Durham E-Theses

On some random walk problems

LO, CHAK,HEI

How to cite:

LO, CHAK,HEI (2017) *On some random walk problems*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/12498/

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full Durham E-Theses policy for further details.

Academic Support Office, Durham University, University Office, Old Elvet, Durham DH1 3HP
e-mail: e-theses.admin@dur.ac.uk Tel: +44 0191 334 6107
http://etheses.dur.ac.uk

On some random walk problems

Chak Hei Lo

Submitted for the degree of Doctor of Philosophy

October 2017

Abstract:

We consider several random walk related problems in this thesis. In the first part, we study a Markov chain on $\mathbb{R}_+ \times S$, where $\mathbb{R}_+$ is the non-negative real numbers and S is a finite set, in which when the $\mathbb{R}_+$-coordinate is large, the S-coordinate of the process is approximately Markov with stationary distribution π_i on S. Denoting by $\mu_i(x)$ the mean drift of the $\mathbb{R}_+$-coordinate of the process at $(x, i) \in \mathbb{R}_+ \times S$, we give an exhaustive recurrence classification in the case where $\sum_i \pi_i \mu_i(x) \to 0$, which is the critical regime for the recurrence-transience phase transition. If $\mu_i(x) \to 0$ for all i, it is natural to study the *Lamperti* case where $\mu_i(x) = O(1/x)$; in that case the recurrence classification is known, but we prove new results on existence and non-existence of moments of return times. If $\mu_i(x) \to d_i$ for $d_i \neq 0$ for at least some i, then it is natural to study the *generalized Lamperti* case where $\mu_i(x) = d_i + O(1/x)$. By exploiting a transformation which maps the generalized Lamperti case to the Lamperti case, we obtain a recurrence classification and an existence of moments result for the former. The generalized Lamperti case is seen to be more subtle, as the recurrence classification depends on correlation terms between the two coordinates of the process.

In the second part of the thesis, for a random walk S_n on $\mathbb{R}^d$ we study the asymptotic behaviour of the associated centre of mass process $G_n = n^{-1} \sum_{i=1}^n S_i$. For lattice distributions we give conditions for a local limit theorem to hold. We prove that if the increments of the walk have zero mean and finite second moment, G_n is recurrent if $d = 1$ and transient if $d \geq 2$. In the transient case we show that G_n has diffusive rate of escape. These results extend work of Grill, who considered simple symmetric random walk. We also give a class of random walks with symmetric heavy-tailed increments for which G_n is transient in $d = 1$.

On some random walk problems

Chak Hei Lo

A Thesis presented for the degree of
Doctor of Philosophy

Probability group
Department of Mathematical Sciences
Durham University
England

October 2017

Contents

List of Assumptions **vi**

List of Figures **viii**

Declaration **x**

Acknowledgements **xi**

1 Introduction **1**

1.1 Random walk 1

1.2 Markov chains and recurrence classification 3

1.3 Simple symmetric random walk 5

1.4 Homogeneous random walk on $\mathbb{R}^d$ 10

1.5 Non-homogeneous random walk on $\mathbb{R}^d$ 13

1.6 Law of large numbers and central limit theorem 16

1.7 Thesis outline 18

Part I: Non-Homogeneous Walks on a Half Strip **21**

2 Notation, preliminaries and prerequisites **22**

2.1 Literature review 22

2.2 The state space Σ 24

2.3 Recurrence classification for the half strip 26

2.4 Assumptions of the model 29

2.5 The Lamperti problem 32

3 Main results **36**

3.1 Lamperti drift on a half strip 36

3.1.1 Recurrence classification 36

3.1.2 Existence and non-existence of moments 38

3.2 Generalized Lamperti drift on a half strip 40

3.2.1 Recurrence classification 40

3.2.2 Existence and non-existence of moments 45

4 Proofs and technical details **47**

4.1 Semi-martingale criteria for recurrence classification 47

4.2 Lyapunov function estimates for the half strip 49

4.2.1 Lyapunov function for constant drift 50

4.2.2 Lyapunov function for Lamperti drift 53

4.3 Some consequences of the Fredholm alternative 59

4.3.1 Fredholm alternative 60

4.3.2 Corollaries 61

4.4 Proof of the constant drift classification 61

4.5 Proofs of results for Lamperti drift 62

4.5.1 Proof of recurrence and transience in the Lamperti drift case . 63

4.5.2 Proof of existence of moments 64

4.5.3 Proof of non-existence of moments 66

4.5.4 Complete classification 68

4.6 Proofs of results for generalized Lamperti drift 69

4.6.1 Transformation to Lamperti drift case 69

4.6.2 Preservation of the recurrence classification after transformation 70

4.6.3 Calculation of the transformed moments 72

4.6.4 Proof of recurrence classification 73

4.6.5 Proofs of existence and non-existence of moments 74

5 Examples, applications and simulations **75**

5.1 Correlated random walk 75

5.2 Proofs of theorems on correlated random walk 77
5.3 Numerical examples and simulations 77

Part II: Centre of Mass of Random Walks 83

6 Notation, preliminaries and prerequisites 84
6.1 Literature review 84
6.2 Lattice distributions and characteristic functions 90
6.3 Hewitt-Savage zero one law 95
6.4 Local limit theorem for random walks 96
6.5 Stable distributions and domains of attraction 98

7 Main results 101
7.1 Law of large numbers and central limit theorem for centre of mass . . 101
7.2 Local limit theorem for centre of mass 103
7.3 Transience and diffusive rate of escape in two or higher dimensions . . 104
7.4 One dimension 105

8 Proofs and technical details 107
8.1 A characteristic function result 107
8.2 Proof of the local limit theorem 108
8.2.1 Integral estimates 109
8.3 Proofs for two or higher dimensions 117
8.4 Proofs for one dimension 119
8.4.1 Recurrence 119
8.5 Case of stable distribution 122
8.5.1 Local limit theorem for stable distribution 122
8.5.2 Transience with stable distribution in one dimension 126

9 Examples and applications 127
9.1 Lazy simple symmetric random walk 127
9.2 Simple symmetric random walk 128

Part III: Conclusions **131**

10 Open problems and conjectures **132**

10.1 Half strip 132

10.2 Centre of mass 133

Bibliography **135**

List of Assumptions

(A) Suppose that (X_n, η_n), $n \in \mathbb{Z}_+$, is a time-homogeneous, irreducible Markov chain on Σ, a locally finite subset of $\mathbb{R}_+ \times S$. Suppose that for each $k \in S$ the line Λ_k is unbounded.

($\mathbf{B}_p$) There exists a constant $C_p < \infty$ such that for all $n \in \mathbb{Z}_+$,

$$\mathbb{E}[|X_{n+1} - X_n|^p \mid X_n = x,\, \eta_n = i] \leq C_p, \text{ for all } (x, i) \in \Sigma.$$

($\mathbf{D_C}$) For each $i \in S$ there exists $d_i \in \mathbb{R}$ such that $\mu_i(x) = d_i + o(1)$ as $x \to \infty$.

($\mathbf{D_G}$) For $i, j \in S$ there exist $d_i \in \mathbb{R}$, $e_i \in \mathbb{R}$, $d_{ij} \in \mathbb{R}$ and $t_i^2 \in \mathbb{R}_+$, with at least one t_i^2 non-zero, such that

(a) for all $i \in S$, $\mu_i(x) = d_i + \frac{e_i}{x} + o(x^{-1})$ as $x \to \infty$;

(b) for all $i \in S$, $\sigma_i^2(x) = t_i^2 + o(1)$ as $x \to \infty$;

(c) for all $i, j \in S$, $\mu_{ij}(x) = d_{ij} + o(1)$ as $x \to \infty$; and

(d) $\sum_{i \in S} \pi_i d_i = 0$.

($\mathbf{D_G^+}$) There exist $\delta_2 \in (0, 1)$, $d_i \in \mathbb{R}$, $e_i \in \mathbb{R}$, $d_{ij} \in \mathbb{R}$ and $t_i^2 \in \mathbb{R}_+$, with at least one t_i^2 non-zero, such that

(a) for all $i \in S$, $\mu_i(x) = d_i + \frac{e_i}{x} + o(x^{-1-\delta_2})$ as $x \to \infty$;

(b) for all $i \in S$, $\sigma_i^2(x) = t_i^2 + o(x^{-\delta_2})$ as $x \to \infty$; and

(c) for all $i, j \in S$, $\mu_{ij}(x) = d_{ij} + o(x^{-\delta_2})$ as $x \to \infty$.

($\mathbf{D_L}$) For each $i \in S$ there exist $c_i \in \mathbb{R}$ and $s_i^2 \in \mathbb{R}_+$, with at least one s_i^2 non-zero, such that, as $x \to \infty$, $\mu_i(x) = \frac{c_i}{x} + o(x^{-1})$ and $\sigma_i^2(x) = s_i^2 + o(1)$.

($\mathbf{D_L^+}$) Suppose that there exist $\delta_1 \in (0,1)$, $c_i \in \mathbb{R}$, and $s_i^2 \in \mathbb{R}_+$, with at least one s_i^2 non-zero, such that for all $i \in S$, as $x \to \infty$, $\mu_i(x) = \frac{c_i}{x} + o(x^{-1-\delta_1})$ and $\sigma_i^2(x) = s_i^2 + o(x^{-\delta_1})$.

(L) Suppose that the minimal subgroup of $\mathbb{R}^d$ associated with X is $L := H\mathbb{Z}^d$ with $h := |\det H| > 0$.

(M) Suppose that $\mathbb{E}[\|X\|^2] < \infty$ and M is positive-definite.

($\mathbf{Q_\infty}$) Suppose that $\lim_{x\to\infty} q_{ij}(x) = q_{ij}$ exists for all $i, j \in S$, and (q_{ij}) is an irreducible stochastic matrix.

($\mathbf{Q_\infty^+}$) Suppose that there exists $\delta_0 \in (0,1)$ such that $\max_{i,j\in S} |q_{ij}(x) - q_{ij}| = O(x^{-\delta_0})$ as $x \to \infty$.

($\mathbf{Q_G}$) For $i, j \in S$ there exist $\gamma_{ij} \in \mathbb{R}$ such that $q_{ij}(x) = q_{ij} + \frac{\gamma_{ij}}{x} + o(x^{-1})$, where (q_{ij}) is a stochastic matrix.

($\mathbf{Q_G^+}$) There exist $\delta_3 \in (0,1)$ and $\gamma_{ij} \in \mathbb{R}$ such that $q_{ij}(x) = q_{ij} + \frac{\gamma_{ij}}{x} + o(x^{-1-\delta_3})$.

(S) Suppose that $X \stackrel{d}{=} -X$ and X is in the domain of normal attraction of a symmetric α-stable distribution with $\alpha \in (0,1)$.

(V) Suppose that $\mathbb{E}[\|\xi\|^2] < \infty$ and write $\Sigma := \mathbb{E}[(\xi - \mu)(\xi - \mu)^\top]$. Here Σ is a nonnegative-definite, symmetric d by d matrix; we write $\sigma^2 := \operatorname{tr}\Sigma = \mathbb{E}[\|\xi - \mu\|^2]$.

($\mathbf{W_\mu}$) Let $d \in \mathbb{N}$, and suppose that $Z, Z_1, Z_2, \ldots$ are i.i.d. random variables with $\mathbb{E}\,\|Z\| < \infty$ and $\mathbb{E}\,Z = \mu \in \mathbb{R}^d$. The random walk $(S_n, n \in \mathbb{Z}_+)$ is the sequence of partial sums $S_n := \sum_{i=1}^n Z_i$ with $S_0 := 0$.

List of Figures

1.1 Three simulated trajectories of 1D SSRW against time. 6

1.2 Three simulated trajectories of 2D SSRW. 6

1.3 Three simulated trajectories of 3D SSRW. 7

1.4 Three simulated paths of two dimensional random walk with drift. . . 12

1.5 Three simulated paths of two dimensional random walk with heavy-tailed distributions. 13

1.6 A 2D elliptic random walk with comparatively large radial component. 15

1.7 A 2D elliptic random walk with comparatively large transversal component. 15

2.1 An illustration of the half strip model. 23

5.1 An example of a one-step correlated random walk. 78

5.2 Two simulations of 10^3 steps of one-step correlated random walks, as an application of the half strip model. Top: $c = 1$, bottom: $c = -1$. . 79

5.3 An example of a two-steps correlated random walk. 80

5.4 Two simulations of 10^3 steps of two-steps correlated random walks, as an application of the half strip model. Top: $c = 1$, bottom: $c = -1$. 81

6.1 An example of a one dimensional random walk (Light blue to blue) with the corresponding centre of mass process (Orange to red). 85

6.2 Two examples of a two dimensional random walk (Light blue to blue) with the corresponding centre of mass process (Orange to red). 86

6.3 An example of a three dimensional random walk (Blue) with the corresponding centre of mass process (Red) at front (top left), top (top right), left (bottom left) and side (bottom right) angle. 87

8.1 Controlling G_n along a subsequence. 119

Declaration

The work in this thesis is based on research carried out in the Department of Mathematical Sciences, at the University of Durham, England. No part of this thesis has been submitted elsewhere for any other degree or qualification. It is all my own work unless referenced to the contrary in the text.

Parts of Chapters 2 to 5 are adapted from joint work with Andrew R. Wade [69].

Parts of Chapters 6 to 9 are adapted from joint work with Andrew R. Wade [70].

Copyright © October 2017 by Chak Hei Lo.

"The copyright of this thesis rests with the author. No quotations from it should be published without the author's prior written consent and information derived from it should be acknowledged."

Acknowledgements

I am extremely thankful to my supervisors, Mikhail Menshikov and Andrew Wade, for all of their guidance and encouragement, which enabled me to develop a deep understanding of the theory of random walks. They gave me much inspiration and patiently listened to many of my presentations throughout the last three years.

I am also grateful to Nicholas Georgiou and Ostap Hryniv for fruitful discussions on the topics of this thesis.

Thank you to everyone in the Probability and Statistics group at Durham for creating an enjoyable working environment. Special thanks should be given to all of the postgraduate students in the department who broadened my view with discussions on a variety of Mathematical topics.

Finally, words alone cannot express the sincere thanks I owe to my parents, for their moral support, especially in helping me to pass through some hard times.

Dedicated to

Wing Ming Lo

and

Vince Ho

Chapter 1

Introduction

Many *stochastic processes* arising in applications exhibit a range of possible behaviours depending upon the values of certain key parameters. Investigating *phase transitions* for such systems leads to interesting and challenging mathematics. Much progress has been made over the years, using various techniques. The most subtle case is when the system is near-critical in some sense (near a phase boundary). This thesis will study a few particular near-critical Markov models, with an aim to extend known criteria for classifying *recurrence* and *transience*.

Now we will start on some background knowledge and classical results on random walk theory, together with some new intuitions.

1.1 Random walk

Random walk is one of the most important models in probability theory. It displays profound mathematical properties and has a wide range of application in many scientific fields and much more. It is a stochastic process which describes the random trajectory of a particle (or random walker) in space. The motion of the particle is explained with a succession of random increments or jumps at discrete instants in time. The *long term asymptotic behaviour* of the particle or walker is of great interest and has stimulated extensive research in this field. It has a long and rich history across a variety of subjects. The classical one-dimensional random walk dates back

to the 'gambler's ruin' problem, addressed a few centuries ago by Fermat and Pascal [92]. The mathematical theory started to formalize as the French mathematician Louis Bachelier gave his insight to his stock prices model using the random walk reasoning in his Ph.D. thesis in 1900 [7]. The popularity of the term 'random walk' gradually increased when a Professor of Economics at Princeton University, Burton Malkiel, published his book, A Random Walk Down Wall Street, in 1973 [72].

For the more general version of the model in *several dimensions*, it was probably first studied in around 1880 in the form of Lord Rayleigh's theory of sound [87]. Shortly after, similar ideas from Albert Einstein's theory of Brownian motion (1905-1908) in statistical physics [28] and English statistician Karl Pearson's theory of random migration of species (1906) in biology [84] arose. The term random walk is first suggested by Pearson in a letter to the journal Nature [83] and it is stated as a path with a succession of random steps, usually on a d-dimensional lattices in classical literature.

In 1920, the Hungarian mathematician George Pólya confirmed the mathematical importance of this indispensable random walk model [85]. Numerous elegant connections and ideas in random walk blossom and propagate to other significant branches of mathematics such as combinatorics, harmonic analysis, potential theory, and spectral theory over the last century. The theory of random walk then continued to proliferate in lively realm of modern science. A broad range of studies can be found in [94].

The popularity of researching the random walk model is due to its vast applications in different subjects such as, but no limited to the following.

- Chemistry: Polymer conformation in molecular chemistry [8, 71, 99];
- Biology: Modelling of microbe locomotion in microbiology [9, 99];
- Economics: Financial systems, modelling stock prices [31];
- Psychology: Human memory search in a semantic network [6].

In this chapter, we will discuss some of the history and motivation behind the study of such random walk problems. We will also give some foundation material

on random walk theory with some personal intuition. Let's discover these hidden gems through the exciting adventures of some random walk problems.

1.2 Markov chains and recurrence classification

The *Markov process*, named after the Russian mathematician Andrey Markov, has a characteristic property that it retains no memory of where it has been in the past. This property is sometimes known as the *Markov property* or the *memorylessness property*. In other words, where the process will go next only depends on the current state of the process. By conditioning on the current state of the process, its future and past states are independent. When the Markov process has a finite or countable set of states in particular, we would call it a *Markov chain*.

Although Andrey Markov studied Markov chains and Markov processes, with his first paper on these topics in 1906, other specific models of Markov processes already existed. Random walk is an example of a Markov chain, and was studied hundreds of years earlier [98].

Compared to the usual use of the term random walk, which suggests that the process is on a regular lattice, Markov chains are usually more general in terms of describing a more complicated state space. As both of them are stochastic processes, we would not distinguish them specifically in the context of this thesis, and will use them interchangeably.

A very important property for Markov chains is the *recurrence classification*. It gives us a general idea of how the process will evolve in the long term. Given a Markov chain $(X_n), n \geq 0$ on a countable state space S, a state $i \in S$ is called *recurrent* if

$$\mathbb{P}(X_n = i \text{ for infinitely many } n | X_0 = i) = 1.$$

A state $i \in S$ is called *transient* if

$$\mathbb{P}(X_n = i \text{ for infinitely many } n | X_0 = i) = 0.$$

Although it is not immediate, standard Markov chain theory shows that any state

can only be either recurrent or transient, see [81, p.26, Theorem 1.5.3].

We can also understand the idea of recurrence and transience by looking at the *return time*, also know as the *first passage time* and the *hitting time*, defined as follows. For $i \in S$,

$$\tau_i = \inf\{n \geq 1 : X_n = i\},$$

with the convention that $\inf \emptyset := \infty$. Intuitively if $X_0 = i$, τ_i is time it takes for the process to come back to its original position. Again, from standard Markov chain theory, we can easily see that a state i is recurrent if and only if $\mathbb{P}(\tau_i < \infty | X_0 = i) = 1$ and it is transient if and only if $\mathbb{P}(\tau_i < \infty | X_0 = i) < 1$.

If a state is recurrent, it implies that the process will come back to this state with probability one, but it does not guarantee that the process will come back in finite time in expectation. Hence we could further classify the recurrent case into positive recurrent or null recurrent. We define a recurrent state i to be *positive recurrent* if

$$\mathbb{E}[\tau_i | X_0 = i] < \infty$$

and *null recurrent* if

$$\mathbb{E}[\tau_i | X_0 = i] = \infty.$$

This time, it is clear that it is a *dichotomous* classification.

In order to understand the recurrence classification for the whole process, we should understand the structure of the walk first. Sometimes, it is possible to break a chain into smaller pieces, so that we can understand the behaviour of each piece separately in a relatively simple way, and group them all back together to get a result for the whole chain. This involves identification of *communication classes* of the chain.

Given a Markov chain $(X_n), n \geq 0$ on a countable state space S, for any states $i, j \in S$ we say that i *leads to* j and write $i \to j$ if

$$\mathbb{P}(X_n = j \text{ for some } n \geq 0 | X_0 = i) > 0.$$

We also say that i *communicates with* j and write $i \leftrightarrow j$ if both $i \to j$ and

$j \to i$. It is clear that $i \leftrightarrow i$ from the definition. Together with the fact that $i \leftrightarrow j$ and $j \leftrightarrow k$ implies $i \leftrightarrow k$ for any states i, j and $k \in S$, we conclude that $\leftrightarrow$ is an equivalence relation on S. So we can partition S into *communicating classes*. If a chain only consists of one class, then it is called *irreducible*.

From standard Markov chain theory, the properties of positive recurrence, null recurrence and transience are all class properties. This means if a state in a certain class is transient, then any state in the class is also transient.

In our context of random walks in this thesis, they are always irreducible Markov chains, hence the recurrence classification for the process (with certain fixed parameters) we considered as a whole is well defined.

Hence when we say recurrence classification in context of this thesis, we want to determine how the parameters in the model will affect the process to be positive recurrent, null recurrent, or transient.

1.3 Simple symmetric random walk

The most comprehensively studied random walk model is the *simple symmetric random walk*. Formally, denote by $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_d\}$ the standard orthonormal basis on $\mathbb{R}^d$, and let $U_d := \{\pm\mathbf{e}_1, \pm\mathbf{e}_2, \cdots, \pm\mathbf{e}_d\}$ be the set of possible jumps of the random walk. Given a sequence of independent identically distributed (i.i.d.) random variables $Z, Z_1, Z_2, \ldots$, with

$$\mathbb{P}(Z = \mathbf{e}) = \frac{1}{2d} \quad \text{for } \mathbf{e} \in U_d, \tag{1.3.1}$$

we define the *simple symmetric random walk* as a discrete-time Markov process $(S_n, n \geq 0)$ on the d-dimensional integer lattice $\mathbb{Z}^d$ by

$$S_n = \sum_{k=1}^{n} Z_k. \tag{1.3.2}$$

Alternatively, we can think about this process in the natural way. To move from a certain point S_n to the next point S_{n+1} in $\mathbb{Z}^d$, we chose, uniformly at random, from all of the $2d$ neighbours of S_n, in other words, all the points which differ from S_n

by exactly ± 1 in a single coordinate. Here are some pictures of simple symmetric random walks in one, two and three dimensions.

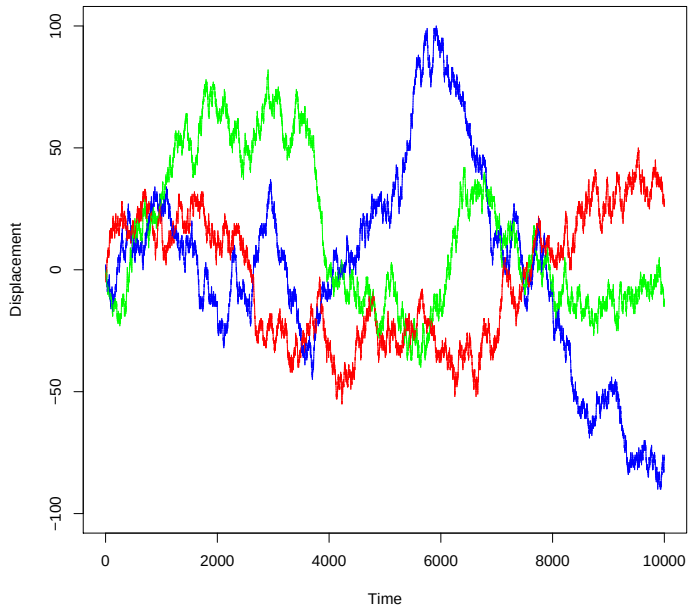

Figure 1.1: Three simulated trajectories of 1D SSRW against time.

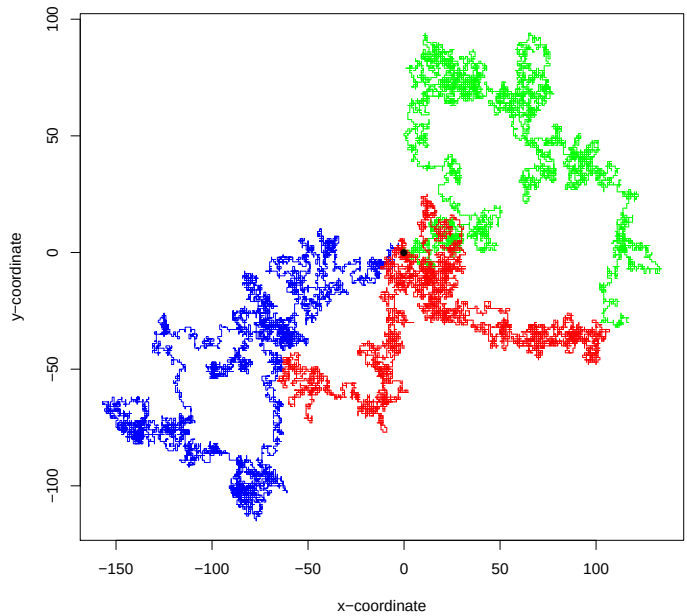

Figure 1.2: Three simulated trajectories of 2D SSRW.

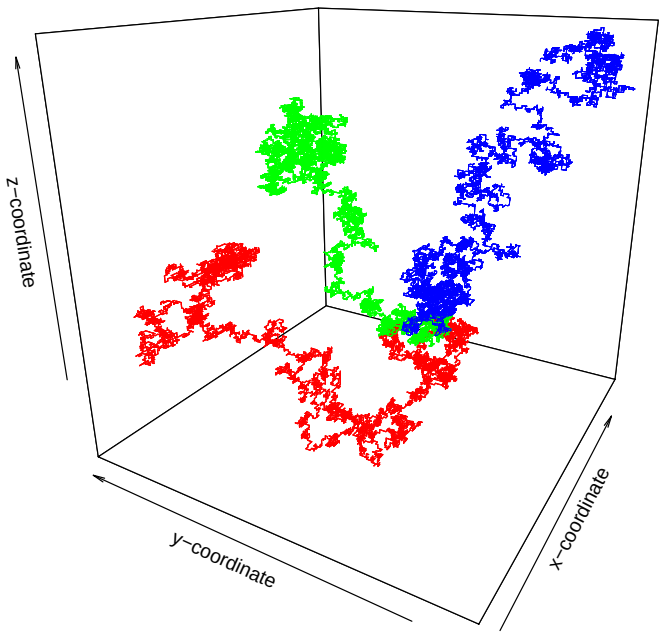

Figure 1.3: Three simulated trajectories of 3D SSRW.

One of the most fundamental properties of a random walk is the recurrence property. The story goes back to 1920s. George Pólya enjoyed to take random running paths in a big park as his daily exercise. Although his paths were completely random, he often met the same couple during his journey, who was also running around the area [85]. He realized that assuming the couple also takes a random path every day, then his relative position to the couple is also a random walk. This can be done by just combining the two steps of the random walks by Pólya and the couple at every time point as one big step. Then they will meet each other whenever the combined random walk visits the origin. Now the real question is, what is the probability that the walk will eventually returns to $\mathbf{0}$? Mathematically, define $\tau_d := \min\{n \geq 1 : S_n = \mathbf{0}\}$ to be the time needed for the first return to the origin. If the walk never comes back, then $\tau_d = \infty$, as with the usual convention that $\min \emptyset := \infty$. Now our interest is in the *Pólya's random walk constant* p_d, defined as

$$p_d := \mathbb{P}(\tau_d < \infty). \tag{1.3.3}$$

Similar to the recurrence classification for Markov chains that discussed in Section 1.2, we call the random walk *recurrent* if $p_d = 1$, and *transient* if $p_d < 1$. Intuitively, a recurrent walk means that the random walk will visit the origin infinitely often with probability one while a transient walk means with probability one, it will only come back to the origin finitely many times, and never return again.

In general, finding this classification is very difficult due to the fact that the intrinsic properties of the state space or the movement of the walk is complicated to quantify for meaningful analysis. However, in the case of simple symmetric random walk, which is a pleasant model to study due to the simple and clean structure, there are a lot of well developed combinatorial techniques based on counting sample paths that give us elegant properties of the walk. We now present the following beautiful result by George Pólya in 1921 [85].

Theorem 1.3.1 (Pólya's Recurrence Theorem)**.** *The simple symmetric random walk on $\mathbb{Z}^d$ is recurrent in one or two dimensions, but transient in three or more dimensions. Equivalently, $p_1 = p_2 = 1$ but $p_d < 1$ for all $d \geq 3$.*

The essence of this theorem can be easily understood by the aphorism credited to Shizuo Kakutani in a UCLA colloquium talk: 'A drunk man will eventually find his way home, but a drunk bird may get lost forever' [27, p.191]. My version to remember the critical dimension is by thinking of the sentence 'Everyone but astronaut drinks'.

More precisely on the value of p_d, Montroll [77] in 1956 showed that for $d \geq 3$, $p_d = 1 - u_d^{-1}$ where

$$u_d = \int_0^\infty \left[I_0\left(\frac{t}{d}\right)\right]^d e^{-t} \mathrm{d}t, \tag{1.3.4}$$

and $I_0(z)$ is the modified Bessel function of the first kind. Numerically, $p_3 \approx 0.340537$, see [21, 24, 41, 97] and $p_4 \approx 0.193206$ [34, 77].

The intuition behind this phenomenon is actually quite difficult to come up with. At first sight, one might think as the dimension increases, the number of points in the lattice increases and also more choices are available at each time point, that is why it is more difficult for the particle or the walker to jump back to the origin. This is not a very convincing argument since if you are away from the origin, you

have many choices in higher dimensions, but a high proportion of them are 'helping' you to get back in terms of shortening the distance from the starting point, then you should still have a lot of tendency to come back. In one dimension, except the starting point, we always have equal tendency to move to or away from the origin. In two dimensions, most of the points on the lattice have equal number of choices to help or not help you to come back, while on the axis there are actually more choices that push you away than those pull you back! However, both one or two dimensions fall into the recurrent case. This argument is unclear from the classification, and there is no hint for why the critical change is from two to three dimensions, but not, say, four to five dimensions.

In fact, Pólya's original argument was based on delicate path counting and is largely combinatorial, which the intuition remains hidden behind. Some other intuition is based on the proof by electric networks and potential theory technique. The end of the proof boils down to the convergence of harmonic series. The increase of dimension changes the convergence to divergence, and thus the critical point emerges from two to three dimensions, algebraically. Again, this is not a very satisfactory explanation due to the lack of explaining the physical meaning of how the dimension affects the series.

If we want to generalize the above methods to more general random walks, they just completely break down due to the complicated structure or long distance correlation. We realized that not only the average drift in the model matters, but the variance of jumps is equally important.

One of the heuristic and intuitive arguments that I came across in the literature is the following. Consider the random walk in $\mathbb{R}^d$ then the probability of the random walk being within distance $O(1)$ of the origin after n steps will become order $O(n^{-\frac{d}{2}})$ from the local limit theorem for random walk, that will be explained in Section 6.5. Now if we consider all possible n and sum the probabilities up, we get an expression $\sum_{n=1}^{\infty} n^{-\frac{d}{2}}$ which is divergent when $d = 2$ and convergent when $d = 3$. By the Borel-Cantelli lemma this gives a sufficient condition for transience. However, this argument does not give both directions, i.e. the divergent sequence does not imply

recurrence directly.

In my own opinion, the best and neatest argument is using the idea of Lyapunov functions which can be found in [75], which involves a version of Lamperti's fundamental recurrence classification [64]. We will delay this argument to Chapter 2.

1.4 Homogeneous random walk on $\mathbb{R}^d$

Simple symmetric random walk is a specific model that is very restrictive to the movement of the walk. It is natural to extend the theory to a more general class of random walks. A famous intermediate extension involves the Pearson-Rayleigh random walk on $\mathbb{R}^d$, which allows the walk to jump to any point on the unit circle/sphere centred at the current position, with uniform probability. Similar results to those for the simple symmetric random walk can also be obtained. In fact, we can do far more than this. Without any particular structure of the jump, we define a random walk as a discrete-time Markov process $(S_n; n \geq 0)$ on an unbounded state space $\Sigma \subseteq \mathbb{R}^d$. Throughout the whole thesis, we always assume the walk is time-homogeneous, i.e. the distribution of S_{n+1} given $(S_0, S_1, \ldots, S_n)$ only depends on S_n but not on n.

A typical type of random walk that was studied extensively in the literature is the *spatially homogeneous random walk*. We can define it as $S_n = \sum_{k=1}^{n} Z_k$ where $Z, Z_1, \ldots, Z_n$ are i.i.d. random variables, taking values in $\mathbb{R}^d$, so the law of the increment does not depend on the current position of the walk.

In the context of the general random walk, there are some results on the generalization of the seminal Pólya's recurrence theorem for the continuous state space $\mathbb{R}^d$. However, we need to reconsider the definition of recurrence and transience again. The original definition of recurrence is not completely clear in a continuous state space. Do we insist of the walk going back to the exact same point or do we allow the walk just come back to a small neighbourhood of the point it visited in the past? These two situation exhibit a very different behaviour in critical situations. Hence we should separate them clearly. Without any specification on the structure of the walk, we will use the following definition.

Definition 1.4.1. A random walk $(S_n; n \geq 0)$ taking values in $\Sigma \subseteq \mathbb{R}^d$ is transient if $\lim_{n\to\infty} \|S_n\| = \infty$, a.s. The walk is recurrent if, for some constant $r_0 \in \mathbb{R}_+$, $\liminf_{n\to\infty} \|S_n\| \leq r_0$.

It is very important to know that the classification of recurrence and transience is not necessarily exhaustive in general, we will look deeper in this later in our specific model. In the special case of spatially homogeneous random walk, one can apply the Hewitt-Savage zero-one law to prove the dichotomy. We will explain in more details in Part II of the thesis. Also, even with these more general definitions, we need to make sure that the walk should not be 'trapped' in part of the state space as the transient definition suggest the walk will go to infinity eventually, but here the walk can just go to a finite limiting point, breaking the dichotomy. So the classification is not properly defined in this case. The easiest way is to assume the state space Σ to be locally finite to get some form of irreducibility so we can avoid the ambiguity on the recurrence classification.

Now we are ready to generalize the influential result of Pólya's recurrence theorem. In 1951, Two mathematicians Kai-lai Chung and Wolfgang Heinrich Johannes Fuchs (see [19] and [58, Chapter 9]) extended the result to non-degenerate homogeneous random walks whose increments have finite second moments as follows.

Theorem 1.4.2 (Chung-Fuchs Theorem)**.** *Let S_n be a random walk in $\mathbb{R}^d$. Then we have the following statements.*

1. *When $d = 1$, if $\mathbb{E}[|Z|] < \infty$ and $\mathbb{E}[Z] = 0$, then S_n is recurrent.*
2. *When $d = 2$, if $\mathbb{E}[Z^2] < \infty$ and $\mathbb{E}[Z] = 0$, then S_n is recurrent.*
3. *If $d \geq 3$ and the random walk is not contained in a lower-dimensional subspace, then it is transient.*

Notably, the Brownian motion, as a continuous version of the simple symmetric random walk, exhibits similar behaviour. However, the proof does not follow by the theorem above.

Compared to the classic path counting proof of Pólya's theorem, the proof of the Chung-Fuchs theorem is based on Fourier analysis. Although the methods are

different, they both retain the unsatisfactory fact that intuition is still hidden behind the calculations.

In the early 1960s, John W. Lamperti made a momentous breakthrough on developing the approach of Lyapunov functions [64]. This method can be applied to a broader variety of random walks than the combinatorial and analytical approaches. Just as importantly, it is probably the first method which clarifies the probabilistic intuition behind the recurrence classification problem. We will see more about this in the next chapter.

At the end of this section we will provide some pictures of homogeneous random walks in two dimensions. The behaviour can vary a lot depending on the properties of the walk.

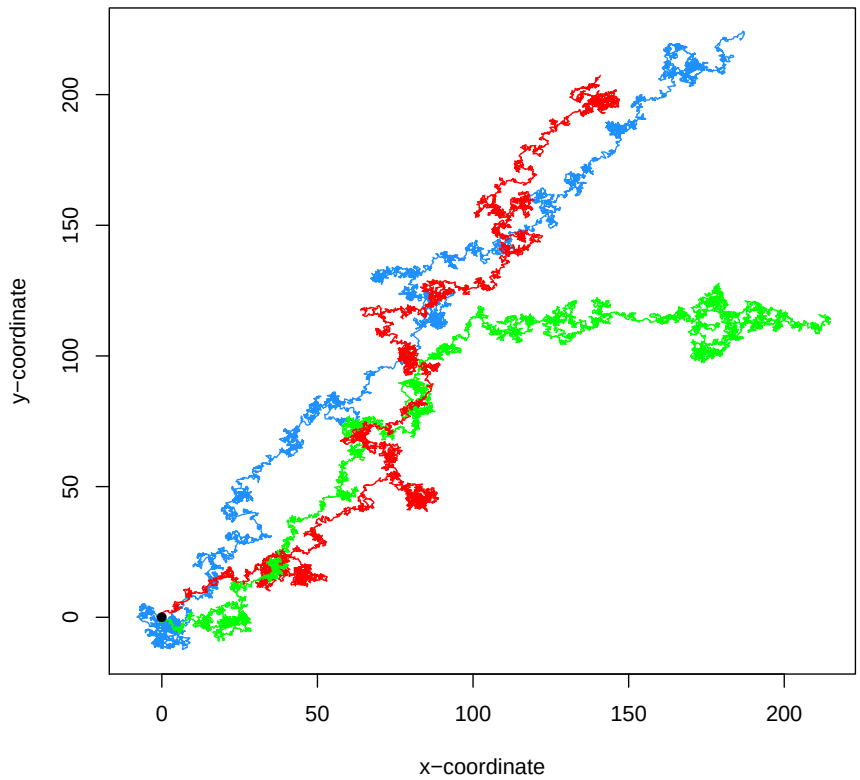

Figure 1.4: Three simulated paths of two dimensional random walk with drift.

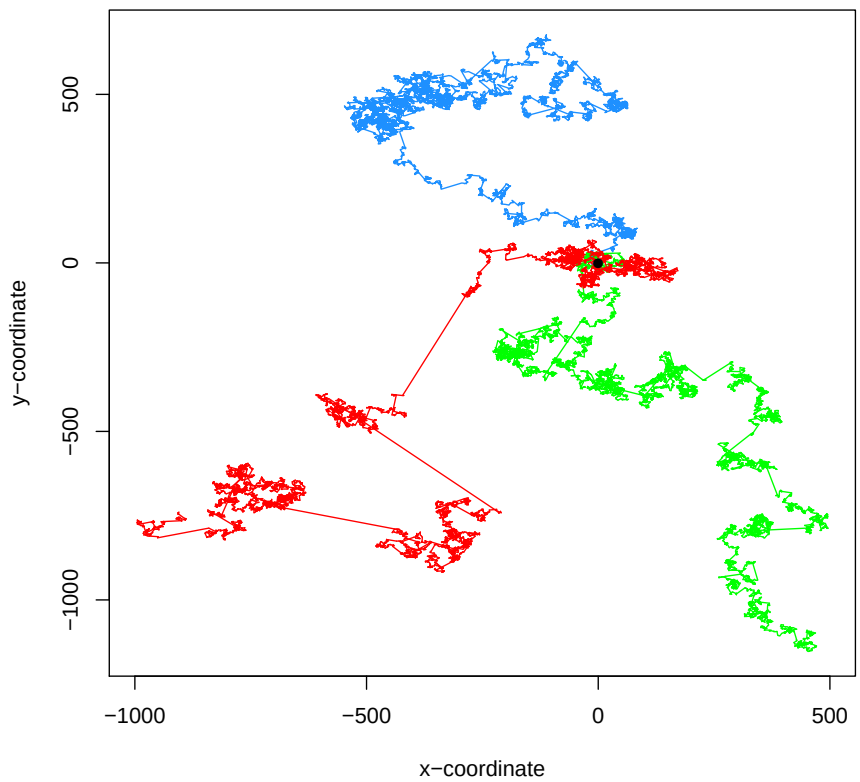

Figure 1.5: Three simulated paths of two dimensional random walk with heavy-tailed distributions.

1.5 Non-homogeneous random walk on $\mathbb{R}^d$

Now we would like to go a step further to ease the restriction of spatial homogeneity. What will happen if we allow the jump distribution to depend on the current location? This means in particular that $\mu(x) := \mathbb{E}[S_{n+1} - S_n | S_n = x]$ becomes a function of the current position $x \in \mathbb{R}^d$. First we should just consider the case that $\mu(x) = \mu$ is a constant (vector) not depending on x. Again if this constant (vector) is not zero (vector), then we will still have the trivial case that the walk will be transient for any dimensions. The interesting case is if we have zero drift. Is this condition enough to determine the recurrence classification? Are we able to draw

similar conclusions as the Chung-Fuchs theorem?

For one dimension, the answer is already quite complicated. See the discussion in [75, p.50]. A zero drift non-homogeneous random walk must be recurrent on $\mathbb{R}_+$, but not on $\mathbb{R}$. Details and a counter example, which is a particular case of Kemperman's oscillating random walk [59], can be found in [90]. The increment law is one of two distributions (with mean zero but finite variance) depending on the walk's present sign. In contrast, for a spatially homogeneous random walk on $\mathbb{R}$, the zero drift condition does imply recurrence, see [58, Chapter 9].

In higher dimensions, the situation is even more subtle. Either recurrent or transient behaviour is possible even for walks with uniformly bounded increments. As a result we quote the following Theorem, as in [75, Theorem 1.5.3].

Theorem 1.5.1. *There exist non-homogeneous random walks with uniformly bounded jumps and $\mu(x) = 0$ for all $x \in \mathbb{R}^d$ that are*

- *transient in $d = 2$;*
- *recurrent in $d \geq 3$.*

A recent paper in 2015 [38] gave some examples with elliptical random walks related to this theorem. They showed that the key property for the recurrence classification is the *increment covariance*. It can be shown that if the increment covariance is *fixed* throughout space, then one recovers the same conclusion as the Chung-Fuchs theorem (recurrence if and only if $d \leq 2$), see Thm 1.5.4 in [75].

Here are some examples of the non-homogeneous elliptic random walks.

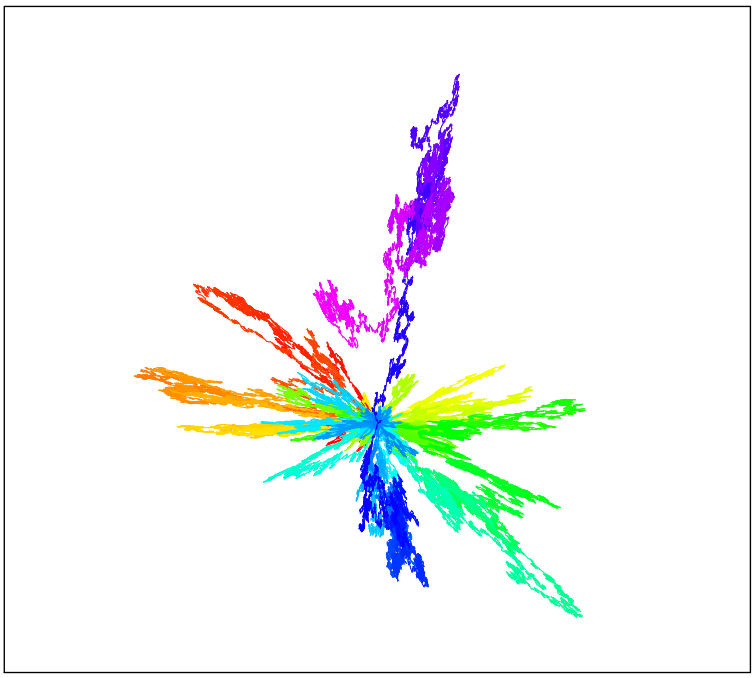

Figure 1.6: A 2D elliptic random walk with comparatively large radial component.

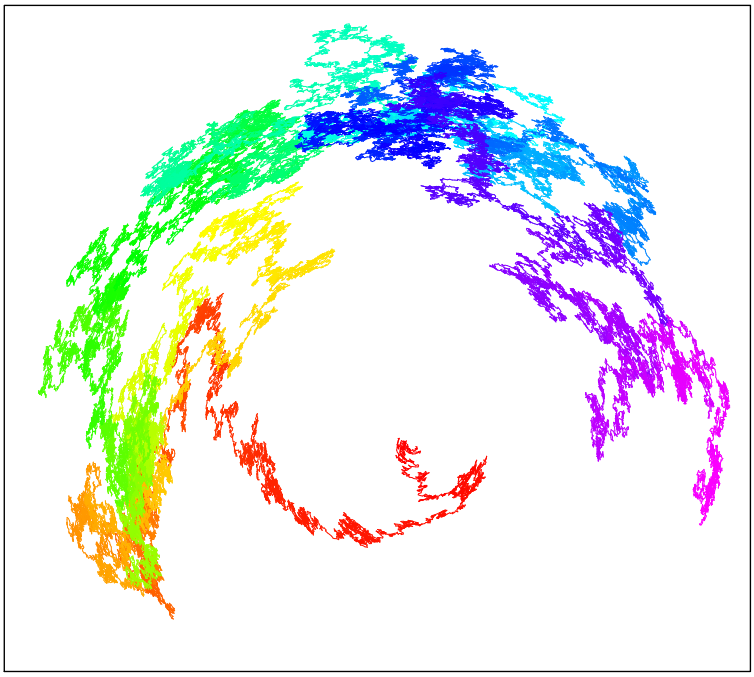

Figure 1.7: A 2D elliptic random walk with comparatively large transversal component.

The general classification for non-homogeneous random walk in $\mathbb{R}^d$ is a long standing open problem. Despite this fact, we are going to present you a full classification on a specially structured state space in Part I of this thesis.

1.6 Law of large numbers and central limit theorem

In this section, we will state the classical results of the law of large numbers and the central limit theorem for homogeneous random walk. This will provide us with a rough idea of how the walk behaves in long term.

In the past, these limit theorems started with the form of a 'law of averages'. It first appeared in a theorem of Bernoulli [10] on the sums of binary random variables, but it was only stated in 1713 over a century after comments of Cardano in his work on dice games [15]. Fifty years later, Halley's treatise of mortality rates [48] clearly expressed a knowledge of decreasing errors in large samples. The term 'law of large numbers' itself wasn't coined until one of Poisson's late works on probability theory in 1837 [80], in which the sum of Bernoulli random variables with varying probabilities of success was shown to converge to the sum of the probabilities; the theorem was only rigorously proved by Chebyshev in 1867 [16].

The first description of a law for more general random variables was produced in 1929 by Kinchin [60] and this became the weak law of large numbers. In the succeeding couple of years, Kolmogorov [61] improved the result to establish the well known strong law, which we will present shortly after in this section.

Now we should formally define the random walk that we are considering and set up the assumptions.

(W) Let $d \in \mathbb{N}$, and suppose that $Z, Z_1, Z_2, \ldots$ are i.i.d. random variables with $\mathbb{E}\,\|Z\| < \infty$ and $\mathbb{E}\,Z = \mu \in \mathbb{R}^d$. The random walk $(S_n, n \in \mathbb{Z}_+)$ is the sequence of partial sums $S_n := \sum_{i=1}^n Z_i$ with $S_0 := 0$.

The first moment condition is not required in the setting of a general homogeneous

walk, but it is necessary for our law of large numbers and central limit theorem to hold. Here is our formal statement for the law of large numbers.

Theorem 1.6.1 (Law of large numbers of a random walk)**.** *Suppose that (**W**) holds, then*

$$\frac{1}{n}(S_n - n\mu) \xrightarrow{\text{a.s.}} 0, \ \ as\ n \to \infty. \tag{1.6.1}$$

The symbol $\xrightarrow{\text{a.s.}}$ stands for almost sure convergence. The proof of this theorem can be found in [27, p.73, Theorem 2.4.1], which follows the classical lines of Etemadi's proof in 1981 [29]. More background material can be found in [100].

To have more control of the walk, in addition to (**W**), we will sometimes assume the following:

(V) Suppose that $\mathbb{E}[\|\xi\|^2] < \infty$. We write $\Sigma := \mathbb{E}[(\xi - \mu)(\xi - \mu)^\top]$ and $\sigma^2 := \operatorname{tr} \Sigma = \mathbb{E}[\|\xi - \mu\|^2$, where Σ is a nonnegative-definite, symmetric d by d matrix.

Again, we may not always have this for the general setting, but have to assume this for the central limit theorem. Now we are ready for another classical result, the Lindeberg–Lévy central limit theorem:

Theorem 1.6.2 (Central limit theorem of a random walk)**.** *Suppose that (**W**) and (**V**) hold; then*

$$\frac{1}{\sqrt{n}}(S_n - n\mu) \xrightarrow{d} \mathcal{N}_d(0, \Sigma), \ \ as\ n \to \infty, \tag{1.6.2}$$

where $\mathcal{N}_d(0, \Sigma)$ is the d-dimensional normal distribution with mean 0 and covariance matrix Σ.

Again, this theorem is an adaptation from [27, p.124, Theorem 3.4.1], and the proof can be found therein.

1.7 Thesis outline

The essence of this thesis consists of three directions of generalization of the classical theory, namely spatial non-homogeneity, structured state space, and derived processes.

First, a considerable amount of literature including books such as [55,67,88,95] is devoted to spatially homogeneous random walks. The spatial homogeneity provides a well behaved model to first consider a difficult problem. However, it restricts the random movement of the particle to be the same in any location in the space, which is often not very realistic due to the underlying environment. This suggests us to study *non-homogeneous random walks*. Compared to the homogeneous random walks, non-homogeneous random walks provide a better understanding of phase transitions and near-critical behaviour. See [75] for a systematic account of non-homogeneous random walks on $\mathbb{R}^d$.

Second, random walks on the standard multidimensional integer lattice are common in the literature. Motivated by certain applications (see Section 2.1), it is also of interest to consider state spaces with additional structure. We include the strip and half strip models, and a generalization of the lattice distribution, in the first and second part of the thesis respectively.

Third, of interest is not only the random walk, but certain other processes derived from the random walk. For example, the Wiener process, also known as the standard Brownian motion, is a limit of random walk. It further expands the universe of random walk to various continuous models including the study of eternal inflation in physical cosmology and the Black-Scholes option pricing model in the mathematical theory of finance [50]. Although Brownian motion has been extensively studied, other simple derived processes remain hidden in the literature as they are very difficult to understand and investigate.

It is a very difficult task to implement all these three new ideas into one model of random walk. Non-homogeneous walks and some derived processes from random walk are quite rarely investigated due to their complexity and difficulty in the treatment of the mathematical structure.

Instead, with these ideas in mind, we hand picked two interesting models in the two main parts of the thesis. The first part will focus on the half strip model. This model consists of the first two elements of generalization of the classical theory. First, instead of the traditional state space on $\mathbb{Z}^d$, we considered a Markov chain on a specially structured state space. This state space gives more useful structure to the model, in particular to apply in certain specific applications, which are impossible to analyse with the traditional state space. Second, instead of restricting the walk to be spatially homogeneous, we allow the walk to be more flexible and only require the walk to converge to a (different) drift on each line. This suggests an extremely general model, to the extent that it is usually more general than all of the situations that most of the applications would need to apply to. Our analysis of the recurrence classification is complete with any sensible parameters for the applications we considered.

The second model is on the centre of mass of homogeneous random walk. It is a simple derived process of the random walk by taking the average of the sum over its past trajectory. Despite the simplicity of the model, almost no literature can be found concerning this process except one in the very special case of simple symmetric random walk.

The material in this thesis is aimed to be as self-contained as possible. After this chapter on general introduction and some basics of random walk theory, this thesis will divided into two parts for two different problems. The first part is about a model with non-homogeneous random walks on an unusual state space called the half strip. Our main focus of this part will be the recurrence classification around the critical region of phase change, and the moment existence or non existence problems of the model, which quantify the degree of recurrence. Our first group of main results includes a complete classification depending on various parameters including the drift and variability of each line, the interactions between the lines, and the probability to change or stay on the same line. The second group of main results provides the necessary and sufficient conditions for the moment existence or non existence depending on the same set of parameters.

The second part of the thesis is about the centre of mass process of the random walk in d-dimensions. We want to investigate the change of the recurrence property when we increase the dimensions. The main results include a local limit theorem, which help us to prove that the process is transient for dimension 2 or higher. Explicitly, we show that the centre of mass process has diffusive rate of escape in the transient case. On the other hand, we proved that the process is recurrent in one dimension. We also give a class of random walks with symmetric heavy-tailed increments for which the centre of mass process is transient in one dimension.

A journey of thoughts starts here.

Part I:

Non-Homogeneous Walks on a Half Strip

Chapter 2

Notation, preliminaries and prerequisites

2.1 Literature review

Markov processes (X_n, η_n) on structured state-spaces Σ contained in $\mathbb{X} \times S$ are of interest in many applications. In this part of the thesis, we are interested in the case where $X_n \in \mathbb{X} = \mathbb{R}_+$ and $\eta_n \in S$ a finite set, in which case Σ is a *half strip*. Motivating applications include

- modulated queues [79], where X_n represents the queue length and η_n tracks the state of a service regime or buffer;
- regime-switching processes in mathematical finance, where η_n tracks a state of the market;
- physical processes with internal degrees of freedom [63], where η_n tracks internal momentum states of a particle.

In much of the literature, η_n is itself a Markov chain; in this case (X_n, η_n) is known as a *Markov-modulated Markov chain* or a *Markov random walk* [2,52]; in the contexts of strips, study of these models goes back to Malyshev [73]. The case where η_n is Markov also includes processes that can be represented as *additive functionals*

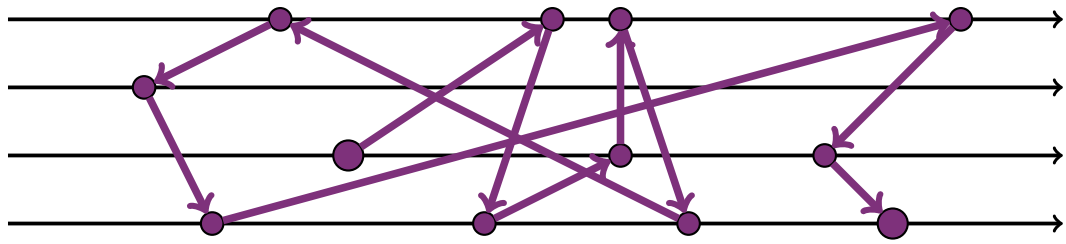

Figure 2.1: An illustration of the half strip model.

of Markov chains [89]. Such models pose a variety of mathematical questions, which have been studied rather deeply over several decades using various techniques that take advantage of the additional Markov structure, and much is now known.

Much less is known when η_n is *not* Markov. In this part of the thesis, following [30, 39], we are interested in the case where η_n is not Markov but is, roughly speaking, approximately Markov when X_n is large, with stationary distribution π_i on S. This relaxation is necessary to probe more intimately the recurrence/transience phase transition for these models. If $\mu_i(x)$ is the mean drift of the $\mathbb{R}_+$-coordinate of the process at $(x, i) \in \Sigma$, then crucial to the asymptotic behaviour of the process are the asymptotics of the μ_i in comparison to the π_i. If $\mu_i(x) \to d_i \in \mathbb{R}$ for each $i \in S$, then the process is transient if $\sum_i \pi_i d_i > 0$ and positive recurrent if $\sum_i \pi_i d_i < 0$ [30, 39]. The critical case $\sum_i \pi_i d_i = 0$ is more subtle, and to investigate the recurrence/transience phase transition it is natural, by analogy with classical work of Lamperti on $\mathbb{R}_+$ [64, 65], to study the case where $\sum_i \pi_i \mu_i(x) = O(1/x)$. In particular, the law of the increments is *non-homogeneous* in X_n, which typically precludes η_n from being Markovian, but admits our weaker conditions.

The *Lamperti drift* case in which every line has $\mu_i(x) = O(1/x)$ was studied in [39], and we will state the results in Section 3.1, with some new techniques to prove the results. The main focus of this part of the thesis is the *generalized Lamperti drift* case where $\mu_i(x) = d_i + O(1/x)$ with $\sum_{i \in S} \pi_i d_i = 0$.

We obtain a recurrence classification for the generalized Lamperti drift case, and in the recurrent case we obtain results on existence and non-existence of passage-

time moments, quantifying the recurrence. We obtain these results by use of a transformation of the process into one with *Lamperti drift*, and so we establish new results on existence and non-existence of passage-time moments in that setting first. Our method is different from that of [39], which relied on an analysis of an embedded Markov chain, in that we make use of some Lyapunov functions for the half-strip model.

2.2 The state space Σ

Let us start with the traditional model in the literature first. We define (X_n, η_n) as a time-homogeneous *irreducible* Markov chain on $\mathbb{Z}_+ \times S$. We need the irreducibility here because we want to keep the recurrence classification as a class property for the whole problem rather than a property in some states. Although all the results in this part of the thesis will be applicable to this model, we would like to first do some modification of the state space. There are technical reasons for this change that we will explain later in Section 3.2, see Remark 3.2.5(a). However, we should now provide some intuition why we should make such a change.

Originally, the Markov chain (X_n, η_n) is on $\mathbb{Z}_+ \times S$. This is very restrictive in terms of the mean drift that you can get from this model. Later in this part, we would like to have a more general non-homogeneous drift. If we stick with this model, then we can only assign a complicated probability on each point in order to achieve the right drift, rather than having the flexibility to assign a simple probability for a point with non-integer horizontal coordinate. In reality it is very tricky to achieve the drift we want: one must carefully pick all those integer-valued jumps to obtain such a subtle drift. This is the reason we want to extend the state space from $\mathbb{Z}_+ \times S$ to Σ, as the following,

- Σ is a locally finite subset of $\mathbb{R}_+ \times S$, where $\mathbb{R}_+$ is the set of positive real numbers and S is a finite and non-empty set.

- $\Lambda_k := \{x \in \mathbb{R}_+ : (x, k) \in \Sigma\}$.

- $\Lambda := \bigcup_{k \in S} \Lambda_k$.
- $S_x := \{i \in S : (x, i) \in \Sigma\}$.

In here, we call Λ_k a line, where $k \in S$ and also Λ as the projection of Σ. S_x stores the information of which line has an accessible state that can project to Λ at a certain horizontal reference point x.

We need to assume Λ_k *unbounded* for each $k \in S$ to make sure that the model is allowed to go to infinity, i.e. be transient, on any line whenever possible to preserve the structure of the model, so that the classification make sense.

Recall that Σ being a locally finite subset of $\mathbb{R}_+ \times S$ means that for any $c \in \mathbb{R}_+$, $\Sigma \cap ([0, c] \times S)$ has finite number of points. Notice here the locally finite property is inherited by each line from the state space.

The local finiteness condition is to ensure that Σ has no finite limit points, so that if (X_n, η_n) is transient, then $X_n \to \infty$. Consider the following example when the local finiteness condition is not satisfied. First we define the state space to be

$$\Sigma = \left(\mathbb{Z}_+ \cup \left\{ \frac{k}{k+1} : k \in \mathbb{Z} \right\} \right) \times \{1\}.$$

Then we assign the transition probabilities as follows,

- $\mathbb{P}(X_{n+1} = \frac{k+1}{k+2} | X_n = \frac{k}{k+1}) = p$, $\mathbb{P}(X_{n+1} = \frac{k-1}{k} | X_n = \frac{k}{k+1}) = 1 - p$ for all $k \in \mathbb{Z}_+$,
- $\mathbb{P}(X_{n+1} = k - 1 | X_n = k) = \mathbb{P}(X_{n+1} = k + 1 | X_n = k) = \frac{1}{2}$ for all $k \in \mathbb{Z}_+$,
- $\mathbb{P}(X_{n+1} = \frac{1}{2} | X_n = 0) = \mathbb{P}(X_{n+1} = 1 | X_n = 0) = \frac{1}{2}$.

When p is close to 1, we can see that whenever the walk goes into the state 0, it has half probability to go to state $\frac{1}{2}$, and then the process has very high tendency not to go back to 0 and keep on increasing, while it does not go to infinity as it would not be greater than 1.

From now we extend and replace the definition of half strips or semi-infinite strips from the state space $\mathbb{Z}_+ \times S$ to Σ unless otherwise specified. Here is our model formally.

(A) Suppose that (X_n, η_n), $n \in \mathbb{Z}_+$, is a time-homogeneous, irreducible Markov chain on Σ, a locally finite subset of $\mathbb{R}_+ \times S$. Suppose that for each $k \in S$ the line Λ_k is unbounded.

Notice that all the results in this part are also applicable to the more restricted state space $\mathbb{Z}_+ \times S$.

2.3 Recurrence classification for the half strip

As described earlier, one of the most important properties to understand for a random walk or a Markov chain is the recurrence classification. Intuitively, as we saw in the introduction, recurrent means that the random walk will always come back to any state in long-run, while transient means the random walk will to go to infinity in some direction and never come back. Some thought is required to see how this applies to the present state space. First, in the vertical direction, S is finite and thus the walk cannot actually escape in this direction. On the other hand, in the horizontal direction $\mathbb{R}_+$, the process cannot escape to the left, but only to the right side. It can escape via any line due to the fact that Λ_k is unbounded for all $k \in S$ when we set up the model. Here is the formal definition for our half strip model.

Lemma 2.3.1. *Let (X_n, η_n) be a time-homogeneous irreducible Markov chain on the state-space Σ. Exactly one of the following holds:*

(i) If (X_n, η_n) is recurrent, then $\mathbb{P}[X_n = x \text{ i.o.}] = 1$ for any $x \in \Lambda$.

(ii) If (X_n, η_n) is transient, then $\mathbb{P}[X_n = x \text{ i.o.}] = 0$ for any $x \in \Lambda$, and $X_n \to \infty$ a.s.

In the former case, we call (X_n) recurrent, and in the latter case, we call (X_n) transient.

Notice that the process (X_n) is not a Markov chain so this is different from our usual definition. This is a lemma but not a definition because it is not trivial that the dichotomy of recurrence and transience holds, i.e. the probability must be 0 or 1 rather than other values. Now we are going to prove Lemma 2.3.1.

Proof. As (X_n, η_n) is an irreducible Markov chain, the states of (X_n, η_n) are either all recurrent or all transient. In the former case, for any $x \in \Lambda$, where $\Lambda = \bigcup_k \Lambda_k$, we have $x \in \Lambda_k$ for some $k \in S$. Then we get $(x, k) \in \Sigma$. That (X_n, η_n) is recurrent means $(X_n, \eta_n) = (x, k)$ *i.o. a.s.*, thus we have $X_n = x$ *i.o. a.s.* This gives $\mathbb{P}(X_n = x \ i.o.) = 1$.

On the other hand, if (X_n, η_n) is transient, for any $x \in \Lambda$, $(X_n, \eta_n) = (x, k)$ only *f.o.* for any k such that $(x, k) \in \Sigma$. Summing over k, of which there are finitely many, we have $X_n = x$ only *f.o.* So we have $\mathbb{P}(X_n = x \ i.o.) = 0$.

This implies $X_n \in R$ *f.o.* for any finite non-empty set $R \in \Lambda$. As Σ is locally finite, we know Λ_k is also locally finite. With the knowledge that S is finite, we get that Λ is locally finite. For any $L \in \mathbb{Z}_+$, denote $R_L = \Lambda \cap [0, L]$, which is finite and non-empty for L large enough. Summing over $X_n = i$ *f.o.* for $i \in R_L$, we have $X_n \in R_L$ *f.o.* as R_L is finite. Hence we have $\liminf_{n\to\infty} X_n \geq L$. As L was arbitrary, we conclude that $\liminf_{n\to\infty} X_n = \infty$. So we have $\lim_{n\to\infty} X_n = \infty$. □

As in the usual random walk, recurrence in the half strip can be further classified as null recurrence or positive recurrence. Again, we have to properly define these concepts due to the complication of the state space. Intuitively, null recurrence means the expected time of return to any point is infinite while it is finite if the random walk is positive recurrent. We also define *null* to be null recurrent or transient. Here are the formal definitions.

Lemma 2.3.2. *Let (X_n, η_n) be a time-homogeneous irreducible Markov chain on the state-space Σ. There exists a unique measure $\nu : \Lambda \to \mathbb{R}_+$ such that*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbf{1}\{X_k = x\} = \nu(x), \quad a.s.$$

Exactly one of the following holds.

(i) If (X_n, η_n) is null, then $\nu(x) = 0$ for all $x \in \Lambda$.

(ii) If (X_n, η_n) is positive recurrent, then $\nu(x) > 0$ for all $x \in \Lambda$ and $\sum_{x\in\Lambda} \nu(x) = 1$.

If X_n is recurrent, then we say that it is null recurrent if (i) holds and positive recurrent if (ii) holds.

This is again a lemma because it is not trivial that the case that $\nu(x) = 0$ for some x and $\nu(x) > 0$ for some other x would not happen. The proof relies on careful separation of the two coordinates of the state space.

Proof. By standard Markov chain theory, e.g. [81], P.35, Theorem 1.7.5 and 1.7.6, there exists a (unique) measure $\phi(x, i) : \Sigma \to \mathbb{R}_+$ such that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbf{1}\{X_k = x, \eta_k = i\} = \phi(x, i), \quad a.s.$$

Define $\nu(x)$ as the projection of $\phi(x, i)$ on the second component, i.e.

$$\nu(x) = \sum_{i\in S_x} \phi(x, i)$$

for any $x \in \Lambda$. Then we get, a.s.,

$$\begin{aligned}
\nu(x) &= \sum_{i\in S_x} \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbf{1}\{X_k = x, \eta_k = i\} \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \sum_{i\in S_x} \mathbf{1}\{X_k = x, \eta_k = i\} \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbf{1}\{X_k = x\}.
\end{aligned}$$

It is very important to notice that the sum for i here is finite so that it can be taken out of the other sum and limit without causing any extra problem. The set S_x is also non-empty because given the fact that $x \in \Lambda$, there exist some $i \in S$ such that $(x, i) \in \Sigma$. So the set $S_x \neq \emptyset$ for $x \in \Lambda$.

Now when (X_n, η_n) is null, then $\phi(x, i) = 0$ for all $(x, i) \in \Sigma$, so $\nu(x) = \sum_{i\in S_x} \phi(x, i) = 0$, always bearing in mind that we are doing a finite sum.

For (X_n, η_n) positive recurrent, $\phi(x, i) > 0$ for all $(x, i) \in \Sigma$ and hence $\nu(x) > 0$ since as $\nu(x) = \sum_{i\in S_x} \phi(x, i) > 0$ and the sum is not empty. With the fact that $\sum_{(x,i)\in\Sigma} \phi(x, i) = 1$, we can separate the sum across the two coordinates and get $\sum_{x\in\Lambda} \sum_{i\in S_x} \phi(x, i) = 1$. This is the same as saying $\sum_{x\in\Lambda} \nu(x) = 1$. Hence all of the claims in the lemma are proved. □

2.4 Assumptions of the model

To solve our recurrence classification problem, we also need the following technical assumptions. First, to be realistic, we first need to assume the displacement of the X-coordinate has bounded p-moments for some $p < \infty$. This is a crucial but weak assumption because without this, there will be no control of the size of jumps. We do not want the walk have an increasing size of boundless jumps when it is at the position far on the right side. In this bad behaviour the walk can suddenly jump back to the far left or have a very big jump on the right in one step, so that all the steps that the walk had before are negligible. So we would like to impose this uniform bound for the walk to get some regularity to predict the long term behaviour.

($\mathbf{B}_p$) There exists a constant $C_p < \infty$ such that for all $n \in \mathbb{Z}_+$,

$$\mathbb{E}[|X_{n+1} - X_n|^p \mid X_n = x,\ \eta_n = i] \leq C_p, \text{ for all } (x, i) \in \Sigma.$$

We will need $p > 2$ most of the time in this part of the thesis, which we sometimes refer to as demanding that 'two moments exist'. However, for some of the results, $p > 1$, i.e. 'one moment exists' is already sufficient.

We define $p(x, i, y, j)$ as the transition probabilities of our irreducible Markov chain $(X_n, \eta_n) \in \Sigma$, i.e.

$$\mathbb{P}[(X_{n+1}, \eta_{n+1}) = (y, j) \mid (X_n, \eta_n) = (x, i)] = p(x, i, y, j).$$

For the sake of reasonable behaviour of the probabilities so that we can have the unique stationary distribution $\boldsymbol{\pi}$ from the embedded process in the vertical, i.e. η, direction, we need to assume that η_n is approximately Markov when X_n is large. First, we define

$$q_{ij}(x) = \sum_{y \in \Lambda_j} p(x, i, y, j) \tag{2.4.1}$$

as we do not need the information of the exact point that the walk is jumping to, but only which line it jumps to and which point it starts from. Here is our assumption:

($\mathbf{Q}_\infty$) Suppose that $\lim_{x\to\infty} q_{ij}(x) = q_{ij}$ exists for all $i, j \in S$, and (q_{ij}) is an irreducible stochastic matrix.

Now if we assume (Q_∞), then we can define a new process (η_n^*), $n \in \mathbb{Z}_+$, as a Markov chain on S with transition probabilities given by q_{ij}. As (η_n^*) is irreducible and finite, we know that there exists a *unique* stationary distribution $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_{|S|})^\top$ on S with $\pi_j > 0$ for all $j \in S$ and satisfying $\pi_j = \sum_{i\in S} \pi_i q_{ij}$ for all $j \in S$. (Q_∞) is very important here because if $\boldsymbol{\pi}$ does not exist, then we cannot define the total average drift of the whole system, which determines the recurrence classification.

Naturally, we want to specify the movement of the chain by the one-step mean (horizontal) drift at each point on each line, i.e., its first moment in the X-coordinate on line i. This is:

$$\mu_i(x) := \mathbb{E}[X_{n+1} - X_n \mid X_n = x, \eta_n = i] = \sum_{j\in S} \sum_{y\in\Lambda_j} (y - x)p(x, i, y, j);$$

notice that $\mu_i(x)$ is finite if (B_p) holds for some $p \geq 1$. In the simplest case, we suppose that each line has an asymptotically *constant drift*, and we assume

($\mathbf{D_C}$) For each $i \in S$ there exists $d_i \in \mathbb{R}$ such that $\mu_i(x) = d_i + o(1)$ as $x \to \infty$.

Although this is called the constant drift, from the term $o(1)$ we actually allow $\mu_i(x)$ to fluctuate around the constants, as long as the fluctuation converges to zero when $x \to \infty$. In some sense, only the behaviour when x is big matters.

Instead of stating the original theorems by Malyshev [73] or Falin [30], we shall state a slightly generalised and polished result in a paper of Georgiou and Wade [39], for the model that we are using now.

Theorem 2.4.1 (Georgiou, Wade, 2014, amended)**.** *Suppose that* (A) *holds, and that* (B_p) *holds for some* $p > 1$*. Suppose also that* (Q_∞) *and* ($\mathrm{D_C}$) *hold. Then the following classification applies.*

- *If* $\sum_{i\in S} d_i\pi_i > 0$*, then* (X_n, η_n) *is transient.*
- *If* $\sum_{i\in S} d_i\pi_i < 0$*, then* (X_n, η_n) *is positive recurrent.*

Theorem 2.4.1 is a minor generalization of Theorem 2.4 of [39], which took $\Sigma = \mathbb{Z}_+ \times S$; the proof there readily extends to the statement here. We give an alternative proof, using Lyapunov functions, in Section 4.4. Earlier versions of the result, which had the extra assumption that $q_{ij}(x) = q_{ij}$ not depending on x, are Theorem 3.1.2 of [32] and the results of [30]. The proof in [39] is based on the investigation of the embedded process (Y_n), which records the X-coordinate of the chain when it returns to a given line. They use increment moment estimates together with some Foster-Lamperti conditions to classify the process (Y_n), and then deduce the classification for (X_n) from the equivalence results.

Intuitively, $\sum_{i\in S} d_i\pi_i$ stands for the total average drift of the system, as it is summing over all lines with the average drift on each line multiplied by the proportion of time spent on the line. So if the total average drift is positive, the walk has the tendency to go to the right on average, thus it is difficult for the process to return to the points on the left in long term, and the walk is transient. On the other hand, if we have a negative total average drift, then the walk will have the tendency to go to the left, and keep coming back to the left boundary, thus the walk is (positive) recurrent.

As you can see, Theorem 2.4.1 has nothing to say about the much more subtle case where $\sum_{i\in S} d_i\pi_i = 0$. One natural guess would just be null recurrence whenever the condition is satisfied but this is not always true. In fact, the model can fall into any classification, i.e., it can be positive recurrent, null recurrent or transient. Here further assumptions are required to reach any conclusion.

One way to achieve $\sum_{i\in S} d_i\pi_i = 0$ is to have $d_i = 0$ for all $i \in S$. In this case, by analogy with the classical one-dimensional work of Lamperti [64, 65], the natural setting in which to probe the recurrence-transience phase transition is that of *Lamperti drift*, as studied in [39], which we present in Section 3.1. In this setting we give new results on existence and non-existence of moments of passage times.

The second possibility and the most subtle case, in which $d_i \neq 0$ for some $i \in S$ but nevertheless $\sum_{i\in S} d_i\pi_i = 0$, leads to what we call *generalized Lamperti drift*, which is the main focus of this part of the thesis and is presented in Section 3.2. Here

we establish a recurrence classification as well as results on passage-time moments.

The proof of the theorems introduced in these sections will be delayed until Chapter 4, after we introduce various techniques related to Lyapunov functions method, martingale theory and some well known linear algebra results.

2.5 The Lamperti problem

For the first step to probe the recurrence classification for the Lamperti drift case in our half strip problem, we should recall the origin of the name, i.e., the Lamperti problem, see [75], Section 1.3 and Chapter 3.

We start again with the simple symmetric random walk S_n on $\mathbb{Z}^d$, and start the walk at the origin. This time instead of going through the standard proof of Pólya's recurrence theorem to get the recurrence classification, we will try a different method. First we reduce this d-dimensional problem into a one dimensional one by the *Lyapunov function*, a transformation process given by

$$X_n := \|S_n\|,$$

where $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^d$. Hence X_n is just the distance between the origin and the particle at time n. So now the stochastic process will take values in $S := \{\|x\| : x \in \mathbb{Z}^d\}$, a countable subset of the half line $\mathbb{R}_+$. Notice that the recurrence classification property will transfer from S_n to X_n, since $S_n = 0$ if and only if $X_n = 0$. However the Markov property was sacrificed for the reduction in dimensionality. One can easily observe, say in two dimension, for the same value of X_n on different positions for S_n may give different distributions, thus the Markov property will not hold for X_n, see the example in [75], Section 1.3. Hence from this point, we need to have a method to find the recurrence classification of X_n, which does not heavily depend on the Markov property.

This topic leads to a more general area called the *Lamperti problem*, introduced by John Lamperti [64, 65] in early 1960s. Informally, let us begin with a discrete-time time-homogeneous Markov process X_n with well-defined increments moment

functions

$$m_k(x) = \mathbb{E}\left[(X_{n+1} - X_n)^k | X_n = x\right]$$

for all $k \geq 0$. Having a uniform bound on the increments can easily guarantee this condition, but is is not necessary. The Lamperti problem is asking if we are given the first few moments, especially the first two, μ_1 and μ_2, how to determine the asymptotic behaviour of X_n. If we indeed impose the uniform bound condition formally,

$$\mathbb{P}\left(|X_{n+1} - X_n| \leq B\right) = 1 \tag{2.5.1}$$

for some $B \in \mathbb{R}_+$, then we can have a slightly modified version of Lamperti's fundamental recurrence classification, see Theorem 1.3.1 of [75].

Theorem 2.5.1 (Lamperti, 1960)**.** *Suppose that X_n is a Markov process on S satisfying* (2.5.1)*. Under mild conditions on irreducibility, the following recurrence classification holds. Let $\varepsilon > 0$.*

- *If $2xm_1(x) + m_2(x) < -\varepsilon$, then X_n is positive recurrent;*
- *If $2x|m_1(x)| \leq m_2(x) + O(x^{-\varepsilon})$, then X_n is null recurrent;*
- *If $2xm_1(x) - m_2(x) > \varepsilon$, then X_n is transient;*

Notice that the null recurrence classification is slightly sharper than Lamperti's original results. This theorem states that if the absolute value of the first moment is large enough compared to the second moment in the tail (infinite side) of the walk, then the process will have enough force to go in the specific direction, left or right, depending on the sign of the drift, resulting in transience or positive recurrence. Otherwise, if the absolute value of (twice) the drift is not large enough compared to the variance, then the walk does not have enough force to go in a specific direction, as the variance dominates the effect of the drift, resulting in the null-recurrent case.

Although this version of the theorem does not directly give us the Pólya's Theorem because of the lost of Markov property stated before, this method is still applicable by slight modification of the definition of μ_k. By computing the first and

second moment of X_n explicitly for this simple symmetric random walk S_n, we get

$$\mathbb{E}\left[X_{n+1} - X_n | S_n = x\right] = \left(\frac{d-1}{2d}\right)\frac{1}{\|x\|} + O(\|x\|^{-2})$$
$$\mathbb{E}\left[(X_{n+1} - X_n)^2 | S_n = x\right] = \frac{1}{d} + O(\|x\|^{-1}).$$

So the corresponding terms in the theorem will be

$$2xm_1(x) = 1 - \frac{1}{d} + O(x^{-1})$$
$$m_2(x) = \frac{1}{d} + O(x^{-1}).$$

Hence using the theorem we get S_n is transient if and only if

$$1 - \frac{1}{d} > \frac{1}{d},$$

which is equivalent to $d > 2$. For the technical details see [75] Section 3.5. As you can see, this is a potent way to prove the Pólya's Theorem. With the sole and elementary computations of the increment moments of X_n using Taylor's theorem, the method can generalize to a broad range of random walks, and does not require any special structure on the original process.

Finally, back to our half strip model, if we take the special case that S, the vertical component of Σ to be a singleton, it reduces back to the model in the Lamperti problem. So one might see the half strip model is actually a generalization of the Lamperti problem. One may think we can easily push the Lamperti's fundamental recurrence classification result through the half strip model. However, the real situation is much more difficult than that. There is no doubt that if all of the lines have the same classification, say transient, then the whole system of the half strip will also be transient, because no matter which line the process is on, we still have the tendency to go to infinity on the right side. However, what if some of the lines are recurrent and some of the lines are transient? Then the result is not clear, as it depends on how much time the process spends on each line and how recurrent or transient each line is. In Section 2.4, we gave the result when we have a non-zero total average drift, and in the next chapter we will discuss the subtle case when we

have zero total average drift, starting with the special case of *Lamperti drift*, and complete the classification with *generalised Lamperti drift*.

Chapter 3

Main results

3.1 Lamperti drift on a half strip

3.1.1 Recurrence classification

For the remainder of this part of the thesis we introduce the following shorthand to simplify notation:

$$\mathbb{E}_{x,i}[\,\cdot\,] = \mathbb{E}[\,\cdot\ \mid X_n = x,\ \eta_n = i].$$

Continuing with our half strip model, we would like to probe the classification in the special case with zero total average drift, i.e. $\sum_{i\in S} d_i\pi_i = 0$. To proceed with more complicated drifts, as in the Lamperti's fundamental recurrence classification, we need to have some control on the variance, i.e. the second moment of the increments. So we define, for $(x, i) \in \Sigma$,

$$\sigma_i^2(x) := \mathbb{E}_{x,i}[(X_{n+1} - X_n)^2];$$

note that $\sigma_i^2(x)$ is finite if (B_p) holds for some $p \geq 2$. The formal definition for the Lamperti drift case of the half strip model is as follows:

($\mathbf{D_L}$) For each $i \in S$ there exist $c_i \in \mathbb{R}$ and $s_i^2 \in \mathbb{R}_+$, with at least one s_i^2 non-zero, such that, as $x \to \infty$, $\mu_i(x) = \frac{c_i}{x} + o(x^{-1})$ and $\sigma_i^2(x) = s_i^2 + o(1)$.

The reason that we named this case the Lamperti drift is because the problem has a very similar structure and result as in the Lamperti problem. And in fact for our

half strip state space Σ, if we take S to be a singleton, it returns to the well-known Lamperti problem. Results in this chapter hence cover the results from Lamperti.

In this case, comparing to (D_C), we have $d_i = 0$ for all $i \in S$. We specify the error in $o(1)$ can be in the natural form $\frac{c_i}{x} + o(x^{-1})$, but it is possible to impose the drift in other forms such as $\frac{c_i}{\sqrt{x}}$. The exact form of the drift does not actually affect the theory here but the calculation would be different. So for the time being we will stick with the traditional drift type coinciding with the representation in the Lamperti problem.

To obtain results at the critical point for the phase transition we will need to strengthen the assumptions (Q_∞) and $(\mathrm{D_L})$ by imposing additional assumptions:

$(\mathbf{Q}_\infty^+)$ Suppose that there exists $\delta_0 \in (0,1)$ such that $\max_{i,j\in S} |q_{ij}(x) - q_{ij}| = O(x^{-\delta_0})$ as $x \to \infty$.

$(\mathbf{D}_\mathbf{L}^+)$ Suppose that there exist $\delta_1 \in (0,1)$, $c_i \in \mathbb{R}$, and $s_i^2 \in \mathbb{R}_+$, with at least one s_i^2 non-zero, such that for all $i \in S$, as $x \to \infty$, $\mu_i(x) = \frac{c_i}{x} + o(x^{-1-\delta_1})$ and $\sigma_i^2(x) = s_i^2 + o(x^{-\delta_1})$.

We need these assumptions in the critical case to have slightly more control on the error terms of the transition probability and the mean and variance of the horizontal increments. In the Lamperti drift setting, we have the following recurrence classification.

Theorem 3.1.1. *Suppose that* (A) *holds, and that* (B_p) *holds for some* $p > 2$. *Suppose also that* (Q_∞) *and* $(\mathrm{D_L})$ *hold. Then the following classification applies.*

- *If* $\sum_{i\in S}(2c_i - s_i^2)\pi_i > 0$, *then* (X_n, η_n) *is transient.*
- *If* $|\sum_{i\in S} 2c_i\pi_i| < \sum_{i\in S} s_i^2\pi_i$, *then* (X_n, η_n) *is null recurrent.*
- *If* $\sum_{i\in S}(2c_i + s_i^2)\pi_i < 0$, *then* (X_n, η_n) *is positive recurrent.*

If, in addition, (Q_∞^+) *and* $(\mathrm{D}_\mathrm{L}^+)$ *hold, then the following condition also applies (yielding an exhaustive classification):*

- *If* $|\sum_{i\in S} 2c_i\pi_i| = \sum_{i\in S} s_i^2\pi_i$, *then* (X_n, η_n) *is null recurrent.*

Theorem 3.1.1 is a slight generalization of Theorem 2.5 of [39], which took $\Sigma = \mathbb{Z}_+ \times S$. The proof in [39], which made use of Lamperti's [64, 65] results applied to the embedded process obtained by observing the X-coordinate on each visit to a reference line, extends readily to the statement here. We give an alternative proof in Section 4.5 of the first three points in the theorem (not the critical case).

We can use similar intuition behind Theorem 2.5.1 to understand the theorem here. Instead of considering only one line, we consider the weighted average of the total drift with the weighted average of the total variance in the system, weighting on the proportion of time spent on each line. If the absolute value of the former is large enough compared to the latter, then it will give the system a strong enough push to a direction either right or left in average, depending on the sign of the drift, resulting in transience or positive recurrence accordingly. However, if the absolute value of the former is not big enough, the walk will not be able to generate enough force to overcome the second moment, thus giving the null-recurrent case.

In the next subsection, we will quantify these two forces from the first and second moment. Comparing the size of these will give us the knowledge of the degree of recurrence of the process.

3.1.2 Existence and non-existence of moments

In the case of recurrence, we can actually quantify how recurrent the process is. Instead of just having the classification of positive recurrent and null recurrent, one way to obtain quantitative information on the nature of recurrence is to study moments of *passage times*. For $x \in \mathbb{R}_+$, define the stopping time

$$\tau_x := \min\{n \geq 0 : X_n \leq x\}. \tag{3.1.1}$$

In the positive-recurrent situation, we have that $\mathbb{E}[\tau_x] < \infty$ a.s., for all x sufficiently large. In the case of null, $\mathbb{E}[\tau_x^s] = \infty$ a.s., for all $s \geq 1$, and x sufficiently large.

First we state a result that gives conditions for $\mathbb{E}[\tau_x^s]$ to be finite.

Theorem 3.1.2. *Suppose that* (A) *holds, and that* (B_p) *holds for some* $p > 2$.

Suppose also that (Q_∞) *and* $(\mathrm{D_L})$ *hold. If for some* $\theta > 0$,

$$\sum_{i\in S}\left[2c_i + (2\theta-1)s_i^2\right]\pi_i < 0, \tag{3.1.2}$$

then for any $s \in \left[0, \theta \wedge \frac{p}{2}\right]$, *we have* $\mathbb{E}[\tau_x^s] < \infty$ *for all* x *sufficiently large.*

We have the following result in the other direction.

Theorem 3.1.3. *Suppose that* (A) *holds, and that* (B_p) *holds for some* $p > 2$. *Suppose also that* (Q_∞) *and* $(\mathrm{D_L})$ *hold. If for some* $\theta \in (0, \frac{p}{2}]$,

$$\sum_{i\in S}\left[2c_i + (2\theta-1)s_i^2\right]\pi_i > 0, \tag{3.1.3}$$

then for any $s \in \left[\theta, \frac{p}{2}\right]$, *we have* $\mathbb{E}[\tau_x^s] = \infty$ *for all* x *sufficiently large.*

In the case where S is a singleton, Theorems 3.1.2 and 3.1.3 reduce to versions of Propositions 1 and 2, respectively, of [5] on passage-time moments for Markov chains on $\mathbb{R}_+$.

Using these two theorems, by plugging in different values of θ in the expression $\sum_{i\in S}\left[2c_i + (2\theta-1)s_i^2\right]\pi_i$, we can pinpoint which moments of the passage times exist or not. In short, if more moments exist then the process is more recurrent, and we should expect a smaller scale of time for the process to return.

We also see that if we put $\theta = 1$ in Theorems 3.1.2, we can see the moments of the passage time exists for all $s \in [0, 1]$, implying that the process is positive recurrent. if we put $\theta \to 0^+$ in Theorems 3.1.3, we can see that the moments of the passage time does not exists for all $s \in [0, \frac{p}{2}]$, implying that the process is null. (This does not directly imply transient unfortunately because some null-recurrent random walk can also have no moment exist, e.g. simple random walk on $\mathbb{Z}^2$.)

Intuitively, these two theorems add an extra parameter θ in the equation, comparing to Theorem 3.1.1, which gives some extra flexibility on how tolerant is the drift size comparing to the variance. For Theorems 3.1.2, the stronger the restriction on c_i, i.e. imposing a larger θ, the more moments you can get from the passage time. This means if there is a larger θ that satisfies the equation in the theorem,

the process is more 'recurrent' in some sense. On the opposite hand, if we impose a smaller θ, giving more flexibility to c_i, you will get fewer moments as a result.

Theorem 3.1.3 is essentially the opposite consideration of Theorem 3.1.2. Its use is to pinpoint the critical value of s which gives you the existence-non-existence phase transition.

The proofs of Theorem 3.1.2 and Theorem 3.1.3 will be presented in Chapter 4, with the use of some specific Lyapunov functions and some semi-martingale methods. Notice that we need to use different functions for the proofs of Theorems 3.1.2 and Theorems 3.1.3, and there is no direct relation between them.

The next section will discuss the most subtle case when $d_i \neq 0$ for some $i \in S$ but nevertheless $\sum_{i\in S} d_i\pi_i = 0$, which is what we call the *generalized Lamperti drift.*

3.2 Generalized Lamperti drift on a half strip

3.2.1 Recurrence classification

Now we turn to the main topic of this part of the thesis. The last case is when some (or all) of the lines have non-zero constant drift, but the total average drift is zero. This case is the most subtle, as it is possible to construct some examples with the same $\mu_i(x)$ and $\sigma_i(x)$ but which fall into different classifications. We will show some explicit examples in Chapter 5. We discovered that the asymptotic properties of the process depend not only on $\mu_i(x)$ and $\sigma_i^2(x)$ but also on the quantities

$$\mu_{ij}(x) := \mathbb{E}_{x,i}\left[(X_{n+1} - X_n)\mathbf{1}\{\eta_{n+1} = j\}\right];$$

this alerts us to the fact that correlations between the components of the increments are now crucial. The case of generalized Lamperti drift is the following. To avoid confusion with the Lamperti drift case, we changed the symbols for c_i and s_i to e_i and t_i.

($\mathbf{D_G}$) For $i, j \in S$ there exist $d_i \in \mathbb{R}$, $e_i \in \mathbb{R}$, $d_{ij} \in \mathbb{R}$ and $t_i^2 \in \mathbb{R}_+$, with at least one t_i^2 non-zero, such that

(a) for all $i \in S$, $\mu_i(x) = d_i + \frac{e_i}{x} + o(x^{-1})$ as $x \to \infty$;

(b) for all $i \in S$, $\sigma_i^2(x) = t_i^2 + o(1)$ as $x \to \infty$;

(c) for all $i, j \in S$, $\mu_{ij}(x) = d_{ij} + o(1)$ as $x \to \infty$; and

(d) $\sum_{i \in S} \pi_i d_i = 0$.

Note that necessarily we have the relation $d_i = \sum_{j \in S} d_{ij}$.

As in the Lamperti drift case, we need to have an additional condition at the phase boundary.

($\mathbf{D_G^+}$) There exist $\delta_2 \in (0,1)$, $d_i \in \mathbb{R}$, $e_i \in \mathbb{R}$, $d_{ij} \in \mathbb{R}$ and $t_i^2 \in \mathbb{R}_+$, with at least one t_i^2 non-zero, such that

(a) for all $i \in S$, $\mu_i(x) = d_i + \frac{e_i}{x} + o(x^{-1-\delta_2})$ as $x \to \infty$;

(b) for all $i \in S$, $\sigma_i^2(x) = t_i^2 + o(x^{-\delta_2})$ as $x \to \infty$; and

(c) for all $i, j \in S$, $\mu_{ij}(x) = d_{ij} + o(x^{-\delta_2})$ as $x \to \infty$.

We also must impose refined forms of the condition (Q_∞), where now further terms come into play.

($\mathbf{Q_G}$) For $i, j \in S$ there exist $\gamma_{ij} \in \mathbb{R}$ such that $q_{ij}(x) = q_{ij} + \frac{\gamma_{ij}}{x} + o(x^{-1})$, where (q_{ij}) is a stochastic matrix.

($\mathbf{Q_G^+}$) There exist $\delta_3 \in (0,1)$ and $\gamma_{ij} \in \mathbb{R}$ such that $q_{ij}(x) = q_{ij} + \frac{\gamma_{ij}}{x} + o(x^{-1-\delta_3})$.

The fact that $\sum_{j \in S} q_{ij}(x) = 1$ implies, after the following calculation, that $\sum_{j \in S} \gamma_{ij} = 0$ for all $i \in S$.

First as the sum of all the transition probabilities on a line is 1, we have

$$\sum_{j \in S} q_{ij}(x) = 1.$$

Plugging in the condition ($\mathbf{Q_G}$), we get

$$\sum_{j \in S} \left(q_{ij} + \frac{\gamma_{ij}}{x} + o(x^{-1}) \right) = 1.$$

Simplifying,

$$\sum_{j\in S}\frac{\gamma_{ij}}{x}=o(x^{-1})$$

for all $x\in\Lambda$. By choosing appropriate $x\in\Lambda$, we have

$$\left|\sum_{j\in S}\gamma_{ij}\right|\leq\epsilon.$$

Since $\epsilon>0$ was arbitrary, we get

$$\sum_{j\in S}\gamma_{ij}=0.$$

The underlying intuition of how many terms we should consider before the error term for each parameter is quite interesting. In principle, we need to take the same order on every basic variable to get the balance of the estimation. That is if we take the first two order terms on the drift of each line, it is sensible to take the first two terms of the transition probabilities. However because the second moment and the interaction between the lines is already on one higher level of the model, as they are like the first level, i.e. pairwise interaction between the basic variables, we only need the first term of the estimation. So now we can have every parameter on the same accuracy of consideration, and it turns out that this accuracy level is enough for determining our classification.

This time, for understanding the statement of our recurrence classification in the generalized Lamperti case, we need the following preliminary result on solutions $\mathbf{a}=(a_1,\ldots,a_{|S|})^\top$ to the system of equations

$$d_i+\sum_{j\in S}(a_j-a_i)q_{ij}=0,\ \text{for all } i\in S; \tag{3.2.1}$$

we say that a solution $\mathbf{a}=(a_1,\ldots,a_{|S|})^\top$ is *unique up to translation* if all solutions $\mathbf{a}'=(a'_1,\ldots,a'_{|S|})^\top$ have a'_j-a_j constant for all $j\in S$.

Lemma 3.2.1. *Let $d_i\in\mathbb{R}$ and (q_{ij}) be an irreducible stochastic matrix with stationary distribution π. Then the following statements are equivalent.*

- $\sum_{i\in S}d_i\pi_i=0$.

- *There exists a solution* $\mathbf{a} = (a_1, \ldots, a_{|S|})^\top$ *to* (3.2.1) *that is unique up to translation.*

For the proof of Lemma 3.2.1, see Section 4.3.

Next we give our main recurrence classification for the model with generalized Lamperti drift. The criteria involve solutions to (3.2.1); as described in Lemma 3.2.1 such solutions are not unique, but nevertheless the expressions in which they appear in Theorem 3.2.2 are invariant under translations (see Remark 3.2.5(c)), and so the statement makes sense.

Theorem 3.2.2. *Suppose that* (A) *holds, and that* (B_p) *holds for some* $p > 2$. *Suppose also that* ($\mathrm{Q_G}$) *and* ($\mathrm{D_G}$) *hold. Define* $\mathbf{a} = (a_1, \ldots, a_{|S|})^\top$ *to be a solution to* (3.2.1) *whose existence is guaranteed by Lemma 3.2.1. Define*

$$U := \sum_{i \in S} \left(2e_i + 2 \sum_{j \in S} a_j \gamma_{ij} \right) \pi_i, \text{ and } V := \sum_{i \in S} \left(t_i^2 + 2 \sum_{j \in S} a_j d_{ij} \right) \pi_i. \tag{3.2.2}$$

Then the following classification applies.

- *If* $U > V$ *then* (X_n, η_n) *is transient.*
- *If* $|U| < V$ *then* (X_n, η_n) *is null recurrent.*
- *If* $U < -V$ *then* (X_n, η_n) *is positive recurrent.*

If, in addition, ($\mathrm{Q_G^+}$) *and* ($\mathrm{D_G^+}$) *hold, then the following condition also applies (yielding an exhaustive classification):*

- *If* $|U| = V$ *then* (X_n, η_n) *is null recurrent.*

From this complicated theorem, you can see that each of the parameters has its own role in controlling the recurrence classification. The a_i's here are actually a key element to the proof of the theorem. They give the shift on each line in the state space, resulting in a transformation to the system. In this way, the system is aligned in a way that the constant term d_i's in the drift are eliminated and we can recover the Lamperti drift after the transformation. When all a_i's are zero, it

actually implies all d_i's are zero, and Theorem 3.2.2 recovers the Lamperti drift case as in Theorem 3.1.1.

After this transformation on a_i's, the effects of d_i's transfer to the a_i's, so similar to the Lamperti drift type, we can just compare the size of the Lamperti component of the drift, e_i's to the second moment t_i's, with the proportion of time spent on each line, given by π_i's, and most importantly, the effect on the shifting of lines. That is the reason why now we have got some extra terms, with the interactions, γ_{ij} and d_{ij} coming into play, depending on the weight that how much we shift the line. Focusing on a single line i, the larger value of γ_{ij} from any point on any line j in the same direction of the Lamperti drift component, e_i's , with the same direction of the shift a_i, (decrease in the other direction) will help to increase the total of the drift, thus giving more force to walk on that line to go either transient or positive recurrent depending on the direction. If the increase on the second term of the transition probability is either opposite to the direction of the Lamperti component of the drift, or the direction of the shift (not both), then they will cancel out each other. So it will have a counter effect on the drift thus lower the force to go through the fluctuation of the variance of the line, giving a higher tendency to go to the case of null recurrence. In the last case that the the transition probability is increases in both the opposite direction of the Lamperti drift component and the direction of shift, these two opposing signs will work together thus increase the force on the line to go to either transient or positive recurrent depending on the direction of the Lamperti drift component. Vice versa for the case of decreasing the transition probabilities.

The other quantity d_{ij}, on the other hand, would affect the power of the second moment of the walk. Again, it depends also on the fact if the sign of a_i is the same as the interacting drift d_{ij} or not. The sign of the variance plays no role here because it is always positive. This means for a specific line, if a_i is positive, i.e., shifting to the right, then if d_{ij} is also positive (same direction), then increasing the interacting drift d_{ij} would also increase the fluctuation of the walk. This will help to increase the corrected variance and the walk on this line will need more drift in order to go pass

the effect of the second moment. So this increase the tendency for the walk to go to the null-recurrent case. The same happens when both a_i and d_{ij} is negative as they also help each other in the same way. On the contrary, if they have a different sign, increasing d_{ij} would decrease the fluctuation of the walk, thus shorten the tolerance gap for small drifts. This would mean that the walk now need a smaller drift to go though the variance and result in transient or positive recurrent, depending on the sign of the Lamperti drift component.

Weighting these tendencies with the right proportion of time spent on each line, it will adjust the right comparison with the corrected drift and variance in the whole system on average, thus giving you the right classification.

The proof of this theorem will be the main focus of Chapter 4.

3.2.2 Existence and non-existence of moments

As in Section 3.1, we quantify the degree of recurrence by establishing existence and non-existence of moments of the passage times τ_x as defined at (3.1.1). First we give conditions for existence of moments.

Theorem 3.2.3. *Suppose that* (A) *holds, and that* (B_p) *holds for some* $p > 2$. *Suppose also that* ($\mathrm{Q_G}$) *and* ($\mathrm{D_G}$) *hold. Define* $\mathbf{a} = (a_1, \ldots, a_{|S|})^\top$ *to be a solution to* (3.2.1) *whose existence is guaranteed by Lemma 3.2.1. If for some* $\theta > 0$, *with* U *and* V *as given by* (3.2.2),

$$U + (2\theta - 1)V < 0, \tag{3.2.3}$$

then for any $s \in \left[0, \theta \wedge \frac{p}{2}\right]$, *we have* $\mathbb{E}[\tau_x^s] < \infty$ *for all* x *sufficiently large.*

Finally, we give conditions for non-existence of moments.

Theorem 3.2.4. *Suppose that* (A) *holds, and that* (B_p) *holds for some* $p > 2$. *Suppose also that* ($\mathrm{Q_G}$) *and* ($\mathrm{D_G}$) *hold. Define* $\mathbf{a} = (a_1, \ldots, a_{|S|})^\top$ *to be a solution to* (3.2.1) *whose existence is guaranteed by Lemma 3.2.1. If for some* $\theta \in (0, \frac{p}{2}]$, *with* U *and* V *as given by* (3.2.2),

$$U + (2\theta - 1)V > 0, \tag{3.2.4}$$

then for any $s \geq \theta$, we have $\mathbb{E}[\tau_x^s] = \infty$ for all sufficiently large $X_0 > x$.

Remarks 3.2.5. (a) The generalization of the state-space Σ from $\mathbb{Z}_+ \times S$ considered in [39] and previous work is not merely for the sake of generalization; it is necessary for the technical approach of the generalized Lamperti drift case, whereby we find a transformation $\phi : \Sigma \to \Sigma'$ such that if (X_n, η_n) has generalized Lamperti drift, then $\phi(X_n, \eta_n)$ has Lamperti drift (i.e., the constant components of the drifts are eliminated). We then apply the results of Section 3.1 to deduce the results in Section 3.2. Even if $\Sigma = \mathbb{Z}_+ \times S$, the state-space Σ' obtained after the transformation ϕ will not be (lines are translated in a certain way).

(b) The local finiteness assumption ensures that transience of the Markov chain (X_n, η_n) is equivalent to $\lim_{n\to\infty} X_n = +\infty$, a.s., and hence all our conditions on $\mu_i(x)$ etc. are asymptotic conditions as $x \to \infty$.

(c) As mentioned above, the non-uniqueness of solutions to (3.2.1) is not a problem for the statement of the theorems in this section, because the quantities in our conditions are unchanged under translation of the a_i. The variables a_i are well defined here in a non-trivial way. Indeed, Lemma 3.2.1 shows that if $(a_i, i \in S)$ is a solution then so is $(c+a_i, i \in S)$ for any $c \in \mathbb{R}$, and, furthermore, every solution is of this form. Moreover, the facts that $\sum_{j\in S} \gamma_{ij} = 0$ and $\sum_{i\in S} \sum_{j\in S} d_{ij}\pi_i = \sum_{i\in S} d_i\pi_i = 0$ guarantee that replacing every a_i by $c + a_i$ does not change the conditions in our theorems. Another way to go around this is to choose a particular line $0 \in S$ and set $a_0 = 0$, then a_i is now forced to be unique. There is no loss of generality if $a_0 \neq 0$, we can also obtain a new set of solutions by a translation $\tilde{a}_i = a_i - a_0$.

Chapter 4

Proofs and technical details

4.1 Semi-martingale criteria for recurrence classification

In this section we will present some of the fundamental results on the semi-martingale criteria for recurrence classification. These results on discrete-time martingales are due to Doob [26]. More of these results and their proofs can also be found in [27,93]. First we recall the definitions of martingales, submartingales and supermartingales.

Definition 4.1.1 (Martingales, submartingales, supermartingales)**.** A real-valued stochastic process X_n adapted to a filtration $\mathcal{F}_n$ is a *martingale* (with respect to the given filtration) if, for all $n \geq 0$,

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

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

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

For the term semimartingale, it does not just includes martingales, submartingales and supermartingales. We will use it in a broader context with some stochastic process which drift is of similar structure, on the whole space or just locally on some tail set.

We use the standard notation

$$x^+ := \max\{0, x\}. \tag{4.1.1}$$

Recall the follow fundamental result from martingale theory.

Theorem 4.1.2 (Martingale convergence theorem)**.** *Assume that X_n is a submartingale such that $\sup_n \mathbb{E}[X_n^+] < \infty$. Then there is an integrable random variable X such that $X_n \to X$ a.s. as $n \to \infty$.*

For the proof please see [27], Therem 5.2.8. Now we give an important corollary to Theorem 4.1.2 and Fatou's lemma.

Theorem 4.1.3 (Convergence of non-negative supermartingales)**.** *Assume that $X_n \geq 0$ is a supermartingale. Then there is an integrable random variable X such that $X_n \to X$ a.s. as $n \to \infty$, and $\mathbb{E}[X] \leq \mathbb{E}[X_0]$.*

For the proof please see [27], Therem 5.2.9. Based on the previous convergence, we give the following recurrence and transience criteria, which are central to our analysis of the half strip model. The statements here are taken from Section 2.5 of [75].

Theorem 4.1.4 (Recurrence criterion)**.** *An irreducible Markov chain X_n on a countably infinite state space Σ is recurrent if and only if there exist a function $f : \Sigma \to \mathbb{R}_+$ and a finite non-empty set $A \subset \Sigma$ such that*

$$\mathbb{E}\left[f(X_{n+1}) - f(X_n) \mid X_n = x\right] \leq 0, \text{ for all } x \in \Sigma \setminus A, \tag{4.1.2}$$

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

Theorem 4.1.5 (Transience criterion)**.** *An irreducible Markov chain X_n on a countably infinite state space Σ is transient if and only if there exist a function $f : \Sigma \to \mathbb{R}_+$ and a non-empty set $A \subset \Sigma$ such that*

$$\mathbb{E}\left[f(X_{n+1}) - f(X_n) \mid X_n = x\right] \leq 0, \text{ for all } x \in \Sigma \setminus A, \tag{4.1.3}$$

and

$$f(y) < \inf_{x \in A} f(x), \text{ for at least one site } y \in \Sigma \setminus A. \tag{4.1.4}$$

These two criterion can be trace back to the work of F.G. Foster [35]. He proved the 'if' part of Theorem 4.1.4 in the case where the exceptional set A is a singleton. For the finite set version for this direction can be found in Pakes [82]. The 'only if' part of Theorem 4.1.4 is due to Mertens *et al.* [76]. Foster [35] also proved Theorem 4.1.5 for the case where A is a single point. The finite set version is due to Harris and Marlin [49] and Mertens *et al.* [76].

4.2 Lyapunov function estimates for the half strip

Recall in Section 2.5 we proved the Pólya's Theorem with a Lyapunov function using the technique of reduction of dimensionality. We took $X_n := \|S_n\|$ as our function and one critical bit to apply the semi-martingale criteria is the calculation of expectations. Although it is pretty straightforward in the model of simple symmetric random walk, it can take a bit of effort in general models.

The main difficulty in applying the theorems in the previous section for the classification is to find a good Lyapunov function which gives suitable $\mathbb{E}\left[f(X_{n+1}) - f(X_n) \mid X_n = x\right]$. Depending on the model, these functions can be sometimes simple and easy to find, while sometimes it is very difficult to come up with the right function and calculate the expectation stated. In our half strip problem, we will give a Lyapunov function for each of the constant drift case and the Lamperti drift case. The formulation and the calculation of the former one is straightforward, while the latter one requires a lot more effort. They show both the strength and weakness of this Lyapunov function method. Although the method is very robust and constructive, it is tricky to start with the right function without any experience. Also, without explicit calculation of the expectation, it is hard to tell if the function that we picked is indeed the right one. The Lyapunov function for a specific model is usually not unique and it can be in various forms. To pick a good Lyapunov function that enables simplier calculation among all those which will satisfy the conditions in the theorems is a skill derived from experience.

4.2.1 Lyapunov function for constant drift

Our analysis for the constant drift case is based on two Lyapunov functions $g : \Sigma \to (0, \infty)$ and $h_\nu : \Sigma \to (0, \infty)$ for $\nu > 0$ for the recurrent case and transient case respectively, defined by

$$g(x, i) := x + b_i \tag{4.2.1}$$

for some $b_i \in \mathbb{R}$, and

$$h_\nu(x, i) := \begin{cases} x^{-\nu} - \nu b_i x^{-\nu-1} & \text{if } x \geq x_0, \\ x_0^{-\nu} - \nu b_i x_0^{-\nu-1} & \text{if } x < x_0, \end{cases} \tag{4.2.2}$$

where $b_i \in \mathbb{R}$ and $x_0 := 1 + 2\nu \max_{i \in S} |b_i|$.

We will need the following increment moment estimates for our Lyapunov function in the constant drift case. For the function g, we have the following lemma.

Lemma 4.2.1. *Suppose that* (A) *holds, and that* (B_p) *holds for some* $p > 1$. *Suppose also that* (Q_∞) *and* (D_C) *hold. Then we have, as* $x \to \infty$,

$$\mathbb{E}_{x,i}\left[g(X_{n+1}, \eta_{n+1}) - g(X_n, \eta_n)\right] = d_i + \sum_{j \in S} (b_j - b_i) q_{ij} + o(1). \tag{4.2.3}$$

Proof. Using the condition (D_C) that $\mathbb{E}_{x,i}\left[X_{n+1} - X_n\right] = d_i + o(1)$, we get

$$\begin{aligned} \mathbb{E}_{x,i}\left[g(X_{n+1}, \eta_{n+1}) - g(X_n, \eta_n)\right] &= \mathbb{E}_{x,i}\left[X_{n+1} - X_n\right] + \mathbb{E}_{x,i}\left[b_{\eta_{n+1}} - b_{\eta_n}\right] \\ &= \left[d_i + o(1)\right] + \sum_{j \in S} q_{ij}(b_j - b_i), \end{aligned}$$

by applying (Q_∞) in the last step. Hence we have the result as stated. □

On the other hand for the function h_ν, we have a slightly more complicated situation.

Lemma 4.2.2. *Suppose that* (A) *holds, and that* (B_p) *holds for some* $p > 1$. *Suppose*

also that (Q_∞) *and* $(\mathrm{D_C})$ *hold. Then for any* $\nu \in (0,p]$*, we have, as* $x \to \infty$*,*

$$\mathbb{E}_{x,i}\left[h_\nu(X_{n+1},\eta_{n+1}) - h_\nu(X_n,\eta_n)\right] = -\nu x^{-1-\nu}\left(d_i + \sum_{j\in S}(b_j - b_i)q_{ij} + o(1)\right). \tag{4.2.4}$$

Proof. Denote $\Delta_n := X_{n+1} - X_n$, and consider the event $E_n := \{|\Delta_n| \le X_n^\zeta\}$ where $\zeta \in (0,1)$. Then we choose $\zeta \in (0,1)$ and $x_1 \in \mathbb{R}_+$ such that $x_1 - x_1^\zeta \ge x_0$; then on the event $E_n \cap \{X_n \ge x_1\}$ we have $X_{n+1} \ge x_1 - x_1^\zeta \ge x_0$. Thus, for all $x \ge x_1$, we may write

$$\begin{aligned}
&\mathbb{E}_{x,i}\left[h_\nu(X_{n+1},\eta_{n+1}) - h_\nu(X_n,\eta_n)\right]\\
&= \mathbb{E}_{x,i}\left[\left(X_{n+1}^{-\nu} - X_n^{-\nu}\right)\mathbf{1}(E_n)\right] - \nu\,\mathbb{E}_{x,i}\left[\left(b_{\eta_{n+1}}X_{n+1}^{-\nu-1} - b_{\eta_n}X_n^{-\nu-1}\right)\mathbf{1}(E_n)\right]\\
&\quad + \mathbb{E}_{x,i}\left[\left(h_\nu(X_{n+1},\eta_{n+1}) - h_\nu(X_n,\eta_n)\right)\mathbf{1}(E_n^{\mathrm{c}})\right]
\end{aligned} \tag{4.2.5}$$

For the first term in equation (4.2.5), we apply Taylor's expansion and get

$$\begin{aligned}
\mathbb{E}_{x,i}\left[\left(X_{n+1}^{-\nu} - X_n^{-\nu}\right)\mathbf{1}(E_n)\right] &= x^{-\nu}\,\mathbb{E}_{x,i}\left[\left(\left(1+\frac{\Delta_n}{X_n}\right)^{-\nu} - 1\right)\mathbf{1}(E_n)\right]\\
&= x^{-\nu}\,\mathbb{E}_{x,i}\left[-\nu\left(\frac{\Delta_n}{X_n}\right)\mathbf{1}(E_n) + Z\right]
\end{aligned}$$

where $|Z| \le CX_n^{-2}|\Delta_n|^2\mathbf{1}(E_n)$, $C \in \mathbb{R}$ a constant. As $|Z| \le CX_n^{-2}|\Delta_n|\cdot|\Delta_n|\mathbf{1}(E_n) \le CX_n^{-2}|\Delta_n|X_n^\zeta$, we have

$$|\mathbb{E}_{x,i}[Z]| \le \mathbb{E}_{x,i}[|Z|] \le Cx^{\zeta-2}\,\mathbb{E}_{x,i}[|\Delta_n|] = O(x^{\zeta-2}),$$

using (B_p) in the last step. Thus we get $\mathbb{E}_{x,i}[Z] = o(x^{-1})$. So we get

$$\mathbb{E}_{x,i}\left[\left(X_{n+1}^{-\nu} - X_n^{-\nu}\right)\mathbf{1}(E_n)\right] = -\nu x^{-1-\nu}\left(\mathbb{E}_{x,i}[\Delta_n\mathbf{1}(E_n)] + o(1)\right).$$

Observing

$$|\Delta_n|\mathbf{1}(E_n^c) = |\Delta_n|^p|\Delta_n|^{1-p}\mathbf{1}(E_n^{\mathrm{c}}) \le |\Delta_n|^p X_n^{\zeta(1-p)},$$

we have

$$|\mathbb{E}_{x,i}[\Delta_n\mathbf{1}(E_n^{\mathrm{c}})]| \le \mathbb{E}_{x,i}[|\Delta_n|\mathbf{1}(E_n^{\mathrm{c}})] \le C_p x^{\zeta(1-p)},$$

by (B_p), where C_p is a constant depending on p. As $\zeta \in (0,1)$ and $p > 1$, we obtain

$$\mathbb{E}_{x,i}\left[\Delta_n \mathbf{1}(E_n)\right] = \mathbb{E}_{x,i}\left[\Delta_n\right] - \mathbb{E}_{x,i}\left[\Delta_n \mathbf{1}(E_n^{\mathrm{c}})\right] = \mathbb{E}_{x,i}\left[\Delta_n\right] + o(1)$$

Using (D_{C}), we get

$$\mathbb{E}_{x,i}\left[\left(X_{n+1}^{-\nu} - X_n^{-\nu}\right)\mathbf{1}(E_n)\right] = -\nu x^{-1-\nu}\left[d_i + o(1)\right]. \tag{4.2.6}$$

For the second term in equation (4.2.5), first observe that

$$\begin{aligned}&\mathbb{E}_{x,i}\left[\left(b_{\eta_{n+1}} X_{n+1}^{-\nu-1} - b_{\eta_n} X_n^{-\nu-1}\right)\mathbf{1}(E_n)\right]\\ &= \mathbb{E}_{x,i}\left[b_{\eta_{n+1}}\left(X_{n+1}^{-\nu-1} - X_n^{-\nu-1}\right)\mathbf{1}(E_n)\right] + \mathbb{E}_{x,i}\left[\left(b_{\eta_{n+1}} - b_{\eta_n}\right) X_n^{-\nu-1}\mathbf{1}(E_n)\right].\end{aligned} \tag{4.2.7}$$

We deal with the two terms on the right-hand side of (4.2.7) separately. First,

$$\left|\mathbb{E}_{x,i}\left[b_{\eta_{n+1}}\left(X_{n+1}^{-\nu-1} - X_n^{-\nu-1}\right)\mathbf{1}(E_n)\right]\right| \le \left(\max_{j\in S}|b_j|\right)\mathbb{E}_{x,i}\left[\left|X_{n+1}^{-\nu-1} - X_n^{-\nu-1}\right|\mathbf{1}(E_n)\right],$$

where, by Taylor's formula, given $X_n = x$,

$$\left|X_{n+1}^{-\nu-1} - X_n^{-\nu-1}\right|\mathbf{1}(E_n) = O(x^{-\nu-2+\zeta}) = o(x^{-\nu-1}).$$

On the other hand,

$$\mathbb{E}_{x,i}\left[\left(b_{\eta_{n+1}} - b_{\eta_n}\right) X_n^{-\nu-1}\mathbf{1}(E_n)\right] = x^{-\nu-1}\sum_{j\in S}(b_j - b_i)\mathbb{P}_{x,i}\left[\{\eta_{n+1} = j\}\cap E_n\right].$$

Here we have

$$\begin{aligned}\left|\mathbb{P}_{x,i}\left[\{\eta_{n+1} = j\}\cap E_n\right] - q_{ij}(x)\right| &\le \mathbb{P}_{x,i}\left[E_n^{\mathrm{c}}\right] = \mathbb{E}_{x,i}\left[|\Delta_n|^p|\Delta_n|^{-p}\mathbf{1}\{E_n^{\mathrm{c}}\}\right]\\ &\le x^{-p\zeta}\,\mathbb{E}_{x,i}\left[|\Delta_n|^p\right] \to 0.\end{aligned}$$

Hence we get

$$\mathbb{E}_{x,i}\left[\left(b_{\eta_{n+1}} X_{n+1}^{-\nu-1} - b_{\eta_n} X_n^{-\nu-1}\right)\mathbf{1}(E_n)\right] = x^{-\nu-1}\sum_{j\in S}(b_j - b_i)q_{ij}(x) + o(x^{-\nu-1}).$$

For the third term in equation (4.2.5), we observe that

$$0 \le h_\nu(x,i) \le C, \text{ for all } x \ge 0 \text{ and all } i \in S.$$

for some C, depending on ν and $(b_i, i \in S)$. As $\nu \in [0, p-1)$. For all x and i, we have

$$\begin{aligned}&|\mathbb{E}_{x,i}\left[(h_\nu(X_{n+1},\eta_{n+1}) - h_\nu(X_n,\eta_n))\,\mathbf{1}(E_n^{\mathrm{c}})\right]| \le C\mathbb{P}_{x,i}\left(E_n^{\mathrm{c}}\right)\\ &= C\,\mathbb{E}_{x,i}\left[|\Delta_n|^p|\Delta_n|^{-p}\mathbf{1}(E_n^{\mathrm{c}})\right] \le Cx^{-p\zeta}\,\mathbb{E}_{x,i}\left[|\Delta_n|^p\right] = O(x^{-p\zeta}),\end{aligned}$$

Since $-\nu > 1-p$, we can choose ζ such that $0 < \frac{1+\nu}{p} < \zeta < 1$, which gives $-\zeta p < -1-\nu$. Finally, grouping all three terms together gives the desired result.

□

4.2.2 Lyapunov function for Lamperti drift

Our analysis for the Lamperti drift case is a lot more complicated. It is based on the Lyapunov function $f_\nu : \Sigma \to (0,\infty)$ defined for $\nu \in \mathbb{R}$ by

$$f_\nu(x,i) := \begin{cases} x^\nu + \frac{\nu}{2}b_i x^{\nu-2} & \text{if } x \ge x_0,\\ x_0^\nu + \frac{\nu}{2}b_i x_0^{\nu-2} & \text{if } x < x_0,\end{cases} \tag{4.2.8}$$

where $b_i \in \mathbb{R}$ and $x_0 := 1 + \sqrt{|\nu| \max_{i\in S}|b_i|}$.

First we want to establishes some bounds on f_ν.

Lemma 4.2.3. *Suppose $\nu \in \mathbb{R}$. There exist positive constants $k_1, k_2 \in (0,\infty)$, depending on ν and $(b_i, i \in S)$, such that*

$$k_1(1+x)^\nu \le f_\nu(x,i) \le k_2(1+x)^\nu, \text{ for all } x \ge 0 \text{ and all } i \in S.$$

Proof. To start with, we consider the case when $x \ge x_0$, with $x_0 = 1+\sqrt{|\nu|B}$, where $B = \max_{i\in S}|b_i|$, we have

$$\left|\frac{\nu}{2}b_i x^{\nu-2}\right| \le \left|\frac{\frac{\nu}{2}b_i x^\nu}{x_0^2}\right| = \left|\frac{\frac{\nu}{2}b_i x^\nu}{\left(1+\sqrt{|\nu|B}\right)^2}\right| \le \frac{\frac{|\nu|}{2}|b_i|\,x^\nu}{|\nu|B} \le \frac{1}{2}x^\nu \tag{4.2.9}$$

So we have

$$\frac{1}{2}x^\nu \le f_\nu(x,i) \le \frac{3}{2}x^\nu \tag{4.2.10}$$

for all $x \geq x_0$. Noticing $x \geq x_0 > 1$, which implies $1 < \frac{x+1}{2} < x$, we have

$$\min(1, 2^{-\nu})(1+x)^\nu \leq x^\nu \leq \max(1, 2^{-\nu})(1+x)^\nu$$

Together with the inequality (4.2.10), we have

$$\frac{1}{2}\min(1, 2^{-\nu})(1+x)^\nu \leq f_\nu(x,i) \leq \frac{3}{2}\max(1, 2^{-\nu})(1+x)^\nu \tag{4.2.11}$$

for all $x \geq x_0$.

On the other hand, Suppose $x < x_0$. Then $f_\nu(x,i) = f_\nu(x_0,i)$, which is a case of (4.2.10) when $x = x_0$, so we have

$$\frac{1}{2}x_0^\nu \leq f_\nu(x,i) \leq \frac{3}{2}x_0^\nu \tag{4.2.12}$$

for all $x < x_0$. Now consider the fact that for $x < x_0$,

$$\min(1, (1+x_0)^{-\nu}) \leq (1+x)^{-\nu} \leq \max(1, (1+x_0)^{-\nu}).$$

Together with the inequality (4.2.12), we get for $x < x_0$,

$$\frac{1}{2}x_0^\nu \min(1, (1+x_0)^{-\nu})(1+x)^\nu \leq f_\nu(x,i) \leq \frac{3}{2}x_0^\nu \max(1, (1+x_0)^{-\nu})(1+x)^\nu. \tag{4.2.13}$$

Hence the proof is completed by taking appropriate positive constants k_1 and k_2 in different cases as just shown. □

The next result, which is central to what follows, provides increment moment estimates for our Lyapunov function in the Lamperti drift case.

Lemma 4.2.4. *Suppose that* (A) *holds, and that* (B_p) *holds for some* $p > 2$. *Suppose also that* (Q_∞) *and* ($\mathrm{D_L}$) *hold. Then for any* $\nu \in (2-p, p]$, *we have, as* $x \to \infty$,

$$\begin{aligned}&\mathbb{E}_{x,i}\left[f_\nu(X_{n+1}, \eta_{n+1}) - f_\nu(X_n, \eta_n)\right] \\ &\quad = \frac{\nu}{2}x^{\nu-2}\left(2c_i + (\nu-1)s_i^2 + \sum_{j\in S}(b_j - b_i)q_{ij} + o(1)\right).\end{aligned} \tag{4.2.14}$$

The rest of this section is devoted to the proof of Lemma 4.2.4. Denote $\Delta_n := X_{n+1} - X_n$, and again consider the event $E_n := \{|\Delta_n| \leq X_n^\zeta\}$ where $\zeta \in (0,1)$. The basic idea behind the proof of Lemma 4.2.4 is to use a Taylor's formula expansion.

Such an expansion is valid only if Δ_n is not too large; to handle various truncation estimates we will thus need the following result.

Lemma 4.2.5. *Suppose that* (B_p) *holds for some* $p > 2$. *Then for any* $\zeta \in (0,1)$ *and any* $q \in [0,p]$, *we have*

$$\mathbb{E}_{x,i}\left[|\Delta_n|^q \mathbf{1}(E_n^{\mathrm{c}})\right] \leq C_p x^{\zeta(q-p)}. \tag{4.2.15}$$

Furthermore, if $\zeta \in (\frac{1}{p-1}, 1)$, *we have*

$$\mathbb{E}_{x,i}\left[\Delta_n \mathbf{1}(E_n)\right] = \mathbb{E}_{x,i}\left[\Delta_n\right] + o\left(x^{-1}\right), \tag{4.2.16}$$

$$\mathbb{E}_{x,i}\left[\Delta_n^2 \mathbf{1}(E_n)\right] = \mathbb{E}_{x,i}\left[\Delta_n^2\right] + o\left(1\right). \tag{4.2.17}$$

Proof. For $q \in [0,p]$,

$$|\Delta_n|^q \mathbf{1}(E_n^{\mathrm{c}}) = |\Delta_n|^p |\Delta_n|^{q-p} \mathbf{1}(E_n^{\mathrm{c}}) \leq |\Delta_n|^p X_n^{\zeta(q-p)} \tag{4.2.18}$$

The inequality follows as $q - p \leq 0$, so under the condition that $|\Delta_n| > X_n^{\zeta}$, we have $|\Delta_n|^{q-p} \leq (X_n^{\zeta})^{(q-p)}$. Taking the conditional expectation on both sides of (4.2.18) and using the condition (B_p), we obtain (4.2.15). For the second statement, we use the fact that for $r \in \{1,2\}$,

$$\mathbb{E}_{x,i}\left[\Delta_n^r\right] = \mathbb{E}_{x,i}\left[\Delta_n^r \mathbf{1}(E_n)\right] + \mathbb{E}_{x,i}\left[\Delta_n^r \mathbf{1}(E_n^{\mathrm{c}})\right],$$

where, by the $q = r$ case of (4.2.15),

$$\left|\mathbb{E}_{x,i}\left[\Delta_n^r \mathbf{1}(E_n^{\mathrm{c}})\right]\right| \leq \mathbb{E}_{x,i}\left[|\Delta_n|^r \mathbf{1}(E_n^{\mathrm{c}})\right] \leq C_p x^{\zeta(r-p)}. \tag{4.2.19}$$

When $r = 1$, we choose $\zeta \in (\frac{1}{p-1}, 1)$, so we have $\zeta(1-p) < -1$, and then (4.2.19) gives (4.2.16). When $r = 2$, we know $r - p < 0$, and then (4.2.19) gives (4.2.17). $\square$

To obtain Lemma 4.2.4, we decompose the increment of f_ν. First note that, for $\zeta \in (0,1)$,

$$\begin{aligned}\mathbb{E}_{x,i}\left[f_\nu(X_{n+1},\eta_{n+1}) - f_\nu(X_n,\eta_n)\right] &= \mathbb{E}_{x,i}\left[(f_\nu(X_{n+1},\eta_{n+1}) - f_\nu(X_n,\eta_n))\mathbf{1}(E_n)\right] \\ &\quad + \mathbb{E}_{x,i}\left[(f_\nu(X_{n+1},\eta_{n+1}) - f_\nu(X_n,\eta_n))\mathbf{1}(E_n^{\mathrm{c}})\right]. \end{aligned} \tag{4.2.20}$$

Choose $\zeta \in (\frac{1}{p-1}, 1)$ and $x_1 \in \mathbb{R}_+$ such that $x_1 - x_1^\zeta \geq x_0$; then on the event $E_n \cap \{X_n \geq x_1\}$ we have $X_{n+1} \geq x_1 - x_1^\zeta \geq x_0$. Thus, for all $x \geq x_1$, we may write

$$\begin{aligned}
&\mathbb{E}_{x,i}\left[(f_\nu(X_{n+1}, \eta_{n+1}) - f_\nu(X_n, \eta_n))\,\mathbf{1}(E_n)\right] \\
&= \mathbb{E}_{x,i}\left[\left(X_{n+1}^\nu - X_n^\nu\right)\mathbf{1}(E_n)\right] + \frac{\nu}{2}\,\mathbb{E}_{x,i}\left[\left(b_{\eta_{n+1}} X_{n+1}^{\nu-2} - b_{\eta_n} X_n^{\nu-2}\right)\mathbf{1}(E_n)\right]. \qquad (4.2.21)
\end{aligned}$$

We proceed to estimate the terms on the right-hand sides of (4.2.20) and (4.2.21) separately, via a series of lemmas.

Lemma 4.2.6. *Suppose that* (B_p) *holds for some* $p > 2$. *Suppose also that* $(\mathrm{D_L})$ *holds. Let* $\zeta \in (\frac{1}{p-1}, 1)$. *Then for any* $r \in \mathbb{R}$, *we have, as* $x \to \infty$,

$$\mathbb{E}_{x,i}\left[\left(X_{n+1}^r - X_n^r\right)\mathbf{1}(E_n)\right] = r x^{r-2}\left(c_i + \frac{r-1}{2} s_i^2 + o(1)\right).$$

Proof. By Taylor's formula we have that

$$\begin{aligned}
\mathbb{E}_{x,i}\left[\left(X_{n+1}^r - X_n^r\right)\mathbf{1}(E_n)\right] &= x^r\,\mathbb{E}_{x,i}\left[\left((1 + x^{-1}\Delta_n)^r - 1\right)\mathbf{1}(E_n)\right] \\
&= x^r\,\mathbb{E}_{x,i}\left[\left(r\left(\frac{\Delta_n}{X_n}\right) + \frac{r(r-1)}{2}\left(\frac{\Delta_n}{X_n}\right)^2\right)\mathbf{1}(E_n) + Z\right], \qquad (4.2.22)
\end{aligned}$$

where $|Z| \leq C X_n^{-3} |\Delta_n|^3 \mathbf{1}(E_n)$ for some fixed constant $C \in \mathbb{R}_+$. To bound the term $\mathbb{E}_{x,i}[Z]$, first we observe that

$$|Z| \leq C X_n^{-3} |\Delta_n|^2 |\Delta_n| \mathbf{1}(E_n) \leq C |\Delta_n|^2 X_n^{\zeta-3}.$$

Taking expectations on both sides of the last inequality, we obtain

$$|\mathbb{E}_{x,i}[Z]| \leq \mathbb{E}_{x,i}\,|Z| \leq C x^{\zeta-3}\,\mathbb{E}_{x,i}[|\Delta_n|^2] = O(x^{\zeta-3}),$$

using (B_p). Since $\zeta < 1$ this implies $\mathbb{E}_{x,i}[Z] = o(x^{-2})$, so the expression (4.2.22) becomes

$$\begin{aligned}
&\mathbb{E}_{x,i}\left[\left(X_{n+1}^r - X_n^r\right)\mathbf{1}(E_n)\right] \\
&= r x^{r-1}\,\mathbb{E}_{x,i}\left[\Delta_n \mathbf{1}(E_n)\right] + \frac{r(r-1)}{2} x^{r-2}\,\mathbb{E}_{x,i}\left[\Delta_n^2 \mathbf{1}(E_n)\right] + o\left(x^{r-2}\right). \qquad (4.2.23)
\end{aligned}$$

Then by Lemma 4.2.5 together with the facts that, under $(\mathrm{D_L})$,

$$\mathbb{E}_{x,i}[\Delta_n] = \mu_i(x) = \frac{c_i}{x} + o(x^{-1}), \text{ and } \mathbb{E}_{x,i}[\Delta_n^2] = \sigma_i^2(x) = s_i^2 + o(1),$$

we obtain

$$\begin{aligned}&\mathbb{E}_{x,i}\left[\left(X_{n+1}^r - X_n^r\right)\mathbf{1}(E_n)\right]\\ &= rx^{r-1}\left(\frac{c_i}{x} + o\left(x^{-1}\right)\right) + \frac{r(r-1)}{2}x^{r-2}\left(s_i^2 + o(1)\right) + o\left(x^{r-2}\right),\end{aligned}$$

from which the statement in the lemma follows. □

Lemma 4.2.7. *Suppose that* (B_p) *holds for some* $p > 0$. *Let* $r \in \mathbb{R}$ *and* $\zeta \in (0,1)$, *and let* $g : S \to \mathbb{R}$. *Then, as* $x \to \infty$,

$$\mathbb{E}_{x,i}\left[\left(g(\eta_{n+1})X_{n+1}^r - g(\eta_n)X_n^r\right)\mathbf{1}(E_n)\right] = x^r \sum_{j\in S}(g(j) - g(i))q_{ij}(x) + o(x^r).$$

Proof. First observe that

$$\begin{aligned}&\mathbb{E}_{x,i}\left[\left(g(\eta_{n+1})X_{n+1}^r - g(\eta_n)X_n^r\right)\mathbf{1}(E_n)\right]\\ &= \mathbb{E}_{x,i}\left[g(\eta_{n+1})\left(X_{n+1}^r - X_n^r\right)\mathbf{1}(E_n)\right] + \mathbb{E}_{x,i}\left[\left(g(\eta_{n+1}) - g(\eta_n)\right)X_n^r\mathbf{1}(E_n)\right]. \quad (4.2.24)\end{aligned}$$

We deal with the two terms on the right-hand side of (4.2.24) separately. First,

$$\left|\mathbb{E}_{x,i}\left[g(\eta_{n+1})\left(X_{n+1}^r - X_n^r\right)\mathbf{1}(E_n)\right]\right| \le \left(\max_{j\in S}|g(j)|\right)\mathbb{E}_{x,i}\left[\left|X_{n+1}^r - X_n^r\right|\mathbf{1}(E_n)\right],$$

where, by Taylor's formula, given $X_n = x$,

$$\left|X_{n+1}^r - X_n^r\right|\mathbf{1}(E_n) = O(x^{r+\zeta-1}) = o(x^r).$$

On the other hand,

$$\mathbb{E}_{x,i}\left[\left(g(\eta_{n+1}) - g(\eta_n)\right)X_n^r\mathbf{1}(E_n)\right] = x^r\sum_{j\in S}(g(j) - g(i))\mathbb{P}_{x,i}\left[\{\eta_{n+1} = j\}\cap E_n\right].$$

Here $\left|\mathbb{P}_{x,i}\left[\{\eta_{n+1} = j\}\cap E_n\right] - q_{ij}(x)\right| \le \mathbb{P}_{x,i}\left[E_n^{\mathrm{c}}\right] \to 0$, by the $q = 0$ case of Lemma 4.2.5. Combining these calculations gives the result. □

Combining the last two results we obtain the following estimate for the first term

on the right-hand side of (4.2.20).

Lemma 4.2.8. *Suppose that* (B_p) *holds for some* $p > 2$. *Suppose also that* $(\mathrm{D_L})$ *and* (Q_∞) *hold. Let* $\zeta \in (\frac{1}{p-1}, 1)$. *Then for any* $r \in \mathbb{R}$, *we have, as* $x \to \infty$,

$$\begin{aligned}&\mathbb{E}_{x,i}\left[(f_r(X_{n+1},\eta_{n+1}) - f_r(X_n,\eta_n))\,\mathbf{1}(E_n)\right]\\&\quad= \frac{r}{2}x^{r-2}\left(2c_i + (r-1)s_i^2 + \sum_{j\in S}(b_j - b_i)q_{ij} + o(1)\right).\end{aligned}$$

Proof. In the equation (4.2.21) we apply Lemma 4.2.6 and Lemma 4.2.7 with $g(y) = b_y$ and $r-2$ in place of r; together with (Q_∞) we obtain the result. □

We have the following estimate for the second term on the right-hand side of (4.2.20).

Lemma 4.2.9. *Suppose that* (B_p) *holds for some* $p > 2$. *Then for any* $r \in (2-p, p]$, *we can choose* $\zeta \in (0,1)$ *for which, as* $x \to \infty$,

$$\mathbb{E}_{x,i}\left[|f_r(X_{n+1},\eta_{n+1}) - f_r(X_n,\eta_n)|\,\mathbf{1}(E_n^{\mathrm{c}})\right] = o(x^{r-2}).$$

Proof. We may suppose throughout this proof that $X_n \geq 1$. First suppose that $r \in (0,p]$. If $|\Delta_n| \leq \frac{X_n}{2}$, then $\frac{X_n}{2} \leq X_n + \Delta_n \leq \frac{3X_n}{2}$. Thus with Lemma 4.2.3 we have, on $\{|\Delta_n| \leq \frac{X_n}{2}\}$,

$$f_r(X_{n+1},\eta_{n+1}) \leq k_2\left(1 + \frac{3X_n}{2}\right)^r \leq C_1\left(1 + X_n\right)^r, \tag{4.2.25}$$

for some constant $C_1 \in \mathbb{R}_+$. On the other hand, if $|\Delta_n| > \frac{X_n}{2}$, then $0 \leq X_{n+1} = X_n + \Delta_n \leq 3|\Delta_n|$. So with Lemma 4.2.3, we have, on $\{|\Delta_n| > \frac{X_n}{2}\}$,

$$f_r(X_{n+1},\eta_{n+1}) \leq k_2\left(1 + 3|\Delta_n|\right)^r \leq C_2\left|\Delta_n\right|^r, \tag{4.2.26}$$

for some constant $C_2 \in \mathbb{R}_+$. Combining the results of (4.2.25) and (4.2.26), we obtain for $r > 0$,

$$f_r(X_{n+1},\eta_{n+1}) \leq C_3\left(1 + X_n\right)^r + C_3|\Delta_n|^r, \tag{4.2.27}$$

for some $C_3 \in \mathbb{R}_+$. Hence, for $r > 0$, for some $C \in \mathbb{R}_+$,

$$\begin{aligned}
&|\mathbb{E}_{x,i}\left[(f_r(X_{n+1},\eta_{n+1}) - f_r(X_n,\eta_n))\mathbf{1}(E_n^{\mathrm{c}})\right]| \\
&\leq f_r(x,i)\mathbb{P}_{x,i}\left[E_n^{\mathrm{c}}\right] + \mathbb{E}_{x,i}\left[|f_r(X_{n+1},\eta_{n+1})|\mathbf{1}(E_n^{\mathrm{c}})\right] \\
&\leq C\,(1+x)^r\,\mathbb{P}_{x,i}\left[E_n^{\mathrm{c}}\right] + C\,\mathbb{E}_{x,i}\left[|\Delta_n|^r\mathbf{1}(E_n^{\mathrm{c}})\right],
\end{aligned}$$

where we have used Lemma 4.2.3 and inequality (4.2.27). Then, by the $q = 0$ and $q = r \in (0,p]$ cases of Lemma 4.2.5 we have

$$|\mathbb{E}_{x,i}\left[(f_r(X_{n+1},\eta_{n+1}) - f_r(X_n,\eta_n))\mathbf{1}(E_n^{\mathrm{c}})\right]| = O(x^{r-p\zeta}) + O(x^{\zeta(r-p)}) = O(x^{r-p\zeta}),$$

since $\zeta < 1$. This last term is $o(x^{r-2})$ provided $r - p\zeta < r - 2$, i.e., $\zeta > \frac{2}{p}$.

Finally, suppose that $r \in (2-p, 0]$. Now by Lemma 4.2.3, we have $0 \leq f(x,i) \leq C$ for some $C \in \mathbb{R}_+$ and all x and i, so that

$$|\mathbb{E}_{x,i}\left[(f_r(X_{n+1},\eta_{n+1}) - f_r(X_n,\eta_n))\,\mathbf{1}(E_n^{\mathrm{c}})\right]| \leq C\mathbb{P}_{x,i}\left[E_n^{\mathrm{c}}\right] = O(x^{-p\zeta}),$$

by the $q = 0$ case of Lemma 4.2.5. Since $r > 2 - p$, we can choose ζ such that $0 < \frac{2-r}{p} < \zeta < 1$, which gives $-\zeta p < r - 2$. □

Now we are ready to complete the proof Lemma 4.2.4.

Proof of Lemma 4.2.4. The expression for the first moment in (4.2.14) is simply a combination of the $r = \nu$ cases of Lemmas 4.2.8 and 4.2.9. □

4.3 Some consequences of the Fredholm alternative

This section serves two purposes. The first aim for this section is to prepare for the proofs in the next section, which we need to understand the term with b_i's in the Lyapunov function estimate for the case of Lamperti drift. The other purpose is to show existence of a_i in Lemma 3.2.1 for our translation in the generalized Lamperti drift case.

4.3.1 Fredholm alternative

The following well-known algebraic result will enable us to show that suitable b_i exist to construct the Lyapunov function f_ν as defined at (4.2.8) under appropriate conditions involving π_j.

In this section, vectors are column vectors on $\mathbb{R}^{|S|}$, $\mathbf{0}$ denotes the column vector whose components are all zero, and I denotes the $|S| \times |S|$ identity matrix. We will need the following well-known algebraic result.

Lemma 4.3.1 (Fredholm alternative)**.** *Given an $|S| \times |S|$ matrix A and a column vector $\mathbf{b}$, the equation $A\mathbf{a} = \mathbf{b}$ has a solution $\mathbf{a}$ if and only if any column vector $\mathbf{y}$ for which $A^\top \mathbf{y} = \mathbf{0}$ satisfies $\mathbf{y}^\top \mathbf{b} = 0$.*

See [86] for other formulations and the proof of this theorem. First of all, we shall give the proof of Lemma 3.2.1.

Proof of Lemma 3.2.1. First we write the system of equations (3.2.1) in matrix form. To this end, denote by $Q = (q_{ij})_{i,j \in S}$ the transition matrix for the Markov chain η_n^* on S, and denote column vectors $\mathbf{a} = (a_1, a_2, \ldots, a_{|S|})^\top$ and $\mathbf{d} = (d_1, d_2, \ldots, d_{|S|})^\top$. Then (3.2.1) is equivalent to

$$(Q - I)\mathbf{a} = -\mathbf{d}.$$

Setting $A = Q - I$ and $\mathbf{b} = -\mathbf{d}$, Lemma 4.3.1 shows that (3.2.1) has a solution $\mathbf{a}$ if and only if any column vector $\mathbf{y}$ such that $(Q - I)^\top \mathbf{y} = \mathbf{0}$ satisfies $\mathbf{y}^\top \mathbf{d} = 0$. But $(Q - I)^\top \mathbf{y} = \mathbf{0}$ is equivalent to $\mathbf{y}^\top Q = \mathbf{y}^\top$, which implies that $\mathbf{y} = \alpha \boldsymbol{\pi}$ ($\alpha \in \mathbb{R}$) is a scalar multiple of the (unique) stationary distribution for Q. Thus (3.2.1) has a solution $\mathbf{a}$ if and only if $\boldsymbol{\pi}^\top \mathbf{d} = 0$, i.e., $\sum_{i \in S} d_i \pi_i = 0$, the special case that $\alpha = 0$ is contributing nothing to the condition.

Finally, we show that any solution $\mathbf{a}$ to (3.2.1) is unique up to translation. Suppose there are two solutions, $\mathbf{a}'$ and $\mathbf{a}''$, so that $(Q - I)\mathbf{a}' = (Q - I)\mathbf{a}'' = -\mathbf{d}$; thus $(Q - I)(\mathbf{a}' - \mathbf{a}'') = \mathbf{0}$. In other words, $Q(\mathbf{a}' - \mathbf{a}'') = \mathbf{a}' - \mathbf{a}''$. As Q is a stochastic matrix, this means that $\mathbf{a}' - \mathbf{a}''$ is a scalar multiple of the eigenvector $(1, 1, \ldots, 1)^\top$ corresponding to eigenvalue 1. Thus the components of $\mathbf{a}'$ and $\mathbf{a}''$ differ by a fixed amount. $\square$

4.3.2 Corollaries

A modification of the above argument yields the following statements, with inequalities instead of equality, which will enable us to show that, under appropriate conditions involving π_j, suitable b_i exist to construct the Lyapunov function f_ν satisfying appropriate supermartingale conditions.

Lemma 4.3.2. *Let $u_i \in \mathbb{R}$ for each $i \in S$.*

(i) Suppose $\sum_{i\in S} u_i\pi_i < 0$. Then there exist $(b_i, i \in S)$ such that

$$u_i + \sum_{j\in S}(b_j - b_i)q_{ij} < 0, \text{ for all } i \in S.$$

(ii) Suppose $\sum_{i\in S} u_i\pi_i > 0$. Then there exist $(b_i, i \in S)$ such that

$$u_i + \sum_{j\in S}(b_j - b_i)q_{ij} > 0, \text{ for all } i \in S.$$

Proof. We prove only part (i); the proof of (ii) is similar. Suppose that $\sum_{i\in S} u_i\pi_i = -\varepsilon$ for some $\varepsilon > 0$. Then taking $\varepsilon_i = \frac{\varepsilon}{|S|\pi_i}$ we get $\sum_{i\in S}(u_i+\varepsilon_i)\pi_i = 0$. An application of Lemma 3.2.1 with $d_i = u_i + \varepsilon_i$ shows that there exist b_i such that

$$u_i + \varepsilon_i + \sum_{j\in S}(b_j - b_i)q_{ij} = 0, \text{ for all } i \in S,$$

which gives the result since $\varepsilon_i > 0$. □

4.4 Proof of the constant drift classification

In this section, we will give a new proof of Theorem 2.4.1 using the method of Lyapunov functions. We will apply Theorem 4.1.4 and Theorem 4.1.5 with the Lyapunov functions stated in (4.2.1) and (4.2.2).

Proof of Theorem 2.4.1. For the recurrence part, we will use the Lyapunov function $g(x, i)$ defined at (4.2.1), with suitably chosen b_i. First we see that $g(x, i) \to \infty$ as $x \to \infty$. Thus Theorem 4.1.4 shows that the process is recurrent if

$$\mathbb{E}_{x,i}\left[g(X_{n+1}, \eta_{n+1}) - g(X_n, \eta_n)\right] \le 0 \tag{4.4.1}$$

for all sufficiently large x. Now suppose $\sum_{i\in S} d_i\pi_i < 0$, then we use Lemma 4.3.2 (i) from our Fredholm alternative corollaries, with $u_i = d_i$ to show that we may choose b_i so that

$$d_i + \sum_{j\in S}(b_j - b_i)q_{ij} < 0.$$

Hence from Lemma 4.2.1 we know the condition (4.4.1) is satisfied for x sufficiently large.

For the transience part, this time we will use the Lyapunov function $h_\nu(x,i)$ defined at (4.2.2) for a small positive value of ν close to 0, and apply Theorem 4.1.5. We see that the condition in equation (4.1.4) is satisfied as $h_\nu(x,i)$ is a decreasing function. Hence the process is transient if

$$\mathbb{E}_{x,i}\left[h_\nu(X_{n+1},\eta_{n+1}) - h_\nu(X_n,\eta_n)\right] \leq 0 \tag{4.4.2}$$

for all sufficiently large x. Now suppose $\sum_{i\in S} d_i\pi_i > 0$, using Lemma 4.3.2 (ii) from our Fredholm alternative corollaries, with $u_i = d_i$ we may choose b_i so that

$$d_i + \sum_{j\in S}(b_j - b_i)q_{ij} > 0.$$

Finally, from Lemma 4.2.2 we know the condition (4.4.2) is satisfied for x sufficiently large as we wanted. This completes the proof of the theorem. □

4.5 Proofs of results for Lamperti drift

The first goal of the section is to give a new proof of the first three points in Theorem 3.1.1 using the method of Lyapunov functions. To prove the whole classification, we should separate the argument in a few parts.

First, in this subsection, we will prove the conditions for recurrence and transience, by applying Theorem 4.1.4 and Theorem 4.1.5.

In the second and the third subsection, we turn our attention to our second and third objectives, the proof of existence and non-existence of moments.

Lastly in the fourth subsection, we will show the conditions for positive recurrence and null, which are in fact special cases for the existence and non-existence of

moments. Combining with the dichotomy in the first subsection will give us the full classification as stated in Theorem 3.1.1, with the exception of the boundary cases.

For the critical case for null recurrence in Theorem 3.1.1, we would need a more delicate treatment with a Lyapunov function which grows slower, like $(\log x)^\theta$, $\theta \in (0,1)$. Some general calculation can be found in the book by Menshikov et. al. [75].

4.5.1 Proof of recurrence and transience in the Lamperti drift case

Here is our formal statement to be proved in this subsection.

Theorem 4.5.1. *Suppose that* (A) *holds, and that* (B_p) *holds for some* $p > 2$. *Suppose also that* (Q_∞) *and* ($\mathrm{D_L}$) *hold. Then the following classification applies.*

- *If* $\sum_{i\in S}(2c_i - s_i^2)\pi_i < 0$, *then* (X_n, η_n) *is recurrent.*
- *If* $\sum_{i\in S}(2c_i - s_i^2)\pi_i > 0$, *then* (X_n, η_n) *is transient.*

Proof. Using the Lyapunov function $f_\nu(x,i)$ stated in (4.2.8), we would like to apply Theorem 4.1.4 to get a condition for recurrence.

Suppose that $\nu > 0$, then by Lemma 4.2.3, $f_\nu(x,i) \to \infty$ as $x \to \infty$. So we know the process is recurrent if

$$\mathbb{E}_{x,i}\left[f_\nu(X_{n+1},\eta_{n+1}) - f_\nu(X_n,\eta_n)\right] \le 0 \tag{4.5.1}$$

for all x sufficiently large. Now suppose that $\sum_{i\in S}(2c_i - s_i^2)\pi_i < 0$, then there exists $\nu > 0$ such that

$$\sum_{i\in S}\left[2c_i + (\nu-1)s_i^2\right]\pi_i < 0.$$

Now we use Lemma 4.3.2 (i) from our Fredholm alternative corollaries, with $u_i = 2c_i + (\nu-1)s_i^2$ to show that we may choose b_i such that

$$2c_i + (\nu-1)s_i^2 + \sum_{j\in S}(b_j - b_i)q_{ij} < 0.$$

Hence we get

$$\frac{\nu}{2}x^{\nu-2}\left(2c_i + (\nu-1)s_i^2 + \sum_{j\in S}(b_j - b_i)q_{ij} + o(1)\right) \leq 0$$

for all x sufficiently large. Finally, apply Lemma 4.2.4 to get our recurrence condition in equation (4.5.1) as desired.

For the transient side, this time we take $\nu < 0$ and apply Theorem 4.1.5. With Lemma 4.2.3, we have $f_\nu(x,i) \to 0$ as $x \to \infty$, hence the condition in equation (4.1.4) is immediately satisfied. So the process is transient if (4.5.1) holds for all x sufficiently large. This time we suppose that $\sum_{i\in S}(2c_i - s_i^2)\pi_i > 0$, then there exists $\nu < 0$ such that

$$\sum_{i\in S}\left[2c_i + (\nu-1)s_i^2\right]\pi_i > 0.$$

Now we use Lemma 4.3.2 (ii) from our Fredholm alternative corollaries, with $u_i = 2c_i + (\nu-1)s_i^2$ to show that we can choose b_i such that

$$2c_i + (\nu-1)s_i^2 + \sum_{j\in S}(b_j - b_i)q_{ij} > 0.$$

Hence we get

$$\frac{\nu}{2}x^{\nu-2}\left(2c_i + (\nu-1)s_i^2 + \sum_{j\in S}(b_j - b_i)q_{ij} + o(1)\right) \leq 0.$$

for all x sufficiently large. Finally, apply Lemma 4.2.4 to get our transience condition in equation (4.5.1) as desired. Hence the proof is completed. □

4.5.2 Proof of existence of moments

To obtain existence of moments of hitting times, we apply the following semimartingale result, which is a reformulation of Theorem 1 from [5], see also [75] Corollary 2.7.3.

Lemma 4.5.2. *Let W_n be an integrable $\mathcal{F}_n$-adapted stochastic process, taking values in an unbounded subset of $\mathbb{R}_+$, with $W_0 = w_0$ fixed. Suppose that there exist $\delta > 0$,*

$w > 0$, and $\gamma < 1$ such that for any $n \geq 0$,

$$\mathbb{E}[W_{n+1} - W_n \mid \mathcal{F}_n] \leq -\delta W_n^\gamma, \text{ on } \{n < \lambda_w\}, \tag{4.5.2}$$

where $\lambda_w = \min\{n \geq 0 : W_n \leq w\}$. Then, for any $s \in [0, \frac{1}{1-\gamma}]$, $\mathbb{E}[\lambda_w^s] < \infty$.

Now we can give the proof of Theorem 3.1.2.

Proof of Theorem 3.1.2. Set $W_n := f_\nu(X_n, \eta_n)$ for $\nu \in (0, p]$; note $W_n \to \infty$ as $X_n \to \infty$. We aim to show that (4.5.2) holds with $\gamma = \frac{\nu-2}{\nu} < 1$. First note that, for X_n sufficiently large,

$$\begin{aligned} W_n^\gamma &= \left(X_n^\nu + \frac{\nu}{2} a_{\eta_n} X_n^{\nu-2}\right)^{\frac{\nu-2}{\nu}} \\ &= X_n^{\nu-2}\left(1 + \frac{\nu}{2} a_{\eta_n} X_n^{-2}\right)^{\frac{\nu-2}{\nu}} \\ &= X_n^{\nu-2} + O\left(X_n^{\nu-4}\right), \end{aligned}$$

using the fact that a_{η_n} is uniformly bounded. In other words, $X_n^{\nu-2} = W_n^\gamma + o(W_n^\gamma)$, so we have from Lemma 4.2.4 that

$$\mathbb{E}[W_{n+1} - W_n \mid X_n, \eta_n] = \frac{\nu}{2} W_n^\gamma \left(2c_i + (\nu - 1)s_i^2 + \sum_{j \in S}(a_j - a_i)q_{ij}\right) + o(W_n^\gamma). \tag{4.5.3}$$

Take $\nu = p \wedge 2\theta$ and set $u_i = 2c_i + (\nu - 1)s_i^2$; then, by (3.1.2),

$$\sum_{i \in S} u_i \pi_i \leq \sum_{i \in S} \left[2c_i + (2\theta - 1)s_i^2\right] \pi_i < 0,$$

so that by Lemma 4.3.2(i) we have that the coefficient of W_n^γ on the right-hand side of (4.5.3) is strictly negative. Hence there exists $\delta > 0$ such that

$$\mathbb{E}[W_{n+1} - W_n \mid X_n, \eta_n] \leq -\delta W_n^\gamma, \text{ on } \{W_n \geq w\},$$

for some w big enough. Note that $\frac{1}{1-\gamma} = \frac{\nu}{2} = \theta \wedge \frac{p}{2}$; thus we may apply Lemma 4.5.2 to conclude that $\mathbb{E}[\lambda_w^s] < \infty$ for all w sufficiently large, for any $s \in \left[0, \theta \wedge \frac{p}{2}\right]$.

It remains to deduce that $\mathbb{E}[\tau_x^s] < \infty$ for all x sufficiently large. But Lemma 4.2.5 shows that $X_n \leq C W_n^{1/\nu}$ for some $C \in \mathbb{R}_+$, so $\{W_n \leq w\}$ implies that

$\{X_n \leq Cw^{1/\nu}\}$. It follows that $\tau_{Cw^{1/\nu}} \leq \lambda_w$, completing the proof of the theorem. □

4.5.3 Proof of non-existence of moments

To obtain non-existence of moments of hitting times, we apply the following semi-martingale result, which is a variation on Theorem 2 from [5], see also [75] Theorem 2.7.4.

Lemma 4.5.3. *Let Z_n be a $\mathcal{F}_n$-adapted stochastic process taking values in an unbounded subset of $\mathbb{R}_+$. Suppose that there exist finite positive constants z, B, and c such that, for any $n \geq 0$,*

$$\mathbb{E}[Z_{n+1} - Z_n \mid \mathcal{F}_n] \geq -\frac{c}{Z_n}, \text{ on } \{Z_n \geq z\}; \tag{4.5.4}$$

$$\mathbb{E}[(Z_{n+1} - Z_n)^2 \mid \mathcal{F}_n] \leq B, \text{ on } \{Z_n \geq z\}. \tag{4.5.5}$$

Suppose in addition that for some $p_0 > 0$, the process $Z_{n\wedge\lambda_z}^{2p_0}$ is a submartingale, where $\lambda_z = \min\{n \geq 0 : Z_n \leq z\}$. Then for any $p > p_0$, we have $\mathbb{E}[\lambda_z^p \mid Z_0 = z_0] = \infty$ for any $z_0 > z$.

We will apply this result with $Z_n := W_n^{1/\nu} = (f_\nu(X_n, \eta_n))^{1/\nu}$. Thus we must establish some estimates on the first and second moments of the increments of Z_n; this is the purpose of the next result.

Lemma 4.5.4. *Suppose that* (A) *holds, and that* (B_p) *holds for some $p > 2$. Suppose also that* (Q_∞) *and* ($\mathrm{D_L}$) *hold. Then for any $\nu \in (0, p]$, we have*

$$\mathbb{E}_{x,i}[Z_{n+1} - Z_n] = \frac{c_i}{x} + \frac{1}{2x}\sum_{j\in S}(b_j - b_i)q_{ij} + o\left(x^{-1}\right); \text{ and}$$

$$\mathbb{E}_{x,i}[(Z_{n+1} - Z_n)^2] \leq B,$$

where B is a constant.

Proof. Again we define the event $E_n := \{|\Delta_n| \leq X_n^\zeta\}$ for $\zeta \in (0, 1)$; then

$$\mathbb{E}_{x,i}[Z_{n+1} - Z_n] = \mathbb{E}_{x,i}\left[(Z_{n+1} - Z_n)\,\mathbf{1}(E_n)\right] + \mathbb{E}_{x,i}\left[(Z_{n+1} - Z_n)\,\mathbf{1}(E_n^{\mathrm{c}})\right]. \tag{4.5.6}$$

For the first term on the right-hand side of (4.5.6), we first notice that for x large enough

$$f_\nu^{1/\nu}(x,i) = x\left(1+\frac{\nu}{2}b_i x^{-2}\right)^{1/\nu} = x + \frac{b_i}{2x} + O(x^{-3}).$$

Also, given $(X_n,\eta_n)=(x,i)$, on E_n we have $|X_{n+1}-X_n|\le x^\zeta$ so that $Z_{n+1}\mathbf{1}(E_n) = X_{n+1}+\frac{b_{\eta_{n+1}}}{2x}+o(x^{-1})$. As a result we get

$$\begin{aligned}
&\mathbb{E}_{x,i}\left[(Z_{n+1}-Z_n)\,\mathbf{1}(E_n)\right]\\
&= \mathbb{E}_{x,i}\left[(X_{n+1}-X_n)\mathbf{1}(E_n)\right] + \frac{1}{2x}\,\mathbb{E}_{x,i}\left[(b_{\eta_{n+1}}-b_{\eta_n})\mathbf{1}(E_n)\right] + o(x^{-1})\\
&= \frac{c_i}{x} + \frac{1}{2x}\sum_{j\in S}(b_j-b_i)q_{ij} + o\left(x^{-1}\right) \qquad (4.5.7)
\end{aligned}$$

where in the last equality we used Lemma 4.2.6 and the $r=0$ case of Lemma 4.2.7 for the first and second term respectively. On the other hand, to bound $\mathbb{E}_{x,i}\left[(Z_{n+1}-Z_n)\,\mathbf{1}(E_n^{\mathrm{c}})\right]$, we can just mimic the proof of Lemma 4.2.9, inserting additional powers of $1/\nu$, to obtain

$$\mathbb{E}_{x,i}\left[(Z_{n+1}-Z_n)\,\mathbf{1}(E_n^{\mathrm{c}})\right] = o(x^{-1}),$$

which combined with (4.5.7) gives the first statement in the lemma. For the second moment, given $(X_n,\eta_n)=(x,i)$ we have

$$\begin{aligned}
(Z_{n+1}-Z_n)^2\mathbf{1}(E_n) &\le (X_{n+1}-X_n)^2\mathbf{1}(E_n) + \frac{|X_{n+1}-X_n|}{x}|b_{\eta_{n+1}}-b_{\eta_n}|\mathbf{1}(E_n) + O(x^{-1})\\
&\le (X_{n+1}-X_n)^2 + o(1).
\end{aligned}$$

Taking expectations, we obtain

$$\mathbb{E}_{x,i}\left[(Z_{n+1}-Z_n)^2\mathbf{1}(E_n)\right] \le C,$$

for some $C\in\mathbb{R}_+$. On the other hand, we follow the proof of Lemma 4.2.9, inserting powers of $2/\nu$ and $1/\nu$ respectively, to get

$$\begin{aligned}
\mathbb{E}_{x,i}\left[(Z_{n+1}-Z_n)^2\mathbf{1}(E_n^{\mathrm{c}})\right] &= \mathbb{E}_{x,i}[(Z_{n+1}^2-Z_n^2)\mathbf{1}(E_n^{\mathrm{c}})] - 2\,\mathbb{E}_{x,i}[Z_n(Z_{n+1}-Z_n)\mathbf{1}(E_n^{\mathrm{c}})]\\
&= o(1).
\end{aligned}$$

Combining these estimates the second statement in the lemma follows. □

Now we can complete the proof of Theorem 3.1.3.

Proof of Theorem 3.1.3. Take $\nu = 2\theta$. We will apply Lemma 4.5.3 with $Z_n = (f_\nu(X_n, \eta_n))^{1/\nu}$ and $p_0 = \frac{\nu}{2}$; we must verify (4.5.4) and (4.5.5), and that $Z^\nu_{n\wedge\lambda_z}$ is a submartingale. For the latter condition, it suffices to show that

$$\mathbb{E}[f_\nu(X_{n+1}, \eta_{n+1}) - f_\nu(X_n, \eta_n) \mid X_n, \eta_n] \geq 0, \text{ on } \{Z_n > z\},$$

which follows from the $\nu = 2\theta$ case of Lemma 4.2.4 with hypothesis (3.1.3).

Of the remaining two conditions, (4.5.5) follows immediately from the second statement in Lemma 4.5.4, provided $\nu \leq p$, i.e., $\theta \leq \frac{p}{2}$. From the first statement in Lemma 4.5.4, we have that for all x sufficiently large

$$\mathbb{E}_{x,i}[Z_{n+1} - Z_n] = \frac{1}{x}\left(c_i + \frac{1}{2}\sum_{j\in S}(b_j - b_i)q_{ij} + o(1)\right) \geq -\frac{C_1}{x}, \tag{4.5.8}$$

for all i and all x sufficiently large, where $C_1 \in \mathbb{R}_+$ depends on the c_i and b_i. Now Lemma 4.2.3 implies that $Z_n \leq C_2 X_n$ for some $C_2 > 0$, so we deduce condition (4.5.4) from (4.5.8). □

4.5.4 Complete classification

To complete the classification in Theorem 3.1.1, we need to classify the different conditions for positive recurrent and null. Hence we should prove the following theorem, which is a simple consequence of the moment existence and non-existence results in the previous subsections.

Theorem 4.5.5. *Suppose that* (A) *holds, and that* (B_p) *holds for some* $p > 2$. *Suppose also that* (Q_∞) *and* ($\mathrm{D_L}$) *hold. Then the following classification applies.*

- *If* $\sum_{i\in S}(2c_i + s_i^2)\pi_i < 0$, *then* (X_n, η_n) *is positive recurrent.*
- *If* $\sum_{i\in S}(2c_i + s_i^2)\pi_i > 0$, *then* (X_n, η_n) *is null.*

Proof. Taking $\theta = 1$ in Theorem 3.1.2, we recover the condition for the first moment to exist, applying a technical Lemma 2.6.1 in [75] will give us the positive-recurrent

part. Take $\theta = 1$ in Theorem 3.1.3, we get the condition for the first moment to not exist, which means the process is null. □

If we group the results from Theorem 4.5.1 and Theorem 4.5.5, we proved Theorem 3.1.1 using the Lyapunov function method, completing the classification of positive recurrence and transience for the Lamperti drift case.

4.6 Proofs of results for generalized Lamperti drift

For this section, we turn our attention to the last and most subtle case with the generalized Lamperti drift, and give a complete classification for all situations for the half strip problem. Our main goal is to prove Theorem 3.2.2.

4.6.1 Transformation to Lamperti drift case

As the structure of (X_n, η_n) is quite complicated, in the first step of the proof, we want to transform X_n to $\widetilde{X}_n$ so that we can have a simpler form for the drifts and conditions. An enlightening transformation would be $(\widetilde{X}_n, \eta_n) = (X_n + a_{\eta_n}, \eta_n)$, where a_i are the solution of the system of equations $d_i + \sum_{j \in S}(a_j - a_i)q_{ij} = 0$ for all $i \in S$. The intuition behind this is that the transformation pulls or pushes each line separately in a way such that the effects of d_i are eliminated, turning all $d_i = 0$ after the transformation, i.e., the constant components of the drifts are eliminated; then we can apply the results in Section 3.1, once we have at hand increment moment estimates for the transformed process $(\widetilde{X}_n, \eta_n)$. A few items of technicality that need to be handled are listed as follows.

1. Existence and uniqueness up to translation of a_i,

2. Preservation of the classification under the transformation,

3. Changed moments.

For the first item, for d_i, $i \in S$ with $\sum_{i\in S} \pi_i d_i = 0$ as specified under assumption ($\mathrm{D_G}$), choose a_i, $i \in S$ as guaranteed by Lemma 3.2.1, so that $d_i + \sum_{j\in S}(a_j - a_i)q_{ij} = 0$; since we are free to translate the a_i by any constant, we may (and do) suppose here that $a_i \geq 0$ for all $i \in S$.

Let $\Sigma' = \{(x + a_i, i) : (x, i) \in \Sigma\}$ denote the state space of the transformed process; then Σ' is a locally finite subset of $\mathbb{R}_+ \times S$ with unbounded lines $\Lambda'_k = \{x \in \mathbb{R}_+ : (x, k) \in \Sigma'\}$. The map $(x, i) \mapsto (x + a_i, i)$ is a bijection, so the Markov chain $(\widetilde{X}_n, \eta_n)$ has precisely the same abstract structure as the Markov chain (X_n, η_n), in particular, the transformed process is (positive) recurrent if and only if the original process is (positive) recurrent, and so on, see Theorem 4.6.1 for a formal statement. Thus to obtain results for the process (X_n, η_n) it is sufficient to prove results for the transformed process $(\widetilde{X}_n, \eta_n)$.

4.6.2 Preservation of the recurrence classification after transformation

To see the preservation of the recurrence classification after the transformation, we want to establish the following theorem.

Suppose we have a Markov chain Z_n on a countable space Σ. Define $\lambda_x = \min\{n \geq 0 : Z_n = x\}$ and $\lambda_A = \min_{x\in A} \lambda_x$. Then we call the Markov chain Z_n s-recurrent if, for every $x \in \Sigma$, and any finite, non-empty $A \subset \Sigma$, $\mathbb{E}_x[\lambda_A^s] < \infty$. The following theorem will give all the facts we need for both the recurrence classification, and the moment existence and non-existence.

Theorem 4.6.1. *For any irreducible Markov chain Z_n on a countable space Σ, and any bijective function $f : \Sigma \to \Sigma'$, denote $Z'_n = f(Z_n)$, then we have all of the following.*

1. *Z'_n is an irreducible Markov chain on Σ',*
2. *Z_n is recurrent if and only if Z'_n is recurrent,*
3. *Z_n is positive recurrent if and only if Z'_n is positive recurrent,*

4. Z_n is s-recurrent if and only if Z'_n is s-recurrent.

Proof. As Z_n is a Markov chain, then $\mathbb{P}(Z_{n+1} = z \mid Z_1 = z_1, Z_2 = z_2, \ldots, Z_n = z_n) = \mathbb{P}(Z_{n+1} = z \mid Z_n = z_n)$. By applying the transformation f on every point on the space, we have $\mathbb{P}(Z'_{n+1} = f(z) \mid Z'_1 = f(z_1), Z'_2 = f(z_2), \ldots, Z'_n = f(z_n)) = \mathbb{P}(Z'_{n+1} = f(z) \mid Z'_n = f(z_n))$. So for any $z', z'_1, z'_2, \ldots, z'_n \in \Sigma'$, we can find $z = f^{-1}(z'), z_1 = f^{-1}(z'_1), z_2 = f^{-1}(z'_2), \ldots, z_n = f^{-1}(z'_n)$ so that the Markov property preserves. Also, for any state $z'_1, z'_2 \in \Sigma'$, we can find $z_1 = f^{-1}(z'_1)$ and $z_2 = f^{-1}(z'_2)$. If Z_n is irreducible, then z_1 and z_2 communicates, and so does z'_1 and z'_2. So Z'_n is also irreducible. The other direction follows as f^{-1} is also a bijection.

Next, for any $y \in \Sigma'$, there exists $x \in \Sigma$ such that $x = f^{-1}(y)$. Then if Z_n is recurrent, we have $\mathbb{P}[Z_n = x \ i.o.] = 1$. This means $\mathbb{P}[Z'_n = y \ i.o.] = 1$ by applying the transformation. So we get Z'_n is recurrent. The other directing follows similarly. Hence we have proved the second statement.

The third claim is in fact a special case of the fourth claim with $s = 1$, so we will just prove the fourth claim instead.

Suppose Z_n is s-recurrent. Let $\lambda'_B = \min\{n \geq 0 : Z'_n \in B\}$. Then

$$\begin{aligned}\lambda'_B &= \min\{n \geq 0 : f(Z_n) \in B\} \\ &= \min\{n \geq 0 : Z_n \in f^{-1}(B)\} \\ &= \lambda_{f^{-1}(B)}\end{aligned}$$

so $\mathbb{E}_y(\lambda'_B)^s = \mathbb{E}_{f^{-1}(y)}\, \lambda^s_{f^{-1}(B)} < \infty$, i.e. Z'_n is s-recurrent. The argument clearly goes both ways.

Hence we have proved all claims in the Theorem. □

Now with the first three claims of Theorem 4.6.1, we know the structure of the Markov chain and the classification of positive recurrent, null recurrent and transient is preserved under the transformation.

4.6.3 Calculation of the transformed moments

For the increment moments of the transformed process, we use the notation

$$\widetilde{\mu}_i(y) := \mathbb{E}[\widetilde{X}_{n+1} - \widetilde{X}_n \mid \widetilde{X}_n = y, \eta_n = i],$$
$$\widetilde{\sigma}_i^2(y) := \mathbb{E}[(\widetilde{X}_{n+1} - \widetilde{X}_n)^2 \mid \widetilde{X}_n = y, \eta_n = i].$$

Lemma 4.6.2. *Suppose that* (A) *holds, and that* (B_p) *holds for some* $p > 2$. *Suppose also that* ($\mathrm{Q_G}$) *and* ($\mathrm{D_G}$) *hold. Define* a_i, $i \in S$ *to be a solution to* (3.2.1) *with* $a_i \geq 0$ *for all* i, *whose existence is guaranteed by Lemma 3.2.1. Either (i) set* $\delta_4 = 0$; *or (ii) suppose that* ($\mathrm{Q_G^+}$) *and* ($\mathrm{D_G^+}$) *hold and set* $\delta_4 = \delta_2 \wedge \delta_3 \in (0,1)$. *For* $i \in S$, *define*

$$c_i := e_i + \sum_{j\in S} a_j \gamma_{ij}, \quad \textit{and} \quad s_i^2 := t_i^2 + 2\sum_{j\in S} a_j d_{ij} + \sum_{j\in S}(a_j^2 - a_i^2) q_{ij}. \tag{4.6.1}$$

Then we have that, as $x \to \infty$,

$$\widetilde{\mu}_i(x) = \frac{c_i}{x} + o(x^{-1-\delta_4}), \quad \textit{and} \quad \widetilde{\sigma}_i^2(x) = s_i^2 + o(x^{-\delta_4}).$$

Proof. For concreteness, we give the proof when ($\mathrm{Q_G^+}$) and ($\mathrm{D_G^+}$) hold; in the other case the argument is the same but with δ_4 set to 0 throughout. For the first moment, we have

$$\begin{aligned}\widetilde{\mu}_i(x) &= \mathbb{E}_{x-a_i,i}[X_{n+1} - X_n] + \sum_{j\in S} \mathbb{E}_{x-a_i,i}\left[\left(a_{\eta_{n+1}} - a_{\eta_n}\right) \mathbf{1}\{a_{\eta_{n+1}} = j\}\right] \\ &= \mu_i(x - a_i) + \sum_{j\in S}(a_j - a_i) q_{ij}(x - a_i).\end{aligned}$$

Now using hypothesis (a) in ($\mathrm{D_G^+}$) and ($\mathrm{Q_G^+}$) we obtain

$$\begin{aligned}\widetilde{\mu}_i(x) &= d_i + \frac{e_i}{x - a_i} + \sum_{j\in S}(a_j - a_i) q_{ij} + \sum_{j\in S}(a_j - a_i)\frac{\gamma_{ij}}{x - a_i} + o((x - a_i)^{-1-\delta_4}) \\ &= d_i + \sum_{j\in S}(a_j - a_i) q_{ij} + \frac{e_i}{x} + \frac{1}{x}\sum_{j\in S} a_j \gamma_{ij} + o(x^{-1-\delta_4}),\end{aligned}$$

since $\sum_{j\in S} \gamma_{ij} = 0$. By hypothesis (d) in ($\mathrm{D_G}$) and choice of the a_i (cf Lemma 3.2.1), the constant term here vanishes, so we obtain the expression for $\widetilde{\mu}_i(x)$ in the lemma.

For the second moment, we have

$$\begin{aligned}\widetilde{\sigma}_i^2(x) &= \mathbb{E}_{x-a_i,i}\left[(X_{n+1}-X_n)^2\right] + 2\,\mathbb{E}_{x-a_i,i}\left[(a_{\eta_{n+1}}-a_{\eta_n})(X_{n+1}-X_n)\right] \\ &\quad + \mathbb{E}_{x-a_i,i}\left[a_{\eta_{n+1}}^2 - a_{\eta_n}^2\right] \\ &= s_i^2 + 2\sum_{j\in S}(a_j-a_i)\mu_{ij}(x-a_i) + \sum_{j\in S}(a_j-a_i)^2 q_{ij}(x-a_i).\end{aligned}$$

Then using hypothesis (c) in ($\mathrm{D}_{\mathrm{G}}^+$) and ($\mathrm{Q}_{\mathrm{G}}^+$) we obtain

$$\begin{aligned}\widetilde{\sigma}_i^2(x) &= s_i^2 + 2\sum_{j\in S}(a_j-a_i)d_{ij} + \sum_{j\in S}(a_j-a_i)^2 q_{ij} + o(x^{-\delta_4}) \\ &= s_i^2 + 2\sum_{j\in S}a_j d_{ij} + \sum_{j\in S}(a_j^2-a_i^2)q_{ij} - 2a_i\left(d_i + \sum_{j\in S}(a_j-a_i)q_{ij}\right) + o(x^{-\delta_4}),\end{aligned}$$

which gives the result after once again using the fact that $d_i + \sum_{j\in S}(a_j-a_i)q_{ij} = 0$. □

4.6.4 Proof of recurrence classification

Armed with our transformation of the process, we can now use the results in Section 3.1 to complete the proofs of the theorems in Section 3.2.

Proof of Theorem 3.2.2. Lemma 4.6.2 shows that if (X_n, η_n) satisfies the conditions of Theorem 3.2.2, then $(\widetilde{X}_n, \eta_n)$ satisfies the conditions of Theorem 3.1.1 with c_i and s_i^2 as given by (4.6.1). Theorem 3.1.1 shows that the process is transient if

$$\begin{aligned}0 < \sum_{i\in S}[2c_i - s_i^2]\pi_i &= \sum_{i\in S}\left[2e_i + 2\sum_{j\in S}a_j\gamma_{ij} - \left(t_i^2 + 2\sum_{j\in S}a_j d_{ij} + \sum_{j\in S}(a_j^2-a_i^2)q_{ij}\right)\right]\pi_i \\ &= \sum_{i\in S}\left(2e_i - t_i^2 + 2\sum_{j\in S}a_j(\gamma_{ij}-d_{ij})\right)\pi_i - \sum_{i\in S}\sum_{j\in S}(a_j^2-a_i^2)q_{ij}\pi_i,\end{aligned}$$

using the expressions at (4.6.1). Note that the final term here vanishes, because

$$\begin{aligned}\sum_{i\in S}\sum_{j\in S}(a_j^2-a_i^2)q_{ij}\pi_i &= \sum_{j\in S}a_j^2\sum_{i\in S}q_{ij}\pi_i - \sum_{i\in S}a_i^2\pi_i\sum_{j\in S}q_{ij} \\ &= \sum_{j\in S}a_j^2\pi_j - \sum_{i\in S}a_i^2\pi_i = 0,\end{aligned}$$

using the fact that $\boldsymbol{\pi}$ is the stationary distribution for (q_{ij}). This gives the condition

for transience stated in Theorem 3.2.2.

Similarly, the condition for positive recurrence is

$$\begin{aligned}
0 > \sum_{i\in S}[2c_i + s_i^2]\pi_i &= \sum_{i\in S}\left[2e_i + 2\sum_{j\in S} a_j\gamma_{ij} + \left(t_i^2 + 2\sum_{j\in S} a_j d_{ij} + \sum_{j\in S}(a_j^2 - a_i^2)q_{ij}\right)\right]\pi_i \\
&= \sum_{i\in S}\left(2e_i + t_i^2 + 2\sum_{j\in S} a_j(\gamma_{ij} + d_{ij})\right)\pi_i + \sum_{i\in S}\sum_{j\in S}(a_j^2 - a_i^2)q_{ij}\pi_i \\
&= \sum_{i\in S}\left(2e_i + t_i^2 + 2\sum_{j\in S} a_j(\gamma_{ij} + d_{ij})\right)\pi_i,
\end{aligned}$$

which gives the condition for positive recurrence in the theorem.

4.6.5 Proofs of existence and non-existence of moments

The case for null recurrence and at the critical point follows accordingly by the same calculation. □

Proof of Theorem 3.2.3. The proof is analogous to the proof of Theorem 3.2.2, this time applying Theorem 3.1.2 to the transformed process. □

Proof of Theorem 3.2.4. This time we apply Theorem 3.1.3 to the transformed process. □

Chapter 5

Examples, applications and simulations

To finish this part of the thesis, we present an application of our results to a simple model of a *correlated random walk* and some more complicated numerical examples to see the delicacy of the phase transition.

5.1 Correlated random walk

Correlated random walk is a type of random walk that remembers a fixed number of previous steps of its past trajectory (except the first few steps, which may affected by less steps due to lack of history). There is a long list of literature in which this model is studied with various names by different researchers: as 'persistent random walks' by Fürth [90], 'correlated random walks' by Gillis [40], 'random walks with restricted reversals' by Domb and Fisher [25], and, recently, 'Newtonian random walks' by Lenci [68]. The correlated random walk also leads to the telegrapher's equation in the scaling limit under suitable rescaling, see Goldstein [45] and Kac [57]. For some recent work on correlated random walk and related models, see [1, 17, 54, 96]. Some applications or motivation for studying these models can be found in [37] on physical Brownian motion and [22] on models for molecular configurations. Some background material can be found in [55].

Now we formally state the model. Suppose that a particle performs a random walk on $\mathbb{Z}_+$ with a short-term memory: the distribution of X_{n+1} depends not only on the current position X_n, but also on the 'direction of travel' $X_n - X_{n-1}$. Formally, $(X_n, X_n - X_{n-1})$ is a Markov chain on $\mathbb{Z}_+ \times S$ with $S = \{-1, +1\}$, with

$$\mathbb{P}[(X_{n+1}, \eta_{n+1}) = (x + j, j) \mid (X_n, \eta_n) = (x, i)] = q_{ij}(x), \text{ for } i, j \in S.$$

Then for $i \in S$,

$$\mu_i(x) = \mathbb{E}[X_{n+1} - X_n \mid (X_n, \eta_n) = (x, i)] = q_{i,+1}(x) - q_{i,-1}(x).$$

The simplest model has $q_{ii}(x) = q > 1/2$ for $x \geq 1$, so the walker has a tendency to continue in its direction of travel.

More generally, suppose that for $q \in (0, 1)$ and constants $c_{-1}, c_{+1} \in \mathbb{R}$ and $\delta > 0$,

$$q_{ij}(x) = \begin{cases} q + \frac{ic_i}{2x} + O(x^{-1-\delta}) & \text{if } j = i; \\ 1 - q - \frac{ic_i}{2x} + O(x^{-1-\delta}) & \text{if } j \neq i. \end{cases} \tag{5.1.1}$$

Here is the recurrence/transience classification for this model, which includes as the special case $q = 1/2$ the recurrence classification in Corollary 3.1 of [39].

Theorem 5.1.1. *Consider the correlated random walk specified by* (5.1.1). *Let* $c = (c_{+1} + c_{-1})/2$. *If* $c < -q$, *then the walk is positive recurrent. If* $c > q$, *then the walk is transient. If* $|c| \leq q$, *then the walk is null recurrent.*

We can also achieve results for moment existence and non-existence for the correlated random walk model.

Theorem 5.1.2. *Consider the correlated random walk specified by* (5.1.1). *Let* $c = (c_{+1} + c_{-1})/2$. *Then* $\mathbb{E}[\tau_x^\theta] < \infty$ *if* $\theta < \frac{1}{2} - \frac{c}{2q}$ *while* $\mathbb{E}[\tau_x^\theta] = \infty$ *if* $\theta > \frac{1}{2} - \frac{c}{2q}$.

5.2 Proofs of theorems on correlated random walk

In this section we complete the proof of Theorem 5.1.1 and Theorem 5.1.2 on correlated random walk given in the last section.

Proof of Theorem 5.1.1. Note first that $q_{ii} = q$ and $q_{ij} = 1 - q$ for $j \neq i$; hence $\pi = (\frac{1}{2}, \frac{1}{2})$. By direct calculation, we get $\mu_i(x) = i(2q-1) + \frac{c_i}{x} + O(x^{-1-\delta})$. This gives $d_i = i(2q-1)$ and $e_i = c_i$. Now we want to solve the system of equations (3.2.1) for a_i, which amounts to

$$1 - 2q + (a_1 - a_{-1})(1 - q) = 0.$$

Since the solution (a_i) is unique up to translation, without loss of generality we can choose $a_{-1} = 0$ and then we get $a_{+1} = \frac{2q-1}{1-q}$. Next we observe that $d_{ii} = q$, while if $i \neq j$ we have $d_{ij} = 1 - q$; also, $\gamma_{ij} = \frac{jc_i}{2}$ and $t_i^2 = 1$. Now we calculate

$$\sum_{i \in S}\left(2e_i + 2\sum_{j \in S} a_j \gamma_{ij}\right)\pi_i = e_{+1} + e_{-1} + a_{+1}\gamma_{+1,+1} + a_{+1}\gamma_{-1,+1} = \frac{c_{+1} + c_{-1}}{2(1-q)}; \text{ and}$$

$$\sum_{i \in S}\left(t_i^2 + 2\sum_{j \in S} a_j d_{ij}\right)\pi_i = 1 + (a_{+1}d_{+1,+1} + a_{+1}d_{-1,+1}) = \frac{q}{1-q}.$$

Then applying Theorem 3.2.2 we obtain the desired result. □

Proof of Theorem 5.1.2. With the calculation in the last proof, apply Theorem 3.2.3 and Theorem 3.2.4 gives the desirable results. □

5.3 Numerical examples and simulations

Now we will give some simulations of the correlated random walk. For the first example, we consider a one-step correlated random walk, with the transition probabilities as in the following Figure 5.1, where the '+1' line represents the case where the last step was going to the right and the line '-1' means the last step was going to the left. In this sense we can model the different behaviour of the transition probabilities (i.e. the correlation).

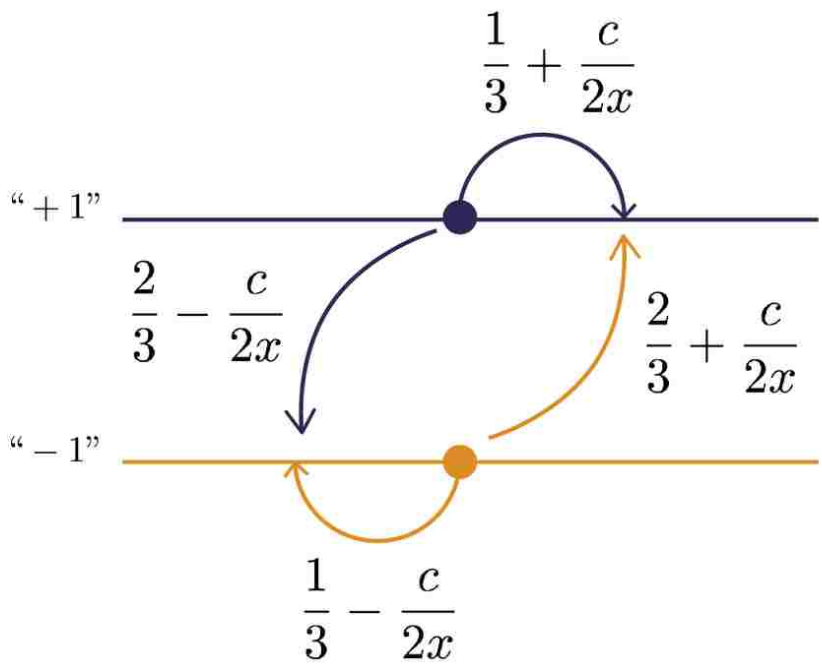

Figure 5.1: An example of a one-step correlated random walk.

We now conduct two simulations of 10^3 steps of the walk, with different values of $c = 1$ (Figure 5.2 left) and $c = -1$ (Figure 5.2 right). The horizontal axis represents the number of steps n and the vertical axis represents the value of X_n. The colour of the line segment represents which line the process is from, with the same colouring scheme shown in Figure 5.1.

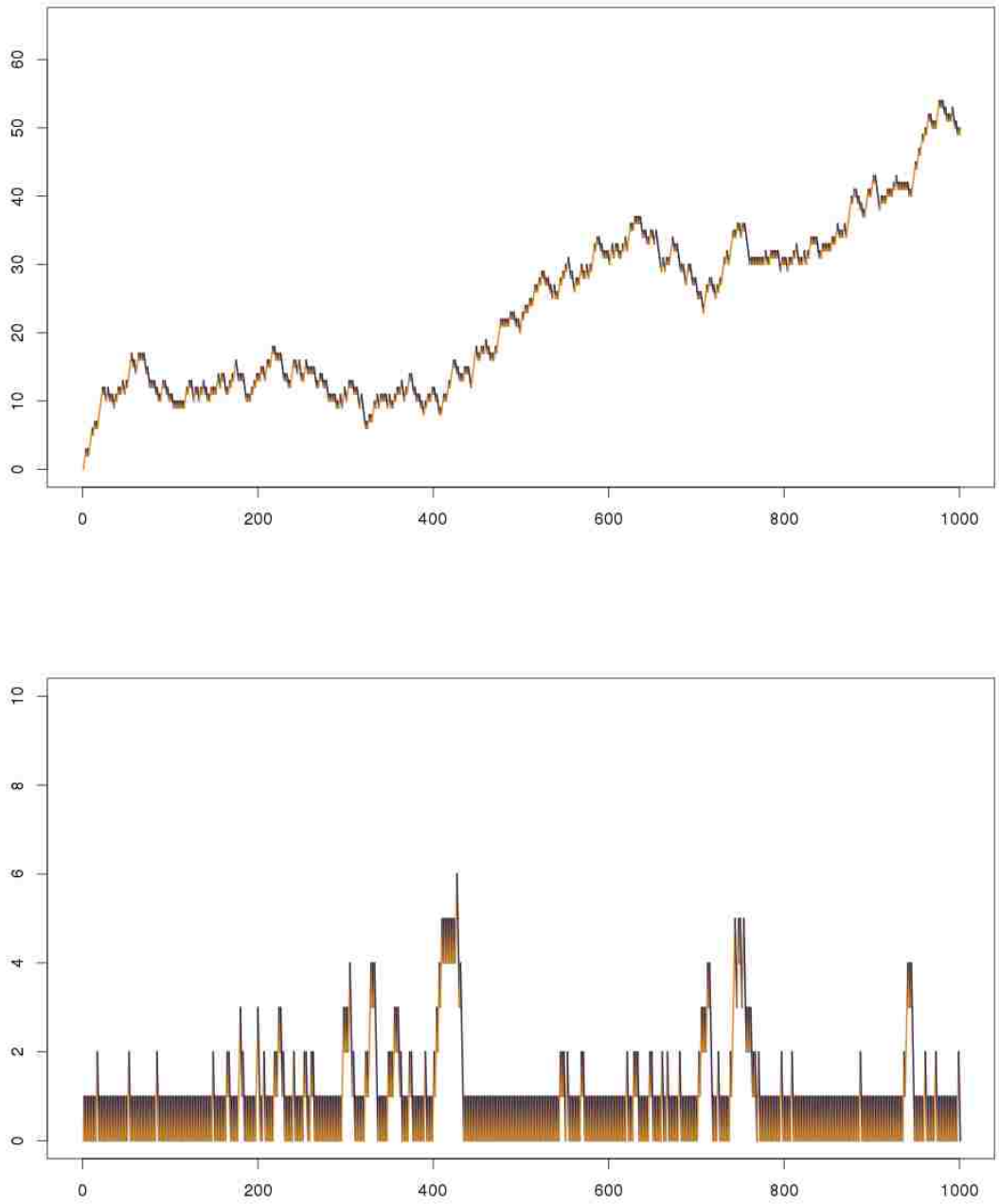

Figure 5.2: Two simulations of 10^3 steps of one-step correlated random walks, as an application of the half strip model. Top: $c = 1$, bottom: $c = -1$.

You can see the two simulations give very different behaviour in this walk. In fact, the left picture in Figure 5.2 shows the transient case and the right one shows the positive-recurrent case. The phase transitions of c are actually as the following, transient for $c > \frac{1}{3}$, positive recurrent for $c < -\frac{1}{3}$ and null recurrent for $|c| \leq \frac{1}{3}$ from the application of Theorem 5.1.1. Moreover, if we apply Theorem 5.1.2, we know $\mathbb{E}[\tau_x^\theta] < \infty$ if $\theta < \frac{1}{2} - \frac{3c}{2}$ while $\mathbb{E}[\tau_x^\theta] = \infty$ if $\theta > \frac{1}{2} - \frac{3c}{2}$.

Here is another example of two-steps correlated random walk, as the transition probabilities as shown in the following Figure 5.3, where this time we will have four lines, spanning all the combinations of the direction of the last two steps.

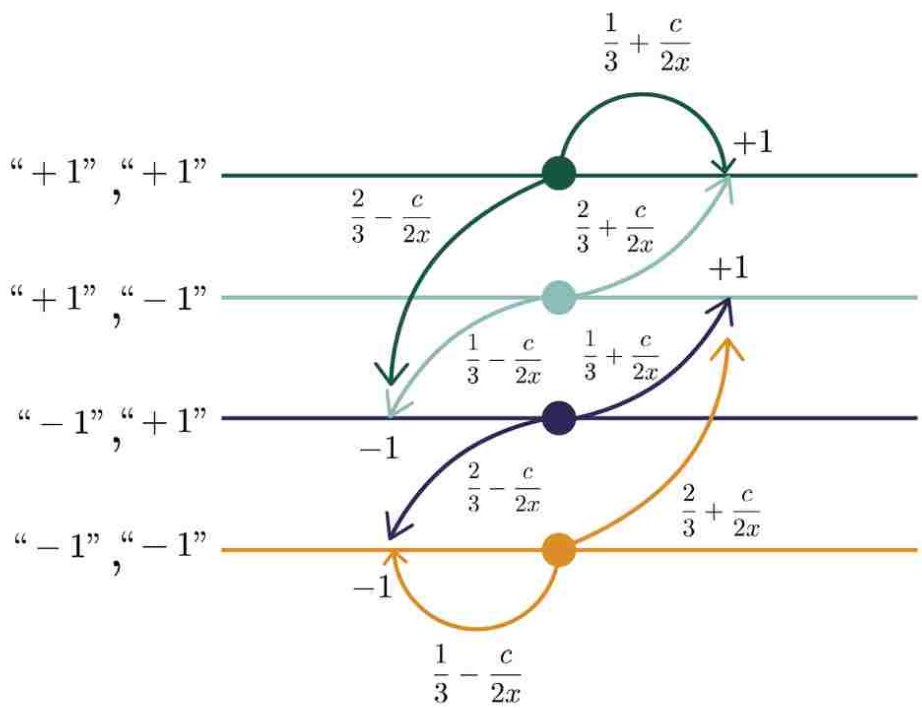

Figure 5.3: An example of a two-steps correlated random walk.

Again, for different value of c, we will get different results of classification.

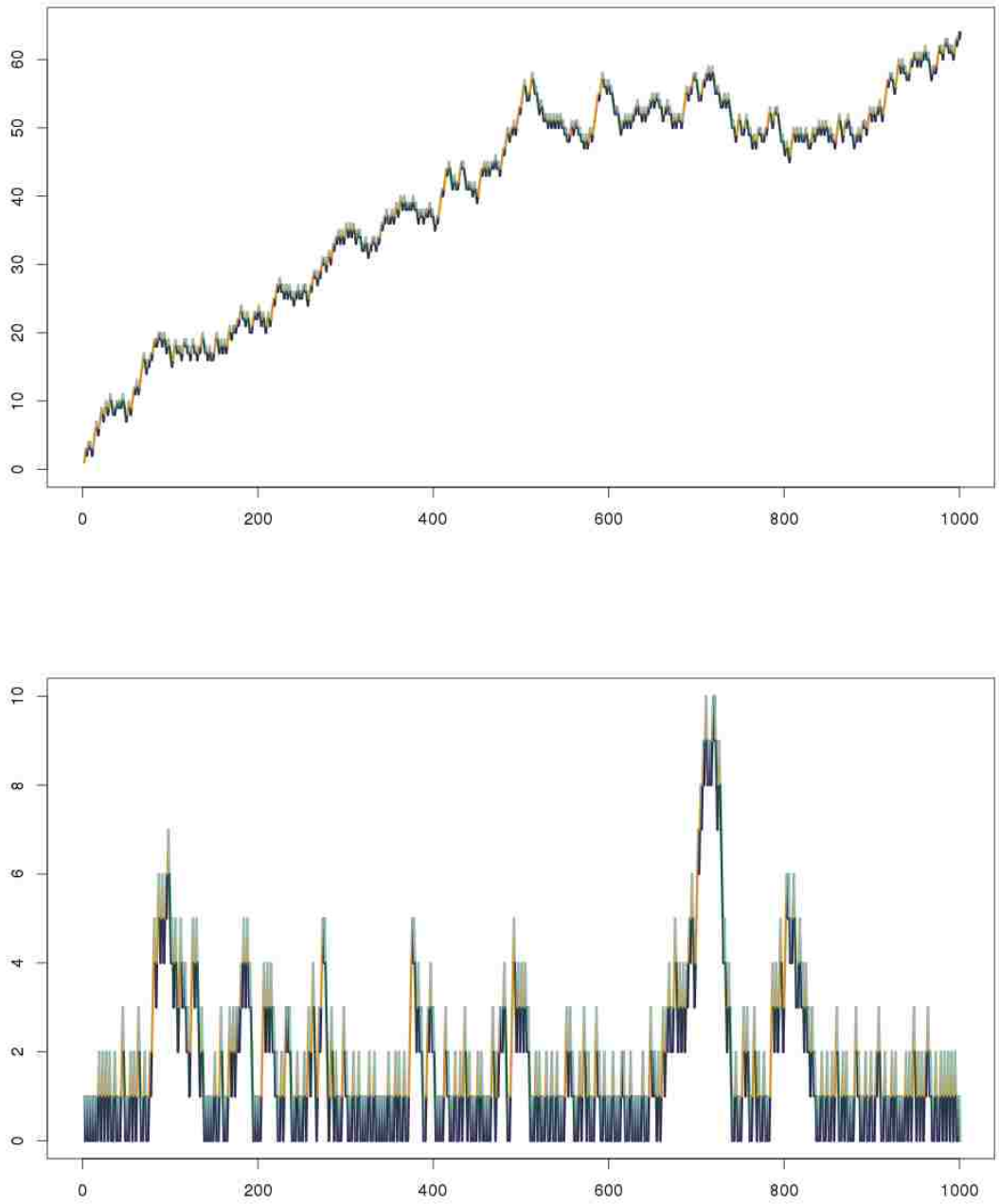

Figure 5.4: Two simulations of 10^3 steps of two-steps correlated random walks, as an application of the half strip model. Top: $c = 1$, bottom: $c = -1$.

The random walk is transient in the left picture in Figure 5.4, with the value $c = 1$ while the walk is positive recurrent for the right picture in Figure 5.4, with the value $c = -1$. The phase transitions of c in this case are transient for $c > \frac{1}{3}$, positive recurrent for $c < -\frac{1}{3}$ and null recurrent for $|c| \leq \frac{1}{3}$, from the application of Theorem 3.2.2. Moreover, if we apply Theorem 3.2.3 and Theorem 3.2.4, we know $\mathbb{E}[\tau_x^\theta] < \infty$ if $\theta < \frac{1}{2} - \frac{3c}{2}$ while $\mathbb{E}[\tau_x^\theta] = \infty$ if $\theta > \frac{1}{2} - \frac{3c}{2}$.

An important observation is that we have the same phase boundary as in the last example with only one step correlated random walk. This is not a mere coincidence. In fact, a careful calculation shows the same result for any n-step correlated random walk for any positive integer n, irrelevant to how we assign the favourableness to stick with or change direction with a certain pattern of the previous steps, for a symmetric and balance design. It would be quite clumsy to state the formal statement here and we shall leave the reader to discover the scintillating calculation.

Part II:

Centre of Mass of Random Walks

Chapter 6

Notation, preliminaries and prerequisites

6.1 Literature review

In physics, the center of mass of a distribution of mass in space is defined as the unique point that the weighted relative position of the distributed mass sums to zero and so the distribution of mass is balanced around the center of mass. We can get the coordinates of the center of mass by calculating the average of the weighted position coordinates of the distribution of mass. It is a very important concept and has a lot of useful applications in physics.

Back to random walk, properties of random walk in $\mathbb{Z}^d$ are undoubtedly a popular subject to study. A vast amount of study has been devoted to the investigation of the recurrence classification. This includes the famous Pólya's and Chung-Fuchs results in Chapter 1.

Now we want to go one step further and combine these two concepts, to consider the centre of mass (or time average, centre of gravity) of a random walk. Let $d \geq 1$. Suppose that $X, X_1, X_2, \ldots$ is a sequence of i.i.d. random variables on $\mathbb{R}^d$. We consider the random walk $(S_n, n \in \mathbb{Z}_+)$ in $\mathbb{R}^d$ defined by $S_0 := \mathbf{0}$ and $S_n := \sum_{i=1}^n X_i$ $(n \geq 1)$.

Our object of interest is the *centre of mass process* $(G_n, n \in \mathbb{Z}_+)$ corresponding

to the random walk, defined by $G_0 := \mathbf{0}$ and $G_n := \frac{1}{n}\sum_{i=1}^{n} S_i$ $(n \geq 1)$. The question of the asymptotic behaviour of G_n was raised by P. Erdős (see [46]).

Lets look at some simulations of how it looks.

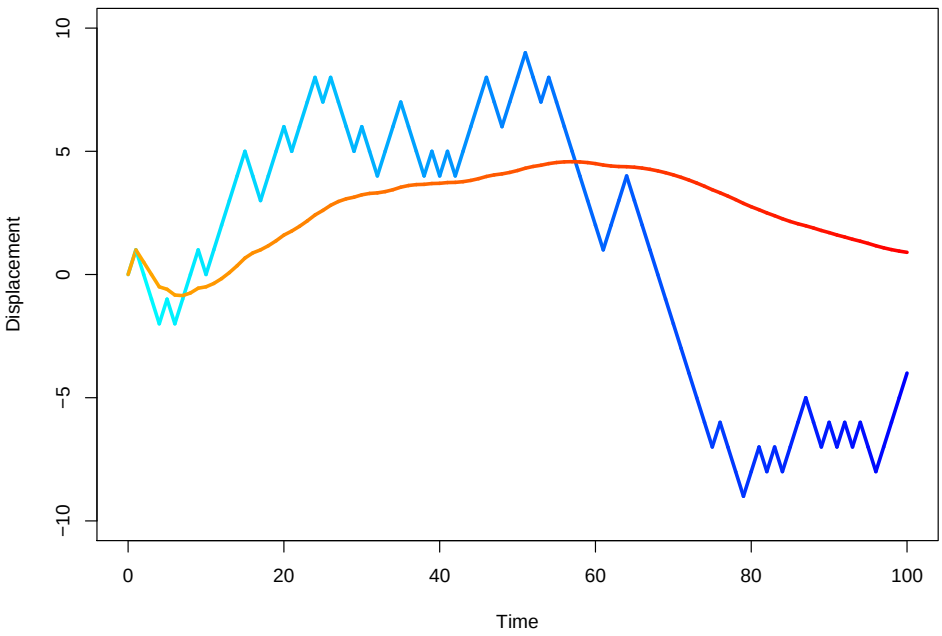

Figure 6.1: An example of a one dimensional random walk (Light blue to blue) with the corresponding centre of mass process (Orange to red).

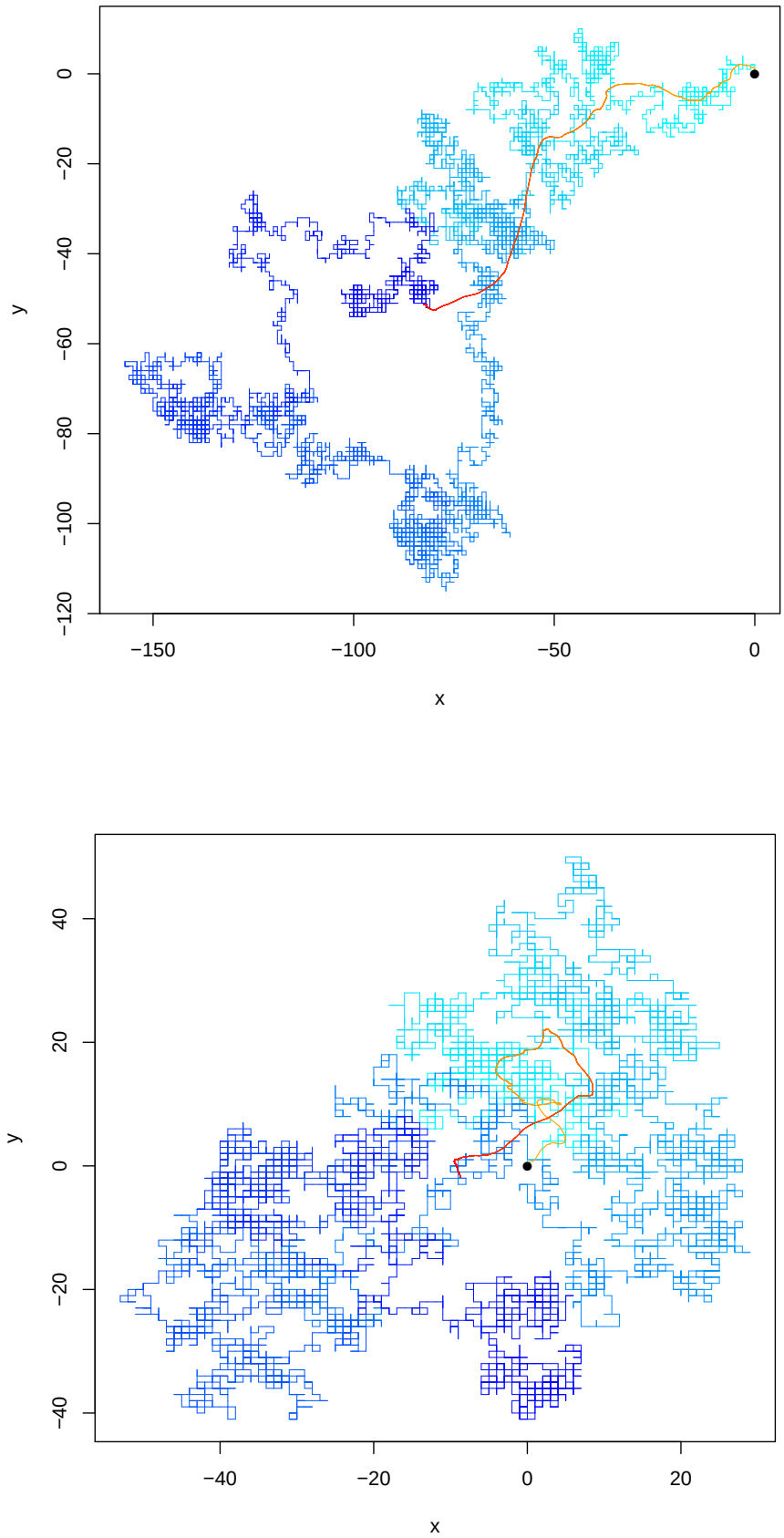

Figure 6.2: Two examples of a two dimensional random walk (Light blue to blue) with the corresponding centre of mass process (Orange to red).

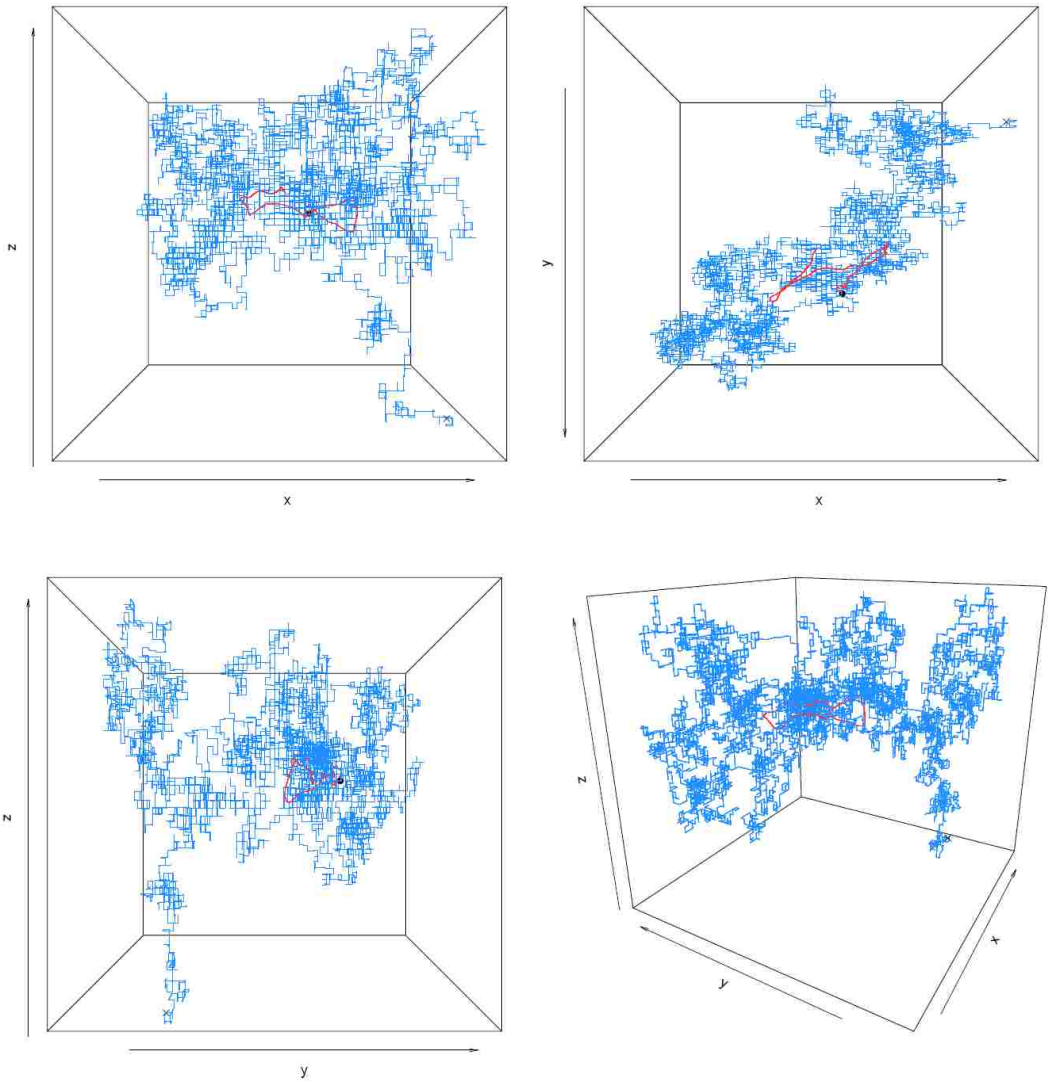

Figure 6.3: An example of a three dimensional random walk (Blue) with the corresponding centre of mass process (Red) at front (top left), top (top right), left (bottom left) and side (bottom right) angle.

Compared to Part I of this thesis, this is a very different model. We used a non homogeneous walk in a very special state space in the former to fit in some applications with the specific structure. However, this is not the only way to put our theory of random walk in action. Instead, we will try to use a common state space on a d-dimensional space, but as general as possible, together with the centre of mass, an averaging process derived from a random walk with as little structure as possible. As a result, we are able to get very general results, which can apply to a broad range of problems.

Random walks can be used to model physical polymer molecules [20,91] in which case the centre of mass is of obvious physical relevance. The random walk can also be used to model animal behaviour, and the motion of both macroscopic and microscopic organisms [53, 84]. In this context the centre of mass is a natural summary statistic of an animal's roaming behaviour.

Despite the fact that it is a interesting and useful model, there is an extremely limited literature which gives any recurrence property of the centre of mass of random walk. The only result related to the properties of this process that we found is due to Karl Grill [46], which gives the fact that for simple symmetric random walk, its center of mass is recurrent for $d = 1$ only and is transient for $d \geq 2$. This answer was originally conjectured by Paul Erdős, and no further generalization or conjecture is found in the literature.

This centre of mass process has a rich meaning in terms of application. For example in [20], the asymptotic behaviour of a d-dimensional self-interacting random walk, which is repelled or attracted by the centre of mass of its previous trajectory, was studied. The walk's trajectory models a random polymer chain in either poor or good solvent.

In this part of the thesis we will deeply investigate this mysterious centre of mass process and prove some of its properties in a general setting. The first goal in this part is to get the asymptotic behaviour of the process. We give the *strong law of large numbers* and the *central limit theorem* for the centre of mass, under minimal assumptions and we will further extend these results in the next part of the thesis.

The first main result is a local central limit theorem. Unlike the usual central limit theorem, it is a much more precise theorem that requires a lot of effort to prove as it is not a direct consequence of the central limit theorem. One crucial difference here is that the usual central limit theorem does not distinguish between continuous and discrete random variables, while the local limit theorem does. Especially in the discrete case, the local limit theorem tells us more on the structure of the process. Roughly speaking, the central limit theorem tell us the process should behave like the normal distribution as a whole, but the local limit theorem tell us that the probability distribution of the process at each point behave like the probability density function of the normal distribution.

The second important goal in this part is to obtain the recurrence classification of the center of mass process, associated with any homogeneous random walk in d-dimensions. If we just focus on the recurrence classification, then the case where the random walk has a non-zero drift will not be very interesting, as the strong law of large numbers that we prove tells us the process must be transient in any dimension. In the case of zero drift, it turns out to be a very difficult problem without any additional assumptions. We acquire the classification result for any lattice random walk on $\mathbb{R}^d$ with finite second moments of the drift, which is a minor but important assumption to control the behaviour of the jumps. We show that under these assumptions the centre of mass is recurrent in one dimension but transient in two or more dimensions. We will also provide more general results in one dimension.

It is notable that the behaviour of the centre of mass of the random walk, especially in 2-dimensions, is counter-intuitive at first sight. If we just observe the central limit theorem of both process, they behave very similarly. In fact, comparing Proposition 7.1.2 with Theorem 1.6.2, we see S_n converge to μ and G_n converge to $\mu/2$. The variance in the normal distribution for G_n is $1/3$ of the one for S_n. Also, after some calculation, we will see that both process have the same magnitude of probabilities to come back to a fixed region. Comparing Theorem 7.2.1 with Theorem 6.5.1 we see that for a fixed ball $\mathcal{B}$, both $\mathbb{P}(S_n \in \mathcal{B})$ and $\mathbb{P}(G_n \in \mathcal{B})$ are

of order $n^{-d/2}$. Hence one might think the recurrence classification of the center of mass process should follow from the random walk. Shockingly, this is not the case because we did not consider the speed of the processes, which will be the key for our proofs of theorems in the next chapter. This once again suggests that the recurrence classification is easy to define and to understand, but to really classify it properly around the critical region is a very subtle problem.

The outline of this part of the thesis is as follows. This chapter will provide the background material we need for the proofs of our main theorems. In the next chapter, we will give a local central limit theorem for the center of mass process, and give the recurrence classification of it in different dimensions. Chapter 8 will provide all the proofs and technical details for these theorems, and we will close up with some examples in Chapter 9.

We view vectors in $\mathbb{R}^d$ as column vectors throughout; $\mathbf{0}$ denotes the zero vector.

6.2 Lattice distributions and characteristic functions

The first main goal for this part of the thesis is a *local* central limit theorem for the centre of mass process. To achieve this we have to know some facts about lattice distributions and characteristic functions.

In our centre of mass model, we assume that X has a non-degenerate d-dimensional lattice distribution. Thus (see [11, Ch. 5]) there is a unique minimal subgroup $L := H\mathbb{Z}^d$ of $\mathbb{R}^d$, where H is a d by d matrix, such that $\mathbb{P}(X \in \mathbf{b} + L) = 1$ for some $\mathbf{b} \in \mathbb{R}^d$, with the property that if $\mathbb{P}(X \in \mathbf{x} + L') = 1$ for some closed subgroup L' of $\mathbb{R}^d$ and $\mathbf{x} \in \mathbb{R}^d$, then $L \subseteq L'$, and with $h := |\det H| \in (0, \infty)$. In other words, we make the following assumption.

(L) Suppose that the minimal subgroup of $\mathbb{R}^d$ associated with X is $L := H\mathbb{Z}^d$ with $h := |\det H| > 0$.

Equivalent conditions to (L) can be formulated in terms of the characteristic function of X or in terms of the maximality of h: see Lemma 6.2.3 below. Note that there may be many matrices H for which $H\mathbb{Z}^d$ is equal to (unique) L, but for all of these $|\det H|$ is the same. Also note that symmetric simple random walk (SSRW) does not satisfy (L) with the obvious choice $H = I$ (the identity), but does satisfy (L) if H has the maximal choice $h = 2$: see Chapter 9 for more details.

We collect some facts about lattice distributions: for reference see [11, Ch. 5] and [95, §7]. Let

$$\mathcal{H} := \{H : \mathbb{P}(X \in \mathbf{b} + H\mathbb{Z}^d) = 1 \text{ for some } \mathbf{b} \in \mathbb{R}^d\}.$$

If X has a lattice distribution, then $\mathcal{H}$ is nonempty, and if X is non-degenerate then any $H \in \mathcal{H}$ has $|\det H| > 0$. (Here and elsewhere, 'non-degenerate' means not supported on any $(d-1)$-dimensional hyperplane.) Let $K := \{|\det H| : H \in \mathcal{H}\}$. The next result gives an upper bound on $h \in K$; note that this bound is sharp in both of the examples in Chapter 9.

Lemma 6.2.1. *Suppose that X has a non-degenerate lattice distribution. Then $K \subseteq (0, \infty)$ is bounded, and* $\inf K = 0$.

Proof. Since X has a non-degenerate lattice distribution, we have that (i) $\mathcal{H}$ is non-empty and $|\det H| > 0$ for all $H \in \mathcal{H}$; and (ii) there exists $\mathcal{X} := \{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_d\}$ such that $\mathbf{x}_0, \ldots, \mathbf{x}_d$ are affinely independent, and $\mathbb{P}(X = \mathbf{x}_i) > 0$ for each i. Statement (i) shows that $K \subseteq (0, \infty)$ is nonempty, and statement (ii) shows that K is bounded. Indeed, for any $H \in \mathcal{H}$ we have that there exists $\mathbf{b}$ such that $\mathcal{X} \subset \mathbf{b} + H\mathbb{Z}^d$, i.e., $H^{-1}(\mathcal{X} - \mathbf{b}) \subset \mathbb{Z}^d$. For $i \in \{1, \ldots, d\}$ let $\lambda_i = \mathbf{x}_i - \mathbf{x}_0$. Then the linearly independent vectors $\lambda_1, \ldots, \lambda_d$ define a parallelepiped P with volume $|\det \Lambda| \in (0, \infty)$, where Λ denotes the $d \times d$ matrix whose columns are $\lambda_1, \ldots, \lambda_d$. Since $H^{-1}(\mathcal{X} - \mathbf{b})$ are points of $\mathbb{Z}^d$, we have that all the vertices of the parallelepiped $P' := H^{-1}(\mathbf{x}_0 + P - \mathbf{b})$ are points of $\mathbb{Z}^d$. Now P' has volume $h^{-1}|\det \Lambda| > 0$, but, as a parallelepiped of positive

volume whose vertices are in $\mathbb{Z}^d$, must have volume at least 1. Thus $h^{-1}|\det\Lambda| \geq 1$, i.e., $h \leq |\det\Lambda| < \infty$. Also, we see that if $H \in \mathcal{H}$, then $H/2 \in \mathcal{H}$ as well, so if $h \in K$ then $h/2^d \in K$ too. □

The next lemma will be on the *characteristic function* of X, which is defined for $\mathbf{t} \in \mathbb{R}^d$ as $\varphi(\mathbf{t}) := \mathbb{E}\,\mathrm{e}^{i\mathbf{t}^\top X}$. We also define $U := \{\mathbf{t} \in \mathbb{R}^d : |\varphi(\mathbf{t})| = 1\}$. Given an invertible d by d matrix H, set $S_H := 2\pi(H^\top)^{-1}\mathbb{Z}^d$. The next result shows that if $H \in \mathcal{H}$, then $S_H \subseteq U$.

Lemma 6.2.2. *Suppose that $H \in \mathcal{H}$. Then $|\varphi(\mathbf{u})| = 1$ for all $\mathbf{u} \in S_H$.*

Proof. First observe that the norm of the characteristic function is invariant under translation by any vector of the form of $2\pi(H^\top)^{-1}\mathbf{k}$ with $\mathbf{k} \in \mathbb{Z}^d$. To see this, note that for any $\mathbf{k} \in \mathbb{Z}^d$,

$$\left|\varphi(\mathbf{t} + 2\pi(H^\top)^{-1}\mathbf{k})\right| = \left|\mathbb{E}\left[\mathrm{e}^{i\mathbf{t}^\top X} \cdot \mathrm{e}^{2\pi i\mathbf{k}^\top H^{-1}X}\right]\right|.$$

Since $H \in \mathcal{H}$, we may write $X = \mathbf{b} + HW$, where $\mathbf{b} \in \mathbb{R}^d$ is constant and $W \in \mathbb{Z}^d$. Hence

$$\left|\varphi(\mathbf{t} + 2\pi(H^\top)^{-1}\mathbf{k})\right| = \left|\mathrm{e}^{2\pi i\mathbf{k}^\top H^{-1}\mathbf{b}}\right| \cdot \left|\mathbb{E}\left[\mathrm{e}^{i\mathbf{t}^\top X} \cdot \mathrm{e}^{2\pi i\mathbf{k}^\top W}\right]\right|,$$

because $\mathbf{k}^\top H^{-1}\mathbf{b}$ is a non-random scalar. Then, since $|\exp\{2\pi i\mathbf{k}^\top H^{-1}\mathbf{b}\}| = 1$ and $\mathbf{k}^\top W \in \mathbb{Z}$, so that $\exp\{2\pi i\mathbf{k}^\top W\} = 1$, it follows that for any $\mathbf{k} \in \mathbb{Z}^d$,

$$\left|\varphi(\mathbf{t} + 2\pi(H^\top)^{-1}\mathbf{k})\right| = |\varphi(\mathbf{t})|. \tag{6.2.1}$$

In particular, the case $\mathbf{t} = \mathbf{0}$ of (6.2.1) shows that $|\varphi(\mathbf{u})| = 1$ if $\mathbf{u} \in S_H$. □

If $\mathbb{P}(X \in \mathbf{b} + H\mathbb{Z}^d) = 1$ and $\mathbb{P}(X = \mathbf{x}) > 0$, then $\mathbf{x} - \mathbf{b} \in H\mathbb{Z}^d$ so that $\mathbf{x} + H\mathbb{Z}^d = \mathbf{b} + H\mathbb{Z}^d$, and so if $H \in \mathcal{H}$ then $\mathbb{P}(X \in \mathbf{x} + H\mathbb{Z}^d) = 1$ for any $\mathbf{x}$ with $\mathbb{P}(X = \mathbf{x}) > 0$.

Lemma 21.4 of [11] shows that there is a unique minimal subgroup L of $\mathbb{R}^d$ such that $\mathbb{P}(X \in \mathbf{x} + L) = 1$ for any $\mathbf{x}$ with $\mathbb{P}(X = \mathbf{x}) > 0$ and if $H \in \mathcal{H}$ then $L \subseteq H\mathbb{Z}^d$. Moreover, the discrete subgroup L is generated by $\{\xi : \mathbb{P}(X = \mathbf{x} + \xi) > 0\}$ for any given $\mathbf{x}$ with $\mathbb{P}(X = \mathbf{x}) > 0$. We have $L = H_0\mathbb{Z}^d$ for some (not necessarily unique) $H_0 \in \mathcal{H}$; let $\mathcal{H}_0 := \{H \in \mathcal{H} : L = H\mathbb{Z}^d\}$.

The next result gives equivalent formulations of the fundamental assumption (L). For $\rho > 0$, define $S_H(\rho) := \cup_{\mathbf{y} \in S_H} B(\mathbf{y}; \rho)$, where $B(\mathbf{y}; \rho)$ is the open Euclidean ball of radius ρ centred at $\mathbf{y} \in \mathbb{R}^d$.

Lemma 6.2.3. *Suppose that X is non-degenerate and $H \in \mathcal{H}$. The following are equivalent.*

(i) $H \in \mathcal{H}_0$.

(ii) $|\det H|$ *is the maximal element of* K.

(iii) $S_H = U$.

Moreover, if any one of these conditions holds then, for any $\rho > 0$, there exists a positive constant c_ρ such that

$$|\varphi(\mathbf{u})| \leq \mathrm{e}^{-c_\rho}, \text{ for any } \mathbf{u} \notin S_H(\rho).$$

Proof. Suppose that $H_0 \in \mathcal{H}_0$ and $H \in \mathcal{H}$. Let $h_0 = |\det H_0|$ and $h = |\det H|$. Then, by minimality, $H_0\mathbb{Z}^d \subseteq H\mathbb{Z}^d$, i.e., $H^{-1}H_0\mathbb{Z}^d \subseteq \mathbb{Z}^d$. Thus $H^{-1}H_0[0,1]^d$ is a parallelepiped whose vertices are all in $\mathbb{Z}^d$, and necessarily this parallelepiped has volume at least 1. Hence $h_0/h \geq 1$, i.e., $h \leq h_0$. Thus if $H \in \mathcal{H}_0$ then $|\det H|$ is maximal. On the other hand, suppose $H \in \mathcal{H} \setminus \mathcal{H}_0$ and $H_0 \in \mathcal{H}_0$. Then $H_0\mathbb{Z}^d \subset H\mathbb{Z}^d$ are not equal, so there is some $\mathbf{x} \in H\mathbb{Z}^d$ with $\mathbf{x} \notin H_0\mathbb{Z}^d$. Thus if $\mathbf{y} = H^{-1}\mathbf{x} \in \mathbb{Z}^d$, we have that $H^{-1}H_0\mathbb{Z}^d \subset \mathbb{Z}^d$ with $\mathbf{y} \notin H^{-1}H_0\mathbb{Z}^d$. For $\mathbf{z} \in \mathbb{Z}^d$ we have $\mathbf{y} = H^{-1}H_0(\mathbf{z} + \alpha)$ where $\alpha \in [0,1]^d$ is not a vertex; but then $\mathbf{y} - H^{-1}H_0\mathbf{z} \in \mathbb{Z}^d$ as well. Thus $\beta = H^{-1}H_0\alpha$ is a point of $\mathbb{Z}^d$ contained in the parallelepiped $P = H^{-1}H_0[0,1]^d$, and moreover all the vertices of P are in $\mathbb{Z}^d$, and β is not a vertex. Hence the parallelepiped P has volume strictly greater than 1 (see [95, p. 69]), and so $h_0/h > 1$. Thus if $H \notin \mathcal{H}_0$ then $|\det H|$ is not maximal. Thus (i) and (ii) are equivalent.

We show that (i) implies (iii). For $H \in \mathcal{H}$ set

$$\begin{aligned} R_H &:= \{\mathbf{t} \in \mathbb{R}^d : \mathbf{x}^\top \mathbf{t} \in 2\pi\mathbb{Z} \text{ for all } \mathbf{x} \in H\mathbb{Z}^d\} \\ &= \{\mathbf{t} \in \mathbb{R}^d : \mathbf{z}^\top H^\top \mathbf{t} \in 2\pi\mathbb{Z} \text{ for all } \mathbf{z} \in \mathbb{Z}^d\}. \end{aligned}$$

It follows that

$$R_H = 2\pi(H^\top)^{-1}\{\mathbf{y} \in \mathbb{R}^d : \mathbf{z}^\top \mathbf{y} \in \mathbb{Z} \text{ for all } \mathbf{z} \in \mathbb{Z}^d\} = 2\pi(H^\top)^{-1}\mathbb{Z}^d = S_H.$$

So $R_H = S_H$ for any $H \in \mathcal{H}$ with $|\det H| > 0$. Moreover, Lemma 21.6 of [11] shows that $R_H = U$ if $H\mathbb{Z}^d$ is minimal. Thus (i) implies (iii).

Next we show that (iii) implies (ii). Let $h_\star := \sup K$, which, by Lemma 6.2.1 is finite and positive. Suppose that $H \in \mathcal{H}$ with $|\det H| = h \in (0, h_\star)$. Then for any $\varepsilon > 0$ sufficiently small, we can find $H_1 \in \mathcal{H}$ with $|\det H_1| = h_1 \in (h, h_\star]$ such that $h_1 > (1 + 2\varepsilon)h$ and $h_1 > (1 - \varepsilon)h_\star$. Let $S = 2\pi(H^\top)^{-1}\mathbb{Z}^d$ and $S_1 = 2\pi(H_1^\top)^{-1}\mathbb{Z}^d$.

Consider $\mathbf{x}$ with $\mathbb{P}(X = \mathbf{x}) > 0$. Then there exist $\mathbf{b}, \mathbf{b}_1 \in \mathbb{R}^d$ (not depending on $\mathbf{x}$) and $\mathbf{z}, \mathbf{z}_1 \in \mathbb{Z}^d$ (depending on $\mathbf{x}$) such that

$$\mathbf{x} = \mathbf{b} + H\mathbf{z} = \mathbf{b}_1 + H_1\mathbf{z}_1,$$

and hence

$$\mathbf{z} = H^{-1}(\mathbf{b}_1 - \mathbf{b}) + H^{-1}H_1\mathbf{z}_1. \tag{6.2.2}$$

Take $\mathbf{s} = 2\pi(H_1^\top)^{-1}\mathbf{z}_1 \in S_1$. Assume, for the purpose of deriving a contradiction, that $S_1 \subseteq S$. Then $\mathbf{s} \in S$, i.e., there exists $\mathbf{z}_2 \in \mathbb{Z}^d$ such that

$$\mathbf{s} = 2\pi(H_1^\top)^{-1}\mathbf{z}_1 = 2\pi(H^\top)^{-1}\mathbf{z}_2.$$

Together with (6.2.2), this implies that

$$\mathbf{z} = H^{-1}H_1H_1^\top(H^\top)^{-1}\mathbf{z}_2 + H^{-1}(\mathbf{b}_1 - \mathbf{b}).$$

It follows that

$$\mathbf{x} = \mathbf{b} + H\mathbf{z} = \mathbf{b}_1 + H_1H_1^\top(H^\top)^{-1}\mathbf{z}_2.$$

Now if we take $\mathbf{b}_2 = \mathbf{b}_1$ and $H_2 = H_1H_1^\top(H^\top)^{-1}$, we have shown that every $\mathbf{x}$ for which $\mathbb{P}(X = \mathbf{x}) > 0$ has $\mathbf{x} \in \mathbf{b}_2 + H_2\mathbb{Z}^d$, i.e., $H_2 \in \mathcal{H}$. But

$$\begin{aligned}|\det H_2| &= |\det H_1| \left|\det H_1^\top\right| \left|\det(H^\top)^{-1}\right| = \frac{h_1^2}{h} \\ &> (1 + 2\varepsilon)(1 - \varepsilon)h_\star > h_\star,\end{aligned}$$

for ε sufficiently small, which contradicts the definition of $h_\star$. Thus there exists some $\mathbf{x} \in S_1$ with $\mathbf{x} \notin S$.

From Lemma 6.2.2, we have $S_1 \subseteq U$; hence there is some $\mathbf{x} \in U$ with $\mathbf{x} \notin S$. In other words, we have shown that if $h \in (0, h_\star)$ then $S \neq U$. Thus if we assume that $S = U$, the only possibility is $h = h_\star \in \mathcal{H}$. Thus (iii) implies (ii).

To prove the final statement in the lemma, we may suppose that (iii) holds. Then $|\varphi(\mathbf{u})| < 1$ if $\mathbf{u} \notin S_H$. To finish the proof of the lemma, it suffices to show that $\sup_{\mathbf{u} \notin S_H(\rho)} |\varphi(\mathbf{u})| < 1$. But, by the periodicity property (6.2.1), we have $\sup_{\mathbf{u} \notin S_H(\rho)} |\varphi(\mathbf{u})| = \sup_{\mathbf{u} \in T_H(\rho)} |\varphi(\mathbf{u})|$ where $T_H(\rho) := 2\pi(H^\top)^{-1}[-\frac{1}{2}, \frac{1}{2}]^d \setminus B(\mathbf{0}; \rho)$. Suppose that $\sup_{\mathbf{u} \in T_H(\rho)} |\varphi(\mathbf{u})| = 1$; then by the continuity of $|\varphi(\mathbf{u})|$, the supremum is attained at a point $\mathbf{u}$ in the compact set $T_H(\rho)$, contradicting the fact that $|\varphi(\mathbf{u})| < 1$ for all $\mathbf{u} \notin S_H$. Hence $\sup_{\mathbf{u} \in T_H(\rho)} |\varphi(\mathbf{u})| < 1$, and the proof is completed. □

Further information about lattice distributions can also be found in [11] and [95].

6.3 Hewitt-Savage zero one law

The *Hewitt-Savage zero-one law* is important for us to establish some recurrence results in one dimensional center of mass process.

It is a theorem similar to *Kolmogorov's zero-one law* and the Borel-Cantelli lemma, that specifies that a particular type of event will happen almost surely or not happen almost surely.

Formally, we will follow the formulation and definitions from Chow and Teicher [18, p.232].

Let $(S, \mathcal{S})$ be a measurable space. Define $S^\infty = S \times S \times S \times \ldots$ and let $\mathcal{S}^\infty$ be the Borel subsets of S^∞. Let $X = (X_1, X_2, \ldots)$ be a sequence of random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in S, so $X \in S^\infty$. A mapping $\pi = (\pi_1, \pi_2, \ldots)$ from $\mathbb{N}$ to $\mathbb{N}$ is called a finite permutation if π is one-to-one and $\pi_n = n$ for all n sufficiently large. Let Π be the set of all finite permutations. For $\pi \in \Pi$ define $\pi X = (X_{\pi_1}, X_{\pi_2}, \ldots)$.

Then

$$\mathcal{E} = \left\{ X^{-1}(\mathcal{B}) : \mathcal{B} \in \mathcal{S}^{\infty}, \mathbb{P}\left(X^{-1}(\mathcal{B}) \triangle (\pi X)^{-1}(\mathcal{B}) \right) = 0 \text{ for all } \pi \in \Pi \right\}$$

is called the exchangeable σ-algebra of X, where $\triangle$ is the symmetric difference. In words, events in $\mathcal{E}$ are those events which are invariant under all finite permutations of the X_i. Now we state the Hewitt-Savage zero-one law as in [18]. For the proof see [18, p.238].

Theorem 6.3.1 (Hewitt-Savage zero-one law)**.** *Let $X_1, X_2, \ldots$ be i.i.d. random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in a measurable space $(S, \mathcal{S})$, and let $\mathcal{E}$ denote the exchangeable σ-algebra. Then $\mathcal{E}$ is trivial, i.e., if $A \in \mathcal{E}$ then $\mathbb{P}(A) \in \{0, 1\}$.*

In the special case of spatially homogeneous random walk, one can apply the Hewitt-Savage zero-one law to prove the dichotomy of the recurrence classification of the walk as the the event $R = \{S_n = 0 \text{ infinitely often}\}$ is invariant under finite permutations of the increments, i.e., exchangeable. Notice that the Kolmogorov's zero-one law cannot directly apply here as R is not a tail event.

6.4 Local limit theorem for random walks

It is always good to start with a similar or classical result in random walk theory and try to compare them with our centre of mass process. Sometimes, it is possible to extract useful bits in the classical proofs of similar theorems. Towards our goal to the local limit theorem of the centre of mass process, we should first establish a local limit theorem for our random walk. Here is the formal setup.

Throughout we will use the notation

$$\boldsymbol{\mu} := \mathbb{E}\, X, \quad M := \mathbb{E}[(X - \boldsymbol{\mu})(X - \boldsymbol{\mu})^{\top}]$$

whenever the expectations exist; when defined, M is a symmetric d by d matrix.

To go further we typically assume the following.

(M) Suppose that $\mathbb{E}[\|X\|^2] < \infty$ and M is positive-definite.

Notice that $\mathbb{P}(X \in \mathbf{b} + H\mathbb{Z}^d) = 1$ implies $\mathbb{P}(S_n \in n\mathbf{b} + H\mathbb{Z}^d) = 1$. For $\mathbf{x} \in \mathbb{R}^d$, define $q_n(\mathbf{x}) := \mathbb{P}(n^{-1/2}S_n = \mathbf{x})$, and

$$m(\mathbf{x}) := \frac{\exp\{-\frac{1}{2}\mathbf{x}^\top M^{-1}\mathbf{x}\}}{(2\pi)^{d/2}\sqrt{\det M}}. \tag{6.4.1}$$

Also define

$$L_n := \left\{n^{-1/2}\left(n\mathbf{b} + H\mathbb{Z}^d\right)\right\}.$$

Here is our local limit theorem for the random walk.

Theorem 6.4.1. *Suppose that* (L) *hold and suppose* $\mathbb{E}[\|X\|^2] < \infty$ *and* M *is positive-definite. Then we have*

$$\lim_{n\to\infty} \sup_{\mathbf{x}\in L_n} \left| \frac{n^{d/2}}{h} q_n(\mathbf{x}) - m\left(\mathbf{x} - n^{1/2}\boldsymbol{\mu}\right)\right| = 0. \tag{6.4.2}$$

Note that $\mathbb{P}(S_n \in \mathcal{B}) = O(n^{-d/2})$ for a fixed ball $\mathcal{B}$ since $n^{-1/2}\mathcal{B}$ contains $O(1)$ lattice points.

This rest of this section is devoted to the proof of Theorem 6.4.1.

Proof of Theorem 6.4.1. With standard Fourier analysis, we know Theorem 6.4.1 in the case where $\mathbf{b} = \mathbf{0}$ and $H = I$ (the identity), see e.g. [67, §2.2–§2.3] for details. The first proof of the one-dimensional result can be trace back to [42].

Now it remains to show that it suffices to establish Theorem 6.4.1 in this special case. To see this, suppose that $X \in \mathbf{b} + H\mathbb{Z}^d$ and set $\tilde{X} = H^{-1}(X - \mathbf{b})$. Then $\tilde{X} \in \mathbb{Z}^d$. By linearity of expectation, we have

$$\tilde{\boldsymbol{\mu}} := \mathbb{E}\,\tilde{X} = H^{-1}(\mu - \mathbf{b}), \text{ and } \tilde{M} := \mathbb{E}[(\tilde{X} - \tilde{\boldsymbol{\mu}})(\tilde{X} - \tilde{\boldsymbol{\mu}})^\top] = H^{-1}M(H^{-1})^\top.$$

Note that $(H^{-1})^\top$ is nonsingular, so $(H^{-1})^\top\mathbf{x} \neq \mathbf{0}$ for all $\mathbf{x} \neq \mathbf{0}$. Hence for $\mathbf{x} \neq \mathbf{0}$, $\mathbf{x}^\top\tilde{M}\mathbf{x} = \mathbf{y}^\top M\mathbf{y}$ where $\mathbf{y} = (H^{-1})^\top\mathbf{x} \neq \mathbf{0}$, so that since M is positive definite we have $\mathbf{x}^\top\tilde{M}\mathbf{x} > 0$; hence $\tilde{M}$ is also positive definite. We also have $\tilde{S}_n := \sum_{i=1}^n \tilde{X}_i = H^{-1}(S_n - n\mathbf{b})$. The assumption that $H\mathbb{Z}^d$ is minimal for X implies that $\mathbb{Z}^d$ is minimal for $\tilde{X}$. Thus the process defined by $\tilde{X}$ satisfies the hypotheses of Theorem 6.4.1 in the case where $\mathbf{b} = \mathbf{0}$ and $H = I$, with mean $\tilde{\boldsymbol{\mu}}$ and covariance $\tilde{M}$, and that result

yields

$$\lim_{n\to\infty} \sup_{\mathbf{x}\in n^{-1/2}\mathbb{Z}^d} \left| n^{d/2}\mathbb{P}(n^{-1/2}\tilde{S}_n = \mathbf{x}) - \tilde{m}\left(\mathbf{x} - n^{1/2}\tilde{\boldsymbol{\mu}}\right)\right| = 0, \tag{6.4.3}$$

where

$$\tilde{m}(\mathbf{z}) := \frac{(\det \tilde{M})^{-1/2}}{(2\pi)^{d/2}} \exp\left\{-\frac{1}{2}\mathbf{z}^\top \tilde{M}^{-1}\mathbf{z}\right\}.$$

But

$$\mathbb{P}(n^{-1/2}\tilde{S}_n = \mathbf{x}) = \mathbb{P}\left(n^{-1/2}S_n = n^{1/2}\mathbf{b} + H\mathbf{x}\right) = \mathbb{P}(n^{-1/2}S_n = \mathbf{y})$$

where $\mathbf{y} = n^{1/2}\mathbf{b} + H\mathbf{x}$ so $\mathbf{y} \in n^{-1/2}(n\mathbf{b} + H\mathbb{Z}^d)$. Also,

$$\begin{aligned}\mathbf{x} - n^{1/2}\tilde{\boldsymbol{\mu}} &= \left(H^{-1}\mathbf{y} - n^{1/2}H^{-1}\mathbf{b}\right) - n^{1/2}H^{-1}(\boldsymbol{\mu} - \mathbf{b}) \\ &= H^{-1}\mathbf{y} - n^{1/2}H^{-1}\boldsymbol{\mu}.\end{aligned}$$

Hence, since $\tilde{M}^{-1} = H^\top M^{-1}H$ and $\det \tilde{M} = h^{-2}\det M$,

$$\begin{aligned}\tilde{m}\left(\mathbf{x} - n^{1/2}\tilde{\boldsymbol{\mu}}\right) &= \frac{(\det \tilde{M})^{-1/2}}{(2\pi)^{d/2}} \exp\left\{-\frac{1}{2}\left(\mathbf{y} - n^{1/2}\boldsymbol{\mu}\right)^\top M^{-1}\left(\mathbf{y} - n^{1/2}\boldsymbol{\mu}\right)\right\} \\ &= hm\left(\mathbf{y} - n^{1/2}\boldsymbol{\mu}\right).\end{aligned}$$

It follows that (6.4.3) is equivalent to

$$\lim_{n\to\infty} \sup_{\mathbf{y}\in n^{-1/2}(n\mathbf{b}+H\mathbb{Z}^d)} \left|\frac{n^{d/2}}{h}\mathbb{P}(n^{-1/2}S_n = \mathbf{y}) - m\left(\mathbf{y} - n^{1/2}\boldsymbol{\mu}\right)\right| = 0,$$

which is the general statement of Theorem 6.4.1. □

6.5 Stable distributions and domains of attraction

In this section, we will recall some basic theory on stable distributions, which will be used in the one dimensional case of the center of mass process. Definitions, proofs of theorems and more background material can be found in [12] and [100].

Formally, a distribution is defined to be stable if any linear combination of two independent random variables with this distribution has the same distribution, up to a shift and rescaling. A random variable is defined to be stable if its distribution

is stable.

We can characterize this type of function using the characteristic function of the distribution, as stated as the next theorem, a rewritten version of [80, Def. 1.6].

Theorem 6.5.1. *A distribution is stable if and only if the logarithm of its characteristic function is of the form*

$$\log(\phi(t)) = i\gamma t - c|t|^\alpha \left(1 + i\beta \frac{t}{|t|}\omega(t,\alpha)\right)$$

where $\gamma \in \mathbb{R}$, $\alpha \in (0,2]$, $\beta \in [-1,1]$, *and*

$$\omega(t,\alpha) = \begin{cases} \tan\left(\frac{\pi}{2}\alpha\right) & \text{if } \alpha \neq 1 \\ \frac{2}{\pi}\log(|t|) & \text{if } \alpha = 1 \end{cases}$$

In the case of $\alpha < 2$, the formula can express in the following way.

$$\begin{aligned} \log(\phi(t)) = i\gamma t &+ c_1 \int_{-\infty}^{0} \left(e^{itu} - 1 - \frac{itu}{1+u^2}\right) \frac{du}{|u|^{1+\alpha}} \\ &+ c_2 \int_{0}^{\infty} \left(e^{itu} - 1 - \frac{itu}{1+u^2}\right) \frac{du}{u^{1+\alpha}} \end{aligned} \tag{6.5.1}$$

where c_1 and c_2 are some positive constants and, again, $\gamma \in \mathbb{R}$.

The case that $\alpha = 2$ is the classical Gaussian distribution, which we can write down the close form of the distribution. We can also find closed forms for the density functions in the cases $\alpha = 1$ (the Cauchy distribution) and $\alpha = 0.5$ (the Lévy distribution). However, there are no closed form in general, meaning that working with the characteristic functions is necessary. For the purpose of this thesis, we will only consider the symmetric case, i.e. $\gamma = \beta = 0$, when $\phi(t) = e^{-c|t|^\alpha}$, $\alpha \in (0,2)$.

We also recall the definition of the domains of attraction as follows.

Definition 6.5.2. A random variable X belongs to the domain of attraction of a stable law G iff there exist $a_n > 0, b_n \in \mathbb{R}$ such that

$$\frac{S_n - b_n}{a_n} \xrightarrow{d} G,$$

where S_n denotes the nth partial sum, $S_n = \sum_{k=1}^n X_k$, and X_k are i.i.d copies of X. We write $X \in \mathcal{D}(G)$, or, in terms of the distribution functions, $F_X \in \mathcal{D}(F_G)$.

The stable distribution is closely related to domain of attraction in the following way.

Theorem 6.5.3. *A distribution or random variable has a domain of attraction if and only if it is stable.*

We will omit the proof here. For a deeper analysis of domains of attraction, and a characterisation of the distributions within them, see, for example, Chapter XVII of [33].

All stable distributions are, by definition, in their own domains of attraction. Intuitively, by considering the distribution of the later terms of the sequence, we expect the limiting distribution to exhibit the same "self-similar" behaviour we required of stable distributions above.

Theorem 6.5.4. *If ξ is in the domain of attraction of a stable distribution with index $\alpha \in (0, 2)$, then*

$$\begin{aligned} \mathbb{E}[|\xi|^r] &< \infty \text{ if } r < \alpha \\ &= \infty \text{ if } r > \alpha. \end{aligned}$$

For $r = \alpha$ either case is possible.

This implies that for $\alpha < 1$, ξ has no mean, and that for $\alpha < 2$, ξ has no variance. The Normal ($\alpha = 2$) distribution is the only stable distribution with finite variance.

Now we should focus on a slightly narrower interest; the domain of *normal* attraction, in which we specify the form of the constants a_n.

Definition 6.5.5. A random variable X belongs to the domain of normal attraction of a stable law G iff there exist $a > 0$ and $b_n \in \mathbb{R}$ such that

$$\frac{S_n - b_n}{an^{\frac{1}{\alpha}}} \xrightarrow{d} G, \tag{6.5.2}$$

where S_n denotes the nth partial sum, $S_n = \sum_{k=1}^{n} X_k$.

We have set $a_n = an^{\frac{1}{\alpha}}$. The Convergence to Types theorem in [12] allows us to assume without loss of generality that $a = 1$.

Chapter 7

Main results

7.1 Law of large numbers and central limit theorem for centre of mass

In this first section we would like to establish some standard properties and theorems in random walk theory for our centre of mass process.

Recall that we will use the notation

$$\boldsymbol{\mu} := \mathbb{E}\, X, \quad M := \mathbb{E}[(X - \boldsymbol{\mu})(X - \boldsymbol{\mu})^\top]$$

whenever the expectations exist; when defined, M is a symmetric d by d matrix.

Suppose $X, X_1, X_2, \ldots$ is a sequence of i.i.d. random variables on $\mathbb{R}^d$. The strong law of large numbers for S_n yields the following strong law for G_n.

Proposition 7.1.1. *Suppose that* $\mathbb{E}\,\|X\| < \infty$. *Then* $n^{-1}G_n \to \frac{1}{2}\boldsymbol{\mu}$, *a.s., as* $n \to \infty$.

Comparing with the strong law of large numbers for S_n, i.e., Theorem 1.6.1, we see G_n behaves very similarly to S_n, up to a constant factor of $\frac{1}{2}$.

Proof of Proposition 7.1.1. By the strong law for S_n, we have that for any $\varepsilon > 0$ there exists N_ε with $\mathbb{P}(N_\varepsilon < \infty) = 1$ such that $\|S_n - n\boldsymbol{\mu}\| \leq n\varepsilon$ for all $n \geq N_\varepsilon$. Then, by the triangle inequality,

$$\|G_n - (n+1)(\boldsymbol{\mu}/2)\| = \frac{1}{n}\left\|\sum_{i=1}^{n}(S_i - i\boldsymbol{\mu})\right\|$$

$$\leq \frac{1}{n}\sum_{i=1}^{N_\varepsilon} \|S_i - i\boldsymbol{\mu}\| + \frac{1}{n}\sum_{i=N_\varepsilon}^{n} \|S_i - i\boldsymbol{\mu}\|$$
$$\leq \frac{1}{n}\sum_{i=1}^{N_\varepsilon} \|S_i - i\boldsymbol{\mu}\| + \frac{1}{n}\sum_{i=1}^{n} i\varepsilon.$$

It follows that

$$\limsup_{n\to\infty} n^{-1}\,\|G_n - (n+1)(\boldsymbol{\mu}/2)\| \leq \varepsilon/2,$$

and since $\varepsilon > 0$ was arbitrary we get the result. □

To go further we typically assume the condition (M) holds. Note that

$$G_n = \sum_{i=1}^{n} \left(\frac{n-i+1}{n}\right) X_i. \tag{7.1.1}$$

The representation (7.1.1) leads via the Lindeberg–Feller theorem for triangular arrays to the following central limit theorem; we write '$\xrightarrow{d}$' for convergence in distribution, and $\mathcal{N}_d(\mathbf{m}, \Sigma)$ for a d-dimensional normal random variable with mean $\mathbf{m}$ and covariance Σ.

Proposition 7.1.2. *If* (M) *holds, then, as* $n \to \infty$,

$$n^{-1/2}\left(G_n - \frac{n}{2}\boldsymbol{\mu}\right) \xrightarrow{d} \mathcal{N}_d(\mathbf{0}, M/3).$$

Comparing with the central limit theorem for S_n, i.e., Theorem 1.6.2, we see G_n behaves very similarly to S_n except the variance in the normal distribution is one third of the corresponding variance for S_n.

Proof of Proposition 7.1.2. For any unit vector $\mathbf{e} \in \mathbb{S}^{d-1}$, $\mathbf{e}\cdot G_n$ is the centre-of-mass associated with the one-dimensional random walk with increments $\mathbf{e}\cdot X_i$; thus, by the Cramer–Wold device (see e.g [27, Theorem 3.9.5]), it suffices to establish the central limit theorem for $d = 1$.

So take $d = 1$ and write $\boldsymbol{\mu} = \mu$, $M = \sigma^2 \in (0, \infty)$. It follows from (7.1.1) that for fixed n, G_n has the same distribution as

$$G'_n := \sum_{i=1}^{n} \left(\frac{i}{n}\right) X_i.$$

It thus suffices to show that $n^{-1/2}(G'_n - \frac{n}{2}\mu)$ converges in distribution to $\mathcal{N}(0, \sigma^2/3)$.

We show that this follows from [14, Corollary 8.4.1]. Define $T_{n,i} := \frac{i}{n^{3/2}}(X_i - \mu)$. Then

$$\sum_{i=1}^{n} \mathbb{V}\mathrm{ar}(T_{n,i}) = \sum_{i=1}^{n} \frac{i^2}{n^3}\sigma^2 \to \frac{\sigma^2}{3}.$$

It remains to verify the Lindeberg condition for triangular arrays: for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \sum_{i=1}^{n} \mathbb{E}\left[T_{n,i}^2 \mathbf{1}\{|T_{n,i}| > \varepsilon\}\right] = 0.$$

But we have that

$$\begin{aligned}\sum_{i=1}^{n} \mathbb{E}\left[T_{n,i}^2 \mathbf{1}\{|T_{n,i}| > \varepsilon\}\right] &\le \sum_{i=1}^{n} \mathbb{E}\left[T_{n,n}^2 \mathbf{1}\{|T_{n,n}| > \varepsilon\}\right] \\ &= \sum_{i=1}^{n} \frac{1}{n} \mathbb{E}\left[(X_n - \mu)^2 \mathbf{1}\{|X_n - \mu| > \varepsilon\sqrt{n}\}\right] \\ &= \mathbb{E}\left[(X - \mu)^2 \mathbf{1}\{|X - \mu| > \varepsilon\sqrt{n}\}\right].\end{aligned}$$

Now $(X - \mu)^2 \mathbf{1}\{|X - \mu| > \varepsilon\sqrt{n}\} \to 0$ a.s. as $n \to \infty$ and $|(X - \mu)^2 \mathbf{1}\{|X - \mu| > \varepsilon\sqrt{n}\}| \le (X - \mu)^2$ which has $\mathbb{E}[(X - \mu)^2] < \infty$. Thus the dominated convergence theorem yields $\mathbb{E}[(X - \mu)^2 \mathbf{1}\{|X - \mu| > \varepsilon\sqrt{n}\}] \to 0$ as $n \to \infty$ and the Lindeberg condition is verified. □

7.2 Local limit theorem for centre of mass

To get the recurrence classification of the centre of mass process, only knowing the law of large numbers and central limit theorem is not enough to control the trajectory of the process. We need a more precise result. Our first main result is a *local* central limit theorem.

Notice that $\mathbb{P}(X \in \mathbf{b} + H\mathbb{Z}^d) = 1$ implies $\mathbb{P}(S_n \in n\mathbf{b} + H\mathbb{Z}^d) = 1$ which again implies $\mathbb{P}(G_n \in n^{-1}(\frac{1}{2}n(n+1)\mathbf{b} + H\mathbb{Z}^d) = 1$. For $\mathbf{x} \in \mathbb{R}^d$, define

$$p_n(\mathbf{x}) := \mathbb{P}(n^{-1/2} G_n = \mathbf{x}),$$

and

$$n(\mathbf{x}) := \frac{\exp\{-\frac{3}{2}\mathbf{x}^\top M^{-1}\mathbf{x}\}}{(2\pi)^{d/2}\sqrt{\det(M/3)}}. \tag{7.2.1}$$

Also define

$$\mathcal{L}_n := \left\{ n^{-3/2} \left(\tfrac{1}{2} n(n+1)\mathbf{b} + H\mathbb{Z}^d \right) \right\}.$$

Here is our local limit theorem.

Theorem 7.2.1. *Suppose that* (L) *and* (M) *hold. Then we have*

$$\lim_{n\to\infty} \sup_{\mathbf{x}\in\mathcal{L}_n} \left| \frac{n^{3d/2}}{h} p_n(\mathbf{x}) - n \left(\mathbf{x} - \frac{(n+1)}{2n^{1/2}} \boldsymbol{\mu} \right) \right| = 0. \tag{7.2.2}$$

Comparing to Theorem 6.4.1, the local limit theorem for the random walk, G_n has a different scaling factor and a different shift in the normal distribution. The formal difference is one of the key for the difference in recurrence classification of S_n and G_n, we will see more in Section 8.3.

Remarks 7.2.2. (i) In the case $d = 1$, versions of Theorem 7.2.1 are given in [78, Lemma 4.3] and in [46, Lemma 1]; the latter result deals only with the special case of SSRW and only bounds $p_n(\mathbf{x})$ up to constant factors. See Chapter 9 for a demonstration that our assumptions are indeed satisfied by SSRW on $\mathbb{Z}^d$ for appropriate choice of H with $h = 2$. The proof in [78] is only a sketch, and the claim that "it is enough to apply the usual analytical methods" [78, p. 515] does not quite tell the whole story, even in the one-dimensional case. Both [46, 78] also give bivariate local limit theorems for (S_n, G_n) (in the case $d = 1$). A related result is [23, Theorem 4.2].

(ii) If $Z_n := S_n - G_n$, then note that $Z_{n+1} = \frac{n}{n+1} \sum_{i=1}^{n} (i/n) X_{i+1}$, which means that $Z_{n+1} \stackrel{d}{=} \frac{n}{n+1} G_n$, where '$\stackrel{d}{=}$' stands for equality in distribution. Thus Theorem 7.2.1 also yields a local limit theorem for Z_n. However, the *processes* Z_n and G_n may behave very differently: see [20, Remark 1.1].

7.3 Transience and diffusive rate of escape in two or higher dimensions

Now we are ready for the recurrence classification for the centre of mass process and we now turn to the almost-sure asymptotic behaviour of G_n. We have the following transience result in dimensions greater than one. In particular, Theorem 7.3.1 says

that $\lim_{n\to\infty} \|G_n\| = +\infty$, a.s., and gives a diffusive rate of escape; in the case of SSRW the result is due to Grill [46, Theorem 1].

Theorem 7.3.1. *Suppose that $d \geq 2$ and that* (L) *and* (M) *hold, and that $\boldsymbol{\mu} = \mathbf{0}$. Then*

$$\lim_{n\to\infty} \frac{\log \|G_n\|}{\log n} = \frac{1}{2}, \text{ a.s.}$$

As said at the start of this part, it is interesting to investigate why the walk itself is recurrent while the centre of mass process is transient in two dimensions. The real reason behind this is because the centre of mass travels much slower than the original process, with steps of the order of $O(n^{-1/2})$ comparing to $O(1)$ respectively. This suggests that the former process is too slow to return to a specific region in a short amount of time. This idea will be formalized and in fact greatly contribute to the last part of our proof of the recurrence classification.

Obtaining necessary and sufficient conditions for recurrence and transience of G_n is an open problem. See Section 10.2 for details.

7.4 One dimension

Next we have a recurrence result in one dimension; in the case of SSRW the fact that G_n returns i.o. (infinitely often) to a neighbourhood of the origin is due to Grill [46, Theorem 1].

Theorem 7.4.1. *Suppose that $d = 1$ and that either of the following two conditions holds.*

(i) Suppose that $\mathbb{E}\,|X| \in (0, \infty)$ and $X \stackrel{d}{=} -X$.

(ii) Suppose that (M) *holds and that $\mathbb{E}\,X = 0$.*

Then $\liminf_{n\to\infty} G_n = -\infty$, $\limsup_{n\to\infty} G_n = +\infty$, and $\liminf_{n\to\infty} |G_n - x| = 0$ for any $x \in \mathbb{R}$.

In contrast to Theorem 7.4.1, we will show that in the case where $\mathbb{E}\,|X| = \infty$, G_n may be transient. The condition we assume is as follows.

(S) Suppose that $X \stackrel{d}{=} -X$ and X is in the domain of normal attraction of a symmetric α-stable distribution with $\alpha \in (0,1)$.

Theorem 7.4.2. *Suppose that $d = 1$ and the lattice condition* (L) *holds, i.e., $\mathbb{P}(X \in b + h\mathbb{Z}) = 1$ for $b \in \mathbb{R}$ and maximal $h > 0$. Suppose also that* (S) *holds. Then $\liminf_{n\to\infty} G_n = -\infty$, $\limsup_{n\to\infty} G_n = +\infty$, and $\lim_{n\to\infty} |G_n| = \infty$.*

Remark 7.4.3. The transience here fails in the natural continuous time version of this model. The analogous continuum model, a symmetric α-stable Lévy process for $\alpha \in (0,1)$, s_t, has centre of mass $g_t = \frac{1}{t}\int_0^t s_u \mathrm{d}u$, and it is surely true that g_t again changes sign i.o., but in this case continuity of g_t implies that $g_t = 0$ i.o.

Chapter 9 verifies our main assumptions for a couple of simple examples. The proof of Theorem 7.2.1 is given in Section 8.2. The proof of Theorem 7.4.1 uses Proposition 7.1.2, some observations following from the Hewitt–Savage zero–one law, and the fact that in the case where $\mathbb{E}\,X = 0$ oscillating behaviour is sufficient for $\liminf_{n\to\infty} |G_n - x| = 0$: see Section 7.4 and Section 8.4. The proof of Theorem 7.4.2 uses another local limit theorem (Theorem 8.5.1) and is also presented in Section 8.4. The proof of Theorem 7.3.1 relies on Theorem 7.2.1: see Section 8.2. Section 8.1 collects auxiliary results on characteristic functions that we need for the proofs of our local limit theorems.

Chapter 8

Proofs and technical details

In this chapter, we will provide the proofs of all the theorems stated in the last chapter. Our main goal is to prove Theorem 7.2.1, Theorem 7.3.1, Theorem 7.4.1 and Theorem 7.4.2.

8.1 A characteristic function result

We will use the following characteristic function estimation based on Taylor expansion a few times throughout our proof.

Lemma 8.1.1. *Suppose that* $\mathbb{E}[\|X\|^2] < \infty$. *For any* $\mathbf{t} \in \mathbb{R}^d$,

$$\varphi(\mathbf{t}) = 1 + i\mathbf{t}^\top \mathbb{E}\, X - \frac{1}{2}\mathbf{t}^\top \mathbb{E}[XX^\top]\mathbf{t} + \|\mathbf{t}\|^2 W(\mathbf{t}), \tag{8.1.1}$$

where for any $\varepsilon > 0$, *there exists* $\delta > 0$ *such that* $|W(\mathbf{t})| \le \varepsilon$ *for all* $\mathbf{t}$ *with* $\|\mathbf{t}\| \le \delta$.

Proof. Applying [27, Lemma 3.3.7] with $x = \mathbf{t}^\top X$, we get that if $\mathbb{E}[\|X\|^n] < \infty$, then

$$\begin{aligned}\left|\mathbb{E}\, \mathrm{e}^{i\mathbf{t}^\top X} - \sum_{m=0}^{n} \mathbb{E}\, \frac{(i\mathbf{t}^\top X)^m}{m!}\right| &\le \mathbb{E}\left|\mathrm{e}^{i\mathbf{t}^\top X} - \sum_{m=0}^{n} \frac{(i\mathbf{t}^\top X)^m}{m!}\right| \\ &\le \mathbb{E} \min\left(\frac{|\mathbf{t}^\top X|^{n+1}}{(n+1)!}, \frac{2|\mathbf{t}^\top X|^n}{n!}\right).\end{aligned}$$

Taking $n = 2$ and rearranging, we get equation (8.1.1), and $|W(\mathbf{t})| \le \mathbb{E}\, Z(t)$, where $Z(\mathbf{t}) = \min\{\|\mathbf{t}\|\|X\|^3, \|X\|^2\}$. Now $|Z(\mathbf{t})| \le \|X\|^2$ and $\mathbb{E}[\|X\|^2] < \infty$. Also we have

$|Z(\mathbf{t})| \leq \|\mathbf{t}\|\|X\|^3 \to 0$ a.s. as $\|\mathbf{t}\| \to 0$. So the dominated convergence theorem implies that $\mathbb{E}\, Z(\mathbf{t}) \to 0$ as $\|\mathbf{t}\| \to 0$. □

8.2 Proof of the local limit theorem

This section is devoted to the proof of Theorem 7.2.1. The outline of the proof mirrors the standard Fourier-analytic proof of the local central limit theorem for the random walk: compare e.g. [42], [44, Ch. 9], [27, §3.5], or [56, Ch. 4] for the one-dimensional case, and [67, §2.2–§2.3] for the case of walks on $\mathbb{Z}^d$. The details of the proof require some extra effort, however.

First we show that it suffices to establish Theorem 7.2.1 in the case where $\mathbf{b} = \mathbf{0}$ and $H = I$ (the identity). To see this, suppose that $X \in \mathbf{b} + H\mathbb{Z}^d$ and set $\tilde{X} = H^{-1}(X - \mathbf{b})$. Then $\tilde{X} \in \mathbb{Z}^d$. By linearity of expectation, we have

$$\tilde{\boldsymbol{\mu}} := \mathbb{E}\, \tilde{X} = H^{-1}(\mu - \mathbf{b}), \text{ and } \tilde{M} := \mathbb{E}[(\tilde{X} - \tilde{\boldsymbol{\mu}})(\tilde{X} - \tilde{\boldsymbol{\mu}})^\top] = H^{-1}M(H^{-1})^\top.$$

Note that $(H^{-1})^\top$ is nonsingular, so $(H^{-1})^\top\mathbf{x} \neq \mathbf{0}$ for all $\mathbf{x} \neq \mathbf{0}$. Hence for $\mathbf{x} \neq \mathbf{0}$, $\mathbf{x}^\top\tilde{M}\mathbf{x} = \mathbf{y}^\top M\mathbf{y}$ where $\mathbf{y} = (H^{-1})^\top\mathbf{x} \neq \mathbf{0}$, so that since M is positive definite we have $\mathbf{x}^\top\tilde{M}\mathbf{x} > 0$; hence $\tilde{M}$ is also positive definite. Also, $\tilde{S}_n := \sum_{i=1}^n \tilde{X}_i = H^{-1}(S_n - n\mathbf{b})$ and $\tilde{G}_n := n^{-1}\sum_{i=1}^n \tilde{S}_i = H^{-1}(G_n - \frac{n+1}{2}\mathbf{b})$. The assumption that $H\mathbb{Z}^d$ is minimal for X implies that $\mathbb{Z}^d$ is minimal for $\tilde{X}$. Thus the process defined by $\tilde{X}$ satisfies the hypotheses of Theorem 7.2.1 in the case where $\mathbf{b} = \mathbf{0}$ and $H = I$, with mean $\tilde{\boldsymbol{\mu}}$ and covariance $\tilde{M}$, and that result yields

$$\lim_{n\to\infty} \sup_{\mathbf{x}\in n^{-3/2}\mathbb{Z}^d} \left| n^{3d/2}\mathbb{P}(n^{-1/2}\tilde{G}_n = \mathbf{x}) - \tilde{n}\left(\mathbf{x} - \frac{(n+1)}{2n^{1/2}}\tilde{\boldsymbol{\mu}}\right)\right| = 0, \tag{8.2.1}$$

where

$$\tilde{n}(\mathbf{z}) := \frac{(\det \tilde{M}/3)^{-1/2}}{(2\pi)^{d/2}} \exp\left\{-\frac{3}{2}\mathbf{z}^\top\tilde{M}^{-1}\mathbf{z}\right\}.$$

But

$$\mathbb{P}(n^{-1/2}\tilde{G}_n = \mathbf{x}) = \mathbb{P}\left(n^{-1/2}G_n = \frac{(n+1)}{2n^{1/2}}\mathbf{b} + H\mathbf{x}\right) = \mathbb{P}(n^{-1/2}G_n = \mathbf{y})$$

where $\mathbf{y} = \frac{(n+1)}{2n^{1/2}}\mathbf{b} + H\mathbf{x}$ so $\mathbf{y} \in n^{-3/2}(\frac{1}{2}n(n+1)\mathbf{b} + H\mathbb{Z}^d)$. Also,

$$\begin{aligned}\mathbf{x} - \frac{(n+1)}{2n^{1/2}}\tilde{\boldsymbol{\mu}} &= \left(H^{-1}\mathbf{y} - \frac{(n+1)}{2n^{1/2}}H^{-1}\mathbf{b}\right) - \frac{(n+1)}{2n^{1/2}}H^{-1}(\boldsymbol{\mu} - \mathbf{b}) \\ &= H^{-1}\mathbf{y} - \frac{(n+1)}{2n^{1/2}}H^{-1}\boldsymbol{\mu}.\end{aligned}$$

Hence, since $\tilde{M}^{-1} = H^\top M^{-1} H$ and $\det \tilde{M} = h^{-2}\det M$,

$$\begin{aligned}&\tilde{n}\left(\mathbf{x} - \frac{(n+1)}{2n^{1/2}}\tilde{\boldsymbol{\mu}}\right) \\ &= \frac{(\det \tilde{M}/3)^{-1/2}}{(2\pi)^{d/2}} \exp\left\{-\frac{3}{2}\left(\mathbf{y} - \frac{(n+1)}{2n^{1/2}}\boldsymbol{\mu}\right)^\top M^{-1}\left(\mathbf{y} - \frac{(n+1)}{2n^{1/2}}\boldsymbol{\mu}\right)\right\} \\ &= hn\left(\mathbf{y} - \frac{(n+1)}{2n^{1/2}}\boldsymbol{\mu}\right).\end{aligned}$$

It follows that (8.2.1) is equivalent to

$$\lim_{n\to\infty} \sup_{\mathbf{y}\in n^{-3/2}(\frac{1}{2}n(n+1)\mathbf{b}+H\mathbb{Z}^d)} \left|\frac{n^{3d/2}}{h}\mathbb{P}(n^{-1/2}G_n = \mathbf{y}) - n\left(\mathbf{y} - \frac{(n+1)}{2n^{1/2}}\boldsymbol{\mu}\right)\right| = 0,$$

which is the general statement of Theorem 7.2.1. Thus for the remainder of this section we suppose that $\mathbf{b} = \mathbf{0}$ and $H = I$; hence $\mathcal{L}_n = n^{-3/2}\mathbb{Z}^d$.

8.2.1 Integral estimates

After the reduction, the next step of the proof will focus on the estimation of the characteristic function of the centre of mass process. This can done by delicate Fourier-type analysis as follows.

Let $Y_n := \sum_{i=1}^n S_i$ and thus $G_n = Y_n/n$. Recall that φ denotes the characteristic function (ch.f.) of X, and let Φ_n be the ch.f. of $n^{-3/2}Y_n$, i.e., for $\mathbf{t} \in \mathbb{R}^d$,

$$\varphi(\mathbf{t}) := \mathbb{E}\,\mathrm{e}^{i\mathbf{t}^\top X} \quad \text{and} \quad \Phi_n(\mathbf{t}) := \mathbb{E}\,\mathrm{e}^{in^{-3/2}\mathbf{t}^\top Y_n}.$$

Denoting the smallest eigenvalue of M by $\lambda_{\min}(M)$ we have that

$$\inf_{\mathbf{t}\neq\mathbf{0}} \hat{\mathbf{t}}^\top M\hat{\mathbf{t}} = \lambda_{\min}(M) > 0, \tag{8.2.2}$$

under assumption (M), where $\hat{\mathbf{t}} := \mathbf{t}/\|\mathbf{t}\|$ for $\mathbf{t} \neq \mathbf{0}$. Define

$$f_n(\mathbf{t}) := \exp\left\{\frac{i(n+1)\mathbf{t}^\top \boldsymbol{\mu}}{2n^{1/2}} - \frac{\mathbf{t}^\top M \mathbf{t}}{6}\right\}. \tag{8.2.3}$$

Our starting point for the proof of the local limit theorem is the following.

Lemma 8.2.1. *Suppose that* (M) *holds and that* $\mathbb{P}(X \in \mathbb{Z}^d) = 1$. *Then*

$$\sup_{\mathbf{x}\in\mathcal{L}_n}\left|n^{3d/2}p_n(\mathbf{x}) - n\left(\mathbf{x} - \frac{(n+1)}{2n^{1/2}}\boldsymbol{\mu}\right)\right| \\ \leq \int_{R(n)} |D_n(\mathbf{t})|\,\mathrm{d}\mathbf{t} + \int_{R^{\mathrm{c}}(n)} \exp\left\{-\frac{\lambda_{\min}(M)}{6}\|\mathbf{t}\|^2\right\}\mathrm{d}\mathbf{t},$$

where $R(n) := [-\pi n^{3/2}, \pi n^{3/2}]^d$, $R^{\mathrm{c}}(n) := \mathbb{R}^d \setminus R(n)$, *and* $D_n(\mathbf{t}) := \Phi_n(\mathbf{t}) - f_n(\mathbf{t})$.

Proof. For a lattice random variable $W \in \mathbb{Z}^d$, by the inversion formula for the characteristic function (see e.g. [67, Corollary 2.2.3, p. 29]) we have that

$$\mathbb{P}(W = \mathbf{y}) = \frac{1}{(2\pi)^d}\int_{[-\pi,\pi]^d} \mathrm{e}^{-i\mathbf{u}^\top\mathbf{y}}\,\mathbb{E}\left[\mathrm{e}^{i\mathbf{u}^\top W}\right]\mathrm{d}\mathbf{u}, \tag{8.2.4}$$

for $\mathbf{y} \in \mathbb{Z}^d$. Now we have for $\mathbf{x} \in \mathcal{L}_n$, $p_n(\mathbf{x}) = \mathbb{P}(Y_n = n^{3/2}\mathbf{x})$, so applying (8.2.4) with $W = Y_n \in \mathbb{Z}^d$, we get for $\mathbf{x} \in \mathcal{L}_n$ that

$$p_n(\mathbf{x}) = \frac{1}{(2\pi)^d}\int_{[-\pi,\pi]^d} \mathrm{e}^{-in^{3/2}\mathbf{u}^\top\mathbf{x}}\,\mathbb{E}\left[\mathrm{e}^{i\mathbf{u}^\top Y_n}\right]\mathrm{d}\mathbf{u}.$$

Using the substitution $\mathbf{u} = n^{-3/2}\mathbf{t}$, we obtain

$$n^{3d/2}p_n(\mathbf{x}) = \frac{1}{(2\pi)^d}\int_{[-\pi n^{3/2},\pi n^{3/2}]^d} \mathrm{e}^{-i\mathbf{t}^\top\mathbf{x}}\Phi_n(\mathbf{t})\mathrm{d}\mathbf{t}. \tag{8.2.5}$$

On the other hand, since the probability density $n(\mathbf{x} - \frac{(n+1)}{2n^{1/2}}\boldsymbol{\mu})$, with $n(\cdot)$ as defined at (7.2.1), corresponds to the ch.f. $f_n(\mathbf{t})$ as defined at (8.2.3), the inversion formula for densities yields

$$n\left(\mathbf{x} - \frac{(n+1)}{2n^{1/2}}\boldsymbol{\mu}\right) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \mathrm{e}^{-i\mathbf{t}^\top\mathbf{x}} f_n(\mathbf{t})\mathrm{d}\mathbf{t}, \tag{8.2.6}$$

for $\mathbf{x} \in \mathbb{R}^d$. Now we subtract (8.2.6) from (8.2.5) to get

$$
\begin{aligned}
&n^{3d/2}p_n(\mathbf{x}) - n\left(\mathbf{x} - \frac{(n+1)}{2n^{1/2}}\boldsymbol{\mu}\right) \\
&= \frac{1}{(2\pi)^d}\int_{R(n)} \mathrm{e}^{-i\mathbf{t}^\top\mathbf{x}} D_n(\mathbf{t})\mathrm{d}\mathbf{t} - \frac{1}{(2\pi)^d}\int_{R^c(n)} \mathrm{e}^{-i\mathbf{t}^\top\mathbf{x}} f_n(\mathbf{t})\mathrm{d}\mathbf{t}.
\end{aligned}
$$

Thus, by the triangle inequality with the estimates $\pi > 1$ and $|\mathrm{e}^{-i\mathbf{t}^\top\mathbf{x}}| \leq 1$, we obtain

$$
\sup_{\mathbf{x}\in\mathcal{L}_n}\left|n^{3d/2}p_n(\mathbf{x}) - n\left(\mathbf{x} - \frac{(n+1)}{2n^{1/2}}\boldsymbol{\mu}\right)\right| \leq \int_{R(n)} |D_n(\mathbf{t})|\,\mathrm{d}\mathbf{t} + \int_{R^c(n)} \exp\left\{-\frac{\mathbf{t}^\top M\mathbf{t}}{6}\right\}\mathrm{d}\mathbf{t},
$$

which with (8.2.2) yields the statement in the lemma. □

To prove Theorem 7.2.1 we must show that the right-hand side of the inequality in Lemma 8.2.1 approaches 0 when $n \to \infty$. To do so, we bound $D_n(\mathbf{t})$ in different regions for $\mathbf{t}$. Observing that $Y_n = \sum_{i=1}^n S_i = \sum_{j=1}^n (n-j+1)X_j$, we see

$$
\Phi_n(\mathbf{t}) = \mathbb{E}\left[\exp\left\{in^{-3/2}\mathbf{t}^\top Y_n\right\}\right] = \mathbb{E}\left[\exp\left\{in^{-3/2}\sum_{j=1}^n (n-j+1)\mathbf{t}^\top X_j\right\}\right].
$$

For fixed n, $\sum_{j=1}^n (n-j+1)\mathbf{t}^\top X_j \stackrel{d}{=} \sum_{j=1}^n j\mathbf{t}^\top X_j$, so that

$$
\Phi_n(\mathbf{t}) = \mathbb{E}\left[\exp\left\{in^{-3/2}\sum_{j=1}^n j\mathbf{t}^\top X_j\right\}\right] = \prod_{j=1}^n \mathbb{E}\left[\exp\left\{in^{-3/2}j\mathbf{t}^\top X_j\right\}\right].
$$

Hence we conclude that for $\mathbf{t} \in \mathbb{R}^d$,

$$
\Phi_n(\mathbf{t}) = \prod_{j=1}^n \varphi(n^{-3/2}j\mathbf{t}). \tag{8.2.7}
$$

To study Φ_n we require certain characteristic function estimates, presented in Section 8.1.

We partition $R(n)$ into four regions defined as follows:

$$
\begin{aligned}
R_1 &:= [-A, A]^d \\
R_2(n) &:= [-\delta\sqrt{n}, \delta\sqrt{n}]^d \setminus R_1 \\
R_3(n) &:= [-\pi\sqrt{n}, \pi\sqrt{n}]^d \setminus (R_1 \cup R_2(n)) \\
R_4(n) &:= R(n) \setminus (R_1 \cup R_2(n) \cup R_3(n))
\end{aligned}
$$

where constants $A \in (0, \infty)$ and $\delta \in (0, \pi)$ will be chosen later. We also denote the corresponding integrals $I_k(n) := \int_{R_k} |D_n(t)|\,\mathrm{d}t$, $k = 1, 2, 3, 4$.

Lemma 8.2.2. *For $\delta > 0$ sufficiently small, the following statements are true.*

(i) *For any $A \in \mathbb{R}_+$, $\lim_{n\to\infty} |I_1(n)| = 0$,*

(ii) $\lim_{A\to\infty} \sup_n |I_2(n)| = 0$,

(iii) $\lim_{n\to\infty} |I_3(n)| = 0$,

(iv) $\lim_{n\to\infty} |I_4(n)| = 0$.

We will combine all the estimates at the end of the argument.

Proof of Lemma 8.2.2. First we aim to show that

$$\lim_{n\to\infty} \sup_{\mathbf{t}\in R_1} |D_n(\mathbf{t})| = 0. \tag{8.2.8}$$

Since $\mathbb{E}\, X = \boldsymbol{\mu}$ and $\mathbb{E}[(X-\boldsymbol{\mu})(X-\boldsymbol{\mu})^\top] = M$, we have $\mathbb{E}[XX^\top] = M + \boldsymbol{\mu}\boldsymbol{\mu}^\top$, so that Lemma 8.1.1 implies, uniformly over $\mathbf{t} \in R_1$, as $n \to \infty$,

$$\prod_{j=1}^n \varphi(n^{-3/2} j\mathbf{t}) = \exp\left\{\sum_{j=1}^n \log\left[1 + A(n,j,\mathbf{t}) + o(n^{-1})\right]\right\},$$

where

$$A(n,j,\mathbf{t}) := in^{-3/2} j\mathbf{t}^\top\boldsymbol{\mu} - \frac{1}{2}n^{-3}j^2\mathbf{t}^\top(M + \boldsymbol{\mu}\boldsymbol{\mu}^\top)\mathbf{t}. \tag{8.2.9}$$

Taylor's theorem for a complex variable shows that for a constant $C < \infty$,

$$\left|\log(1+z) - \left(z - \frac{z^2}{2}\right)\right| \le C|z|^3, \tag{8.2.10}$$

for z in an open disc containing 0. Note from (8.2.9) that

$$A(n,j,\mathbf{t})^2 = -n^{-3}j^2\mathbf{t}^\top\boldsymbol{\mu}\boldsymbol{\mu}^\top\mathbf{t} + \Delta_0(n,j,\mathbf{t}), \tag{8.2.11}$$

where $\max_{1\le j\le n} \sup_{\mathbf{t}\in R_1} |\Delta_0(n,j,\mathbf{t})| = O(n^{-3/2})$. Then, by (8.2.7), (8.2.10), (8.2.11), and the fact that $\max_{1\le j\le n} \sup_{\mathbf{t}\in R_1} |A(n,j,\mathbf{t})| = O(n^{-1/2})$, it follows that

$$\Phi_n(\mathbf{t}) = \exp\left\{\sum_{j=1}^n \left(in^{-3/2}j\mathbf{t}^\top\boldsymbol{\mu} - \frac{1}{2}n^{-3}j^2\mathbf{t}^\top M\mathbf{t}\right) + \Delta_0(n,\mathbf{t})\right\},$$

where $\sup_{\mathbf{t}\in R_1} |\Delta_0(n,\mathbf{t})| \to 0$. Elementary algebra gives $\sum_{j=1}^n j = \frac{1}{2}n(n+1)$ and $\sum_{j=1}^n j^2 = \frac{1}{6}n(n+1)(2n+1)$, so we obtain the estimate

$$\Phi_n(\mathbf{t}) = \exp\left\{\frac{i(n+1)\mathbf{t}^\top \boldsymbol{\mu}}{2n^{1/2}} - \frac{\mathbf{t}^\top M\mathbf{t}}{6} + \Delta_1(n,\mathbf{t})\right\},$$

where $\sup_{\mathbf{t}\in R_1} |\Delta_1(n,\mathbf{t})| \to 0$ as $n \to \infty$. Hence, by (8.2.3),

$$|D_n(\mathbf{t})| = |\Phi_n(\mathbf{t}) - f_n(\mathbf{t})| \le |1 - \exp\{\Delta_1(n,\mathbf{t})\}|,$$

which establishes (8.2.8) and proves part (i) of the lemma. For part (ii), suppose that $\mathbf{t} \in [0, \delta n^{1/2}]^d$. Fix $\varepsilon > 0$. Then for $1 \le j \le n$, we have $\|n^{-3/2} j\mathbf{t}\| \le \delta d^{1/2}$. Thus, from Lemma 8.1.1,

$$\varphi(n^{-3/2} j\mathbf{t}) = 1 + A(n,j,\mathbf{t}) + \Delta_1(n,j,\mathbf{t}),$$

where $A(n,j,\mathbf{t})$ is as defined at (8.2.9), and $|\Delta_1(n,j,\mathbf{t})| \le \varepsilon n^{-1}\|\mathbf{t}\|^2$ for all $\mathbf{t} \in [0, \delta n^{1/2}]^d$ and δ sufficiently small. Also note that $|A(n,j,\mathbf{t})| \le Cn^{-1/2}\|\mathbf{t}\|$, so that

$$|A(n,j,\mathbf{t})|^3 \le Cn^{-3/2}\|\mathbf{t}\|^3 \le C'\delta n^{-1}\|\mathbf{t}\|^2 \le \varepsilon n^{-1}\|\mathbf{t}\|^2, \tag{8.2.12}$$

for δ sufficiently small; here C and C' are constants that do not depend on δ. Thus we may apply (8.2.10) to obtain

$$\begin{aligned}\prod_{j=1}^n \varphi(n^{-3/2} j\mathbf{t}) &= \exp\left\{\sum_{j=1}^n \log\left[1 + A(n,j,\mathbf{t}) + \Delta_1(n,j,\mathbf{t})\right]\right\} \\ &= \exp\left\{\sum_{j=1}^n \left(A(n,j,\mathbf{t}) - \frac{1}{2}A(n,j,\mathbf{t})^2\right) + \Delta_1(n,\mathbf{t})\right\},\end{aligned}$$

where $|\Delta_1(n,\mathbf{t})| \le \varepsilon\|\mathbf{t}\|^2$ for δ sufficiently small. Here (8.2.11) holds, where now, for all $\mathbf{t} \in [0, \delta n^{1/2}]^d$, similarly to (8.2.12), $|\Delta_0(n,j,\mathbf{t})| \le \varepsilon n^{-1}\|\mathbf{t}\|^2$ for δ sufficiently small. So, for δ sufficiently small, for $\mathbf{t} \in R_2(n)$,

$$\prod_{j=1}^n \varphi(n^{-3/2} j\mathbf{t}) = \exp\left\{\frac{i(n+1)\mathbf{t}^\top \mu}{2n^{1/2}} - \frac{\mathbf{t}^\top M\mathbf{t}}{6} + \Delta_2(n,\mathbf{t})\right\},$$

where $|\Delta_2(n,\mathbf{t})| \le \varepsilon\|\mathbf{t}\|^2$ for all n sufficiently large. Suppose $\varepsilon \in (0, \lambda_{\min}(M)/12)$,

so that, by (8.2.2), $\mathbf{t}^\top M\mathbf{t} \geq 12\varepsilon\|\mathbf{t}\|^2$. Then

$$|\Phi_n(\mathbf{t})| = \left|\exp\left\{-\frac{\mathbf{t}^\top M\mathbf{t}}{6} + \Delta_2(n,\mathbf{t})\right\}\right| \leq \exp\{-\varepsilon\|\mathbf{t}\|^2\}.$$

So we have

$$\begin{aligned} I_2(n) &\leq \int_{R_2(n)} |\Phi_n(\mathbf{t})|\mathrm{d}\mathbf{t} + \int_{R_2(n)} |f_n(\mathbf{t})|\mathrm{d}t \\ &\leq 2\int_{\mathbb{R}^d\setminus R_1} \exp\left\{-\varepsilon\|\mathbf{t}\|^2\right\}\mathrm{d}\mathbf{t}, \end{aligned}$$

for δ sufficiently small and n sufficiently large. This yields part (ii) of the lemma.

Now we proceed to estimate $I_3(n)$. First note that, by (8.2.7),

$$|\Phi_n(\mathbf{t})| = \prod_{j=1}^{n} |\varphi(n^{-3/2}j\mathbf{t})| \leq \prod_{j=\lceil n/2\rceil}^{n} |\varphi(n^{-3/2}j\mathbf{t})|. \tag{8.2.13}$$

For any $\mathbf{t} \in R_3(n)$, we have $n^{-3/2}j\mathbf{t} \in [-\pi j/n, \pi j/n]^d \setminus [-\delta j/n, \delta j/n]^d$. In particular

$$\bigcup_{j=\lceil n/2\rceil}^{n} \{n^{-3/2}j\mathbf{t}\} \subset [-\pi,\pi]^d \setminus [-\delta/2,\delta/2]^d.$$

Thus we may apply the final statement in Lemma 6.2.3 for some ρ sufficiently small to obtain

$$\sup_{\mathbf{t}\in R_3(n)} \sup_{\lceil n/2\rceil \leq j \leq n} |\varphi(n^{-3/2}j\mathbf{t})| \leq \mathrm{e}^{-c_\rho},$$

for some $c_\rho > 0$. Hence from (8.2.13) we have

$$\sup_{\mathbf{t}\in R_3(n)} |\Phi_n(\mathbf{t})| \leq \mathrm{e}^{-nc_\rho/2}.$$

It follows that

$$\begin{aligned} |I_3(n)| &\leq \int_{[-\pi\sqrt{n},\pi\sqrt{n}]^d} \mathrm{e}^{-nc_\rho/2} + \int_{\mathbb{R}^d\setminus[-\delta\sqrt{n},\delta\sqrt{n}]^d} \exp\left\{-\frac{\mathbf{t}^\top M\mathbf{t}}{6}\right\}\mathrm{d}\mathbf{t} \\ &\leq (2\pi)^d n^{d/2}\mathrm{e}^{-nc_\rho/2} + \int_{\mathbb{R}^d\setminus[-\delta\sqrt{n},\delta\sqrt{n}]^d} \exp\left\{-\frac{\lambda_{\min}(M)}{6}\|\mathbf{t}\|^2\right\}\mathrm{d}\mathbf{t}, \end{aligned}$$

using (8.2.2). This gives part (iii) of the lemma.

It remains to estimate $I_4(n)$. Fix $\mathbf{t} \in R_4(n)$, and consider sets

$$\Lambda_n(\mathbf{t}) = \left\{n^{-3/2}j\mathbf{t} : j \in \{1,2,\ldots,n\}\right\}, \text{ and } L_n(\mathbf{t}) = \left\{n^{-3/2}u\mathbf{t} : 1 \leq u \leq n\right\}.$$

Recall that $S_H := 2\pi\mathbb{Z}^d$ in the case $H = I$, and, for $\rho > 0$, define $S_H(\rho) := \cup_{\mathbf{y}\in S} B(\mathbf{y};\rho)$, where $B(\mathbf{y};\rho)$ is the open Euclidean ball of radius ρ centred at $\mathbf{y} \in \mathbb{R}^d$. Define $N_n(\mathbf{t}) := |\Lambda_n(\mathbf{t}) \setminus S_H(\rho)|$. Lemma 6.2.3 and (8.2.7) show that

$$|\Phi_n(\mathbf{t})| = \prod_{j=1}^{n} |\varphi(n^{-3/2} j\mathbf{t})| \leq \exp\{-c_\rho N_n(\mathbf{t})\}, \tag{8.2.14}$$

for some positive constant c_ρ. We aim to show that $N_n(\mathbf{t})$ is bounded below by a constant times n. To do this we use a counting argument related to one used in [23, Lemma 4.4].

Let $K_n(\mathbf{t})$ be the number of $\mathbf{x} \in S_H$ such that $B(\mathbf{x};\rho) \cap L_n(\mathbf{t}) \neq \emptyset$. Set $\nu := n^{-3/2}\|\mathbf{t}\|$. As $\mathbf{t} \in [-\pi n^{3/2}, \pi n^{3/2}]^d \setminus [-\pi n^{1/2}, \pi n^{1/2}]^d$, we have

$$\frac{\pi}{n} \leq \nu \leq \pi\sqrt{d}. \tag{8.2.15}$$

Take $\rho = \pi/8$. We claim that between any two balls of $S_H(\rho)$ that intersect $L_n(\mathbf{t})$ there is at least one point of $\Lambda_n(\mathbf{t})$. Write $\mathbf{y}_j = n^{-3/2} j\mathbf{t}$ for $j \in \{1, \ldots, n\}$. Suppose $i_1, i_2 \in \{1, \ldots, n\}$ with $i_1 < i_2$ and $\mathbf{x}_1, \mathbf{x}_2 \in S_H$ with $\mathbf{x}_1 \neq \mathbf{x}_2$ are such that $\mathbf{y}_{i_1} \in B(\mathbf{x}_1;\rho)$ and $\mathbf{y}_{i_2} \in B(\mathbf{x}_2,\rho)$. To prove the claim we need to show that there exists j with $i_1 < j < i_2$ such that $\mathbf{y}_j \notin S_H(\rho)$. First note that since $n^{-3/2}\mathbf{t} \in [-\pi,\pi]^d$ and $\mathbf{y}_{i_1} \in B(\mathbf{x}_1;\rho)$, the point $\mathbf{y}_{i_1+1}$ must lie in the box $Q(\mathbf{x}_1) = \mathbf{x}_1 + [-9\pi/8, 9\pi/8]^d$. As $9\pi/8 < 15\pi/8 = 2\pi - \rho$, the box $Q(\mathbf{x}_1)$ does not intersect any balls in $S_H(\rho)$ other than $B(\mathbf{x}_1;\rho)$. There are two cases. Either (i) $\mathbf{y}_{i_1+1} \notin B(\mathbf{x}_1;\rho)$, or (ii) $\mathbf{y}_{i_1+1} \in B(\mathbf{x}_1;\rho)$. In case (i) the claim is proved. In case (ii), we have $\nu \leq 2\rho$, and since $B(\mathbf{x}_1;3\rho)$ is contained in $Q(\mathbf{x}_1)$, there is some j with $i_1 + 1 < j < i_2$ such that $\mathbf{y}_j \notin S_H(\rho)$, proving the claim. Hence

$$N_n(\mathbf{t}) \geq K_n(\mathbf{t}) - 1. \tag{8.2.16}$$

The total length of $L_n(\mathbf{t})$ is less than νn, and each segment of $L_n(\mathbf{t})$ between neighbouring balls that intersect $L_n(\mathbf{t})$ has length at least $2\pi - 2\rho$, so $(K_n(\mathbf{t}) - 1)(2\pi - 2\rho) \leq \nu n$, or, equivalently,

$$K_n(\mathbf{t}) \leq \frac{4\nu n}{7\pi} + 1. \tag{8.2.17}$$

Moreover, each ball of $S_H(\rho)$ that intersects $L_n(\mathbf{t})$ contains at most $2\rho/\nu + 1$ points

of $\Lambda_n(\mathbf{t})$, so that the number of points in $\Lambda_n(\mathbf{t}) \cap S_H(\rho)$ satisfies

$$n - N_n(\mathbf{t}) \leq K_n(\mathbf{t}) \left(\frac{\pi}{4\nu} + 1 \right). \tag{8.2.18}$$

Let $\varepsilon > 0$ be a constant. We consider the following two cases.

Case 1: $K_n(\mathbf{t}) \leq \varepsilon n \nu$. In this case we have from (8.2.18) and (8.2.15) that

$$N_n(\mathbf{t}) \geq n - \frac{\pi\varepsilon}{4} n - \varepsilon n \nu \geq n - \frac{\pi}{4} \varepsilon n - \varepsilon n \pi \sqrt{d} \geq \varepsilon n,$$

for ε small enough.

Case 2: $K_n(\mathbf{t}) > \varepsilon n \nu$. If $\nu \geq \frac{1}{2}$, then we have from (8.2.16) that,

$$N_n(\mathbf{t}) \geq K_n(\mathbf{t}) - 1 \geq (\varepsilon/3) n,$$

for n sufficiently large. On the other hand, if $\nu < \frac{1}{2}$, then (8.2.18) and (8.2.17) show that

$$\begin{aligned} N_n(\mathbf{t}) &\geq n - \left(\frac{4\nu n}{7\pi} + 1 \right) \left(\frac{\pi}{4\nu} + 1 \right) \\ &= \frac{6n}{7} - \frac{\pi}{4\nu} - \frac{4\nu n}{7\pi} - 1 \\ &\geq \frac{6n}{7} - \frac{n}{4} - \frac{2n}{7\pi} - 1, \end{aligned}$$

by (8.2.15). Thus we have shown that, in any case, $N_n(\mathbf{t}) \geq \varepsilon n$ for some constant $\varepsilon > 0$ and all n sufficiently large. Thus from (8.2.14) we conclude that

$$\begin{aligned} |I_4(n)| &\leq \int_{R_4(n)} |\Phi_n(\mathbf{t})| \mathrm{d}\mathbf{t} + \int_{R_4(n)} |f_n(\mathbf{t})| \mathrm{d}\mathbf{t} \\ &\leq (2\pi n^{3/2})^d \exp\{-\varepsilon c_\rho n\} + \int_{\mathbb{R}^d \setminus [-\pi n^{1/2}, \pi n^{1/2}]^d} \exp\left\{ -\frac{\lambda_{\min}(M)}{6} \|\mathbf{t}\|^2 \right\} \mathrm{d}\mathbf{t}. \end{aligned}$$

Hence we have proved the last statement in Lemma 8.2.2. □

Now we can gather all our estimates and complete the proof of Theorem 7.2.1.

Proof of Theorem 7.2.1. We have from Lemma 8.2.1 that

$$\sup_{\mathbf{x} \in \mathcal{L}_n} \left| \frac{n^{3d/2}}{h} p_n(\mathbf{x}) - n(\mathbf{x}) \right| \leq \sum_{k=1}^{4} I_k(n) + \int_{R^c(n)} \exp\left\{ -\frac{\lambda_{\min}(M)}{6} \|\mathbf{t}\|^2 \right\} \mathrm{d}\mathbf{t}. \tag{8.2.19}$$

Clearly the integral term tends to 0 as $n \to \infty$, while Lemma 8.2.2 shows that

$|I_3(n) + I_4(n)| \to 0$. Lemma 8.2.2 also shows that for any $\varepsilon > 0$, we can choose A large enough so that $|I_2(n)| \le \varepsilon$ for all n, and hence $\limsup_{n\to\infty} |I_1(n) + I_2(n)| \le \varepsilon$. Hence $|I_1(n) + I_2(n)| \to 0$ as well. This completes the proof of the theorem. □

8.3 Proofs for two or higher dimensions

This section is devoted to the proof of Theorem 7.3.1. The idea is to use the local limit theorem to control (via Borel–Cantelli) the visits of G_n to a growing ball, along a subsequence of times suitably chosen so that the slow movement of the centre of mass controls the trajectory between the times of the subsequence as well. Here is our estimate on the deviations.

Lemma 8.3.1. *Suppose that* (M) *holds and that* $\boldsymbol{\mu} = \mathbf{0}$. *Let* $a_n = \lceil n^\beta \rceil$ *for some* $\beta > 1$. *Then, for any* $\varepsilon > 0$, *a.s. for all but finitely many* n,

$$\max_{a_n \le m \le a_{n+1}} \|G_m - G_{a_n}\| \le n^{\frac{\beta}{2}-1+\varepsilon}.$$

Proof. We use the crude bound that for any $\varepsilon > 0$, $\|S_n\| \le n^{(1/2)+\varepsilon}$ all but f.o., a.s. It follows from the triangle inequality that

$$\|G_n\| \le \frac{1}{n}\sum_{i=1}^{n} \|S_i\| \le \max_{1\le i\le n} \|S_i\| \le n^{(1/2)+\varepsilon}, \tag{8.3.1}$$

all but f.o., a.s. Next, by the triangle inequality again, for any $\varepsilon > 0$, a.s., all but f.o.,

$$\|G_{n+1} - G_n\| = \left\|\frac{S_{n+1} - G_n}{n+1}\right\| \le \frac{\|S_{n+1}\|}{n+1} + \frac{\|G_n\|}{n+1} \le n^{-(1/2)+\varepsilon}. \tag{8.3.2}$$

It follows that for any $\varepsilon > 0$, a.s., all but f.o.,

$$\max_{a_n \le m \le a_{n+1}} \|G_m - G_{a_n}\| \le (a_{n+1} - a_n) \max_{a_n \le m \le a_{n+1}-1} \|G_{m+1} - G_m\|,$$

where $a_{n+1} - a_n \le (n+1)^\beta - n^\beta + 1 = O(n^{\beta-1})$, and, a.s., all but f.o., by (8.3.2),

$$\max_{a_n \le m \le a_{n+1}-1} \|G_{m+1} - G_m\| \le a_n^{-(1/2)+\varepsilon} = O(n^{-(\beta/2)+\beta\varepsilon}).$$

Since $\varepsilon > 0$ was arbitrary, the result follows. □

Now we are ready to prove Theorem 7.3.1.

Proof of Theorem 7.3.1. First, given the upper bound in equation (8.3.1), we only need to show that for any $\varepsilon > 0$, a.s., for all but finitely many n,

$$\|G_n\| \geq n^{(1/2)-\varepsilon}. \tag{8.3.3}$$

Let $B(r)$ denote the closed Euclidean ball, centred at the origin, of radius $r > 0$. We show that for any $\gamma \in (0, 1/2)$, G_n will return to the ball $B(n^\gamma)$ only f.o. To do this, we show that along a suitable subsequence $a_n = \lceil n^\beta \rceil$, $\beta > 1$, G_{a_n} returns to the ball $B(2a_n^\gamma)$ only f.o., and Lemma 8.3.1 controls the trajectory between the instants of the subsequence.

First, we claim that

$$\mathbb{P}(G_n \in B(2n^\gamma)) \leq Cn^{d\left(\gamma - \frac{1}{2}\right)}, \tag{8.3.4}$$

for sufficiently large n and some constant C. Then

$$\sum_{n=1}^{\infty} \mathbb{P}(G_{a_n} \in B(2a_n^\gamma)) \leq C \sum_{n=1}^{\infty} n^{\beta d\left(\gamma - \frac{1}{2}\right)}.$$

Assuming that

$$\beta > \frac{2}{d(1-2\gamma)} \tag{8.3.5}$$

this sum converges, so the Borel–Cantelli lemma shows that $G_{a_n} \notin B(2a_n^\gamma)$ for all but finitely many n, a.s. It then follows from Lemma 8.3.1 that between any a_n and a_{n+1} with n sufficiently large, the trajectory deviates by at most $n^{(\beta/2)-1+\varepsilon}$. In particular, the trajectory between times a_n and a_{n+1} will not visit $B(a_n^\gamma)$ if we ensure that $n^{(\beta/2)-1+\varepsilon} < a_n^\gamma$. (See Figure 8.1.) The latter condition can be achieved (for sufficiently small choice of ε) if $(\beta/2) - 1 < \beta\gamma$, i.e., $\beta < (\frac{1}{2} - \gamma)^{-1}$. Combined with (8.3.5) we see that we must choose $\beta > 1$ such that

$$\frac{2}{d(1-2\gamma)} < \beta < \frac{2}{(1-2\gamma)},$$

which is possible for any $\gamma \in (0, 1/2)$, provided $d \geq 2$.

Consider n such that $a_m \leq n < a_{m+1}$; then we have shown that a.s., for all but

finitely many n,

$$\|G_n\| \geq a_m^\gamma \geq m^{\beta\gamma} \geq \left(\frac{m^{\beta\gamma}}{2(m+1)^{\beta\gamma}}\right) a_{m+1}^\gamma.$$

In particular, for all n sufficiently large, $\|G_n\| \geq (1/4)n^\gamma$, which establishes (8.3.3).

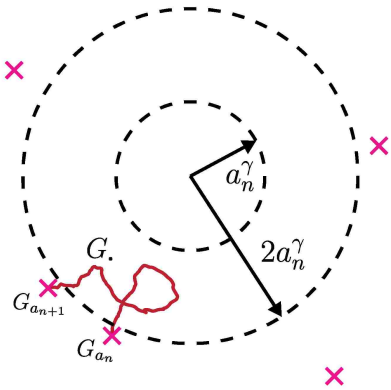

Figure 8.1: Controlling G_n along a subsequence.

It remains to prove the claim (8.3.4); here we use our local limit theorem. First note that

$$\mathbb{P}(G_n \in B(2n^\gamma)) = \mathbb{P}(n^{-1/2}G_n \in n^{-1/2}B(2n^\gamma)).$$

The ball $n^{-1/2}B(2n^\gamma)$ has radius $O(n^{\gamma-\frac{1}{2}})$, and the lattice spacing of $\mathcal{L}_n$ is of order $n^{-3/2}$, so $n^{-1/2}B(2n^\gamma)$ contains $O(n^{d(\gamma+1)})$ lattice points. From Theorem 7.2.1, we also know that for all $x \in \mathcal{L}_n$, $\mathbb{P}(n^{-1/2}G_n = x) = O(n^{-3d/2})$. Summing up over all $x \in n^{-1/2}B(2n^\gamma)$ we get

$$\mathbb{P}(n^{-1/2}G_n \in n^{-1/2}B(2n^\gamma)) = O\left(n^{-3d/2} \times n^{d(\gamma+1)}\right) = O\left(n^{d\left(\gamma-\frac{1}{2}\right)}\right),$$

establishing (8.3.4). This completes the proof. □

8.4 Proofs for one dimension

8.4.1 Recurrence

We will start with a couple of general observations that help the proof of Theorem 7.4.1 in Section 8.4. Recall the definition of an exchangeable event from Section 6.3.

For our one dimensional centre of mass process, we have the following result.

Lemma 8.4.1. *Let $d = 1$. For any $x \in \mathbb{R}$, the event $\{\limsup_{n\to\infty} G_n \geq x\}$ is exchangeable.*

Proof. For any $x \in \mathbb{R}$, we notice that for any fixed positive integer k,

$$\begin{aligned}
&\left\{\limsup_{n\to\infty} G_n \geq x\right\} \\
&= \left\{\limsup_{n\to\infty} \left[\frac{1}{n}(S_1 + S_2 + \cdots + S_k) + \frac{1}{n}(S_{k+1} + S_{k+2} + \cdots + S_n)\right] \geq x\right\} \\
&= \left\{\limsup_{n\to\infty} \frac{1}{n}(S_{k+1} + S_{k+2} + \cdots + S_n) \geq x\right\},
\end{aligned}$$

which is invariant under permutations of $X_1, X_2, \ldots, X_k$. □

The next result shows that G_n can either be trivial, transient, or oscillating.

Lemma 8.4.2. *Let $d = 1$. One and only one of the following will occur with probability 1.*

(i) $G_n = 0$ *for all* n.

(ii) $G_n \to \infty$.

(iii) $G_n \to -\infty$.

(iv) $-\infty = \liminf_{n\to\infty} G_n < \limsup_{n\to\infty} G_n = \infty$.

Proof. We adapt the proof of Theorem 4.1.2 in [27]. Lemma 8.4.1 and the Hewitt–Savage zero–one law (Theorem 6.3.1) imply $\limsup_{n\to\infty} G_n = \ell$, a.s., for some $\ell \in [-\infty, \infty]$. Let $G'_n := \frac{n+1}{n}(G_{n+1} - X_1) = \sum_{i=1}^{n} \frac{n-i+1}{n} X_{i+1}$. Recalling (7.1.1), we see the sequence (G'_n) has the same distribution as (G_n). So taking $n \to \infty$ in $\frac{n}{n+1} G'_n = G_{n+1} - X_1$ we obtain $\ell = \ell - X_1$, a.s., implying $X_1 = 0$ a.s. if ℓ is finite, which is case (i). Otherwise, $\ell = -\infty$ or $+\infty$. A similar argument applies to $\liminf_{n\to\infty} G_n$. The 3 possible combinations ($\limsup_{n\to\infty} G_n = -\infty$ and $\liminf_{n\to\infty} G_n = \infty$ being impossible) give (ii), (iii), and (iv). □

Clearly cases (ii) and (iii) of Lemma 8.4.2 are transient; case (iv), when the walk oscillates, is the most interesting case. The next result shows that oscillating behaviour is enough to ensure recurrence provided that $\mathbb{E}\, X = 0$.

Lemma 8.4.3. *Suppose that* $d = 1$ *and* $\mathbb{E}\,X = 0$. *Suppose that* $\limsup_{n\to\infty} G_n = +\infty$ *and* $\liminf_{n\to\infty} G_n = -\infty$. *Then, for any* $x \in \mathbb{R}$, $\liminf_{n\to\infty} |G_n - x| = 0$, *a.s.*

Proof. Fix $\varepsilon > 0$. Since $S_n/n \to 0$ a.s. and, by Proposition 7.1.1, $G_n/n \to 0$ a.s., we have

$$G_{n+1} - G_n = \frac{S_{n+1} - G_n}{n+1} \to 0, \text{ a.s.}$$

Hence $|G_{n+1} - G_n| < \varepsilon$ all but f.o. (finitely often). For any $x \in \mathbb{R}$, $\limsup_{n\to\infty} G_n = +\infty$ and $\liminf_{n\to\infty} G_n = -\infty$ implies that there are infinitely many n for which $G_n - x$ and $G_{n+1} - x$ have opposite signs. Hence $|G_n - x| < \varepsilon$ infinitely often. □

The next result shows that G_n does oscillate when (M) holds.

Lemma 8.4.4. *Suppose that* $d = 1$, *that* (M) *holds, and that* $\mathbb{E}\,X = 0$. *Then* $\limsup_{n\to\infty} G_n = +\infty$ *and* $\liminf_{n\to\infty} G_n = -\infty$.

Proof. For any $x \in \mathbb{R}$, we have that

$$\begin{aligned}
\mathbb{P}\left(\limsup_{n\to\infty} G_n \geq x\right) &\geq \mathbb{P}\left(G_n \geq x \text{ i.o.}\right) \\
&= \mathbb{P}\left(\bigcap_{m=1}^{\infty} \bigcup_{n\geq m} \{G_n \geq x\}\right) \\
&= \lim_{m\to\infty} \mathbb{P}\left(\bigcup_{n\geq m} \{G_n \geq x\}\right) \\
&\geq \lim_{m\to\infty} \mathbb{P}\left(G_m \geq x\right) \\
&= \frac{1}{2},
\end{aligned}$$

by the central limit theorem, Proposition 7.1.2. With Lemma 8.4.1 and the Hewitt–Savage zero–one law (Theorem 6.3.1), it follows that $\limsup_{n\to\infty} G_n \geq x$, a.s., and since $x \in \mathbb{R}$ was arbitrary, we get $\limsup_{n\to\infty} G_n = +\infty$. A similar argument gives $\liminf_{n\to\infty} G_n = -\infty$. □

Now we are ready to prove our main recurrence theorem in one dimension.

Proof of Theorem 7.4.1. Under the conditions in part (i) of the theorem, the process (G_n) has the same distribution as the process $(-G_n)$, and so we must be in either

case (i) or (iv) of Lemma 8.4.2. The trivial case (i) is ruled out since $\mathbb{E}\,|X| > 0$. Thus case (iv) applies, and G_n changes sign i.o., so by Lemma 8.4.3 we obtain the desired conclusion.

Under the conditions in part (ii), Lemma 8.4.4 applies, so (iv) applies again, and the same argument gives the result. □

8.5 Case of stable distribution

8.5.1 Local limit theorem for stable distribution

For the remainder of this section we work towards a proof of Theorem 7.4.2. The proof rests on the following local limit theorem. We use the notation

$$\mathcal{L}_n := \left\{ n^{-1-1/\alpha} \left(\tfrac{1}{2} n(n+1)b + h\mathbb{Z} \right) \right\},$$

and $p_n(x) := \mathbb{P}(G_n = n^{1/\alpha} x)$.

Theorem 8.5.1. *Suppose that $d = 1$ and* (L) *holds, i.e., $\mathbb{P}(X \in b + h\mathbb{Z}) = 1$ for $b \in \mathbb{R}$ and $h > 0$ maximal. Suppose also that* (S) *holds. Then*

$$\lim_{n\to\infty} \sup_{\mathbf{x}\in\mathcal{L}_n} \left| \frac{n^{1+1/\alpha}}{h} p_n(x) - (\alpha+1)^{1/\alpha} g\left((\alpha+1)^{1/\alpha} x \right) \right| = 0, \tag{8.5.1}$$

where $g(x)$ is the density of the stable distribution in (S).

Proof. The proof is similar to that of Theorem 7.2.1, and can also be compared to the proof of the local limit theorem for sums of i.i.d. random variables in the domain of attraction of a stable law: see [56, §4.2].

Assumption (S) implies that $n^{-1/\alpha} S_n$ converges in distribution to a (constant multiple of) a random variable with characteristic function $\nu(t) = \mathrm{e}^{-c|t|^\alpha}$, where $c > 0$ and $\alpha \in (0,1)$; see Theorems 2.2.2 and 2.6.7 of [56]. It also follows, by an examination of the statements of Theorems 2.6.1 and 2.6.7 of [56] and the proof of Theorem 2.6.5 of [56], that for t in a neighbourhood of 0,

$$\log \varphi(t) = -c|t|^\alpha \left(1 + \varepsilon(t)\right), \tag{8.5.2}$$

where $|\varepsilon(t)| \to 0$ as $t \to 0$.

Define $Y_n = \sum_{i=1}^n S_i$ and let

$$\Phi_n(t) := \mathbb{E}\, \mathrm{e}^{in^{-1-1/\alpha} t Y_n}.$$

Using the $d = 1$ case of the inversion formula (8.2.4) with $W = \left(Y_n - \frac{n(n+1)}{2} b\right)/h \in \mathbb{Z}$, we get

$$p_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{e}^{-\frac{iu}{h}\left(n^{1+1/\alpha} x - \frac{n(n+1)}{2} b\right)} \mathbb{E}\left[\mathrm{e}^{\frac{iu}{h}\left(Y_n - \frac{n(n+1)}{2} b\right)}\right] \mathrm{d}u, \text{ for } x \in \mathcal{L}_n.$$

Using the substitution $t = un^{1+1/\alpha}/h$, we obtain

$$\frac{n^{1+1/\alpha}}{h} p_n(x) = \frac{1}{2\pi} \int_{-\pi n^{1+1/\alpha}/h}^{\pi n^{1+1/\alpha}/h} \mathrm{e}^{-itx} \Phi_n(t) \mathrm{d}t. \tag{8.5.3}$$

On the other hand, from the inversion formula for densities we have that

$$g(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-itx} \nu(t) \mathrm{d}t,$$

where g is the density corresponding to ν. It follows that

$$\begin{aligned}(\alpha+1)^{1/\alpha} g\left((\alpha+1)^{1/\alpha} x\right) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} (\alpha+1)^{1/\alpha} \mathrm{e}^{-it(\alpha+1)^{1/\alpha} x} \nu(t)\, \mathrm{d}t \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-isx} \nu\left(\frac{s}{(\alpha+1)^{1/\alpha}}\right) \mathrm{d}s,\end{aligned}$$

using the substitution $s = (\alpha+1)^{1/\alpha} t$. Since $\nu(t) = \mathrm{e}^{-c|t|^\alpha}$, we get

$$(\alpha+1)^{1/\alpha} g\left((\alpha+1)^{1/\alpha} x\right) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-itx - \frac{c|t|^\alpha}{\alpha+1}} \mathrm{d}t. \tag{8.5.4}$$

Subtracting equation (8.5.4) from equation (8.5.3) we obtain

$$\left|\frac{n^{1+1/\alpha}}{h} p_n(x) - (\alpha+1)^{1/\alpha} g\left((\alpha+1)^{1/\alpha} x\right)\right| \le \sum_{k=1}^{4} J_k(n) + J_5,$$

where

$$\begin{aligned}J_1(n) &:= \int_{-A}^{A} \left|\Phi_n(t) - \mathrm{e}^{-\frac{c|t|^\alpha}{\alpha+1}}\right| \mathrm{d}t \\ J_2(n) &:= \int_{A \le |t| \le \delta n^{1/\alpha}} |\Phi_n(t)|\, \mathrm{d}t \\ J_3(n) &:= \int_{\delta n^{1/\alpha} \le |t| \le \pi n^{1/\alpha}/h} |\Phi_n(t)|\, \mathrm{d}t\end{aligned}$$

$$J_4(n) := \int_{\pi n^{1/\alpha}/h \le |t| \le \pi n^{1+1/\alpha}/h} |\Phi_n(t)| \, \mathrm{d}t$$
$$J_5 := \int_{|t|>A} \left| \mathrm{e}^{-\frac{c|t|^\alpha}{\alpha+1}} \right| \mathrm{d}t$$

for some constants A and δ to be determined later. The statement of the theorem will follow once we show that

$$\lim_{A\to\infty} \lim_{n\to\infty} \left(\sum_{k=1}^{4} J_k(n) + J_5 \right) = 0.$$

Thus it remains to establish this fact.

Since Y_n has the same distribution as $\sum_{j=1}^n jX_j$, we get

$$\begin{aligned} \log \Phi_n(t) &= \log \prod_{j=1}^n \varphi\left(\frac{jt}{n^{1+1/\alpha}}\right) = \sum_{j=1}^n \log \varphi\left(\frac{jt}{n^{1+1/\alpha}}\right) \\ &= -\frac{c|t|^\alpha}{n^{\alpha+1}} \sum_{j=1}^n j^\alpha \left(1 + \varepsilon\left(\frac{jt}{n^{1+1/\alpha}}\right)\right), \end{aligned} \tag{8.5.5}$$

using (8.5.2). Since $|\varepsilon(t)| \to 0$ as $t \to 0$, we have

$$\lim_{n\to\infty} \sup_{t\in[-A,A]} \max_{j\in\{1,2,\cdots,n\}} \varepsilon\left(\frac{jt}{n^{1+1/\alpha}}\right) = 0. \tag{8.5.6}$$

A simple consequence of the fact that $\sum_{k=0}^{n-1} k^\alpha \le \int_0^n u^\alpha \mathrm{d}u \le \sum_{k=1}^n k^\alpha$ for $\alpha > 0$ is

$$\sum_{j=1}^n j^\alpha = \frac{n^{\alpha+1}}{\alpha+1} + O(n^\alpha). \tag{8.5.7}$$

It follows from (8.5.5), (8.5.6) and (8.5.7), that uniformly over $t \in [-A, A]$, as $n \to \infty$,

$$\log \Phi_n(t) = -\frac{c|t|^\alpha}{1+\alpha}(1 + o(1)).$$

It follows that $\lim_{n\to\infty} J_1(n) = 0$ for any $A \in \mathbb{R}_+$.

For $J_2(n)$, we see that

$$\lim_{\delta\to 0} \sup_n \sup_{t\in[-\delta n^{1/\alpha}, \delta n^{1/\alpha}]} \max_{j\in\{1,2,\cdots,n\}} \varepsilon\left(\frac{jt}{n^{1+1/\alpha}}\right) = 0.$$

So by (8.5.7) we may choose δ small enough so that for $t \in [-\delta n^{1/\alpha}, \delta n^{1/\alpha}]$,

$$\log \Phi_n(t) \le -\frac{c|t|^\alpha}{n^{\alpha+1}} \left(1 - \frac{1}{4}\right) \left(\frac{1}{\alpha+1} n^{\alpha+1} + O(n^\alpha)\right).$$

Hence for sufficiently large n, for all $t \in [-\delta n^{1/\alpha}, \delta n^{1/\alpha}]$,

$$\log \Phi_n(t) \le -\frac{1}{2} \frac{c|t|^\alpha}{\alpha+1}.$$

It follows that

$$\sup_n J_2(n) \le \int_{|t| \ge A} \mathrm{e}^{-\frac{1}{2}\frac{c|t|^\alpha}{\alpha+1}} \mathrm{d}t,$$

which tends to 0 as $A \to \infty$.

Next we consider $J_3(n)$. First observe that

$$|\Phi_n(t)| = \prod_{j=1}^n \left|\varphi\left(\frac{jt}{n^{1+1/\alpha}}\right)\right| \le \prod_{j=\lceil n/2 \rceil}^n \left|\varphi\left(\frac{jt}{n^{1+1/\alpha}}\right)\right|.$$

Now for any $\delta n^{1/\alpha} \le |t| \le \pi n^{1/\alpha}/h$ and any $\lceil n/2 \rceil \le j \le n$, we have

$$\frac{\delta}{2} \le \left|\frac{jt}{n^{1+1/\alpha}}\right| \le \frac{\pi}{h}.$$

We can take ρ sufficiently small so that

$$\rho < \frac{\delta}{2} \le \left|\frac{jt}{n^{1+1/\alpha}}\right| \le \frac{\pi}{h} < \frac{2\pi}{h} - \rho.$$

So an application of the $d = 1$ case of Lemma 6.2.2 gives, for all n,

$$\sup_{\delta n^{1/\alpha} \le |t| \le \pi n^{1/\alpha}/h} \sup_{\lceil n/2 \rceil \le j \le n} \left|\varphi\left(\frac{jt}{n^{1+1/\alpha}}\right)\right| \le \mathrm{e}^{-c_\rho},$$

for some $c_\rho > 0$. Hence we have

$$\sup_{\delta n^{1/\alpha} \le |t| \le \pi n^{1/\alpha}/h} |\Phi_n(t)| \le \mathrm{e}^{-nc_\rho/2},$$

and hence

$$J_3(n) = \int_{\delta n^{1/\alpha} \le |t| \le \pi n^{1/\alpha}/h} |\Phi_n(t)| \, \mathrm{d}t \le \frac{\pi}{h} n^{1/\alpha} \mathrm{e}^{-nc_\rho/2} \to 0,$$

as $n \to \infty$.

For $J_4(n)$, we follow essentially the same counting argument as that used for

$I_4(n)$ in Section 8.2. Let $t' = t/h$. Define

$$\Lambda'(t') := \left\{n^{-1-1/\alpha} jt' : j \in \{1, 2, \ldots, n\}\right\} \text{ and } L'_n(t') := \left\{n^{-1-1/\alpha} ut' : 1 \le u \le n\right\}$$

Let $\nu := n^{-1-1/\alpha}|t'|$ denote the spacing of the points of $\Lambda'(t')$. Since $\pi n^{1/\alpha} \le |t'| \le \pi n^{1+1/\alpha}$, we have

$$\frac{\pi}{n} \le \nu \le \pi,$$

which is just the $d = 1$ case of (8.2.15). Since the counting argument is based on the fact that there are n points with spacing satisfying (8.2.15), the rest of the argument goes through unchanged and we get

$$J_4(n) = \int_{\pi n^{1/\alpha}/h \le |t| \le \pi n^{1+1/\alpha}/h} |\Phi_n(t)| \, \mathrm{d}t \le \frac{\pi}{h} n^{1+1/\alpha} \exp\{-\varepsilon c_\rho n\} \to 0,$$

as $n \to 0$.

Finally, it is clear that $\lim_{A\to\infty} \sup_n J_5 = 0$. □

8.5.2 Transience with stable distribution in one dimension

Using the local limit theorem we just proved, we are ready for the proof of the transient case in one dimension.

Proof of Theorem 7.4.2. Fix $x \in (0, \infty)$ and consider the interval $I = (-x, x)$. Then $\mathbb{P}(G_n \in I) = \mathbb{P}(n^{-1/\alpha} G_n \in n^{-1/\alpha} I)$. Since the lattice spacing of $\mathcal{L}_n$ is of order $n^{-1-1/\alpha}$, the interval $n^{-1/\alpha} I$ contains $O(n)$ lattice points of $\mathcal{L}_n$. Theorem 8.5.1 shows that each such lattice point is associated with probability $O(n^{-1-1/\alpha})$. So we get $\mathbb{P}(G_n \in I) = O(n^{-1/\alpha})$, which is summable for $\alpha \in (0, 1)$. Hence the Borel–Cantelli lemma implies that $\liminf_{n\to\infty} |G_n| \ge x$, a.s., and since x was arbitrary the result follows. □

Chapter 9

Examples and applications

In this chapter, we will give a few examples to illustrate the complication on the lattice distribution. In order to apply the local limit theorems in Chapter 7, one has to carefully find the maximal span h to interpret the theorem in the right lattice. This is not always immediate even for some classical random walks.

Recall the notation that $\varphi(\mathbf{t}) := \mathbb{E}[\mathrm{e}^{i\mathbf{t}^\top X}]$ for the characteristic function of X. Also, we set $U := \{\mathbf{t} \in \mathbb{R}^d : |\varphi(\mathbf{t})| = 1\}$, and given an invertible d by d matrix H, set $S_H := 2\pi(H^\top)^{-1}\mathbb{Z}^d$.

9.1 Lazy simple symmetric random walk

It is remarkable that the trivial choice of lattice distribution for simple symmetric random walk, i.e. $\mathbf{b} = \mathbf{0}$ and $H = I$, the d by d identity matrix, is not the right one as the span h in this case is not maximal. The right choice is actually quite a hassle to obtain. We will discuss that in the next section. Before that, there is actually an elementary walk that has the trivial choice as the right choice. It is the lazy simple symmetric random walk.

Example 9.1.1 (Lazy SSRW on $\mathbb{Z}^d$)**.** Let $\mathbf{e}_1, \ldots, \mathbf{e}_d$ be the standard orthonormal basis vectors of $\mathbb{R}^d$, and suppose that $\mathbb{P}(X = \mathbf{e}_i) = \mathbb{P}(X = -\mathbf{e}_i) = \frac{1}{4d}$ for all i, and $\mathbb{P}(X = \mathbf{0}) = \frac{1}{2}$. Then for $\mathbf{b} = \mathbf{0}$ and $H = I$, the d by d identity matrix, we have $\mathbb{P}(X \in \mathbb{Z}^d) = 1$. To verify that $L = \mathbb{Z}^d$ is minimal, it is sufficient (see Lemma 6.2.3)

to check that $U = S_H = 2\pi\mathbb{Z}^d$. If $\mathbf{t} = (t_j) \in \mathbb{R}^d$,

$$\varphi(\mathbf{t}) = \frac{1}{2} + \frac{1}{4d}\sum_{j=1}^{d}\left(\mathrm{e}^{it_j} + \mathrm{e}^{-it_j}\right) = \frac{1}{2} + \frac{1}{2d}\sum_{j=1}^{d}\cos t_j.$$

Thus $\mathbf{t} \in U$ if and only if $\cos t_j = 1$ for all j, i.e., $U = 2\pi\mathbb{Z}^d = S_H$, as required. Note that we could alternatively see the upper bound and apply Lemma 6.2.1 to check that $h = 1$ is maximal.

Maybe this walk is just too lazy to bother with a complicated choice of a lattice distribution.

9.2 Simple symmetric random walk

Without delay, we will show a right choice of lattice distribution for simple symmetric random walk.

Example 9.2.1 (SSRW on $\mathbb{Z}^d$)**.** Suppose that $\mathbb{P}(X = \mathbf{e}_i) = \mathbb{P}(X = -\mathbf{e}_i) = \frac{1}{2d}$ for all i. For SSRW the construction of H for which (L) holds is non-trivial. For $d = 1$, we take $b = -1$ and $h = 2$. In general $d \geq 2$, we take $H = (h_{ij})$ and $\mathbf{b} = (b_i)$ defined as follows. If $d = 2n - 1$ for $n \geq 2, n \in \mathbb{Z}$, we take

$$b_i = -1 \quad \text{for all } i = 1, 2, \ldots, d;$$

$$h_{ij} = \begin{cases} 1 & \text{if } i - j \equiv 0 \text{ or } n \pmod{2n-1}, \\ 0 & \text{otherwise.} \end{cases}$$

If $d = 2n$ for $n \geq 1, n \in \mathbb{Z}$, we take

$$b_i = \begin{cases} 0 & \text{if } i = 2n, \\ -1 & \text{otherwise;} \end{cases}$$

$$h_{ij} = \begin{cases} -1 & \text{if } (i,j) = (2n, 1), \\ 1 & \text{if } j - i \equiv 0 \text{ or } 1 \pmod{2n} \text{ and } (i,j) \neq (2n,1), \\ 0 & \text{otherwise.} \end{cases}$$

For example, for $d = 2$ we have

$$\mathbf{b} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad \text{and} \quad H = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

For $d = 3$, we have

$$\mathbf{b} = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} \quad \text{and} \quad H = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

For $d = 4$, we have

$$\mathbf{b} = \begin{pmatrix} -1 \\ -1 \\ -1 \\ 0 \end{pmatrix} \quad \text{and} \quad H = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ -1 & 0 & 0 & 1 \end{pmatrix}.$$

For $d = 5$, we have

$$\mathbf{b} = \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \\ -1 \end{pmatrix} \quad \text{and} \quad H = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix},$$

and so on. Note that $h = 2$ for all such H. It is elementary to verify that $\mathbb{P}(X \in \mathbf{b} + H\{0,1\}^d = 1)$. It suffices to check that $H^{-1}(\mathbf{x} - \mathbf{b}) \in \{0,1\}^d$ for any $\mathbf{x} = \pm\mathbf{e}_i$. For example, in the case $d = 2n - 1$ we have that H^{-1} has elements h_{ij}^{-1} given by

$$h_{ij}^{-1} = \begin{cases} \frac{1}{2} & \text{if } i - j = 0, 1, \ldots, n-1 \pmod{2n-1}, \\ -\frac{1}{2} & \text{otherwise,} \end{cases}$$

and then one checks that, for example, $H^{-1}(\mathbf{e}_i - \mathbf{b}) = \mathbf{a}$ where $\mathbf{a}$ has all components zero apart from $a_i = \cdots = a_{i+n-1} = 1$ (for $i < n - 1$). The other cases are similar.

We show that (L) holds for SSRW with this choice of H, by checking (see

Lemma 6.2.3) that $U = S_H$. Since Lemma 6.2.2 shows that $S_H \subseteq U$, it suffices to show that $U \subseteq S_H$. For SSRW on $\mathbb{Z}^d$, if $\mathbf{t} = (t_j) \in \mathbb{R}^d$,

$$\varphi(t) = \frac{1}{2d} \sum_{j=1}^{d} \left(\mathrm{e}^{it_j} + \mathrm{e}^{-it_j} \right) = \frac{1}{d} \sum_{j=1}^{d} \cos t_j.$$

So $\mathbf{t} \in U$ if and only if $|\sum_{j=1}^{d} \cos t_j| = d$, which occurs if and only if either (i) $\cos t_j = 1$ for all j, or (ii) $\cos t_j = -1$ for all j. Case (i) is equivalent to $\mathbf{t} \in 2\pi\mathbb{Z}^d$ and case (ii) is equivalent to $\mathbf{t} \in \pi\mathbf{1} + 2\pi\mathbb{Z}^d$, where $\mathbf{1}$ is the vector of all 1s. Hence

$$U = (2\pi\mathbb{Z}^d) \cup (\pi\mathbf{1} + 2\pi\mathbb{Z}^d).$$

Consider $\mathbf{x} \in U$. Then for some $\mathbf{a} \in \mathbb{Z}^d$, either (i) $\mathbf{x} = 2\pi\mathbf{a}$, or (ii) $\mathbf{x} = \pi\mathbf{1} + 2\pi\mathbf{a}$. In case (i), let $\mathbf{z} = H^\top \mathbf{a}$; then since all entries in H are integers, we have $\mathbf{z} \in \mathbb{Z}^d$ and $2\pi(H^\top)^{-1}\mathbf{z} = 2\pi\mathbf{a} = \mathbf{x}$, so $\mathbf{x} \in S_H$. In case (ii), let $\mathbf{z} = H^\top(\frac{1}{2}\mathbf{1} + \mathbf{a})$. Note that if d is odd then $\frac{1}{2}H^\top\mathbf{1} = \mathbf{1}$ while if d is even, $\frac{1}{2}H^\top\mathbf{1} = (0, 1, 1, \ldots, 1)^\top$; in any case it follows that $\mathbf{z} \in \mathbb{Z}^d$. Then $2\pi(H^\top)^{-1}\mathbf{z} = \pi\mathbf{1} + 2\pi\mathbf{a} = \mathbf{x}$, so $\mathbf{x} \in S_H$. Thus $U \subseteq S_H$.

In general, it is very difficult to find a right lattice distribution for a specific random walk which has the maximal span. It requires onerous effort for trial and error. Even if we can find a possible candidate, it is tricky to prove that indeed it is the right choice. The author feels like it is more an unsolved discrete geometry problem, rather than anything in the field of probability, and we should proceed to our next part of our thesis, leaving the pleasure for the reader to discover more with lattice distribution.

Part III:

Conclusions

Chapter 10

Open problems and conjectures

In this chapter, we are going to give some open problems and conjectures for further research topics.

10.1 Half strip

There are a few models related to the continuum analogue of the half strip model. One of the popular model is called 'Markov-modulated diffusions' [79]. Consider $(X_t, \eta_t) \in \mathbb{R} \times S$, where S is a finite set. The reason why taking the whole real lines instead of the half lines is because we need to define the movement of the walk using stochastic differential equation. Assume that η_t is a continuous-time irreducible Markov chain on S with stationary distribution π. Let $\mu : S \to \mathbb{R}$ and $\sigma : S \to (0, \infty)$, then the movement of the horizontal direction can be determined by the following stochastic differential equation

$$\mathrm{d}X_t = \mu(\eta_t)\mathrm{d}t + \sigma(\eta_t)\mathrm{d}W_t, \tag{10.1.1}$$

where W_t is a Brownian motion on $\mathbb{R}$.

For a full continuum analogue for the (half) strip model, define $\mathcal{A}$ as a manifold in $\mathbb{R}^d$ without boundary. Then our point of interest is $(X_t, \eta_t) \in \mathbb{R} \times \mathcal{A}$, a diffusion process.

On $\mathbb{R} \times \mathcal{A}$, we assume η_t is e.g. a Brownian motion on $\mathcal{A} = \mathcal{S}^{d-1}$ and also assume

there exist a stationary distribution π such that $\eta_t \to \pi$ on $\mathcal{A}$. Again we have the following stochastic differential equation,

$$\mathrm{d}X_t = \mu(\eta_t)\mathrm{d}t + \sigma(\eta_t)\mathrm{d}W_t, \tag{10.1.2}$$

where W_t is a Brownian motion on $\mathbb{R}$. We also assume $\mu : \mathcal{A} \to \mathbb{R}$ is bounded and continuous and σ bounded uniformly from $\{0, \infty\}$, e.g. $\sigma = 1$. We conjecture the following recurrence classification, similar to our discrete model.

Conjecture 10.1.1. *If $\int_{\mathcal{A}} \pi(\mathrm{d}x)\mu(x) > 0$, then $X_t \to \infty$. If $\int_{\mathcal{A}} \pi(\mathrm{d}x)\mu(x) < 0$, then $X_t \to -\infty$.*

10.2 Centre of mass

In $d = 1$, we believe that for any zero drift random walk, G_n is recurrent:

Conjecture 10.2.1. *Suppose $d = 1$. If $\mathbb{E}\, X = 0$, then G_n is recurrent.*

For $d \geq 2$, we believe that G_n is always 'at least as transient' as the situation in Theorem 7.3.1:

Conjecture 10.2.2. *Suppose that* $\operatorname{supp} X$ *is not contained in a one-dimensional subspace of $\mathbb{R}^d$. Then*

$$\liminf_{n\to\infty} \frac{\log \|G_n\|}{\log n} \geq \frac{1}{2} \qquad \text{a.s.}$$

Also, is there an analogue of Chung-Fuchs Theorem, i.e., criterion for recurrence using characteristic function of X?

There is also a continuum analogue for the centre of mass process. Suppose we start with B_t, a Brownian motion on $\mathbb{R}^d$. Consider the following stochastic differential equation,

$$\mathrm{d}X_t = B_t dt. \tag{10.2.1}$$

The process X_t is also known as the integrated Brownian motion, which can also be written as

$$X_t = \int_0^t B_s ds. \tag{10.2.2}$$

The process (B_t, X_t) is the famous Kolmogorov diffusion process [62], which satisfies the Markov property. The joint distribution of the continuum analogue for the random walk and its centre of mass process, $(B_t, \frac{1}{t}X_t)$, however, behaves very differently. It is expected that $\frac{1}{t}X_t$ is recurrent if and only if $d = 1$.

Bibliography

[1] P.C. Allaart and M.G. Monticino, Optimal stopping rules for directionally reinforced processes, *Adv. in Appl. Probab.* **33** (2001) 483–504.

[2] G. Alsmeyer, Recurrence theorems for Markov random walks, *Probability and Mathematical Statistics* **21** Fasc. 1 (2001) 123–134.

[3] R. B. Ash, C. A. Doleans-Dade, *Probability & Measure Theory*, 2nd ed., Academic Press, (1999).

[4] S. Asmussen, *Applied Probability and Queues*, 2nd ed., Springer (2003).

[5] S. Aspandiiarov, R. Iasnogorodski and M. Menshikov, Passage-time moments for nonnegative stochastic processes and an application to reflected random walks in a quadrant, *Ann. Probab.* **24** (1996), 932–960.

[6] J. L. Austerweil, J. T. Abbott and T. L. Griffiths, Human memory search as a random walk in a semantic network, *Advances in Neural Information Processing Systems* **25**, (2012) 3050–3058.

[7] L. Bachelier, Théorie de la spéculation, *Ann. Sci. École Norm. Sup.* **17** (1900), 21–86.

[8] M. N. Barber and B. W. Ninham, *Random and Restricted Walks: Theory and Applications*, Gordon and Breach, New York (1970).

[9] H. C. Berg, *Random Qalks in Biology*, expanded edition ed., Princeton University Press, New Jersey (1993).

[10] J. Bernoulli, *Ars conjectandi. Impensis Thurnisiorum*, fratrum (1713).

[11] R.N. Bhattacharya and R.R. Rao, *Normal Approximation and Asymptotic Expansions*, updated reprint of the 1986 edition, Society for Industrial and Applied Mathematics, Philadelphia (2010).

[12] P. Billingsley, *Convergence of Probability Measures*, 2nd ed. John Wiley & Sons, Inc., New York (1999).

[13] P. Billingsley, *Probability and Measure*, 3rd ed, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons Inc., New York (1995).

[14] A.A. Borovkov, *Probability Theory*, Springer, London (2013).

[15] G. Cardano, *Liber de ludo aleae* (1564).

[16] P. L. Chebyshev. On mean quantities, *Mat. Sbornik* **2** (1867).

[17] A. Y. Chen and E. Renshaw, The general correlated random walk, *J. Appl. Probab.* **31** (1994) 869–884.

[18] Y. S. Chow, H. Teicher, *Probability Theory: Independence, Interchangeability, Martingales* 3rd ed., Springer Texts in Statistics (2003).

[19] K. L. Chung and W. H. J. Fuchs, On the distribution of values of sums of random variables, *Mem. Amer. Math. Soc.* **6** (1951) 12.

[20] F. Comets, M.V. Menshikov, S. Volkov, and A.R. Wade, Random walk with barycentric self-interaction, *J. Stat. Phys.* **143** (2011) 855–888.

[21] W. H. McCrea, and F. J. W. Whipple, Random Paths in Two and Three Dimensions, *Proc. Roy. Soc. Edinburgh* **60** (1940) 281–298.

[22] H.E. Daniels, The statistical theory of stiff chains, *Proc. Roy. Soc. Edinburgh* **63** (1952) 290–311.

[23] R. Dobrushin and O. Hryniv, Fluctuations of shapes of large areas under paths of random walks, *Probab. Theory and Related Fields* **105** (1996) 423–458.

[24] C. Domb, On multiple returns in the random-walk problem, *Proc. Cambridge Philos. Soc.* **50** (1954) 586–591.

[25] C. Domb and M.E. Fisher, On random walks with restricted reversals, *Proc. Camb. Phil. Soc.* **54** (1958) 48–59.

[26] J. L. Doob, *Stochastic Processes*, John Wiley & Sons Inc., New York (1953).

[27] R. Durrett, *Probability: Theory and Examples*, Fourth ed., Cambridge University Press (2010).

[28] A. Einstein, *Investigations on the Theory of the Brownian Movement*, Dover Publications Inc., New York (1956) Edited with notes by R. Fürth, translated by A. D. Cowper.

[29] N. Etemadi, An elementry proof of the strong law of large numbers, *Z. Warsch. verw. Gebiete.* **55** (1981) 119–122.

[30] G. I. Falin, *Ergodicity of Random Walks in the Half-Strip*, *Math. Notes* **44** (1988) 606–608; translated from *Mat. Zametki* **44** (1988) 225–230 [in Russian].

[31] E. F. Fama, Random walks in stock market prices, *Financial Analysts Journal* **51** (1995), 75–80.

[32] G. Fayolle, V. A. Malyshev, and M. V. Menshikov, *Topics in the Constructive Theory of Countable Markov Chains*, Cambridge University Press, Cambridge (1995).

[33] W. Feller, *An Introduction to Probability Theory and its Applications. Vol. II*, 2nd ed., Wiley, New York (1971).

[34] S. R. Finch, *Mathematical Constants*, Cambridge University Press (2003) Section 5.9, 322–331.

[35] F. G. Foster, On the stochastic matrices associated with certain queuing processes, *Ann. Math. Statistics* **24** (1953), 355–360.

[36] E.I. Fredholm, Sur une classe d'equations fonctionnelles, *Acta Math.* **27** (1903) 365–390.

[37] R. Fürth, Die Brownsche Bewegung bei Berücksichtigung einer Persistenz der Bewegungsrichtung, *Zeits. f. Phys* **2** (1920) 244–256.

[38] N. Georgiou, M. V. Menshikov, A. Mijatović and A. R. Wade, Anomalous recurrence properties of many-dimensional zero-drift random walks, *Advances in Applied Probability* **48A** (2016) 99–118.

[39] N. Georgiou and A. R. Wade, Non-homogeneous random walks on a semi-infinite strip, *Stochastic Processes and their Applications*, **124** (2014) 3179–3205.

[40] J. Gillis, Correlated random walk, *Proc. Camb. Phil. Soc.* **51** (1955) 639–651.

[41] M. L. Glasser, and I. J. Zucker, Extended Watson integrals for the cubic lattices, *Proc. Nat. Acad. Sci. U.S.A.* **74** (1977) 1800–1801.

[42] B. V. Gnedenko, On a local limit theorem of the theory of probability, *Uspekhi Mat. Nauk*, **3**:3(25) (1948), 187–194.

[43] B.V. Gnedenko, *The Theory of Probability*, Translated from the fourth Russian edition by B.D. Seckler, Chelsea Publishing Co., New York (1967).

[44] B.V. Gnedenko and A.N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley, Reading, Mass. (1954).

[45] S. Goldstein, On diffusion by discontinuous movements, and on the telegraph equation, *Quart. J. Mech.* **4** (1951) 129–156.

[46] K. Grill, On the average of a random walk, *Statist. Probab. Lett.* **6** (1988) 357–361.

[47] A. Gut, *Probability: A Graduate Course*, 2nd ed., Springer (2012).

[48] E. Halley, An estimate of the degrees of the mortality of mankind, *Philosophical Transactions*, **196** (1693) 596–610.

[49] C. M. Harris and P. G. Marlin, A note on feedback queues with bulk service, *J. Assoc. Comput. Mach.* **19** (1972), 727–733.

[50] K. Hagen, *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets*, 4th ed., World Scientific, Singapore (2004).

[51] B. Hayes, First links in the Markov chain, *American Scientist* **101** (2013) 92–96.

[52] L. Hervé and F. Pène, On the recurrence set of planar Markov random walks, *J. Theoret. Probab.* **26** (2013) 169–197.

[53] P. Holgate, Random walk models for animal behavior, *Statistical Ecology* **2** (1971), 1–12.

[54] M. Holmes and A. Sakai, Senile reinforced random walks, *Stochastic Process. Appl.* **117** (2007) 1519–1539.

[55] B. D. Hughes, *Random Walks and Random Environments*, Vol. 1, Oxford Science Publications, The Clarendon Press Oxford University Press, New York (1995).

[56] I.A. Ibragimov and Y.V. Linnik, *Independent and Stationary Sequences of Random Variables*, Wolters-Noordhoff, Groningen, The Netherlands (1971).

[57] M. Kac, A stochastic model related to the telegrapher's equation, *Rocky Mountain J. Math.* **4** (1974) 497–509.

[58] O. Kallenberg, *Foundations of Modern Probability*, second ed., Springer-Verlag, New York (2002).

[59] J. H. B. Kemperman, The oscillating random walk, *Stochastic Processes Appl.* **2** (1974), 1–29.

[60] A. Y. Khinchin, Sur la loi des grandes nombres, *Comptes Rendus de l'Academie des Sciences* **188** (1929) 477–479.

[61] A. N. Kolmogorov, Sur la loi forte des grands nombres, *Comptes Rendus de l'Academie des Sciences* **191** (1930) 910–912.

[62] A. N. Kolmogorov, Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung, *Mathematische Annalen* **104** (1931) 415–458.

[63] A. Krámli and D. Szász, Random walks with internal degrees of freedom, I. Local limit theorems, *Z. Wahrsch. Verw. Gebiete* **63** (1983) 85–95.

[64] J. Lamperti, Criteria for the recurrence and transience of stochastic processes I, *J. Math. Anal. Appl.* **1** (1960) 314–330.

[65] J. Lamperti, Criteria for stochastic processes II: passage-time moments, *J. Math. Anal. Appl.* **7** (1960) 127–145.

[66] J. Lamperti, *Stochastic Processes: A Survey of the Mathematical Theory*, 2nd ed., Springer (1997).

[67] G.F. Lawler and V. Limic, *Random Walk: A Modern Introduction*, Cambridge University Press, Cambridge (2010).

[68] M. Lenci, Recurrence for persistent random walks in two dimensions, *Stoch. Dyn.* **7** (2007) 53–74.

[69] C. H. Lo and A. R. Wade, Non-homogeneous random walks on a half strip with generalized Lamperti drifts, *Markov Processes and Related Fields* **23** (2017) 125–146.

[70] C. H. Lo and A. R. Wade, On the centre of mass of a random walk, submitted.

[71] G. G. Lowry, *Markov Chains and Monte Carlo Calculations in Polymer Science*, Marcel Dekker, New York (1970).

[72] B. G. Malkiel, *A Random Walk Down Wall Street*, 6th ed., W.W. Norton & Company, Inc (1973).

[73] V.A. Malyshev, Homogeneous random walks on the product of finite set and a halfline, *Veroyatnostnye Metody Issledovania (Probability Methods of Investigation)* **41** (1972) 5–13 [in Russian], ed. A.N. Kolmogorov, Moscow State University, Moscow.

[74] M. V. Menshikov, I. M. Asymont, and R. Iasnogorodskii, Markov processes with asymptotically zero drifts, *Problems of Information Transmission* **31** (1995) 248–261; translated from *Problemy Peredachi Informatsii* **31** (1995) 60–75 [in Russian].

[75] M. Menshikov, S. Popov, A. R. Wade, *Non-homogeneous Random Walks: Lyapunov Function Methods for Near-critical Stochastic Systems*, Cambridge University Press (2016).

[76] J.-F. Mertens, E. Samuel-Cahn, and S. Zamir, Necessary and sufficient conditions for recurrence and transience of Markov chains, in terms of inequalities, *J. Appl. Probab.* **15** (1978), 848–851.

[77] E. W. Montroll, Random walks in multidimensional spaces, especially on periodic lattices, *J. Soc. Indust. Appl. Math.* **4** (1956) 241–260.

[78] T. Mountford, L.P.R. Pimentel, and G. Valle, Central limit theorem for the self-repelling random walk with directed edges, *ALEA, Lat. Am. J. Probab. Math. Stat.* **11** (2014) 503–517.

[79] M. F. Neuts, *Structured Stochastic Matrices of M/G/1 Type and Their Applications*, Marcel Dekker, Inc., New York (1989).

[80] J. Nolan, *Stable Distributions: Models for Heavy-tailed Data*, Birkhäuser (2017).

[81] J. R. Norris, *Markov Chains*, Cambridge University Press (1998).

[82] A. G. Pakes, Some conditions for ergodicity and recurrence of Markov chains, *Operations Res.* **17** (1969) 1058–1061.

[83] K. Pearson, The Problem of the Random Walk, *Nature.* **72** (1905) 294.

[84] K. Pearson and J. Blakeman, *A Mathematical Theory of Random Migration*, Drapers' Company Research Memoirs Biometric Series, Dulau and co., London (1906).

[85] G. Pólya, Über eine aufgabe betreffend die irrfahrt im strassennetz, *Math. Ann.* **84** (1921) 149–160.

[86] A. G. Ramm, A simple proof of the Fredholm alternative and a characterization of the Fredholm operators, *Amer. Math. Monthly* **108** (2001) 855–860.

[87] L. Rayleigh, On the resultant of a large number of vibrations of the same pitch and of arbitrary phase, *Phil. Mag. Ser. 5* **10** (1880), 73–78.

[88] Pál Révész, *Random Walk in Random and Non-random Environments*, 3rd ed., World Scientic Publishing Co. Pte. Ltd., Hackensack, NJ (2013).

[89] L.C.G. Rogers, Recurrence of additive functionals of Markov chains, *Sankhyā* **47** (1985) 47–56.

[90] B. A. Rogozin and S. G. Foss, The recurrence of an oscillating random walk, *Theor. Probability Appl.* **23** (1978), no. 1, 155–162, translated from *Teor. Verojatnost. i Primenen.* **23** (1978) 161–169 [in Russian].

[91] M. Rubinstein and R.H. Colby, *Polymer Physics*, Oxford University Press (2003).

[92] I. Todhunter, *A History of the Mathematical Theory of Probability*, New York, Chelsea (1865), Reprinted 1965.

[93] A. N. Shiryaev, *Probability*, second ed., Graduate Texts in Mathematics, vol. 95, Springer-Verlag, New York (1996), translated from the first (1980) Russian edition by R. P. Boas.

[94] M. F. Shlesinger and B. J. West (eds.), *Random Walks and Their Applications in the Physical and Biological Sciences*, American Institute of Physics, New York (1984).

[95] F. Spitzer, *Principles of Random Walk*, second ed., Springer-Verlag, New York (1976).

[96] D. Szász and B. Tóth, Persistent random walks in a one-dimensional random environment, *J. Stat. Phys.* **37** (1984) 27–38.

[97] G. N. Watson, Three triple integrals, *Quart. J. Math., Oxford Ser. 2* **10** (1939) 266–276.

[98] G. H. Weiss, *Encyclopedia of Statistical Sciences: 1*, Wiley (2006).

[99] G. H. Weiss, Random walks and their applications, *American Scientist* **71** (1983), 65–71.

[100] W. Whitt, *Stochastic-Process Limits*, Springer-Verlag, New York (2002).