infty\}$; $f(x) = o(g(x))$ as $x \to a$ means that $\lim_{x\to a} \frac{f(x)}{g(x)} = 0$.

Euclidean spaces and vectors

- $\|x\|$ is the Euclidean norm of $x \in \mathbb{R}^d$ or $x \in \mathbb{Z}^d$.
- $x \cdot y$ is the scalar product of $x, y \in \mathbb{R}^d$.
- We write $x \sim y$ if x and y are neighbours in $\mathbb{Z}^d$ (i.e., $x, y \in \mathbb{Z}^d$ and $\|x - y\| = 1$).
- $(e_k, k = 1, \dots, d)$ are the canonical coordinate vectors in $\mathbb{R}^d$ or $\mathbb{Z}^d$.
- For $A, B \subset \mathbb{R}^d$ or $\mathbb{Z}^d$, $\operatorname{dist}(A, B) = \inf_{x\in A, y\in B} \|x - y\|$, and $\operatorname{dist}(x, A) := \operatorname{dist}(\{x\}, A)$; also, $\operatorname{diam}(A) = \sup_{x,y\in A} \|x - y\|$.

General probability and stochastic processes

- $(\mathcal{F}_n, n \geq 0)$ is a filtration (a nondecreasing sequence of sigma-algebras).
- a.s. stands for “almost surely” (with probability 1).
- $\mathbf{1}\{\text{event}\}$ is the indicator function of event $\{\text{event}\}$.
- $(p(x, y), x, y \in \Sigma)$ are transition probabilities of a Markov chain on a state space Σ, and $(p_n(x, y), x, y \in \Sigma)$ are the n-step transition probabilities.
- $\mathbb{P}_x$ and $\mathbb{E}_x$ are probability and expectation for a process (normally, a random walk – the one that we are considering at the moment) starting from x.
- SRW is an abbreviation for “simple random walk”.
- $L_n(z) = \sum_{k=1}^{n} \mathbf{1}\{X_k = z\}$ is the local time at z of the process X at time n; we write $L_n^X(z)$ in case when there might be an ambiguity about which process we are considering.[3]
- $\mathcal{G}_n(z)$ is the soft local time of the process at time n at site z.

Simple random walk

- $(S_n, n \geq 0)$ is the simple random walk in $\mathbb{Z}^d$.
- $\tau_A \geq 0$ and $\tau_A^+ \geq 1$ are entrance and hitting times of A by the SRW.
- $\operatorname{Es}_A(x) = \mathbb{P}_x[\tau_A^+ = \infty]\mathbf{1}\{x \in A\}$ is the escape probability from $x \in A$ for SRW in dimensions $d \geq 3$.

[3] We use different notation for local times of SRW and conditioned SRW; see the following notations for (conditioned) simple random walk.

- $G(\cdot,\cdot)$ is Green's function for the SRW in three or more dimensions, $G_\Lambda(\cdot,\cdot)$ is Green's function restricted on Λ.
- $a(\cdot)$ is the potential kernel for the two-dimensional SRW.
- $\mathrm{hm}_A(x)$ is the value of the harmonic measure in $x \in A$.
- $\mathrm{cap}(A)$ is the capacity of A (in two or more dimensions).
- $N_x = \sum_{j=0}^{\infty} \mathbf{1}\{S_j = x\}$ is the total number of visits to x, and $N_x^{(k)} = \sum_{j=0}^{k} \mathbf{1}\{S_k = x\}$ is the total number of visits to x up to time k (i.e., the local time at x at time k).
- Let A be a fixed subset of $\mathbb{Z}^2$; then $N_x^\flat = \sum_{j=0}^{\tau_A^+-1} \mathbf{1}\{S_j = x\}$ is the number of visits to x *before* the first return to A, and by $N_x^\sharp = \sum_{j=\tau_A^+}^{\infty} \mathbf{1}\{S_j = x\}$ the number of visits to x *after* the first return to A (with $N_x^\sharp = 0$ on $\{\tau_A^+ = \infty\}$).

Conditioned simple random walk

- $(\widehat{S}_n, n \geq 0)$ is the simple random walk in $\mathbb{Z}^2$ conditioned on never hitting the origin.
- $\hat{\tau}_A \geq 0$ and $\hat{\tau}_A^+ \geq 1$ are entrance and hitting times of A by the conditioned SRW.
- $\widehat{\mathrm{Es}}_A(x) = \mathbb{P}_x[\hat{\tau}_A^+ = \infty]\mathbf{1}\{x \in A\}$ is the escape probability from $x \in A$ for the conditioned SRW in two dimensions.
- $\widehat{\mathrm{hm}}_A(x)$ is the harmonic measure for the conditioned walk.
- $\widehat{\mathrm{cap}}(A) = \mathrm{cap}(A \cup \{0\})$ is the capacity of set A with respect to the conditioned SRW.
- $\widehat{G}(x,y) = a(y)(a(x) + a(y) - a(x-y))/a(x)$ is Green's function of the conditioned walk.
- $\hat{\ell}(x,y) = 1 + \frac{a(y)-a(x-y)}{a(x)} = \frac{\widehat{G}(x,y)}{a(y)}$, and $\hat{g}(x,y) = \frac{\widehat{G}(x,y)}{a^2(y)} = \hat{g}(y,x)$ is the "symmetrized" $\widehat{G}$.
- $\widehat{N}_x, \widehat{N}_x^{(k)}, \widehat{N}_x^\flat, \widehat{N}_x^\sharp$ are defined just as $N_x, N_x^{(k)}, N_x^\flat, N_x^\sharp$, but with the conditioned walk $\widehat{S}$ instead of SRW S.

Random interlacements

- $\mathrm{RI}(\alpha)$ is the random interlacement process on level $\alpha > 0$.
- $\mathbb{E}^u$ is the expectation for random interlacements on level $u > 0$.
- $\mathcal{I}^u$ 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 α, so these become $\mathcal{I}^\alpha$ and $\mathcal{V}^\alpha$.