Symmetric Random Walk

Gregory F. Lawler and Vlada Limic

Contents

1
Introduction

1.1 Basic definitions

We will define the random walks that we consider in this book. We focus our attention on random walks in $\mathbb{Z}^d$ that have bounded symmetric increment distributions although we occasionally discuss results for wider classes of walks. We also impose an irreducibility criterion to guarantee that all points in the lattice $\mathbb{Z}^d$ can be reached.

We start by setting some basic notation. We use x, y, z to denote points in the integer lattice $\mathbb{Z}^d = \{(x^1, \ldots, x^d) : x^j \in \mathbb{Z}\}$. We use superscripts to denote components, and we use subscripts to enumerate elements. For example, $x_1, x_2, \ldots$ represents a sequence of points in $\mathbb{Z}^d$, and the point x_j can be written in component form $x_j = (x_j^1, \ldots, x_j^d)$. We write $\mathbf{e}_1 = (1, 0, \ldots, 0), \ldots, \mathbf{e}_d = (0, \ldots, 0, 1)$ for the standard basis of unit vectors in $\mathbb{Z}^d$. The prototypical example is (discrete time) simple random walk starting at $x \in \mathbb{Z}^d$. This process can be considered either as a sum of a sequence of independent, identically distributed random variables

$$S_n = x + X_1 + \cdots + X_n$$

where $\mathbb{P}\{X_j = \mathbf{e}_k\} = \mathbb{P}\{X_j = -\mathbf{e}_k\} = 1/(2d), k = 1, \ldots, d$, or it can be considered as a Markov chain with state space $\mathbb{Z}^d$ and transition probabilities

$$\mathbb{P}\{S_{n+1} = z \mid S_n = y\} = \frac{1}{2d}, \quad z - y \in \{\pm\mathbf{e}_1, \ldots \pm \mathbf{e}_d\}.$$

We call $V = \{x_1, \ldots, x_l\} \subset \mathbb{Z}^d \setminus \{0\}$ a *(finite) generating set* if each $y \in \mathbb{Z}^d$ can be written as $k_1x_1 + \cdots + k_lx_l$ for some $k_1, \ldots, k_l \in \mathbb{Z}$. We let $\mathcal{G}$ denote the collection of generating sets V with the property that if $x = (x^1, \ldots, x^d) \in V$ then the first nonzero component of x is positive. An example of such a set is $\{\mathbf{e}_1, \ldots, \mathbf{e}_d\}$. A *(finite range, symmetric, truly d-dimensional) random walk* is given by specifying a $V = \{x_1, \ldots, x_l\} \in \mathcal{G}$ and a function $\kappa : V \to (0, 1]^l$ with $\kappa(x_1) + \cdots + \kappa(x_l) \leq 1$. Associated to this is the symmetric probability distribution on $\mathbb{Z}^d$

$$p(x_k) = p(-x_k) = \frac{1}{2}\,\kappa(x_k), \qquad p(0) = 1 - \sum_{x \in V} \kappa(x).$$

We let $\mathcal{P}_d$ denote the set of such distributions p on $\mathbb{Z}^d$ and $\mathcal{P} = \cup_{d \geq 1} \mathcal{P}_d$. Given p the corresponding random walk S_n can be considered as the time-homogeneous Markov chain with state space $\mathbb{Z}^d$ and transition probabilities

$$p(y, z) := \mathbb{P}\{S_{n+1} = z \mid S_n = y\} = p(z - y).$$

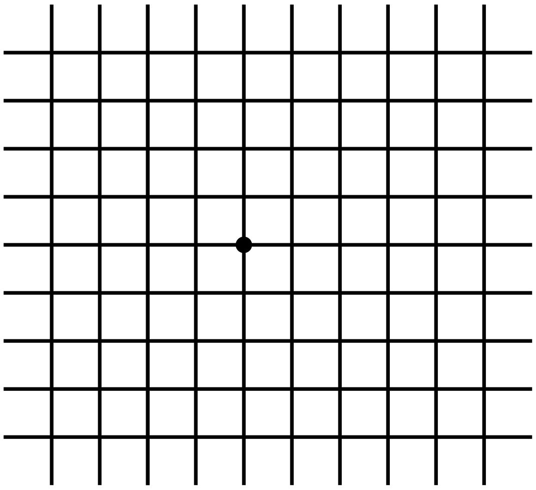

Fig. 1.1. The square lattice $\mathbb{Z}^2$

We can also write

$$S_n = S_0 + X_1 + \cdots + X_n$$

where $X_1, X_2, \ldots$ are independent random variables, independent of S_0, with distribution p. (Most of the time we will choose S_0 to have a trivial distribution.) We will use the phrase $\mathcal{P}$-walk or $\mathcal{P}_d$-walk for such a random walk. We will use the term *simple random walk* for the particular p with

$$p(\mathbf{e}_j) = p(-\mathbf{e}_j) = \frac{1}{2d}, \quad j = 1, \ldots, d.$$

Given $p \in \mathcal{P}$, we write p_n for the n-step distribution

$$p_n(x, y) = \mathbb{P}\{S_n = y \mid S_0 = x\}$$

and $p_n(x) = p_n(0, x)$. Note that $p_n(\cdot)$ is the distribution of $X_1 + \cdots + X_n$ where $X_1, \ldots, X_n$ are independent with increment distribution p.

Comment. In many ways the main focus of this book is simple random walk, and a first-time reader might find it useful to consider this example thoughout. We have chosen to generalize this slightly, because it does not complicate the arguments much and allows the results to be extended to other examples. One particular example is simple random walk on other regular lattices such as the planar triangular lattice. In Section 1.3, we show that walks on other d-dimensional lattices are isomorphic to p-walks on $\mathbb{Z}^d$.

If $S_n = (S_n^1, \ldots, S_n^d)$ is a $\mathcal{P}$-walk with $S_0 = 0$, then $\mathbb{P}\{S_{2n} = 0\} > 0$ for every even integer n; this follows from the easy estimate $\mathbb{P}\{S_{2n} = 0\} \geq [\mathbb{P}\{S_2 = 0\}]^n \geq p(x)^{2n}$ for every $x \in \mathbb{Z}^d$. We will call the walk *bipartite* if $p_n(0,0) = 0$ for every odd n, and we will call it *aperiodic* otherwise. In the latter case, $p_n(0,0) > 0$ for all n sufficiently large (in fact, for all $n \geq k$ where k is the first odd integer with $p_k(0,0) > 0$). Simple random walk is an example of a bipartite walk since $S_n^1 + \cdots + S_n^d$ is odd for odd n and even for even n. If p is bipartite, then we can partition $\mathbb{Z}^d = (\mathbb{Z}^d)_e \cup (\mathbb{Z}^d)_o$ where $(\mathbb{Z}^d)_e$ denotes the points that can be reached from the origin in an even number of steps and $(\mathbb{Z}^d)_o$ denotes the set of points that can be reached in an odd number of steps. In algebraic language, $(\mathbb{Z}^d)_e$ is an additive subgroup of $\mathbb{Z}^d$ of index 2 and $(\mathbb{Z}^d)_o$ is the nontrival coset. Note that if $x \in (\mathbb{Z}^d)_o$, then $(\mathbb{Z}^d)_o = x + (\mathbb{Z}^d)_e$.

Comment. It would suffice and would perhaps be more convenient to restrict our attention to aperiodic walks. Results about bipartite walks can easily be deduced from them. However, since our main example, simple random walk, is bipartite, we have chosen to allow such p.

If $p \in \mathcal{P}$ and $j_1, \ldots, j_d$ are nonnegative integers, the $(j_1, \ldots, j_d)$ moment is given by

$$\mathbb{E}[(X_1^1)^{j_1} \cdots (X_1^d)^{j_d}] = \sum_{x \in \mathbb{Z}^d} (x^1)^{j_1} \cdots (x^d)^{j_d}\, p(x).$$

We let Γ denote the *covariance matrix*

$$\Gamma = \Big[\, \mathbb{E}[X_1^j X_1^k] \,\Big]_{1 \le j,k \le d}.$$

The covariance matrix is symmetric and positive definite. Since the random walk is truly d-dimensional, it is easy to verify (see Proposition 1.1.1 (a)) that the matrix Γ is invertible. There exists a symmetric positive definite matrix Λ such that $\Gamma = \Lambda\, \Lambda^T$ (see Section 11.3). There is a (not unique) orthonormal basis $u_1, \ldots, u_d$ of $\mathbb{R}^d$ such that we can write

$$\Gamma x = \sum_{j=1}^d \sigma_j^2\, (x \cdot u_j)\, u_j, \qquad \Lambda x = \sum_{j=1}^d \sigma_j\, (x \cdot u_j)\, u_j.$$

For future use, we define norms $\mathcal{J}^*, \mathcal{J}$ by

$$\mathcal{J}^*(x)^2 = |x \cdot \Gamma^{-1} x| = |\Lambda^{-1} x|^2 = \sum_{j=1}^d \sigma_j^{-2}\, (x \cdot u_j)^2, \qquad \mathcal{J}(x) = d^{-1/2}\, \mathcal{J}^*(x). \tag{1.1}$$

For simple random walk in $\mathbb{Z}^d$,

$$\Gamma = d^{-1}\, I, \qquad \mathcal{J}^*(x) = d^{1/2}\, |x|, \qquad \mathcal{J}(x) = |x|.$$

We will use $\mathcal{B}_n$ to denote the discrete ball of radius n,

$$\mathcal{B}_n = \{x \in \mathbb{Z}^d : |x| < n\},$$

and $\mathcal{C}_n$ to denote discrete ball under the norm $\mathcal{J}$,

$$\mathcal{C}_n = \{x \in \mathbb{Z}^d : \mathcal{J}(x) < n\} = \{x \in \mathbb{Z}^d : \mathcal{J}^*(x) < d^{1/2}\, n\}.$$

We choose to use $\mathcal{J}$ in the definition of $\mathcal{C}_n$ so that for simple random walk, $\mathcal{C}_n = \mathcal{B}_n$. We will write $\mathrm{R} = \mathrm{R}_p = \max\{|x| : p(x) > 0\}$ and we will call R the *range* of p. The following is very easy, but it is important enough to state as a proposition.

Proposition 1.1.1 *Suppose $p \in \mathcal{P}_d$.*

(a) There exists an $\epsilon > 0$ such that for every unit vector $u \in \mathbb{R}^d$,

$$\mathbb{E}[(X_1 \cdot u)^2] \ge \epsilon.$$

(b) If $j_1, \ldots, j_d$ are nonnegative integers with $j_1 + \cdots + j_d$ odd, then

$$\mathbb{E}[(X_1^1)^{j_1} \cdots (X_1^d)^{j_d}] = 0.$$

(c) There exists a $\delta > 0$ such that for all x,

$$\delta\, \mathcal{J}(x) \leq |x| \leq \delta^{-1}\, \mathcal{J}(x).$$

In particular,

$$\mathcal{C}_{\delta n} \subset \mathcal{B}_n \subset \mathcal{C}_{n/\delta}.$$

We note for later use that we can construct a random walk with increment distribution $p \in \mathcal{P}$ from a collection of independent one-dimensional simple random walks and an independent multinomial process. To be more precise, let $V = \{x_1, \ldots, x_l\} \in \mathcal{G}$ and let $\kappa : V \to (0,1]^l$ be as in the definition of $\mathcal{P}$. Suppose that on the same probability space we have defined l independent one-dimensional simple random walks $S_{n,1},\ S_{n,2},\ \ldots\ , S_{n,l}$ and an independent multinomial process $L_n = (L_n^1, \ldots, L_n^l)$ with probabilities $\kappa(x_1), \ldots, \kappa(x_l)$. In other words,

$$L_n = \sum_{j=1}^{n} Y_j,$$

where $Y_1, Y_2, \ldots$ are independent $\mathbb{Z}^l$-valued random variables with

$$\mathbb{P}\{Y_k = (1, 0, \ldots, 0)\} = \kappa(x_1),\ \ldots\ , \mathbb{P}\{Y_k = (0, 0, \ldots, 1)\} = \kappa(x_l)$$

and $\mathbb{P}\{Y_k = (0, 0, \ldots, 0)\} = 1 - [\kappa(x_1) + \cdots + \kappa(x_l)]$. It is easy to verify that the process

$$S_n := x_1\, S_{L_n^1,1} + x_2\, S_{L_n^2,2} + \cdots + x_l\, S_{L_n^l,l} \tag{1.2}$$

has the distribution of the random walk with increment distribution p. Essentially what we have done is to split the decision as to how to jump at time n into two decisions: first, to choose an element $x_j \in \{x_1, \ldots, x_l\}$ and then to decide whether to move by $+x_j$ or $-x_j$.

1.2 Continuous-time random walk

It is often more convenient to consider random walks in $\mathbb{Z}^d$ indexed by positive real times. Given V, κ, p as in the previous section, the *continuous-time random walk with increment distribution* p is the continuous-time Markov chain $\tilde{S}_t$ with rates p. In other words, for each $x, y \in \mathbb{Z}^d$,

$$\mathbb{P}\{\tilde{S}_{t+\Delta t} = y \mid \tilde{S}_t = x\} = p(y-x)\, \Delta t + o(\Delta t), \quad y \neq x,$$

$$\mathbb{P}\{\tilde{S}_{t+\Delta t} = x \mid \tilde{S}_t = x\} = 1 - \left[\sum_{y \neq x} p(y-x)\right] \Delta t + o(\Delta t).$$

Let $\tilde{p}_t(x,y) = \mathbb{P}\{\tilde{S}_t = y \mid \tilde{S}_0 = x\}$, and $\tilde{p}_t(y) = \tilde{p}_t(0,y) = \tilde{p}_t(x, x+y)$. Then the expressions above imply

$$\frac{d}{dt}\, \tilde{p}_t(x) = \sum_{y \in \mathbb{Z}^d} p(y)\, [\tilde{p}_t(x-y) - \tilde{p}_t(x)].$$

There is a very close relationship between the discrete time and continuous time random walks with the same increment distribution. We state this as a proposition which we leave to the reader to verify.

Proposition 1.2.1 *Suppose S_n is a (discrete-time) random walk with increment distribution p and N_t is an independent Poisson process with parameter 1. Then $\tilde{S}_t := S_{N_t}$ has the distribution of a continuous-time random walk with increment distribution p.*

There are various technical reasons why continuous-time random walks are sometimes easier to handle than discrete-time walks. One reason is that in the continuous setting there is no periodicity. If $p \in \mathcal{P}_d$, then $\tilde{p}_t(x) > 0$ for every $t > 0$ and $x \in \mathbb{Z}^d$. Another advantage can be found in the following proposition which gives an analogous, but nicer, version of (1.2). We leave the proof to the reader.

Proposition 1.2.2 *Suppose $p \in \mathcal{P}_d$ with generating set $V = \{x_1, \ldots, x_l\}$ and suppose $\tilde{S}_{t,1}, \ldots, \tilde{S}_{t,l}$ are independent one-dimensional continuous-time random walks with increment distibution $q_1, \ldots,$ q_l where $q_j(\pm 1) = p(x_j)$. Then*

$$\tilde{S}_t := x_1\, \tilde{S}_{t,1} + x_2\, \tilde{S}_{t,2} + \cdots + x_l\, \tilde{S}_{t,l} \tag{1.3}$$

has the distribution of a continuous-time random walk with increment distribution p.

If p is the increment distribution for simple random walk, we call the corresponding walk $\tilde{S}_t$ the *continuous-time simple random walk in* $\mathbb{Z}^d$. From the previous proposition, we see that the coordinates of the continuous-time simple random walk are independent — this is clearly not true for the discrete-time simple random walk. In fact, we get the following. Suppose $\tilde{S}_{t,1}, \ldots, \tilde{S}_{t,d}$ are independent one-dimensional continuous-time simple random walks. Then,

$$\tilde{S}_t := (\tilde{S}_{t/d,1}, \ldots, \tilde{S}_{t/d,d})$$

is a continuous time simple random walk in $\mathbb{Z}^d$. In particular, if $\tilde{S}_0 = 0$, then

$$\mathbb{P}\{\tilde{S}_t = (y^1, \ldots, y^d)\} = \mathbb{P}\{\tilde{S}_{t/d,1} = y^1\} \cdots \mathbb{P}\{\tilde{S}_{t/d,l} = y^l\}.$$

Remark. To verify that a discrete-time process S_n is a random walk with distribution $p \in \mathcal{P}_d$ starting at the origin, it suffices to show for all positive integers $j_1 < j_2 < \cdots < j_k$ and $x_1, \ldots, x_k \in \mathbb{Z}^d$,

$$\mathbb{P}\{S_{j_1} = x_1, \ldots, S_{j_k} = x_k\} = p_{j_1}(x_1)\, p_{j_2 - j_1}(x_2 - x_1) \cdots p_{j_k - j_{k-1}}(x_k - x_{k-1}).$$

To verify that a continuous-time process $\tilde{S}_t$ is a continuous-time random walk with distribution p starting at the origin, it suffices to show that the paths are right-continuous with probability one, and that for all real $t_1 < t_2 < \cdots < t_k$ and $x_1, \ldots, x_k \in \mathbb{Z}^d$,

$$\mathbb{P}\{\tilde{S}_{t_1} = x_1, \ldots, \tilde{S}_{t_k} = x_k\} = \tilde{p}_{t_1}(x_1)\, \tilde{p}_{t_2 - t_1}(x_2 - x_1) \cdots \tilde{p}_{t_k - t_{k-1}}(x_k - x_{k-1}).$$

1.3 Other lattices

A *lattice* $\mathbb{L}$ is a discrete additive subgroup of $\mathbb{R}^d$. The term discrete means that there is a real neighborhood of the origin whose intersection with $\mathbb{L}$ is just the origin. While this book will focus on the lattice $\mathbb{Z}^d$, we will show in this section that this also implies results for symmetric, bounded random walks on other lattices. We start by giving a proposition that classifies all lattices.

Proposition 1.3.1 *If $\mathbb{L}$ is a lattice in $\mathbb{R}^d$, then there exists an integer $k \le d$ and elements $x_1, \ldots, x_k \in \mathbb{L}$ that are linearly independent as vectors in $\mathbb{R}^d$ such that*

$$\mathbb{L} = \{j_1\, x_1 + \cdots + j_k\, x_k, \quad j_1, \ldots, j_k \in \mathbb{Z}\}.$$

In this case we call $\mathbb{L}$ a k-dimensional lattice.

Proof Suppose first that $\mathbb{L}$ is contained in a one-dimensional subspace of $\mathbb{R}^d$. Choose $x_1 \in \mathbb{L} \setminus \{0\}$ with minimal distance from the origin. Clearly $\{jx_1 : j \in \mathbb{Z}\} \subset \mathbb{L}$. Also, if $x \in \mathbb{L}$, then $jx_1 \le x < (j+1)x_1$ for some $j \in \mathbb{Z}$, but if $x > jx_1$, then $x - jx_1$ would be closer to the origin than x_1. Hence $\mathbb{L} = \{jx_1 : j \in \mathbb{Z}\}$.

More generally, suppose we have chosen linearly independent $x_1, \ldots, x_j$ such that the following holds: if $\mathbb{L}_j$ is the subgroup generated by $x_1, \ldots, x_j$, and V_j is the real subspace of $\mathbb{R}^d$ generated by the vectors $x_1, \ldots, x_j$, then $\mathbb{L} \cap V_j = \mathbb{L}_j$. If $\mathbb{L} = \mathbb{L}_j$, we stop. Otherwise, let $w_0 \in \mathbb{L} \setminus \mathbb{L}_j$ and let

$$\begin{aligned} U &= \{tw_0 : t \in \mathbb{R}, tw_0 + y_0 \in \mathbb{L} \text{ for some } y_0 \in V_j\} \\ &= \{tw_0 : t \in \mathbb{R}, tw_0 + t_1x_1 + \cdots + t_jx_j \in \mathbb{L} \text{ for some } t_1, \ldots, t_j \in [0,1]\}. \end{aligned}$$

The second equality uses the fact that $\mathbb{L}$ is a subgroup. Using the first description, we can see that U is a subgroup of $\mathbb{R}^d$ (although not necessarily contained in $\mathbb{L}$). We claim that the second description shows that there is a neighborhood of the origin whose intersection with U is exactly the origin. Indeed, the intersection of $\mathbb{L}$ with every bounded subset of $\mathbb{R}^d$ is finite (why?), and hence there are only a finite number of lattice points of the form

$$tw_0 + t_1x_1 + \cdots + t_jx_j$$

with $0 < t \le 1$; and $0 \le t_1, \ldots, t_j \le 1$. Hence there is an $\epsilon > 0$ such that there are no such lattice points with $0 < |t| \le \epsilon$. Therefore U is a one-dimensional lattice, and hence there is a $w \in U$ such that $U = \{kw : k \in \mathbb{Z}\}$. By definition, there exists a $y_1 \in V_j$ (not unique, but we just choose one) such that $x_{j+1} := w + y_1 \in \mathbb{L}$. Let $\mathbb{L}_{j+1}, V_{j+1}$ be as above using $x_1, \ldots, x_j, x_{j+1}$. Note that V_{j+1} is also the real subspace generated by $\{x_1, \ldots, x_j, w_0\}$. We claim that $\mathbb{L} \cap V_{j+1} = \mathbb{L}_{j+1}$. Indeed, suppose that $z \in \mathbb{L} \cap V_{j+1}$, and write $z = s_0w_0 + y_2$ where $y_2 \in V_j$. Then $s_0w_0 \in U$, and hence $s_0w_0 = lw$ for some integer l. Hence, we can write $z = lx_{j+1} + y_3$ with $y_3 = y_2 - ly_1 \in V_j$. But, $z - lx_{j+1} \in V_j \cap \mathbb{L} = \mathbb{L}_j$. Hence $z \in \mathbb{L}_{j+1}$. □

Comment. The proof above seems a little complicated. At first glance it seems that one might be able to simplify the argument as follows. Using the notation in the proof, we start by choosing x_1 to be a nonzero point in $\mathbb{L}$ at minimal distance from the origin, and then inductively to choose x_{j+1} to be a nonzero point in $\mathbb{L} \setminus \mathbb{L}_j$ at minimal distance from the origin. This selection method produces linearly independent $x_1, \ldots, x_k$; however, it is not always the case that

$$\mathbb{L} = \{j_1x_1 + \cdots + j_kx_k : j_1, \ldots, j_k \in \mathbb{Z}\}.$$

As an example, suppose $\mathbb{L}$ is the 5-dimensional lattice generated by

$$2\mathbf{e}_1, \ 2\mathbf{e}_2, \ 2\mathbf{e}_3, \ 2\mathbf{e}_4, \ \mathbf{e}_1 + \mathbf{e}_2 + \cdots + \mathbf{e}_5.$$

Note that $2\mathbf{e}_5 \in \mathbb{L}$ and the only nonzero points in $\mathbb{L}$ that are within distance two of the origin are $\pm 2\mathbf{e}_j, j = 1, \ldots, 5$. Therefore this selection method would choose (in some order) $\pm 2\mathbf{e}_1, \ldots, \pm 2\mathbf{e}_5$. But, $\mathbf{e}_1 + \cdots + \mathbf{e}_5$ is not in the subgroup generated by these points.

It follows from the proposition that if $k \leq d$ and $\mathbb{L}$ is a k-dimensional lattice in $\mathbb{R}^d$, then we can find a linear tranformation $A : \mathbb{R}^d \to \mathbb{R}^k$ that is an isomorphism of $\mathbb{L}$ onto $\mathbb{Z}^k$. Indeed, we define A by $A(x_j) = \mathbf{e}_j$ where $x_1, \ldots, x_k$ is a basis for $\mathbb{L}$ as in the proposition. If S_n is a bounded, symmetric, irreducible random walk taking values in $\mathbb{L}$, then $S_n^* := AS_n$ is a random walk with increment distribution $p \in \mathcal{P}_k$. Hence, results about walks on $\mathbb{Z}^k$ immediately translate to results about walks on $\mathbb{L}$. If $\mathbb{L}$ is a k-dimensional lattice in $\mathbb{R}^d$ and A is the corresponding transformation, we will call $|\det A|$ the *density* of the lattice. The term comes from the fact that as $r \to \infty$, the cardinality of the intersection of the lattice and ball of radius r in $\mathbb{R}^d$ is asymptotically equal to $|\det A|\, r^k$ times the volume of the unit ball in $\mathbb{R}^k$. In particular, if $j_1, \ldots, j_k$ are positive integers, then $(j_1\mathbb{Z}) \times \cdots \times (j_k\mathbb{Z})$ has density $(j_1 \cdots j_k)^{-1}$.

Examples.

- The *triangular lattice*, considered as a subset of $\mathbb{C} = \mathbb{R}^2$ is the lattice generated by 1 and $e^{i\pi/3}$,

$$\mathbb{L}_{\mathrm{T}} = \{k_1 + k_2\, e^{i\pi/3} : k_1, k_2 \in \mathbb{Z}\}.$$

Note that $e^{2i\pi/3} = e^{i\pi/3} - 1 \in \mathbb{L}_{\mathrm{T}}$. The triangular lattice is also considered as a graph with the above vertices and with edges connecting points that are Euclidean distance one apart. In this case, the origin has six nearest neighbors, the six sixth roots of unity. Simple random walk on the triangular lattice is the process that chooses among these six nearest neighbors equally likely. Note that this is a symmetric walk with bounded increments. The matrix

$$A = \begin{bmatrix} 1 & -1/\sqrt{3} \\ 0 & 2/\sqrt{3} \end{bmatrix}.$$

maps $\mathbb{L}_{\mathrm{T}}$ to $\mathbb{Z}^2$ sending $\{1, e^{i\pi/3}, e^{2i\pi/3}\}$ to $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_2 - \mathbf{e}_1\}$. The transformed random walk gives probability $1/6$ to the following vectors: $\pm\mathbf{e}_1, \pm\mathbf{e}_2, \pm(\mathbf{e}_2 - \mathbf{e}_1)$. Note that our transformed walk has lost some of the symmetry of the original walk.

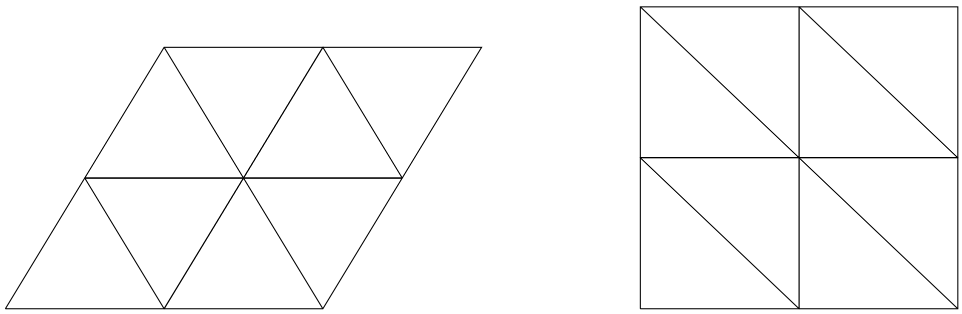

Fig 1.2. The triangular lattice $\mathbb{L}_{\mathrm{T}}$ and its transformation $A\mathbb{L}_{\mathrm{T}}$

- The *hexagonal or honeycomb lattice* is not a lattice in our sense but rather a dual graph to the triangular lattice. It can be constructed in a number of ways. One way is to start with the triangular lattice $\mathbb{L}_T$. The lattice partitions the plane into triangular regions, of which some point up and some point down. We add a vertex in the center of each triangle pointing down. The edges of this graph are the line segments from the center points to the vertices of these triangles (see figure).

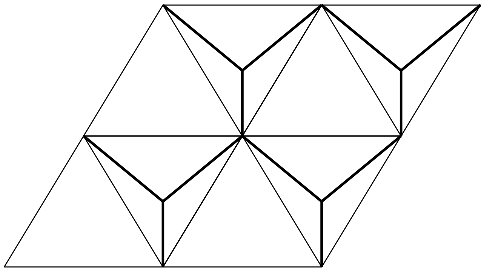

Fig 1.3. The hexagons within $\mathbb{L}_T$

Simple random walk on this graph is the process that at each time step moves to one of the three nearest neighbors. This is not a random walk in our strict sense because the increment distribution depends on whether the current position is a "center" point or a "vertex" point. However, if we start at a vertex in $\mathbb{L}_T$, the two-step distribution of this walk is the same as the walk on the triangular lattice with step distribution $p(\pm 1) = p(\pm e^{i\pi/3}) = p(\pm e^{2i\pi/3}) = 1/9; p(0) = 1/3$.

When studying random walks on other lattices $\mathbb{L}$, we can map the walk to another walk on $\mathbb{Z}^d$. However, since this might lose useful symmetries of the walk, it is sometimes better to work on the original lattice. We let $\mathcal{P}(\mathbb{L})$ be the collection of symmetric, bounded random walk distributions. If $p \in \mathcal{P}(\mathbb{L})$ and $A : \mathbb{L} \to \mathbb{Z}^d$ is an isomorphism, we will write $Ap \in \mathcal{P}_d$ for the corresponding increment distribution,

$$Ap(y) = p(A^{-1}y).$$

1.4 Generator

If $f : \mathbb{Z}^d \to \mathbb{R}$ is a function and $x \in \mathbb{Z}^d$, we define the first and second difference operators in x by

$$\nabla_x f(y) = f(y + x) - f(y),$$

$$\nabla_x^2 f(y) = \frac{1}{2} f(y + x) + \frac{1}{2} f(y - x) - f(y).$$

Note that $\nabla_x^2 = \nabla_{-x}^2$. We will sometimes write just ∇_j, ∇_j^2 for $\nabla_{\mathbf{e}_j}, \nabla_{\mathbf{e}_j}^2$. If $p \in \mathcal{P}_d$ with generator set V, then the generator $\mathcal{L} = \mathcal{L}_p$ is defined by

$$\mathcal{L}f(y) = \sum_{x \in \mathbb{Z}^d} p(x)\, \nabla_x f(y) = \sum_{x \in V} \kappa(x)\, \nabla_x^2 f(y) = -f(y) + \sum_{x \in V} p(x)\, f(x + y).$$

In the case of simple random walk, the generator is often called the *discrete Laplacian* and we will represent it by Δ_D,

$$\Delta_D f(y) = \frac{1}{d} \sum_{j=1}^{d} \nabla_j^2 f(y).$$

Remark. We have defined the discrete Laplacian in the standard way for probability. In graph theory, the discrete Laplacian of f is defined by

$$2d\Delta_D f(y) = \sum_{|x-y|=1} [f(x) - f(y)].$$

The generator of a random walk is very closely related to the walk. We will write $\mathbb{E}^x, \mathbb{P}^x$ to denote expectations and probabilities for random walk (both discrete and continuous time) assuming that $S_0 = x$ or $\tilde{S}_0 = x$. Then, it is easy to check that

$$\mathcal{L}f(y) = \mathbb{E}^y[f(S_1)] - f(y) = \frac{d}{dt}\mathbb{E}^y[f(\tilde{S}_t)]\bigg|_{t=0}.$$

(In the continuous-time case, some restrictions on the growth of f at infinity are needed.) Also, the transition probabilities $p_n(x), \tilde{p}_t(x)$ satisfy the following “heat equations”:

$$p_{n+1}(x) - p_n(x) = \mathcal{L}p_n(x), \qquad \frac{d}{dt}\tilde{p}_t(x) = \mathcal{L}\tilde{p}_t(x).$$

The derivation of these equations uses the symmetry of p. For example to derive the first, we write

$$\begin{aligned} p_{n+1}(x) &= \sum_{y\in\mathbb{Z}^d} \mathbb{P}\{S_1 = y; S_{n+1} - S_1 = x - y\} \\ &= \sum_{y\in\mathbb{Z}^d} p(y)\, p_n(x-y) \\ &= \sum_{y\in\mathbb{Z}^d} p(-y)\, p_n(x-y) = p_n(x) + \mathcal{L}p_n(x). \end{aligned}$$

The generator $\mathcal{L}$ is also closely related to a second order differential operator. If $u \in \mathbb{R}^d$ is a unit vector, let us write ∂_u^2 for the second partial derivative in the direction u. Let $\hat{\mathcal{L}}$ be the operator

$$\hat{\mathcal{L}}f(y) = \sum_{x\in V} \kappa(x)\, |x|^2\, \partial^2_{x/|x|} f(y).$$

In the case of simple random walk, $\hat{\mathcal{L}} = (2d)^{-1}\,\Delta$, where Δ denotes the usual Laplacian,

$$\Delta f(x) = \sum_{j=1}^d \partial_{x_j x_j} f(y);$$

Taylor's theorem shows that there is a c such that if f is C^4 and $y \in \mathbb{Z}^d$,

$$|\mathcal{L}f(y) - \hat{\mathcal{L}}f(y)| \le c\,\mathrm{R}^4\, M_4, \tag{1.4}$$

where R is the range of the walk and $M_4 = M_4(f, y)$ is the maximal absolute value of a fourth derivative of f for $|x - y| \le \mathrm{R}$. If the covariance matrix Γ is diagonalized,

$$\Gamma x = \sum_{j=1}^d \sigma_j^2\, (x \cdot u_j)\, u_j,$$

where $u_1, \ldots, u_d$ is an orthonormal basis, then

$$\hat{\mathcal{L}} f(y) = \frac{1}{2} \sum_{j=1}^{d} \sigma_j^2 \, \partial_{u_j}^2 \, f(y).$$

For future reference, we note that if $y \neq 0$,

$$\hat{\mathcal{L}}[\log \mathcal{J}^*(y)^2] = \hat{\mathcal{L}}[\log \mathcal{J}(y)^2] = \hat{\mathcal{L}} \left[\log \sum_{j=1}^{d} \sigma_j^{-2} \, (y \cdot u_j)^2 \right] = \frac{d-2}{\mathcal{J}^*(y)^2} = \frac{d-2}{d\, \mathcal{J}(y)^2}. \tag{1.5}$$

Comment. The estimate (1.4) uses the symmetry of p. If p is mean zero and finite range, but not necessarily symmetric, we can relate its generator to a (purely) second order differential operator, but the error involves the third derivatives of f. This only requires f to be C^3 and hence can be useful in the symmetric case as well.

1.5 Filtrations and strong Markov property

The basic property of a random walk is that the increments are independent and identically distributed. It is useful to set up a framework that allows more "information" at a particular time than just the value of the random walk. This will not affect the distribution of the random walk provided that this extra information is independent of the future increments of the walk.

A *(discrete-time) filtration* $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots$ is an increasing sequence of σ-algebras. If $p \in \mathcal{P}_d$, then we say that S_n is a random walk with increment distribution p with respect to $\{\mathcal{F}_n\}$ if:

- for each n, S_n is $\mathcal{F}_n$-measurable;
- for each $n > 0$, $S_n - S_{n-1}$ is independent of $\mathcal{F}_{n-1}$ and $\mathbb{P}\{S_n - S_{n-1} = x\} = p(x)$.

Similarly, we define a *(right continous, continuous-time) filtration* to be an increasing collection of σ-algebras $\mathcal{F}_t$ satisfying $\mathcal{F}_t = \cap_{\epsilon>0} \mathcal{F}_{t+\epsilon}$. If $p \in \mathcal{P}_d$, then we say that $\tilde{S}_t$ is a continuous-time random walk with increment distribution p with respect to $\{\mathcal{F}_t\}$ if:

- for each t, $\tilde{S}_t$ is $\mathcal{F}_t$-measurable;
- for each $s < t$, $\tilde{S}_t - \tilde{S}_s$ is independent of $\mathcal{F}_s$ and $\mathbb{P}\{\tilde{S}_t - \tilde{S}_s = x\} = \tilde{p}_{t-s}(x)$.

We let $\mathcal{F}_\infty$ denote the σ-algebra generated by the union of the $\mathcal{F}_t$ for $t > 0$.

If S_n is a random walk with respect to $\mathcal{F}_n$, and T is a random variable independent of $\mathcal{F}_\infty$, then we can add information about T to the filtration and still retain the properties of the random walk. We will describe one example of this in detail here; later on, we will do similar adding of information without being explicit. Suppose T has an exponential distribution with parameter λ, i.e., $\mathbb{P}\{T > \lambda\} = e^{-\lambda}$. Let $\mathcal{F}_n'$ denote the σ-algebra generated by $\mathcal{F}_n$ and the events $\{T \leq t\}$ for $t \leq n$. Then $\{\mathcal{F}_n'\}$ is a filtration, and S_n is a random walk with respect to $\mathcal{F}_n'$. Also, given $\mathcal{F}_n'$, then on the event $\{T > n\}$, the random variable $T - n$ has an exponential distribution with parameter λ. We can do similarly for the continuous-time walk $\tilde{S}_t$.

We will discuss stopping times and the strong Markov property. We will only do the slightly more difficult continuous-time case leaving the discrete-time analogue to the reader. If $\{\mathcal{F}_t\}$ is a filtration, then a *stopping time* with respect to $\{\mathcal{F}_t\}$ is a $[0, \infty]$-valued random variable τ such that for each t, $\{\tau \leq t\} \in \mathcal{F}_t$. Associated to the stopping time τ is a σ-algebra $\mathcal{F}_\tau$ consisting of all events A such that for each t, $A \cap \{\tau \leq t\} \in \mathcal{F}_t$. (It is straightforward to check that the set of such A is a σ-algebra.)

Theorem 1.5.1 (Strong Markov Property) *Suppose $\tilde{S}_t$ is a continuous-time random walk with increment distribution p with respect to the filtration $\{\mathcal{F}_t\}$. Suppose τ is a stopping time with respect to the process. Then on the event $\{\tau < \infty\}$ the process*

$$Y_t = \tilde{S}_{t+\tau} - \tilde{S}_\tau,$$

is a continuous-time random walk with increment distribution p independent of $\mathcal{F}_\tau$.

Proof (sketch) We will assume for ease that $\mathbb{P}\{\tau < \infty\} = 1$. Note that with probability one Y_t has right-continuous paths. We first suppose that there exists $0 = t_0 < t_1 < t_2 < \ldots$ such that with probability one $\tau \in \{t_0, t_1, \ldots\}$. Then, the result can be derived immediately, by considering the countable collection of events $\{\tau = t_j\}$. For more general τ, let τ_n be the smallest dyadic rational $l/2^n$ that is greater than τ. Then, τ_n is a stopping time and the result holds for τ_n. But,

$$Y_t = \lim_{n\to\infty} \tilde{S}_{t+\tau_n} - \tilde{S}_{\tau_n}.$$

$\square$

We will use the strong Markov property throughout this book often without being explicit about its use.

Proposition 1.5.2 (Reflection Principle.) *Suppose S_n (resp., $\tilde{S}_t$) is a random walk (resp., continuous-time random walk) with increment distribution $p \in \mathcal{P}_d$ starting at the origin.*

(a) If $u \in \mathbb{R}^d$ is a unit vector and $b > 0$,

$$\mathbb{P}\{\max_{0\leq j\leq n} S_j \cdot u \geq b\} \leq 2\,\mathbb{P}\{S_n \cdot u \geq b\},$$

$$\mathbb{P}\{\sup_{s\leq t} \tilde{S}_s \cdot u \geq b\} \leq 2\,\mathbb{P}\{\tilde{S}_t \cdot u \geq b\}.$$

(b) If $b > 0$,

$$\mathbb{P}\{\max_{0\leq j\leq n} |S_j| \geq b\} \leq 2\,\mathbb{P}\{|S_n| \geq b\},$$

$$\mathbb{P}\{\sup_{0\leq s\leq t} |\tilde{S}_t| \geq b\} \leq 2\,\mathbb{P}\{|\tilde{S}_t| \geq b\}.$$

Proof We will do the continuous-time case. To prove (a), fix $t > 0$ and a unit vector u and let $A_n = A_{n,t,b}$ be the event

$$A_n = \left\{\max_{j=1,\ldots,2^n} \tilde{S}_{jt2^{-n}} \cdot u \geq b\right\}.$$

The events A_n are increasing in n and right continuity implies that w.p.1,

$$\lim_{n\to\infty} A_n = \left\{\sup_{s\leq t} \tilde{S}_s \cdot u \geq b\right\}.$$

Hence, it suffices to show that for each n, $\mathbb{P}(A_n) \le 2\,\mathbb{P}\{\tilde S_t \cdot u \ge b\}$. Let $\tau = \tau_{n,t,b}$ be the smallest j such that $\tilde S_{jt2^{-n}} \cdot u \ge b$. Note that

$$\bigcup_{j=1}^{2^n} \left\{\tau = j; (\tilde S_t - \tilde S_{jt2^{-n}}) \cdot u \ge 0\right\} \subset \{\tilde S_t \cdot u \ge b\}.$$

Since $p \in \mathcal{P}$, symmetry implies that for all t, $\mathbb{P}\{\tilde S_t \cdot u \ge 0\} \ge 1/2$. Therefore, using independence, $\mathbb{P}\{\tau = j; (\tilde S_t - \tilde S_{jt2^{-n}}) \cdot u \ge 0\} \ge (1/2)\,\mathbb{P}\{\tau = j\}$, and hence

$$\mathbb{P}\{\tilde S_t \cdot u \ge b\} \ge \sum_{j=1}^{2^n} \mathbb{P}\left\{\tau = j; (\tilde S_t - \tilde S_{jt2^{-n}}) \cdot u \ge 0\right\} \ge \frac{1}{2} \sum_{j=1}^{2^n} \mathbb{P}\{\tau = j\} = \frac{1}{2}\,\mathbb{P}(A_n).$$

Part (b) is done similarly, by letting τ be the smallest j with $\{|\tilde S_{jt2^{-n}}| \ge b\}$ and writing

$$\bigcup_{j=1}^{2^n} \left\{\tau = j; (\tilde S_t - \tilde S_{jt2^{-n}}) \cdot \tilde S_{jt2^{-n}} \ge 0\right\} \subset \{|\tilde S_t| \ge b\}.$$

$\square$

Remark. The only fact about the distribution p that is used in the proof is that it is symmetric about the origin.

1.6 Other walks

Although we will focus primarily on $p \in \mathcal{P}$, there are times where we will want to look at more general walks. There are two classes of distributions we will be considering. We let $\mathcal{P}_d^*$ denote the set of p which generate aperiodic, irreducible walks supported on $\mathbb{Z}^d$, i.e., the set of p such that for all $x, y \in \mathbb{Z}^d$ there exists an N such that $p_n(x,y) > 0$ for $n \ge N$. We let $\mathcal{P}_d'$ be the collection of $p \in \mathcal{P}_d^*$ with mean zero and finite variance. We write $\mathcal{P}^* = \cup_d \mathcal{P}_d^*, \mathcal{P}_d' = \cup \mathcal{P}_d'$.

Note that under our definition $\mathcal{P}$ is not a subset of $\mathcal{P}'$ since $\mathcal{P}$ contains bipartite walks. However, if $p \in \mathcal{P}$ is aperiodic, then $p \in \mathcal{P}'$.

1.7 A word about constants

Throughout this book c will denote a positive constant that depends on the dimension d and the increment distribution p but does not depend on any other constants. We write

$$f(n,x) = g(n,x) + O(h(n)),$$

to mean that there exists a constant c such that for all n,

$$|f(n,x) - g(n,x)| \le c\,|h(n)|.$$

Similarly, we write

$$f(n,x) = g(n,x) + o(h(n))$$

if for every $\epsilon > 0$ there is an N such that

$$|f(n,x) - g(n,x)| \le \epsilon\,|h(n)|, \quad n \ge N.$$

Note that implicit in the definition is the fact that c, N can be chosen uniformly for all x. If f, g are positive functions, we will write

$$f(n,x) \asymp g(n,x), \qquad n \to \infty,$$

if there exists a c (again, independent of x) such that for all n, x,

$$c^{-1}\, g(n,x) \leq f(n,x) \leq c\, g(n,x).$$

We will write similarly for asymptotics of $f(t,x)$ as $t \to 0$.

As an example, let $f(z) = \log(1-z), |z| < 1$, where log denotes the branch of the complex logarithm function with $\log 1 = 0$. Then f is analytic in the unit disk with Taylor series expansion

$$\log(1-z) = -\sum_{j=1}^{\infty} \frac{z^j}{j}.$$

By the remainder estimate, for every $\epsilon > 0$,

$$\left| \log(1-z) + \sum_{j=1}^{k} \frac{z^j}{j} \right| \leq \frac{|z|^{k+1}}{\epsilon^k\,(k+1)}, \qquad |z| \leq 1-\epsilon.$$

For a fixed value of k we can write this as

$$\log(1-z) = -\left[\sum_{j=1}^{k} \frac{z^j}{j}\right] + O(|z|^{k+1}), \qquad |z| \leq 1/2, \tag{1.6}$$

or

$$\log(1-z) = -\left[\sum_{j=1}^{k} \frac{z^j}{j}\right] + O_\epsilon(|z|^{k+1}), \qquad |z| \leq 1-\epsilon, \tag{1.7}$$

where we write O_ϵ to indicate that the constant in the error term depends on ϵ.

Exercises

Exercise 1.1 Show that there are exactly $2^d - 1$ additive subgroups of $\mathbb{Z}^d$ of index 2. Describe them and show that they all can arise from some $p \in \mathcal{P}$. (A subgroup G of $\mathbb{Z}^d$ has index two if $G \neq \mathbb{Z}^d$ but $G \cup (x+G) = \mathbb{Z}^d$ for some $x \in \mathbb{Z}^d$.)

Exercise 1.2 Show that if $p \in \mathcal{P}_d$, n is a positive integer, and $x \in \mathbb{Z}^d$, then $p_{2n}(0) \geq p_{2n}(x)$.

Exercise 1.3 Show that if $p \in \mathcal{P}_d^*$, then there exists a finite set $\{x_1, \ldots, x_k\}$ such that:

- $p(x_j) > 0, \quad j = 1, \ldots, k$,
- For every $y \in \mathbb{Z}^d$, there exist (strictly) positive integers $n_1, \ldots, n_k$ with

$$n_1\, x_1 + \cdots + n_k\, x_k = y. \tag{1.8}$$

(Hint: first write each unit vector $\pm \mathbf{e}_j$ in the above form with perhaps different sets $\{x_1, \ldots, x_k\}$. Then add the equations together.)

Use this to show that there exists $\epsilon > 0, q \in \mathcal{P}_d', q' \in \mathcal{P}_d^*$ such that q has finite support and

$$p = \epsilon\, q + (1-\epsilon)\, q'.$$

Note that (1.8) is used with $y = 0$ to guarantee that q has zero mean.

Exercise 1.4 Suppose that $S_n = X_1 + \cdots + X_n$ where $X_1, X_2, \ldots$ are independent $\mathbb{R}^d$-valued random variables with mean zero and covariance matrix Γ. Show that

$$M_n := |S_n|^2 - (\mathrm{tr}\Gamma)\, n$$

is a martingale.

Exercise 1.5 Let $\mathbb{L}$ be a 2-dimensional lattice contained in $\mathbb{R}^d$ and suppose $x_1, x_2 \in \mathbb{L}$ are points such that

$$|x_1| = \min\{|x| : x \in \mathbb{L} \setminus \{0\}\},$$

$$|x_2| = \min\{|x| : x \in \mathbb{L} \setminus \{jx_1 : j \in \mathbb{Z}\}\}.$$

Show that

$$\mathbb{L} = \{j_1 x_1 + j_2 x_2 : j_1, j_2 \in \mathbb{Z}\}.$$

You may wish to compare this to the remark after Proposition 1.3.1.

Exercise 1.6 Let S_n^1, S_n^2 be independent simple random walks in $\mathbb{Z}$ and let

$$Y_n = \left(\frac{S_n^1 + S_n^2}{2}, \quad \frac{S_n^1 - S_n^2}{2}\right),$$

Show that Y_n is a simple random walk in $\mathbb{Z}^2$.

Exercise 1.7 Suppose S_n is a random walk w ith increment distribution $p \in \mathcal{P}^* \cup \mathcal{P}$. Show that there exists an $\epsilon > 0$ such that for every unit vector $\theta \in \mathbb{R}^d$, $\mathbb{P}\{S_1 \cdot \theta \geq \epsilon\} \geq \epsilon$.

2
Local Central Limit Theorem

2.1 Introduction

If $X_1, X_2, \ldots$ are independent, identically distributed random variables in $\mathbb{R}$ with mean zero and variance σ^2, then the central limit theorem (CLT) states that the distribution of

$$\frac{X_1 + \cdots + X_n}{\sqrt{n}} \tag{2.1}$$

approaches that of a normal distribution with mean zero and variance σ^2. In other words, for $-\infty < r < s < \infty$,

$$\lim_{n\to\infty} \mathbb{P}\left\{ r \le \frac{X_1 + \cdots + X_n}{\sqrt{n}} \le s \right\} = \int_r^s \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{y^2}{2\sigma^2}}\, dy.$$

If $p \in \mathcal{P}_1$ is aperiodic with variance σ^2, we can use this to motivate the following approximation:

$$p_n(k) = \mathbb{P}\{S_n = k\} = \mathbb{P}\left\{ \frac{k}{\sqrt{n}} \le \frac{S_n}{\sqrt{n}} < \frac{k+1}{\sqrt{n}} \right\}$$

$$\approx \int_{k/\sqrt{n}}^{(k+1)/\sqrt{n}} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{y^2}{2\sigma^2}}\, dy \approx \frac{1}{\sqrt{2\pi\sigma^2 n}} \exp\left\{ -\frac{k^2}{2\sigma^2 n} \right\}.$$

Similarly, if $p \in \mathcal{P}_1$ is bipartite, we can conjecture that

$$p_n(k) + p_n(k+1) \approx \int_{k/\sqrt{n}}^{(k+2)/\sqrt{n}} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{y^2}{2\sigma^2}}\, dy \approx \frac{2}{\sqrt{2\pi\sigma^2 n}} \exp\left\{ -\frac{k^2}{2\sigma^2 n} \right\}.$$

The local central limit theorem (LCLT) justifies this approximation.

Comment. One gets a better approximation by writing

$$\mathbb{P}\{S_n = k\} = \mathbb{P}\left\{ \frac{k-\frac{1}{2}}{\sqrt{n}} \le \frac{S_n}{\sqrt{n}} < \frac{k+\frac{1}{2}}{\sqrt{n}} \right\} \approx \int_{(k-\frac{1}{2})/\sqrt{n}}^{(k+\frac{1}{2})/\sqrt{n}} \frac{1}{\sqrt{2\pi\sigma^2 n}}\, e^{-\frac{y^2}{2\sigma^2}}\, dy.$$

If $p \in \mathcal{P}_d$ with covariance matrix $\Gamma = \Lambda\,\Lambda^T = \Lambda^2$, then the normalized sums (2.1) approach a joint normal random variable with covariance matrix Γ, i.e., a random variable with density

$$f(x) = \frac{1}{(2\pi)^{d/2}\,(\det\Lambda)}\, e^{-|\Lambda^{-1}x|^2/2} = \frac{1}{(2\pi)^{d/2}\,\sqrt{\det\Gamma}}\, e^{-(x\cdot\Gamma^{-1}x)/2}.$$

(See Section 11.3 for a review of the joint normal distribution.) A similar heuristic argument can be given for $p_n(x)$. Recall from (1.1) that $\mathcal{J}^*(x)^2 = x \cdot \Gamma^{-1} x$. Let $\overline{p}_n(x)$ denote the estimate of $p_n(x)$ that one obtains by the central limit theorem argument,

$$\overline{p}_n(x) = \frac{1}{(2\pi n)^{d/2}\sqrt{\det \Gamma}}\, e^{-\frac{\mathcal{J}^*(x)^2}{2n}} = \frac{1}{(2\pi)^d\, n^{d/2}} \int_{\mathbb{R}^d} e^{i\frac{s\cdot x}{\sqrt{n}}}\, e^{-\frac{s\cdot \Gamma s}{2}}\, d^d s. \tag{2.2}$$

The second equality is a straightforward computation, see (11.14). We define $\overline{p}_t(x)$ for real $t > 0$ in the same way. The LCLT states that for large n, $p_n(x)$ is approximately $\overline{p}_n(x)$. To be more precise, we will say that an aperiodic p satisfies the *LCLT* if

$$\lim_{n\to\infty} n^{d/2} \sup_{x\in\mathbb{Z}^d} |p_n(x) - \overline{p}_n(x)| = 0.$$

A bipartite p satisfies the LCLT if

$$\lim_{n\to\infty} n^{d/2} \sup_{x\in\mathbb{Z}^d} |p_n(x) + p_{n+1}(x) - 2\overline{p}_n(x)| = 0.$$

In this weak form of the LCLT we have not made any estimate of the error term $|p_n(x) - \overline{p}_n(x)|$ other than that it goes to zero faster than $n^{-d/2}$ uniformly in x. Note that $\overline{p}_n(x)$ is bounded by $c\,n^{-d/2}$ uniformly in x. This is the correct order of magnitude for $|x|$ of order $\sqrt{n}$ but $\overline{p}_n(x)$ is much smaller for larger $|x|$. We will prove a LCLT for any mean zero distribution with finite second moment. However, the LCLT we state now for $p \in \mathcal{P}_d$ includes error estimates that do not hold for all $p \in \mathcal{P}'_d$.

Theorem 2.1.1 (Local Central Limit Theorem) *If $p \in \mathcal{P}_d$ is aperiodic, and $\overline{p}_n(x)$ is as defined in (2.2), then there is a c and for every integer $k \geq 4$ there is a $c(k) < \infty$ such that for all integers $n > 0$ and $x \in \mathbb{Z}^d$ the following holds where $z = x/\sqrt{n}$:*

$$|p_n(x) - \overline{p}_n(x)| \leq \frac{c(k)}{n^{(d+2)/2}} \left[(|z|^k + 1)\, e^{-\frac{\mathcal{J}^*(z)^2}{2}} + \frac{1}{n^{(k-3)/2}}\right], \tag{2.3}$$

$$|p_n(x) - \overline{p}_n(x)| \leq \frac{c}{n^{(d+2)/2}\,|z|^2}. \tag{2.4}$$

We will prove this result in a number of steps in Section 2.3. Before doing so, let us consider what the theorem states. Plugging $k = 4$ into (2.3) implies that

$$|p_n(x) - \overline{p}_n(x)| \leq \frac{c}{n^{(d+2)/2}}. \tag{2.5}$$

For "typical" x with $|x| \leq \sqrt{n}$, $\overline{p}_n(x) \asymp n^{-d/2}$. Hence (2.5) implies

$$p_n(x) = \overline{p}_n(x)\left[1 + O\left(\frac{1}{n}\right)\right], \qquad |x| \leq \sqrt{n}.$$

The error term in (2.5) is uniform over x, but as $|x|$ grows, the ratio between the error term and $p_n(x)$ grows. The inequalities (2.3) and (2.4) are improvements on the error term for $|x| \geq \sqrt{n}$. Since $p_n(x) \sim n^{-d/2}\, e^{-\mathcal{J}^*(x)^2/2}$, (2.3) implies

$$p_n(x) = \overline{p}_n(x)\left[1 + \frac{O_k(|x/\sqrt{n}|^k)}{n}\right] + O_k\left(\frac{1}{n^{(d+k-1)/2}}\right), \qquad |x| \geq \sqrt{n},$$

where we write O_k to emphasize that the constant in the error term depends on k.

An even better improvement is established in Section 2.3.1 for $d = 1$ where it is shown that

$$p_n(x) = \overline{p}_n(x)\ \exp\left\{O\left(\frac{1}{n} + \frac{|x|^4}{n^3}\right)\right\}.$$

Although Theorem 2.1.1 is not as useful for atypical x, simple large deviation results as given in the next propositions often suffice to estimate probabilities.

Proposition 2.1.2

- *Suppose $p \in \mathcal{P}_d'$ and S_n is a p-walk starting at the origin. Suppose k is a positive integer such that $\mathbb{E}[|X_1|^{2k}] < \infty$. There exists $c < \infty$ such that for all $s > 0$*

$$\mathbb{P}\left\{\max_{0\le j\le n} |S_j| \ge s\sqrt{n}\right\} \le c\, s^{-2k}. \tag{2.6}$$

- *Suppose $p \in \mathcal{P}_d$ and S_n is a p-walk starting at the origin. There exist $\beta > 0$ and $c < \infty$ such that for all n and all $s > 0$,*

$$\mathbb{P}\left\{\max_{0\le j\le n} |S_j| \ge s\sqrt{n}\right\} \le c\, e^{-\beta s^2}. \tag{2.7}$$

Proof It suffices to prove the results for one-dimensional walks. See Corollaries 11.2.6 and 11.2.7. □

Comment. The statement of the LCLT given here is stronger than is needed for many applications. For example, to determine whether the random walk is recurrent or transient, we only need the following corollary. If $p \in \mathcal{P}_d$ is aperiodic, then there exist $0 < c_1 < c_2 < \infty$ such that for all x, $p_n(x) \le c_2\, n^{-d/2}$, and for $|x| \le \sqrt{n}$, $p_n(x) \ge c_1\, n^{-d/2}$. The exponent $d/2$ is important to remember and can be understood easily. In n steps, the random walk tends to go distance $\sqrt{n}$. In $\mathbb{Z}^d$, there are of order $n^{d/2}$ points within distance $\sqrt{n}$ of the origin. Therefore, the probability of being at a particular point should be of order $n^{-d/2}$.

The proof of Theorem 2.1.1 in Section 2.2 will use the characteristic function. We discuss LCLTs for $p \in \mathcal{P}_d'$, where, as before, $\mathcal{P}_d'$ denotes the set of aperiodic, irreducible increment distributions p in $\mathbb{Z}^d$ with mean zero and bounded second moment. In the proof of Theorem 2.1.1, we will see that we do not need to assume that the increments are bounded. For fixed $k \ge 4$, (2.3) holds for $p \in \mathcal{P}_d'$ provided that $\mathbb{E}[|X|^{k+1}] < \infty$ and the third moments of p vanish. The inequalities (2.5) and (2.4) need only finite fourth moments and vanishing third moments. If $p \in \mathcal{P}_d'$ has finite third moments that are nonzero, we can prove a weaker version of (2.3). Suppose $k \ge 3$, and $\mathbb{E}[|X_1|^{k+1}] < \infty$. There exists $c(k) < \infty$ such that

$$|p_n(x) - \overline{p}_n(x)| \le \frac{c(k)}{n^{(d+1)/2}}\left[(|z|^k + 1)\, e^{-\frac{\mathcal{J}^*(z)^2}{2}} + \frac{1}{n^{(k-2)/2}}\right].$$

Also, for any $p \in \mathcal{P}_d'$ with $\mathbb{E}[|X_1|^3] < \infty$,

$$|p_n(x) - \overline{p}_n(x)| \le \frac{c}{n^{(d+1)/2}}, \qquad |p_n(x) - \overline{p}_n(x)| \le \frac{c}{n^{(d-1)/2}\,|x|^2}.$$

We focus our discussion in Section 2.2 on aperiodic, discrete-time walks, but the next theorem shows that we can deduce the results for bipartite and continuous-time walks from LCLT for aperiodic, discrete-time walks. We state the analogue of (2.3); the analogue of (2.4) can be proved similarly.

Theorem 2.1.3 *If $p \in \mathcal{P}_d$ and $\overline{p}_n(x)$ is as defined in (2.2), then for every $k \geq 4$ there is a $c = c(k) < \infty$ such that the follwing holds for all $x \in \mathbb{Z}^d$.*

- *If n is a positive integer and $z = x/\sqrt{n}$, then*

$$|p_n(x) + p_{n+1}(x) - 2\,\overline{p}_n(x)| \leq \frac{c}{n^{(d+2)/2}} \left[(|z|^k + 1)\, e^{-\mathcal{J}^*(z)^2/2} + \frac{1}{n^{(k-3)/2}}\right]. \tag{2.8}$$

- *If f $t > 0$ and $z = x/\sqrt{t}$,*

$$|\tilde{p}_t(x) - \overline{p}_t(x)| \leq \frac{c}{t^{(d+2)/2}} \left[(|z|^k + 1)\, e^{-\mathcal{J}^*(y)^2/2} + \frac{1}{t^{(k-3)/2}}\right]. \tag{2.9}$$

Proof (assuming Theorem 2.1.1) We only sketch the proof. If $p \in \mathcal{P}_d$ is bipartite, then $S_n^* := S_{2n}$ is an aperiodic walk on the lattice $\mathbb{Z}_e^d$. We can establish the result for S_n^* by mapping $\mathbb{Z}_e^d$ to $\mathbb{Z}^d$ as described in Section 1.3. This gives the asymptotics for $p_{2n}(x), x \in \mathbb{Z}_e^d$ and for $x \in \mathbb{Z}_o^d$, we know that

$$p_{2n+1}(x) = \sum_{y \in \mathbb{Z}^d} p_{2n}(x-y)\, p(y).$$

The continuous-time walk viewed at integer times is the discrete-time walk with increment distribution $\tilde{p} = \tilde{p}_1$. Since $\tilde{p}$ satisfies all the moment conditions, (2.3) holds for $\tilde{p}_n(x), n = 0, 1, 2, \ldots$. If $0 < t < 1$, we can write

$$\tilde{p}_{n+t}(x) = \sum_{y \in \mathbb{Z}^d} \tilde{p}_n(x-y)\, \tilde{p}_t(y),$$

and deduce the result for all t.

□

2.2 Characteristic Functions and LCLT

2.2.1 Characteristic functions of random variables in $\mathbb{R}^d$

One of the most useful tools for studying the distribution of the sums of independent random variables is the characteristic function. If $X = (X^1, \ldots, X^d)$ is a random variable in $\mathbb{R}^d$, then its *characteristic function* $\phi = \phi_X$ is the function from $\mathbb{R}^d$ into $\mathbb{C}$ given by

$$\phi(\theta) = \mathbb{E}[\exp\{i\theta \cdot X\}].$$

Proposition 2.2.1 *Suppose $X = (X^1, \ldots, X^d)$ is a random variable in $\mathbb{R}^d$ with characteristic function ϕ.*

(a) ϕ is a uniformly continuous function with $\phi(0) = 1$ and $|\phi(\theta)| \leq 1$ for all $\theta \in \mathbb{R}^d$.

(b) If $\theta \in \mathbb{R}^d$ then $\phi_{X,\theta}(s) := \phi(s\theta)$ is the characteristic function of the one-dimensional random variable $X \cdot \theta$.

(c) Suppose $d = 1$ and m is a positive integer with $\mathbb{E}[|X|^m] < \infty$. Then $\phi(s)$ is a C^m function of s; in fact,

$$\phi^{(m)}(s) = i^m \, \mathbb{E}[X^m e^{isX}].$$

(d) If m is a positive integer and $\mathbb{E}[|X|^m] < \infty$ and $|u| = 1$, then

$$\left| \phi(su) - \sum_{j=0}^{m-1} \frac{i^j \, \mathbb{E}[(X \cdot u)^j]}{j!} \, s^j \right| \leq \frac{\mathbb{E}[|X \cdot u|^m]}{m!} \, |s|^m.$$

(e) If $X_1, X_2, \ldots, X_n$ are independent random varaibles in $\mathbb{R}^d$, with characteristic functions $\phi_{X_1}, \ldots, \phi_{X_n}$, then

$$\phi_{X_1+\cdots+X_n}(\theta) = \phi_{X_1}(\theta) \cdots \phi_{X_n}(\theta).$$

In particular, if $X_1, X_2, \ldots$ are independent, identically distributed with the same distribution as X, then the characteristic function of $S_n = X_1 + \cdots + X_n$ is given by

$$\phi_{S_n}(\theta) = [\phi(\theta)]^n.$$

Proof To prove uniform continuity, note that

$$|\phi(\theta_1 + \theta) - \phi(\theta)| = |\mathbb{E}[e^{iX(\theta_1+\theta)} - e^{iX\theta}]| \leq \mathbb{E}[|e^{iX\theta_1} - 1|],$$

and the dominated convergence theorem implies that

$$\lim_{\theta_1 \to 0} \mathbb{E}[|e^{iX\theta_1} - 1|] = 0.$$

The other statements in (a) and (b) are immediate. Part (c) is derived by differentiating; the condition $\mathbb{E}[|X|^m] < \infty$ is needed to justify the differentiation using the dominated convergence theorem (details omitted). Part (d) follows from (b), (c), and Taylor's theorem with remainder. Part (e) is immediate from the product rule for expectations of independent random variables. □

We will write $P_m(\theta)$ for the m-th order Taylor series approximation of ϕ about the origin. Then the last proposition implies that if $\mathbb{E}[|X|^m] < \infty$, then

$$\phi(\theta) = P_m(\theta) + o(|\theta|^m), \quad \theta \to 0. \tag{2.10}$$

Note that if $\mathbb{E}[X] = 0$ and $\mathbb{E}[|X|^2] < \infty$, then

$$P_2(\theta) = 1 - \frac{1}{2} \sum_{j=1}^{d} \sum_{k=1}^{d} \mathbb{E}[X^j X^k] \, \theta^j \, \theta^k = 1 - \frac{\theta \cdot \Gamma\theta}{2} = 1 - \frac{\mathbb{E}[(X \cdot \theta)^2]}{2}.$$

Here Γ denotes the covariance matrix for X. If $\mathbb{E}[|X|^m] < \infty$, we write

$$P_m(\theta) = 1 - \frac{\theta \cdot \Gamma\theta}{2} + \sum_{j=3}^{m} q_j(\theta), \tag{2.11}$$

where q_j are homogeneous polynomials of degree three determined by the moments of X. Also, if all the third moments of X exist and equal zero, $q_3 \equiv 0$. If X has a symmetric distribution, then $q_j \equiv 0$ for all odd j for which $\mathbb{E}[|X_1|^j] < \infty$.

2.2.2 Characteristic functions of random variables in $\mathbb{Z}^d$

If $X = (X^1, \ldots, X^d)$ is a $\mathbb{Z}^d$-valued random variable, then its characteristic function has period 2π in each variable, i.e., if $k_1, \ldots, k_d$ are integers,

$$\phi(\theta^1, \ldots, \theta^d) = \phi(\theta^1 + 2k_1\pi, \ldots, \theta^d + 2k_d\pi).$$

The characteristic function determines the distribution of X; in fact, the next proposition gives a simple inversion formula. Here, and for the remainder of this section, we will write $d\theta$ for $d\theta^1 \cdots d\theta^d$.

Proposition 2.2.2 *If $X = (X^1, \ldots, X^d)$ is a $\mathbb{Z}^d$-valued random variable with characteristic function ϕ, then for every $x \in \mathbb{Z}^d$,*

$$\mathbb{P}\{X = x\} = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \phi(\theta)\, e^{-ix\cdot\theta}\, d\theta.$$

Proof Since

$$\phi(\theta) = \mathbb{E}[e^{iX\cdot\theta}] = \sum_{y\in\mathbb{Z}^d} e^{iy\cdot\theta}\, \mathbb{P}\{X = y\},$$

we get

$$\int_{[-\pi,\pi]^d} \phi(\theta)\, e^{-ix\cdot\theta}\, d\theta = \sum_{y\in\mathbb{Z}^d} \mathbb{P}\{X = y\} \int_{[-\pi,\pi]^d} e^{i(y-x)\cdot\theta}\, d\theta.$$

(The dominated convergence theorem justifes the interchange of the sum and the integral.) But, if $x, y \in \mathbb{Z}^d$,

$$\int_{[-\pi,\pi]^d} e^{i(y-x)\cdot\theta}\, d\theta = \begin{cases} (2\pi)^d, & y = x \\ 0, & y \neq x. \end{cases}$$

□

Corollary 2.2.3 *Suppose $X_1, X_2, \ldots$ are independent, identically distributed random variables in $\mathbb{Z}^d$ with characteristic function ϕ. Let $S_n = X_1 + \cdots + X_n$. Then, for all $x \in \mathbb{Z}^d$,*

$$\mathbb{P}\{S_n = x\} = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \phi^n(\theta)\, e^{-ix\cdot\theta}\, d\theta.$$

2.3 LCLT — characteristic function approach

In some sense, Corollary 2.2.3 completely solves the problem of determining the distribution of a random walk at a particular time n. However, the integral is generally hard to evaluate and estimation of oscillatory integrals is tricky. Fortunately, we can use this corollary as a starting point for deriving the local central limit theorem. We will consider $p \in \mathcal{P}'$ in this section. Here, as before, we write $p_n(x)$ for the distribution of $S_n = X_1 + \cdots + X_n$ where $X_1, \ldots, X_n$ are independent with distribution p. We also write $\tilde{S}_t$ for a continuous time walk with rates p. We let ϕ denote the characteristic function of p,

$$\phi(\theta) = \sum_{x\in\mathbb{Z}^d} e^{i\theta\cdot x}\, p(x).$$

We have noted that the characteristic function of S_n is ϕ^n.

Lemma 2.3.1 *The characteristic function of $\tilde{S}_t$ is*

$$\phi_{\tilde{S}_t}(\theta) = \exp\{t[\phi(\theta) - 1]\}.$$

Proof Since $\tilde{S}_t$ has the same distribution as S_{N_t} where N_t is an independent Poisson process with parameter 1, we get

$$\phi_{\tilde{S}_t}(\theta) = \mathbb{E}[e^{i\theta\cdot\tilde{S}_t}] = \sum_{j=0}^{\infty} e^{-t}\,\frac{t^j}{j!}\,\mathbb{E}[e^{i\theta\cdot S_j}] = \sum_{j=0}^{\infty} e^{-t}\,\frac{t^j}{j!}\,\phi(\theta)^j = \exp\{t[\phi(\theta) - 1]\}.$$

□

Corollary 2.2.3 gives the formulas

$$p_n(x) = \frac{1}{(2\pi)^d}\int_{[-\pi,\pi]^d} \phi^n(\theta)\, e^{-i\theta\cdot x}\, d\theta, \tag{2.12}$$

$$\tilde{p}_t(x) = \frac{1}{(2\pi)^d}\int_{[-\pi,\pi]^d} e^{t[\phi(\theta)-1]}\, e^{-i\theta\cdot x}\, d\theta.$$

Lemma 2.3.2 *Suppose $p \in \mathcal{P}'$.*

(a) For every $\epsilon > 0$,

$$\sup\Big\{|\phi(\theta)| : \theta \in [-\pi,\pi]^d, |\theta| \geq \epsilon\Big\} < 1.$$

(b) There is a $b > 0$ such that for all $\theta \in [-\pi,\pi]^d$,

$$|\phi(\theta)| \leq 1 - b|\theta|^2. \tag{2.13}$$

In particular for all $\theta \in [-\pi,\pi]^d$, and $r > 0$,

$$|\phi(\theta)|^r \leq \left[1 - b|\theta|^2\right]^r \leq \exp\left\{-\frac{r}{b|\theta|^2}\right\}. \tag{2.14}$$

Proof By continuity and compactness, to prove (a) it suffices to prove that $|\phi(\theta)| < 1$ for all $\theta \in [-\pi,\pi]^d \setminus \{0\}$. To see this, suppose that $|\phi(\theta)| = 1$. Then $|\phi(\theta)^n| = 1$ for all positive integers n. Since

$$\phi(\theta)^n = \sum_{z\in\mathbb{Z}^d} p_n(z)\, e^{iz\cdot\theta},$$

and for each fixed z, $p_n(z) > 0$ for all sufficiently large n, we see that $e^{iz\cdot\theta} = 1$ for all $z \in \mathbb{Z}^d$. (Here we use the fact that if $w_1, w_2, \ldots \in \mathbb{C}$ with $|w_1 + w_2 + \cdots| = 1$ and $|w_1| + |w_2| + \cdots = 1$, then there is a ψ such that $w_j = r_j e^{i\psi}$ with $r_j \geq 0$.) The only $\theta \in [-\pi,\pi]^d$ that satisfies this is $\theta = 0$. Using (a), it suffices to prove (2.13) in a neighborhood of the origin, and this follows from the second-order Taylor series expansion (2.10). □

Comment. The last lemma does not hold for bipartite p, since in this case $\phi(\pi i, \pi i, \ldots, \pi i) = -1$.

In order to illustrate the proof of the local central limit theorem using the characteristic function, we will consider the one-dimensional case with $p(1) = p(-1) = 1/4$ and $p(0) = 1/2$. Note that this increment distribution is the same as the two-step distribution of (1/2 times) the usual simple random walk. The characteristic function for p is

$$\phi(\theta) = \frac{1}{2} + \frac{1}{4}\, e^{i\theta} + \frac{1}{4}\, e^{-i\theta} = \frac{1}{2} + \frac{1}{2}\, \cos\theta = 1 - \frac{\theta^2}{4} + O(\theta^4).$$

The inversion formula (2.12) tells us that

$$p_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ix\theta}\, \phi(\theta)^n \, d\theta = \frac{1}{2\pi\sqrt{n}} \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} e^{-i(x/\sqrt{n})s}\, \phi(s/\sqrt{n})^n \, ds.$$

The second equality follows from the substitution $s = \theta\sqrt{n}$. For $|s| \le \pi\sqrt{n}$, we can write

$$\phi\left(\frac{s}{\sqrt{n}}\right) = 1 - \frac{s^2}{4n} + O\left(\frac{s^4}{n^2}\right) = 1 - \frac{(s^2/4) + O(s^4/n)}{n}.$$

We can find $\delta > 0$ such that if $|s| \le \delta\sqrt{n}$,

$$\left| \frac{s^2}{4} + O\left(\frac{s^4}{n}\right) \right| \le \frac{n}{2}.$$

Therefore, using (11.3), if $|s| \le \delta\sqrt{n}$,

$$\phi\left(\frac{s}{\sqrt{n}}\right)^n = \left[1 - \frac{s^2}{4n} + O\left(\frac{s^4}{n^2}\right)\right]^n = e^{-s^2/4}\, e^{g(s,n)},$$

where

$$|g(s,n)| \le c\,\frac{s^4}{n}.$$

If $\epsilon = \min\{\delta, 1/\sqrt{8c}\}$ we also have

$$|g(s,n)| \le \frac{s^2}{8}, \qquad |s| \le \epsilon\sqrt{n}.$$

For $\epsilon\sqrt{n} < |s| \le \pi\sqrt{n}$, (2.13) shows that $|e^{-i(x/\sqrt{n})s}\, \phi(s/\sqrt{n})^n| \le e^{-\beta n}$ for some $\beta > 0$. Hence, up to an error that is exponentially small in n, $p_n(x)$ equals

$$\frac{1}{2\pi\sqrt{n}} \int_{-\epsilon\sqrt{n}}^{\epsilon\sqrt{n}} e^{-i(x/\sqrt{n})s}\, e^{-s^2/4}\, e^{g(s,n)} \, ds.$$

We now use

$$|e^{g(s,n)} - 1| \le \begin{cases} c\, s^4/n, & |s| \le n^{1/4} \\ e^{s^2/8}, & n^{1/4} < |s| \le \epsilon\sqrt{n} \end{cases}$$

to bound the error term as follows:

$$\left| \frac{1}{2\pi\sqrt{n}} \int_{-\epsilon\sqrt{n}}^{\epsilon\sqrt{n}} e^{-i(x/\sqrt{n})s}\, e^{-s^2/4}\, [e^{g(s,n)} - 1] \, ds \right| \le \frac{c}{\sqrt{n}} \int_{-\epsilon\sqrt{n}}^{\epsilon\sqrt{n}} e^{-s^2/4}\, |e^{g(s,n)} - 1| \, ds,$$

$$\int_{-n^{-1/4}}^{n^{1/4}} e^{-s^2/4}\,|e^{g(s,n)}-1|\,ds \le \frac{c}{n}\int_{-\infty}^{\infty} s^4\,e^{-s^2/4}\,ds \le \frac{c}{n},$$

$$\int_{n^{1/4}\le|s|\le\epsilon\sqrt{n}} e^{-s^2/4}\,|e^{g(s,n)}-1|\,ds \le \int_{|s|\ge n^{1/4}} e^{-s^2/8}\,ds = o(n^{-1}).$$

Hence we have

$$\begin{aligned} p_n(x) &= O\left(\frac{1}{n^{3/2}}\right) + \frac{1}{2\pi\sqrt{n}}\int_{-\epsilon\sqrt{n}}^{\epsilon\sqrt{n}} e^{-i(x/\sqrt{n})s}\,e^{-s^2/4}\,ds \\ &= O\left(\frac{1}{n^{3/2}}\right) + \frac{1}{2\pi\sqrt{n}}\int_{-\infty}^{\infty} e^{-i(x/\sqrt{n})s}\,e^{-s^2/4}\,ds. \end{aligned}$$

The last term equals $\overline{p}_n(x)$, see (2.2), and so we have shown that

$$p_n(x) = \overline{p}_n(x) + O\left(\frac{1}{n^{3/2}}\right).$$

We will follow this basic line of proof for theorems in this subsection. Before proceeding, it will be useful to outline the main steps.

- Expand $\log\phi(\theta)$ in a neighborhood $|\theta| < \epsilon$ about the origin.
- Use this expansion to approximate $[\phi(\theta/\sqrt{n})]^n$, which is the characteristic function of $S_n/\sqrt{n}$.
- Use the change of variables by $s = \theta\sqrt{n}$ in the inversion formula to yields an integral over $[-\pi\sqrt{n},\pi\sqrt{n}]^d$. Use Lemma 2.3.2 to show that the integral over $|\theta| \ge \epsilon\sqrt{n}$ is exponentially small.
- Use the approximation of $[\phi(\theta/\sqrt{n})]^n$ to compute the dominant term and to give an expression for the error term that needs to be estimated.
- Estimate the error term.

Our first lemma discusses the approximation of the characteristic function of $S_n/\sqrt{n}$ by an exponential. We state the lemma for all $p \in \mathcal{P}_d'$, and then give examples to show how to get sharper results if one makes assumptions on the moments.

Lemma 2.3.3 *Suppose $p \in \mathcal{P}_d'$ with covariance matrix Γ and characteristic function ϕ that we write as*

$$\phi(\theta) = 1 - \frac{\theta\cdot\Gamma\theta}{2} + h(\theta),$$

where $h(\theta) = o(|\theta|^2)$ as $\theta \to 0$. There exists $\epsilon > 0, c < \infty$ such that for all positive integers n and all $|\theta| \le \epsilon\sqrt{n}$, we can write

$$\left[\phi\left(\frac{\theta}{\sqrt{n}}\right)\right]^n = \exp\left\{-\frac{\theta\cdot\Gamma\theta}{2} + g(\theta,n)\right\} = e^{-\frac{\theta\cdot\Gamma\theta}{2}}\,[1+F_n(\theta)], \tag{2.15}$$

where $F_n(\theta) = e^{g(\theta,n)} - 1$ and

$$|g(\theta,n)| \le \min\left\{\frac{\theta\cdot\Gamma\theta}{4},\ n\left|h\left(\frac{\theta}{\sqrt{n}}\right)\right| + \frac{c|\theta|^4}{n^2}\right\}. \tag{2.16}$$

In particular,

$$|F_n(\theta)| \le e^{\frac{\theta\cdot\Gamma\theta}{4}} + 1.$$

Proof Choose $\delta > 0$ such that

$$|\phi(\theta) - 1| \leq \frac{1}{2}, \quad |\theta| \leq \delta.$$

For $|\theta| \leq \delta$, we can write

$$\log \phi(\theta) = -\frac{\theta \cdot \Gamma\theta}{2} + h(\theta) - \frac{(\theta \cdot \Gamma\theta)^2}{8} + O(|h(\theta)|\,|\theta|^2) + O(|\theta|^6). \tag{2.17}$$

Define $g(\theta, n)$ by

$$n \log \phi\left(\frac{\theta}{\sqrt{n}}\right) = -\frac{\theta \cdot \Gamma\theta}{2} + g(\theta, n),$$

so that (2.15) holds. Note that

$$|g(\theta, n)| \leq n \left| h\left(\frac{\theta}{\sqrt{n}}\right)\right| + O\left(\frac{|\theta|^4}{n}\right).$$

Since $n\, h(\theta/\sqrt{n}) = o(|\theta|^2)$, we can find $0 < \epsilon \leq \delta$ such that for $|\theta| \leq \epsilon\sqrt{n}$,

$$|g(\theta, n)| \leq \frac{\theta \cdot \Gamma\theta}{4}.$$

□

Comment. The proofs will require estimates for $F_n(\theta)$. The inequality $|e^z - 1| \leq O(|z|)$ is valid if z is restricted to a bounded set. Hence, the basic strategy is to find $c_1, r(n) \leq O(n^{1/4})$ such that

$$n \left| h\left(\frac{\theta}{\sqrt{n}}\right)\right| \leq c_1, \quad |\theta| \leq r(n)$$

Since $O(|\theta|^4/n) \leq O(1)$ for $|\theta| \leq n^{1/4}$, (2.16) implies

$$|F_n(\theta)| = \left|e^{g(\theta,n)} - 1\right| \leq c\,|g(\theta, n)| \leq c\left[n\left|h\left(\frac{\theta}{\sqrt{n}}\right)\right| + O\left(\frac{|\theta|^4}{n^2}\right)\right], \quad |\theta| \leq r(n),$$

$$|F_n(\theta)| \leq e^{\frac{\theta\cdot\Gamma\theta}{4}} + 1, \quad |\theta| \geq r(n).$$

Examples

We give some examples with different moment assumptions. In the discussion below, ϵ is as in the lemma and θ is restricted to $|\theta| \leq \epsilon\sqrt{n}$.

- If $\mathbb{E}[|X_1|^4] < \infty$, then by (2.11),

  $$h(\theta) = q_3(\theta) + O(|\theta|^4),$$

  and

  $$\log \phi(\theta) = -\frac{\theta \cdot \Gamma\theta}{2} + f_3(\theta) + O(|\theta|^4),$$

  where $f_3 = q_3$ is a homogeneous polynomial of degree three. In this case

  $$g(\theta, n) = n\, f_3\left(\frac{\theta}{\sqrt{n}}\right) + \frac{O(|\theta|^4)}{n}, \tag{2.18}$$

and there exists $c < \infty$ such that

$$|g(\theta,n)| \leq \min\left\{\frac{\theta\cdot\Gamma\theta}{4}, \frac{c\,|\theta|^3}{\sqrt{n}}\right\}.$$

We use here and below the fact that $|\theta|^3/\sqrt{n} \geq |\theta|^4/n$ for $|\theta| \leq \epsilon\sqrt{n}$.

- If $\mathbb{E}[X_1|^6] < \infty$ and all the third and fifth moments of X_1 vanish, then

$$h(\theta) = q_4(\theta) + O(|\theta|^6),$$

$$\log\phi(\theta) = -\frac{\theta\cdot\Gamma\theta}{2} + f_4(\theta) + O(|\theta|^6),$$

where $f_4(\theta) = q_4(\theta) - (\theta\cdot\Gamma\theta)^2/8$ is a homogeneous polynomial of degree four. In this case

$$g(\theta,n) = n\,f_4\left(\frac{\theta}{\sqrt{n}}\right) + \frac{O(|\theta|^6)}{n^2}, \tag{2.19}$$

and there exists $c < \infty$ such that

$$|g(\theta,n)| \leq \min\left\{\frac{\theta\cdot\Gamma\theta}{4}, \frac{c\,|\theta|^4}{n}\right\}.$$

- More generally, suppose that $k \geq 3$ is a positive integer such that $\mathbb{E}[|X_1|^{k+1}] < \infty$. Then

$$h(\theta) = \sum_{j=3}^{k} q_j(\theta) + O(|\theta|^{k+1}),$$

$$\log\phi(\theta) = -\frac{\theta\cdot\Gamma\theta}{2} + \sum_{j=3}^{k} f_j(\theta) + O(|\theta|^{k+1}),$$

where f_j are homogeneous polynomials of degree j that are determined by $\Gamma, q_3, \ldots, q_j$. In this case,

$$g(\theta,n) = \sum_{j=3}^{k} n\,f_j\left(\frac{\theta}{\sqrt{n}}\right) + \frac{O(|\theta|^{k+1})}{n^{(k-1)/2}}, \tag{2.20}$$

Moreover, if j is odd and all the odd moments of X of degree less than or equal to j vanish, then $f_j \equiv 0$. Also,

$$|g(\theta,n)| \leq \min\left\{\frac{\theta\cdot\Gamma\theta}{4}, \frac{c\,|\theta|^{2+\alpha}}{n^{\alpha/2}}\right\},$$

where $\alpha = 2$ if the third moments vanish and otherwise $\alpha = 1$.

- Suppose $\mathbb{E}[e^{b\cdot X}] < \infty$ for all b in a real neighborhood of the origin. Then $z \mapsto \phi(z) = e^{iz\cdot X_1}$ i s a holomorphic function from a neighborhood of the origin in $\mathbb{C}^n$ to $\mathbb{C}$. Hence, we can choose ϵ so that $\log\phi(z)$ is holomorphic for $|z| < \epsilon$ and hence $z \mapsto g(z,n)$ and $z \mapsto F_n(z)$ are holomorphic for $|z| < \epsilon\sqrt{n}$.

The next lemma computes the dominant term and isolates the integral that needs to be estimated in order to obtain error bounds.

Lemma 2.3.4 *Suppose $p \in \mathcal{P}_d'$ with covariance matrix Γ. Let ϕ, ϵ, F_n be as in Lemma 2.3.3. There exist $c < \infty, \zeta > 0$ such that for all $0 \le r \le \epsilon\sqrt{n}$, if we define $v_n(x,r)$ by*

$$p_n(x) = \overline{p}_n(x) + v_n(x,r) + \frac{1}{(2\pi)^d\, n^{d/2}} \int_{|\theta|\le r} e^{-\frac{ix\cdot\theta}{\sqrt{n}}}\, e^{-\frac{\theta\cdot\Gamma\theta}{2}}\, F_n(\theta)\, d\theta,$$

then

$$|v_n(x,r)| \le c\, n^{-d/2}\, e^{-\zeta r^2}.$$

Proof The inversion formula (2.12) gives

$$p_n(x) = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \phi(\theta)^n\, e^{-ix\cdot\theta}\, d\theta = \frac{1}{(2\pi)^d\, n^{d/2}} \int_{[-\sqrt{n}\pi,\sqrt{n}\pi]^d} \phi\left(\frac{s}{\sqrt{n}}\right)^n e^{-iz\cdot s}\, ds,$$

where $z = x/\sqrt{n}$. Lemma 2.3.2 implies that there is a $\beta > 0$ such that $|\phi(\theta)| \le e^{-\beta}$ for $|\theta| \ge \epsilon$. Therefore,

$$\frac{1}{(2\pi)^d\, n^{d/2}} \int_{[-\sqrt{n}\pi,\sqrt{n}\pi]^d} \phi\left(\frac{s}{\sqrt{n}}\right)^n e^{-iz\cdot s}\, ds = O(e^{-\beta n}) + \frac{1}{(2\pi)^d\, n^{d/2}} \int_{|\theta|\le\epsilon\sqrt{n}} \phi\left(\frac{s}{\sqrt{n}}\right)^n e^{-iz\cdot s}\, ds.$$

For $|s| \le \epsilon\sqrt{n}$, we write

$$\phi\left(\frac{s}{\sqrt{n}}\right)^n = e^{-\frac{s\cdot\Gamma s}{2}} + e^{-\frac{s\cdot\Gamma s}{2}}\, F_n(s).$$

By (2.2) we have

$$\frac{1}{(2\pi)^d\, n^{d/2}} \int_{\mathbb{R}^d} e^{-iz\cdot s}\, e^{-\frac{s\cdot\Gamma s}{2}}\, ds = \overline{p}_n(x). \tag{2.21}$$

Also,

$$\left| \frac{1}{(2\pi)^d\, n^{d/2}} \int_{|s|\ge\epsilon\sqrt{n}} e^{-iz\cdot s}\, e^{-\frac{s\cdot\Gamma s}{2}}\, ds \right| \le \frac{1}{(2\pi)^d\, n^{d/2}} \int_{|s|\ge\epsilon\sqrt{n}} e^{-\frac{s\cdot\Gamma s}{2}} \le O(e^{-\beta n}),$$

for perhaps a different β. Therefore,

$$p_n(x) = \overline{p}_n(x) + O(e^{-\beta n}) + \frac{1}{(2\pi)^d\, n^{d/2}} \int_{|\theta|\le\epsilon\sqrt{n}} e^{\frac{-ix\cdot\theta}{\sqrt{n}}}\, e^{-\frac{\theta\cdot\Gamma\theta}{2}}\, F_n(\theta)\, d\theta.$$

This gives the result for $r = \epsilon\sqrt{n}$. For other values of r, we use the estimate

$$|F_n(\theta)| \le e^{\frac{\theta\cdot\Gamma\theta}{4}} + 1,$$

to see that

$$\left| \int_{r\le|\theta|\le\epsilon\sqrt{n}} e^{\frac{-ix\cdot\theta}{\sqrt{n}}}\, e^{-\frac{\theta\cdot\Gamma\theta}{2}}\, F_n(\theta)\, d\theta \right| \le 2 \int_{|\theta|\ge r} e^{-\frac{\theta\cdot\Gamma\theta}{4}}\, d\theta = O(e^{-\zeta r^2}).$$

□

The next theorem establishes LCLTs for $p \in \mathcal{P}'$ with finite third moment and $p \in \mathcal{P}'$ with finite fourth moment and vanishing third moments. It gives an error term that is uniform over all $x \in \mathbb{Z}^d$. The estimate is good for typical x, but is not very sharp for atypically large x.

Theorem 2.3.5 *Suppose $p \in \mathcal{P}'$ with $\mathbb{E}[|X_1|^3] < \infty$. Then, there exist a $c < \infty$ such for all n, x,*

$$|p_n(x) - \overline{p}_n(x)| \leq \frac{c}{n^{(d+1)/2}}. \tag{2.22}$$

If $\mathbb{E}[|X_1|^4] < \infty$ and all the third moments of X_1 are zero, then there is a c such that for all n, x,

$$|p_n(x) - \overline{p}_n(x)| \leq \frac{c}{n^{(d+2)/2}}. \tag{2.23}$$

Proof We use the notations of Lemmas 2.3.3 and 2.3.4. Letting $r = n^{1/8}$ in Lemma 2.3.4, we see that,

$$p_n(x) = \overline{p}_n(x) + O(e^{-\beta n^{1/4}}) + \frac{1}{(2\pi)^d\, n^{d/2}} \int_{|\theta| \leq n^{1/8}} e^{-\frac{ix\cdot\theta}{\sqrt{n}}}\, e^{-\frac{\theta\cdot\Gamma\theta}{2}}\, F_n(\theta)\, d\theta.$$

Note that $|h(\theta)| = O(|\theta|^{2+\alpha})$ where $\alpha = 1$ under the weaker assumption and $\alpha = 2$ under the stronger assumption. For $|\theta| \leq n^{1/8}$, $|g(\theta, n)| \leq c\, |\theta|^{2+\alpha}/n^{\alpha/2}$, and hence

$$|F_n(\theta)| \leq c\, \frac{|\theta|^{2+\alpha}}{n^{\alpha/2}}.$$

This implies

$$\left| \int_{|\theta| \leq n^{1/8}} e^{-\frac{ix\cdot\theta}{\sqrt{n}}}\, e^{-\frac{\theta\cdot\Gamma\theta}{2}}\, F_n(\theta)\, d\theta \right| \leq \frac{c}{n^{\alpha/2}} \int_{\mathbb{R}^d} |\theta|^{2+\alpha}\, e^{-\frac{\theta\cdot\Gamma\theta}{2}}\, d\theta \leq \frac{c}{n^{\alpha/2}}.$$

Comment. The choice $r = n^{1/8}$ in the proof above was somewhat arbitary. The value r was chosen sufficiently large so that the error term $v_n(x, r)$ from Lemma 2.3.4 decays faster than any power of n but sufficiently small so that $|g(\theta, n)|$ is uniformly bounded for $|\theta| \leq r$. We could just as well have chosen $r(n) = n^{\kappa}$ for any $0 < \kappa \leq 1/8$.

□

The estimate (2.5) is a special case of (2.23). We have shown that (2.5) holds for any symmetric $p \in \mathcal{P}'$ with $\mathbb{E}[|X_1|^4] < \infty$. One can obtain a difference estimate for $p_n(x)$ from (2.5). However, we will give another proof below that does requires only third moments of the increment distribution. This theorem also gives a uniform bound on the error term.

Comment. If $\alpha \neq 0$ and

$$f(n) = n^{\alpha} + O(n^{\alpha-1}), \tag{2.24}$$

then

$$f(n+1) - f(n) = [(n+1)^{\alpha} - n^{\alpha}] + [O((n+1)^{\alpha-1}) - O(n^{\alpha-1})].$$

This shows that $f(n+1) - f(n) = O(n^{\alpha-1})$, but the best that we can write about the error terms is $O((n+1)^{\alpha-1}) - O(n^{\alpha-1}) = O(n^{\alpha-1})$, which is as large as the dominant term. Hence an expression such as (2.24) is not sufficient to give good asymptotics on differences of f. One strategy for proving difference estimates is to go back to the derivation of (2.24) to see if the difference of the errors can be estimated. This is the approach used in the next theorem.

Theorem 2.3.6 *Suppose $p \in \mathcal{P}_d'$ with $\mathbb{E}[|X_1|^3] < \infty$. Let ∇_y denote the differences in the x variable,*

$$\nabla_y p_n(x) = p_n(x+y) - p_n(x),; \quad \nabla_y \overline{p}_n(x) = \overline{p}_n(x+y) - \overline{p}_n(x),$$

and $\nabla_j = \nabla_{\mathbf{e}_j}$.

- *There exists $c < \infty$ such that for all x, n, y,*

$$|\nabla_y p_n(x) - \nabla_y \overline{p}_n(x)| \leq \frac{c\,|y|}{n^{(d+2)/2}}.$$

- *If $\mathbb{E}[|X_1|^4] < \infty$ and all the third moments of X_1 vanish, there exists $c < \infty$ such that for all x, n, y,*

$$|\nabla_y p_n(x) - \nabla_y \overline{p}_n(x)| \leq \frac{c\,|y|}{n^{(d+3)/2}}.$$

Proof By the triangle inequality, it suffices to prove the result for $y = \mathbf{e}_j, j = 1, \ldots, d$. Let $\alpha = 1$ under the weaker assumptions and $\alpha = 2$ under the stronger assumptions. As in the proof of Theorem 2.3.5, we see that

$$\nabla_j p_n(x) = \nabla_j \overline{p}_n(x) + O(e^{-\beta n^{1/4}}) + \frac{1}{(2\pi)^d\, n^{d/2}} \int_{|\theta| \leq n^{1/8}} \left[e^{-\frac{i(x+\mathbf{e}_j)\cdot\theta}{\sqrt{n}}} - e^{-\frac{ix\cdot\theta}{\sqrt{n}}} \right] e^{-\frac{\theta\cdot\Gamma\theta}{2}}\, F_n(\theta)\, d\theta.$$

Note that

$$\left| e^{-\frac{i(x+\mathbf{e}_j)\cdot\theta}{\sqrt{n}}} - e^{-\frac{ix\cdot\theta}{\sqrt{n}}} \right| = \left| e^{-\frac{i\mathbf{e}_j\cdot\theta}{\sqrt{n}}} - 1 \right| \leq \frac{|\theta|}{\sqrt{n}},$$

and hence

$$\left| \int_{|\theta| \leq n^{1/8}} \left[e^{-\frac{i(x+\mathbf{e}_j)\cdot\theta}{\sqrt{n}}} - e^{-\frac{ix\cdot\theta}{\sqrt{n}}} \right] e^{-\frac{\theta\cdot\Gamma\theta}{2}}\, F_n(\theta)\, d\theta \right| \leq \frac{1}{\sqrt{n}} \int_{|\theta| \leq n^{1/8}} |\theta|\, e^{-\frac{\theta\cdot\Gamma\theta}{2}}\, |F_n(\theta)|\, d\theta.$$

The estimate

$$\int_{|\theta| \leq n^{1/8}} |\theta|\, e^{-\frac{\theta\cdot\Gamma\theta}{2}}\, |F_n(\theta)|\, d\theta \leq \frac{c}{n^{\alpha/2}},$$

where $\alpha = 1$ under the weaker assumption and $\alpha = 2$ under the stronger assumption, is done as in the previous theorem. □

The next theorem improves the LCLT by giving a better error bound for larger x. The basic strategy is to write $F_n(\theta)$ as the sum of a dominant term and an error term. This requires a stronger moment condition.

Theorem 2.3.7 *Suppose $p \in \mathcal{P}_d'$.*

- *If $\mathbb{E}[|X_1|^4] < \infty$, there exists $c < \infty$ such that*

$$\left| p_n(x) - \overline{p}_n(x) - \frac{u_3(x/\sqrt{n})}{n^{(d+1)/2}} \right| \leq \frac{c}{n^{(d+2)/2}}, \tag{2.25}$$

where

$$u_3(z) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-is\cdot z}\, f_3(s)\, e^{-\frac{s\cdot\Gamma s}{2}}\, ds,$$

and f_3 is the homogeneous polynomial of degree three defined in (2.18).

- *If* $\mathbb{E}[|X_1|^5] < \infty$ *and the third moments of* X_1 *vanish there exists* $c < \infty$ *such that*
$$\left| p_n(x) - \overline{p}_n(x) - \frac{u_4(x/\sqrt{n})}{n^{(d+2)/2}} \right| \le \frac{c}{n^{(d+3)/2}}, \tag{2.26}$$
where
$$u_4(z) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-is\cdot z}\, f_4(s)\, e^{-\frac{s\cdot\Gamma s}{2}}\, ds,$$
and f_4 *is the homogeneous polynomial of degree four defined in (2.18).*

Moreover, if $k \ge 3$ *is a positive integer such that* $\mathbb{E}[|X_1|^k] < \infty$ *and* f_k *is as defined in (2.20), then there is a* $c(k)$ *such that*
$$|u_k(z)| \le c(k)\,(|z|^k + 1)\, e^{-\frac{\mathcal{J}^*(z)^2}{2}}.$$

Proof Let $\alpha = 1$ under the weaker assumptions and $\alpha = 2$ under the stronger assumptions. As in Theorem 2.3.5,
$$n^{d/2}\,[p_n(x) - \overline{p}_n(x)] = O(e^{-\beta n^{1/4}}) + \frac{1}{(2\pi)^d} \int_{|\theta| \le n^{1/8}} e^{-\frac{ix\cdot\theta}{\sqrt{n}}}\, e^{-\frac{\theta\cdot\Gamma\theta}{2}}\, F_n(\theta)\, d\theta. \tag{2.27}$$

Recalling (2.20), we can see that for $|\theta| \le n^{1/8}$,
$$F_n(\theta) = \frac{f_{2+\alpha}(\theta)}{n^{\alpha/2}} + \frac{O(|\theta|^{3+\alpha})}{n^{(\alpha+1)/2}}.$$

Up to an error of $O(e^{-\beta n^{1/4}})$, the right-hand side of (2.27) equals
$$\frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-\frac{ix\cdot\theta}{\sqrt{n}}}\, e^{-\frac{\theta\cdot\Gamma\theta}{2}}\, \frac{f_{2+\alpha}(\theta)}{n^{(\alpha/2)}}\, d\theta + \frac{1}{(2\pi)^d} \int_{|\theta|\le n^{1/8}} e^{-\frac{ix\cdot\theta}{\sqrt{n}}}\, e^{-\frac{\theta\cdot\Gamma\theta}{2}} \left[F_n(\theta) - \frac{f_{2+\alpha}(\theta)}{n^{\alpha/2}} \right] d\theta.$$

The second integral can be bounded as before
$$\left| \int_{|\theta|\le n^{1/8}} e^{-\frac{ix\cdot\theta}{\sqrt{n}}}\, e^{-\frac{\theta\cdot\Gamma\theta}{2}} \left[F_n(\theta) - \frac{f_{2+\alpha}(\theta)}{n^{\alpha/2}} \right] d\theta \right| \le \frac{c}{n^{(\alpha+1)/2}} \int_{\mathbb{R}^d} |\theta|^{3+\alpha}\, e^{-\frac{\theta\cdot\Gamma\theta}{2}}\, d\theta \le \frac{c}{n^{(\alpha+1)/2}}.$$

To estimate u_k we first note that the change of variables $r = \Lambda s$ yields
$$(2\pi)^d\,(\det\Lambda)\, u_k(z) = \int_{\mathbb{R}^d} g(r)\, e^{-|r|^2/2}\, e^{i(r\cdot\Lambda^{-1}z)}\, dr,$$
where $g(\theta) = f_k(\Lambda^{-1}\theta)$. By completing the square we see that the quantity on the right equals
$$e^{-\frac{\mathcal{J}^*(z)^2}{2}} \int g(r)\, \exp\left\{ \frac{1}{2}(ir + \Lambda^{-1}z)\cdot(ir + \Lambda^{-1}z) \right\} dr.$$
The substitution $s = r - i\Lambda^{-1}z$ converts this to
$$e^{-\frac{\mathcal{J}^*(z)^2}{2}} \int g(s + i\Lambda^{-1}z)\, e^{-|s|^2/2}\, ds.$$

Finally,
$$|g(s - i\Lambda^{-1}z)| \le c\,(|z|^k + 1)\,(|s|^k + 1),$$

which gives

$$|u_k(z)| \le c(k)\,(|z|^k+1)\,e^{-\frac{\mathcal{J}^*(z)^2}{2}} \int_{\mathbb{R}^d} (|s|^k+1)\,e^{-|s|^2/2}\,ds \le c(k)\,(|z|^k+1)\,e^{-\frac{\mathcal{J}^*(z)^2}{2}}.$$

□

The next theorem is proved in the same way as Theorem 2.3.7 starting with (2.20), and we omit it. A special case of this theorem is (2.3). The theorem shows that (2.3) holds for all symmetric $p \in \mathcal{P}_d'$ with $\mathbb{E}[|X_1|^6] < \infty$. The results stated for $n \ge |x|^2$ are just a restatement of Theorem 2.3.5.

Theorem 2.3.8 *Suppose $p \in \mathcal{P}_d'$ and $k \ge 3$ is a positive integer such that $\mathbb{E}[|X_1|^{k+1}] < \infty$. There exists $c = c(k)$ such that*

$$\left| p_n(x) - \overline{p}_n(x) - \sum_{j=3}^{k} \frac{u_j(z)}{n^{(d+j-2)/2}} \right| \le \frac{c}{n^{(d+k-1)/2}},$$

where $z = x/\sqrt{n}$ and u_j are as defined in Theorem 2.3.7.

In particular,

$$|p_n(x) - \overline{p}_n(x)| \le \frac{c}{n^{(d+1)/2}} \left[|z|^k\, e^{-\frac{\mathcal{J}^*(z)^2}{2}} + \frac{1}{n^{(k-2)/2}} \right], \quad n \le |x|^2,$$

$$|p_n(x) - \overline{p}_n(x)| \le \frac{c}{n^{(d+1)/2}}, \quad n \ge |x|^2.$$

If the third moments of X_1 vanish (e.g., if p is symmetric) then $u_3 \equiv 0$ and

$$|p_n(x) - \overline{p}_n(x)| \le \frac{c}{n^{(d+2)/2}} \left[|z|^k\, e^{-\frac{\mathcal{J}^*(z)^2}{2}} + \frac{1}{n^{(k-3)/2}} \right], \quad n \le |x|^2,$$

$$|p_n(x) - \overline{p}_n(x)| \le \frac{c}{n^{(d+2)/2}}, \quad n \ge |x|^2.$$

The next theorem is the LCLT assuming only a finite second moment.

Theorem 2.3.9 *Suppose $p \in \mathcal{P}_d'$. Then there exists a sequence $\delta_n \to 0$ such that for all n, x,*

$$|p_n(x) - \overline{p}_n(x)| \le \frac{\delta_n}{n^{d/2}}. \tag{2.28}$$

Proof By Lemma 2.3.4, there exist c, ζ such that for all n, x and $r > 0$,

$$n^{d/2}\,|p_n(x) - \overline{p}_n(x)| \le c\left[e^{-\zeta r^2} + \int_{\theta \le r} |F_n(\theta)|\,d\theta \right] \le c\left[e^{-\zeta r^2} + r^d \sup_{\theta \le r} |F_n(\theta)| \right].$$

We now refer to Lemma 2.3.3. Since $h(\theta) = o(|\theta|^2)$,

$$\lim_{n\to\infty} \sup_{|\theta| \le r} |g(\theta, n)| = 0,$$

and hence

$$\lim_{n\to\infty} \sup_{|\theta| \le r} |F_n(\theta)| = 0.$$

In particular, for all n suffciently large,

$$n^{d/2}\,|p_n(x) - \overline{p}_n(x)| \le 2\,c\,e^{-\zeta r^2}.$$

□

The next theorem improves on this for $|x|$ larger than n^2. The proof uses an integration by parts. One advantage of this theorem is that it does not need any extra moment conditions. However, if we impose extra moment conditions we get a stronger result.

Theorem 2.3.10 *Suppose $p \in \mathcal{P}'_d$. Then there exists a sequence $\delta_n \to 0$ such that for all n, x,*

$$|p_n(x) - \overline{p}_n(x)| \le \frac{\delta_n}{|x|^2\, n^{(d-2)/2}}. \tag{2.29}$$

Moreover,

- *If $\mathbb{E}[|X_1|^3] < \infty$, then δ_n can be chosen $O(n^{-1/2})$.*
- *If $\mathbb{E}[|X_1|^4] < \infty$ and the third moments of X_1 vanish, then δ_n can be chosen $O(n^{-1})$.*

Proof It follows from the inversion formula and Green's theorem (i.e., integration by parts) that

$$p_n(-x) = \frac{1}{(2\pi)^d}\int_{[-\pi,\pi]^d} e^{ix\cdot\theta}\,\psi(\theta)\,d\theta = -\frac{1}{|x|^2\,(2\pi)^d}\int_{[-\pi,\pi]^d} e^{ix\cdot\theta}\,\Delta\psi(\theta)\,d\theta,$$

where $\psi(\theta) = \phi(\theta)^n$. Therefore,

$$\frac{|x|^2}{n}\,p_n(-x) = -\frac{1}{(2\pi)^d}\int_{[-\pi,\pi]^d} e^{ix\cdot\theta}\,\phi(\theta)^{n-2}\,\left[(n-1)\,\lambda(\theta) + \phi(\theta)\,\Delta\phi(\theta)\right]\,d\theta,$$

where

$$\lambda(\theta) = \sum_{j=1}^{d}\left[\partial_j\phi(\theta)\right]^2.$$

The first and second derivatives of ϕ are uniformly bounded. Hence by (2.14), we can see that

$$\left|\phi(\theta)^{n-2}\,\left[(n-1)\,\lambda(\theta) + \phi(\theta)\,\Delta\phi(\theta)\right]\right| \le c\,[1 + n|\theta|^2]\,e^{-n|\theta|^2/b} \le c\,e^{-\beta n|\theta|^2},$$

where $0 < \beta < 1/b$. Hence, we can write

$$\frac{|x|^2}{n}\,p_n(-x) - O(e^{-\beta r^2}) = -\frac{1}{(2\pi)^d}\int_{|\theta|\le r/\sqrt{n}} e^{ix\cdot\theta}\,\phi(\theta)^{n-2}\,\left[(n-1)\,\lambda(\theta) + \phi(\theta)\,\Delta\phi(\theta)\right]\,d\theta.$$

The usual change of variables shows that the right-hand side equals

$$-\frac{1}{(2\pi)^d\,n^{d/2}}\int_{|\theta|\le r} e^{iz\cdot\theta}\,\phi\left(\frac{\theta}{\sqrt{n}}\right)^{n-2}\left[(n-1)\,\lambda\left(\frac{\theta}{\sqrt{n}}\right) + \phi\left(\frac{\theta}{\sqrt{n}}\right)\,\Delta\phi\left(\frac{\theta}{\sqrt{n}}\right)\right]\,d\theta,$$

where $z = x/\sqrt{n}$.

Note that

$$\Delta\left[e^{-\frac{\theta\cdot\Gamma\theta}{2}}\right] = e^{-\frac{\theta\cdot\Gamma\theta}{2}}\,[|\Gamma\theta|^2 - \operatorname{tr}(\Gamma)].$$

We define $\hat{F}_n(\theta)$ by

$$\phi\left(\frac{\theta}{\sqrt{n}}\right)^{n-2}\left[(n-1)\,\lambda\left(\frac{\theta}{\sqrt{n}}\right)+\phi\left(\frac{\theta}{\sqrt{n}}\right)\,\Delta\phi\left(\frac{\theta}{\sqrt{n}}\right)\right]=e^{-\frac{\theta\cdot\Gamma\theta}{2}}\left[|\Gamma\theta|^2-\mathrm{tr}(\Gamma)-\hat{F}_n(\theta)\right].$$

A straightforward calculation using Green's theorem shows that

$$\overline{p}_n(-x)=\frac{1}{(2\pi)^d\,n^{d/2}}\int_{\mathbb{R}^d}e^{i(x/\sqrt{n})\cdot\theta}\,e^{-\frac{\theta\cdot\Gamma\theta}{2}}\,d\theta=-\frac{n}{|x|^2\,(2\pi)^2}\int_{\mathbb{R}^d}e^{i(x/\sqrt{n})\cdot\theta}\,\Delta[e^{-\frac{\theta\cdot\Gamma\theta}{2}}]\,d\theta.$$

Therefore (wtih perhaps a different β),

$$\frac{|x|^2}{n}\,p_n(-x)=\frac{|x|^2}{n}\,\overline{p}_n(-x)+O(e^{-\beta r^2})-\frac{1}{(2\pi)^d\,n^{d/2}}\int_{|\theta|\leq r}e^{iz\cdot\theta}\,e^{-\frac{\theta\cdot\Gamma\theta}{2}}\,\hat{F}_n(\theta)\,d\theta. \tag{2.30}$$

The remaining task is to estimate $\hat{F}_n(\theta)$. Recalling the definition of $F_n(\theta)$ from Lemma 2.3.3, we can see that

$$\phi\left(\frac{\theta}{\sqrt{n}}\right)^{n-2}\left[(n-1)\,\lambda\left(\frac{\theta}{\sqrt{n}}\right)+\Delta\phi\left(\frac{\theta}{\sqrt{n}}\right)\right]=$$
$$e^{-\frac{\theta\cdot\Gamma\theta}{2}}\,[1+F_n(\theta)]\left[(n-1)\,\frac{\lambda(\theta/\sqrt{n})}{\phi(\theta/\sqrt{n})^2}+\frac{\Delta\phi(\theta/\sqrt{n})}{\phi(\theta/\sqrt{n})}\right].$$

We make three possible assumptions:

- $p\in\mathcal{P}_d'$.
- $p\in\mathcal{P}_d'$ with $\mathbb{E}[|X_1|^3]<\infty$.
- $p\in\mathcal{P}_d'$ with $\mathbb{E}[|X_1|^4]<\infty$ and vanishing third moments.

We set $\alpha=0,1,2$, respectively, in these three cases. Then we can write

$$\phi(\theta)=1-\frac{\theta\cdot\Gamma\theta}{2}+q_{2+\alpha}(\theta)+o(|\theta|^{2+\alpha}),$$

where $q_2\equiv 0$ and q_3,q_4 are homomegenous polynomials of degree 3 and 4 respectively. Because ϕ is $C^{2+\alpha}$ and we know the values of the derivatives at the origin, we can write

$$\partial_j\phi(\theta)=-\partial_j\frac{\theta\cdot\Gamma\theta}{2}+\partial_j q_{2+\alpha}(\theta)+o(|\theta|^{1+\alpha}),$$

$$\partial_{jj}\phi(\theta)=-\partial_{jj}\frac{\theta\cdot\Gamma\theta}{2}+\partial_{jj}q_{2+\alpha}(\theta)+o(|\theta|^{\alpha}).$$

Using this, we see that

$$\frac{\sum_{j=1}^d[\partial_j\phi(\theta)]^2}{\phi(\theta)^2}=|\Gamma\theta|^2+\tilde{q}_{2+\alpha}(\theta)+o(|\theta|^{2+\alpha}),$$

$$\frac{\Delta\phi(\theta)}{\phi(\theta)}=-\mathrm{tr}(\Gamma)+\hat{q}_\alpha(\theta)+o(|\theta|^\alpha),$$

where $\tilde{q}_{2+\alpha}$ is a homogeneous polyonomial of degree $2+\alpha$ with $\tilde{q}_2\equiv 0$, and $\hat{q}_\alpha$ is a homogeneous polynomial of degree α with $\hat{q}_0=0$. Therefore, for $|\theta|\leq n^{1/8}$,

$$(n-1)\,\frac{\lambda(\theta/\sqrt{n})}{\phi(\theta/\sqrt{n})^2}+\frac{\Delta\phi(\theta/\sqrt{n})}{\phi(\theta/\sqrt{n})}=|\Gamma\theta|^2-\mathrm{tr}(\Gamma)+\frac{\tilde{q}_{2+\alpha}(\theta)+\hat{q}_\alpha(\theta)}{n^{\alpha/2}}+o\left(\frac{|\theta|^\alpha+|\theta|^{\alpha+2}}{n^{\alpha/2}}\right),$$

which establishes that for $\alpha = 1, 2$

$$|\hat{F}_n(\theta)| = O\left(\frac{1 + |\theta|^{2+\alpha}}{n^{\alpha/2}}\right), \qquad |\theta| \le n^{1/16},$$

and for $\alpha = 0$, for each $r < \infty$,

$$\lim_{n\to\infty} \sup_{|\theta|\le r} |\hat{F}_n(\theta)| = 0.$$

The remainder of the argument follows the proofs of Theorem 2.3.5 and 2.3.9. For $\alpha = 1, 2$ we can choose $r = n^{7/16}$ in (2.30) while for $\alpha = 0$ we choose r independent of n and then let $r \to \infty$.

□

2.3.1 Exponential moments

The estimation of the integral in the proof of the LCLT can be done more precisely if we assume that the moment generating function of the random variable exists in a neighborhood of the origin. We demonstrate this in the case $d = 1$.

Theorem 2.3.11 *Suppose $p \in \mathcal{P}_1'$ such that for some $b > 0$,*

$$\mathbb{E}\left[e^{b|X_1|}\right] < \infty. \tag{2.31}$$

Then there exists $\rho > 0$ such that for all $n \ge 0$ and all integers x with $|x| < \rho n$,

$$p_n(x) = \overline{p}_n(x) \exp\left\{O\left(\frac{1}{\sqrt{n}} + \frac{|x|^3}{n^2}\right)\right\}.$$

Moreover, if $\mathbb{E}[X^3] = 0$,

$$p_n(x) = \overline{p}_n(x) \exp\left\{O\left(\frac{1}{n} + \frac{|x|^4}{n^3}\right)\right\}.$$

Heuristic Note that the conclusion of the theorem can be written

$$|p_n(x) - \overline{p}_n(x)| \le c\,\overline{p}_n(x)\left[\frac{1}{\sqrt{n}} + \frac{|x|^{2+\alpha}}{n^{1+\alpha}}\right], \quad |x| \le n^{\frac{1+\alpha}{2+\alpha}},$$

$$|p_n(x) - \overline{p}_n(x)| \le c\,\overline{p}_n(x) \exp\left\{O\left(\frac{|x|^{2+\alpha}}{n^{1+\alpha}}\right)\right\},$$

where $\alpha = 1$ under the weaker hypothesis and $\alpha = 2$ under the stronger hypothesis. In the proof below, we will let $s = x/\sqrt{n}$, in which case

$$\frac{|x|^{2+\alpha}}{n^{1+\alpha}} = \frac{|s|^{2+\alpha}}{n^{\alpha/2}}.$$

Proof For ease we will assume $\sigma^2 = \mathbb{E}[X_1^2] = 1$; the general case is done similarly. Without loss of generality, we may assume $x \le 0$. As noted in the examples after Lemma 2.3.3, we can choose the ϵ

in that lemma sufficiently small so that $z \mapsto F_n(z)$ is an analytic function for $\{z \in \mathbb{C} : |z| < \epsilon\sqrt{n}\}$. Also,

$$|g(z,n)| \le \max\left\{\frac{z^2}{4}, \frac{c\,|z|^{2+\alpha}}{n^{\alpha/2}}\right\},$$

where $\alpha = 1$ under the weaker assumptions and $\alpha = 2$ under the stronger assumptions.

We set $\rho = \delta/2 = \epsilon/4$ and assume that $|x| \le \rho n$. Lemma 2.3.4 implies that

$$p_n(x) = \overline{p}_n(x) + O(e^{-\beta n}) + \frac{1}{2\pi\sqrt{n}} \int_{|\theta| \le \delta\sqrt{n}} e^{is\theta}\, e^{-\theta^2/2}\, F_n(\theta)\, d\theta,$$

where $s = -x/\sqrt{n} \in [0, \rho\sqrt{n}]$. The change of variable $z = \theta - is$ gives

$$\int_{|\theta| \le \delta\sqrt{n}} e^{is\theta}\, e^{-\theta^2/2}\, F_n(\theta)\, d\theta = e^{-s^2/2} \int_{-\delta\sqrt{n}-is}^{\delta\sqrt{n}-is} F_n(z+is)\, e^{-z^2/2}\, dz.$$

Since $\overline{p}_n(x) \asymp n^{-1/2}\, e^{-x^2/(2n)} = n^{-1/2}\, e^{-s^2/2}$,

$$\left|\frac{1}{2\pi\sqrt{n}} \int_{|\theta| \le \delta\sqrt{n}} e^{is\theta}\, e^{-\theta^2/2}\, F_n(\theta)\, d\theta\right| \le c\,\overline{p}_n(x) \left|\int_{-\delta\sqrt{n}-is}^{\delta\sqrt{n}-is} F_n(z+is)\, e^{-z^2/2}\, dz\right|$$

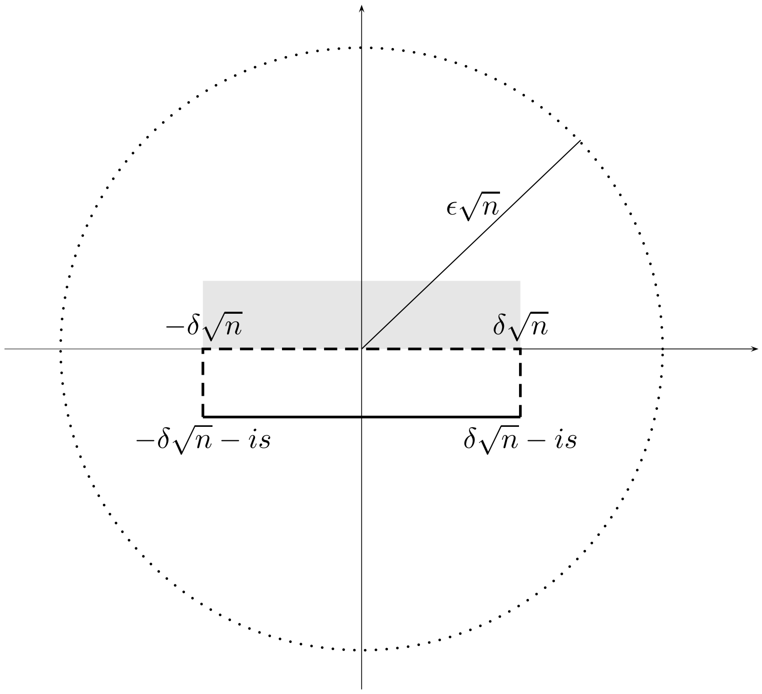

Figure 2.1: The rectangle R and paths of integration

The rectangle

$$\mathcal{R} := \{z : -\delta\sqrt{n} \le \mathrm{Re}(z) \le \delta\sqrt{n},\ 0 \le \mathrm{Im}(z) \le s\},$$

is contained in the closed ball of radius $\delta\sqrt{5n}/2$ about the origin and hence $F_n(z+is)\, e^{-z^2/2}$ is analytic on an open set containing $\mathcal{R} - is$. Therefore, we can evaluate this integral along a different contour, writing

$$\int_{-\delta\sqrt{n}-is}^{\delta\sqrt{n}-is} = \int_{-\delta\sqrt{n}-is}^{-\delta\sqrt{n}} + \int_{-\delta\sqrt{n}}^{\delta\sqrt{n}} + \int_{\delta\sqrt{n}}^{\delta\sqrt{n}-is} .$$

Note that

$$|e^{-z^2/2}| = \exp\left\{-\mathrm{Re}(z^2)/2\right\} = \exp\left\{\frac{-\mathrm{Re}(z)^2 + \mathrm{Im}(z)^2}{2}\right\},$$

and hence if $z \in [-\delta\sqrt{n} - is, -\delta\sqrt{n}] \cup [\delta\sqrt{n} - is, \delta\sqrt{n}]$,

$$|e^{-z^2/2}| \leq e^{-3\delta^2 n/8},$$

and

$$|F_n(z + is)| \leq |e^{g_n(z+is)}| + 1 \leq e^{5\delta^2 n/16} + 1.$$

Using this we can see that

$$\left| \int_{-\delta\sqrt{n}-is}^{-\delta\sqrt{n}} F_n(z + is)\, e^{-z^2/2}\, dz + \int_{\delta\sqrt{n}}^{\delta\sqrt{n}-is} F_n(z + is)\, e^{-z^2/2}\, dz \right|,$$

decays exponentially in n, uniformly over $|s| \leq \rho\sqrt{n}$. Therefore,

$$\left| \int_{-\delta\sqrt{n}-is}^{\delta\sqrt{n}-is} F_n(z + is)\, e^{-z^2/2}\, dz \right| \leq O(e^{-\beta n}) + \int_{-\delta\sqrt{n}}^{\delta\sqrt{n}} |F_n(t + is)|\, e^{-t^2/2}\, dt.$$

Let $q = \alpha/(4 + 2\alpha)$. Suppose $n^q \leq s \leq \rho\sqrt{n}$. If $|t| \geq 2s$, then $|t + is|^2 \leq 5t^2/4$ and hence

$$|F_n(t + is)| \leq e^{5t^2/16} + 1, \qquad \int_{2s \leq |t| \leq \delta\sqrt{n}} |F_n(t + is)|\, e^{-t^2/2}\, dt \leq c.$$

For $|t| \leq 2s$, $|F_n(t + is)| \leq \exp\{O(s^{2+\alpha}/n^{\alpha/2})\}$. Therefore,

$$\int_{0 \leq |t| \leq 2s} |F_n(t + is)|\, e^{-t^2/2}\, dt \leq \exp\{O(s^{2+\alpha}/n^{\alpha/2})\} \int_{-\infty}^{\infty} e^{-t^2/2}\, dt = \exp\{O(s^{2+\alpha}/n^{\alpha/2})\}.$$

Now suppose $0 \leq s < n^q$, and hence $\exp\{O(s^{2+\alpha}/n^{\alpha/2})\} = O(|s|^{2+\alpha}/n^{\alpha/2})$. Then,

$$\int_{2n^q \leq t \leq \delta\sqrt{n}} |F_n(t + is)|\, e^{-t^2/2}\, dt \leq \int_{|t| \geq 2n^q} [e^{5t^2/16} + 1]\, e^{-t^2/2}\, dt = O(e^{-\beta n^{2q}}),$$

and

$$\int_{|t| \leq 2n^q} |F_n(t + is)|\, e^{-t^2/2}\, dt \leq \frac{c}{n^{\alpha/2}} \int_{|t| \leq 2n^q} |t + is|^{2+\alpha}\, e^{-t^2/2}\, dt \leq c\, \frac{1 + s^{2+\alpha}}{n^{\alpha/2}}.$$

□

2.4 Some corollaries of the LCLT

Proposition 2.4.1 *If $p \in \mathcal{P}'$ with bounded support, there is a c such*

$$\sum_{z\in\mathbb{Z}^d} |p_n(z) - p_n(z+y)| \le c\,|y|\,n^{-1/2}.$$

Proof By the triangle inequality, it suffices to prove the result for $y = \mathbf{e} = \mathbf{e}_j$. Let $\delta = 1/2d$. By Theorem 2.3.6,

$$p_n(z+\mathbf{e}) - p_n(z) = \nabla_j \overline{p}_n(z) + O\left(\frac{1}{n^{(d+2)/2}}\right).$$

Also Corollary 11.2.7 shows that

$$\sum_{|z|\ge n^{(1/2)+\delta}} |p_n(z) - p_n(z+\mathbf{e})| \le \sum_{|z|\ge n^{(1/2)+\delta}} [p_n(z) + p_n(z+\mathbf{e})] = o(n^{-1/2}).$$

But,

$$\begin{aligned}\sum_{|z|\le n^{(1/2)+\delta}} |p_n(z) - p_n(z+\mathbf{e})| &\le \sum_{z\in\mathbb{Z}^d} |\overline{p}_n(z) - \overline{p}_n(z+\mathbf{e})| + \sum_{|z|\le n^{(1/2)+\delta}} O\left(\frac{1}{n^{(d+2)/2}}\right)\\ &\le O(n^{-1/2}) + \sum_{z\in\mathbb{Z}^d} |\nabla_j \overline{p}_n(z)|.\end{aligned}$$

A straightforward estimate which we omit gives

$$\sum_{z\in\mathbb{Z}^d} |\nabla_j \overline{p}_n(z)| = O(n^{-1/2}).$$

$\square$

The last proposition holds with much weaker assumptions on the random walk. Recall that $\mathcal{P}^*$ is the set of increment distributions p with the property that for each $x \in \mathbb{Z}^d$, there is an N_x such that $p_n(x) > 0$ for all $n \ge N_x$.

Proposition 2.4.2 *If $p \in \mathcal{P}^*$, there is a c such*

$$\sum_{z\in\mathbb{Z}^d} |p_n(z) - p_n(z+y)| \le c\,|y|\,n^{-1/2}.$$

Proof In Exercise 1.3 it was shown that we can write

$$p = \epsilon\, q + (1-\epsilon) q',$$

where $q \in \mathcal{P}'$ with bounded support and $q' \in \mathcal{P}^*$. By considering the process of first choosing q or q' and then doing the jump, we can see that

$$p_n(x) = \sum_{j=0}^{n} \binom{n}{j} \epsilon^j\,(1-\epsilon)^{n-j} \sum_{z\in\mathbb{Z}^d} q_j(x-z)\, q'_{n-j}(z). \tag{2.32}$$

Therefore,

$$\sum_{x\in\mathbb{Z}^d} |p_n(x) - p_n(x+y)| \leq$$

$$\sum_{j=0}^{n} \binom{n}{j} \epsilon^j\,(1-\epsilon)^{n-j} \sum_{x\in\mathbb{Z}^d} q'_{n-j}(x) \sum_{z\in\mathbb{Z}^d} |q_j(x-z) - q_j(x+y-z)|.$$

We split the first sum into the sum over $j < (\epsilon/2)n$ and $j \geq (\epsilon/2)n$. Standard exponential estimates for the binomial (see Lemma 11.2.9) give

$$\sum_{j<(\epsilon/2)n} \binom{n}{j} \epsilon^j\,(1-\epsilon)^{n-j} \sum_{x\in\mathbb{Z}^d} q'_{n-j}(x) \sum_{z\in\mathbb{Z}^d} |q_j(x-z) - q_j(x+y-z)|$$

$$\leq 2 \sum_{j<(\epsilon/2)n} \binom{n}{j} \epsilon^j\,(1-\epsilon)^{n-j} = O(e^{-\alpha n}),$$

for some $\alpha = \alpha(\epsilon) > 0$. By Proposition 2.4.1,

$$\sum_{j\geq(\epsilon/2)n} \binom{n}{j} \epsilon^j\,(1-\epsilon)^{n-j} \sum_{x\in\mathbb{Z}^d} q'_{n-j}(x) \sum_{z\in\mathbb{Z}^d} |q_j(x-z) - q_j(x+y-z)|$$

$$\leq c\,n^{-1/2}\,|y| \sum_{j\geq(\epsilon/2)n} \binom{n}{j} \epsilon^j\,(1-\epsilon)^{n-j} \sum_{x\in\mathbb{Z}^d} q'_{n-j}(x) \leq c\,n^{-1/2}\,|y|.$$

□

The last proposition has the following useful lemma as a corollary. Since this is essentially a result about Markov chains in general, we leave the proof to the appendix, see Theorem 11.4.6.

Lemma 2.4.3 *Suppose $p \in \mathcal{P}_d^*$. There is a $c < \infty$ such that if $x, y \in \mathbb{Z}^d$, we can define S_n, S_n^* on the same probability space such that: S_n has the distribution of a random walk with increment p with $S_0 = x$; S_n^* has the distribution of a random walk with increment p with $S_0 = y$; and such that for all n,*

$$\mathbb{P}\{S_m = S_m^* \text{ for all } m \geq n\} \geq 1 - \frac{c\,|x-y|}{\sqrt{n}}.$$

Comment. While the proof of this last lemma is a little messy to write out in detail, there really is not a lot of content to it once we have Proposition 2.4.2. Suppose that p, q are two probability distributions on $\mathbb{Z}^d$ with

$$\frac{1}{2}\sum_{z\in\mathbb{Z}^d} |p(z) - q(z)| = \epsilon.$$

Then there is an easy way to define random variables X, Y on the same probability space such that X has distribution p, Y has distribution q and $\mathbb{P}\{X \neq Y\} = \epsilon$. Indeed, if we let $f(z) = \min\{p(z), q(z)\}$ we can let the probability space be $\mathbb{Z}^d \times \mathbb{Z}^d$ and define μ by

$$\mu(z,z) = f(z)$$

and for $x \neq y$,

$$\mu(x,y) = \epsilon^{-1}\,[p(x) - f(x)]\,[q(y) - f(y)].$$

If we let $X(x,y) = x, Y(x,y) = y$, it is easy to check that the marginal of X is p, the marginal of Y is q and $\mathbb{P}\{X = Y\} = 1 - \epsilon$. The more general fact is not much more complicated than this.

Proposition 2.4.4 *Suppose $p \in \mathcal{P}_d^*$. There is a $c < \infty$ such that for all n, x,*

$$p_n(x) \leq \frac{c}{n^{d/2}}. \tag{2.33}$$

Proof If $p \in \mathcal{P}_d'$ with bounded support this follows immediately from (2.22). For general $p \in \mathcal{P}_d^*$, write $p = \epsilon\, q + (1-\epsilon)\, q'$ with $q \in \mathcal{P}_d', q' \in \mathcal{P}_d^*$ as in the proof of Proposition 2.22. Then $p_n(x)$ is as in (2.32). The sum over $j < (\epsilon/2)n$ is $O(e^{-\alpha n})$ and for $j \geq (\epsilon/2)n$, we have the bound $q_j(x-z) \leq c\, n^{-d/2}$. □

The central limit theorem implies that it takes $O(n^2)$ steps to go distance n. This proposition gives some bounds on large deviations for the number of steps.

Proposition 2.4.5 *Suppose S is a random walk with increment distribution $p \in \mathcal{P}_d$ and let*

$$\tau_n = \min\{k : |S_k| \geq n\}, \quad \xi_n = \min\{k : \mathcal{J}^*(S_k) \geq n\}.$$

There exist $t > 0$ and $c < \infty$ such that for all n and all $r > 0$,

$$\mathbb{P}\{\tau_n \leq rn^2\} + \mathbb{P}\{\xi_n \leq rn^2\} \leq c\, e^{-t/r}, \tag{2.34}$$

$$\mathbb{P}\{\tau_n \geq rn^2\} + \mathbb{P}\{\xi_n \geq rn^2\} \leq c\, e^{-rt}. \tag{2.35}$$

Proof There exists a $\tilde{c}$ such that $\xi_{\tilde{c}n} \leq \tau_n \leq \xi_{n/\tilde{c}}$ so it suffices to prove the estimates for τ_n. It also suffices to prove the result for n sufficiently large. The central limit theorem implies that there is an integer k such that for all n sufficiently large,

$$\mathbb{P}\{|S_{kn^2}| \geq 2n\} \geq \frac{1}{2}.$$

By the strong Markov property, this implies for all l

$$\mathbb{P}\{\tau_n > kn^2 + l \mid \xi_n > l\} \leq \frac{1}{2},$$

and hence

$$\mathbb{P}\{\tau_n > jkn^2\} \leq (1/2)^j = e^{-j\log 2} = e^{-jk(\log 2/k)}.$$

This gives (2.35). The estimate (2.34) on τ_n can be written as

$$\mathbb{P}\left\{\max_{1\leq j\leq rn^2} |S_j| \geq n\right\} = \mathbb{P}\left\{\max_{1\leq j\leq rn^2} |S_j| \geq (1/\sqrt{r})\,\sqrt{rn^2}\right\} \leq c\, e^{-t/r},$$

which follows from (2.7). □

Comment. The upper bound (2.35) for τ_n does not need any assumptions on the distribution of the increments other than they be nontrivial, see Exercise 2.7.

Theorem 2.3.10 implies that for all $p \in \mathcal{P}_d'$, $p_n(x) \leq c\, n^{-d/2}\, (\sqrt{n}/|x|)^2$. The next proposition extends this to real $r > 2$ under the assumption that $\mathbb{E}[|X_1|^r] < \infty$.

Proposition 2.4.6 *Suppose $p \in \mathcal{P}_d^*$. There is a c such that for all n, x,*

$$p_n(x) \le \frac{c}{n^{d/2}} \max_{0 \le j \le n} \mathbb{P}\{|S_j| \ge |x|/2\}.$$

In particular, if $r > 2$, $p \in \mathcal{P}_d'$ and $\mathbb{E}[|X_1|^r] < \infty$, then there exists $c < \infty$ such that for all n, x,

$$p_n(x) \le \frac{c}{n^{d/2}} \left(\frac{\sqrt{n}}{|x|}\right)^r. \tag{2.36}$$

Proof Let $m = n/2$ if n is even and $m = (n+1)/2$ if n is odd. Then,

$$\{S_n = x\} = \{S_n = x, |S_m| \ge |x|/2\} \;\cup\; \{S_n = x, |S_n - S_m| \ge |x|/2\}.$$

Hence it suffices to estimate the probabilities of the events on the right-hand side. Using (2.33) we get

$$\begin{aligned}\mathbb{P}\{S_n = x, |S_m| \ge |x|/2\} &= \mathbb{P}\{|S_m| \ge |x|/2\}\,\mathbb{P}\{S_n = x \mid |S_m| \ge |x|/2\} \\ &\le c\,n^{-d/2}\,\mathbb{P}\{|S_m| \ge |x|/2\}.\end{aligned}$$

The other probability can be estimated similarly since

$$\mathbb{P}\{S_n = x, |S_n - S_m| \ge |x|/2\} = \mathbb{P}\{S_n = x, |S_{n-m}| \ge |x|/2\}.$$

We claim that if $p \in \mathcal{P}_d'$, $r \ge 2$, and $\mathbb{E}[|X_1|^r] < \infty$, then there is a c such that $\mathbb{E}[|S_n|^r] \le c\,n^{r/2}$. Once we have this, the Chebyshev inequality gives for $m \le n$,

$$\mathbb{P}\{|S_m| \ge |x|\} \le \frac{c\,n^{r/2}}{|x|^r}.$$

The claim is easier when r is an even integer (for then we can estimate the expectation by expanding $(X_1 + \cdots + X_n)^r$), but we give a proof for all $r \ge 2$. Without loss of generality, assume $d = 1$. For a fixed n define

$$T_1 = \min\{j : |S_j| \ge c_1\sqrt{n}\},$$

and for $l > 1$,

$$T_l = \min\left\{j > T_{l-1} : \left|S_j - S_{T_{l-1}}\right| \ge c_1\sqrt{n}\right\},$$

where c_1 is chosen sufficiently large so that

$$\mathbb{P}\{T_1 > n\} \ge \frac{1}{2}.$$

The existence of such a c_1 follows from (2.6) applied with $k = 1$.

Let $Y_1 = |S_{T_1}|$ and for $l > 1$, $Y_l = \left|S_{T_l} - S_{T_{l-1}}\right|$. Note that $(T_1, Y_1), (T_2, Y_2), (T_3, Y_3), \ldots$ are independent, identically distributed random variables taking values in $\{1, 2, \ldots\} \times \mathbb{R}$. Let ξ be the smallest $l \ge 1$ such that $T_l > n$. Then one can readily check from the triangle inequality that

$$\begin{aligned}|S_n| &\le Y_1 + Y_2 + \cdots + Y_{\xi - 1} + c_1\sqrt{n} \\ &= c_1\sqrt{n} + \sum_{l=1}^{\infty} \hat{Y}_l,\end{aligned}$$

where $\hat{Y}_l = Y_l \, 1\{T_l \le n\} \, 1\{\xi > l-1\}$. Note that

$$\begin{aligned}
\mathbb{P}\{Y_1 \ge c_1 \sqrt{n} + t; T_1 \le n\} &\le \mathbb{P}\{|X_j| \ge t \text{ for some } 1 \le j \le n\} \\
&\le n \, \mathbb{P}\{|X_1| \ge t\}.
\end{aligned}$$

Letting $Z = |X_1|$, we get

$$\begin{aligned}
\mathbb{E}[\hat{Y}_1^r] = \mathbb{E}[Y_1^r; T_l \le n] &= c \int_0^\infty s^{r-1} \, \mathbb{P}\{Y_1 \ge s; T_1 \le n\} \, ds \\
&\le c \left[n^{r/2} + \int_{(c_1+1)\sqrt{n}}^\infty s^{r-1} \, n \, \mathbb{P}\{Z \ge s - c_1 \sqrt{n}\} \, ds \right] \\
&= c \left[n^{r/2} + \int_{\sqrt{n}}^\infty (s + \sqrt{n})^{r-1} \, n \, \mathbb{P}\{Z \ge s\} \, ds \right] \\
&\le c \left[n^{r/2} + 2^{r-1} \int_{\sqrt{n}}^\infty s^{r-1} \, n \, \mathbb{P}\{Z \ge s\} \, ds \right] \\
&\le c \left[n^{r/2} + 2^{r-1} \, n \, \mathbb{E}[Z^r] \right] \le c \, n^{r/2}.
\end{aligned}$$

For $l > 1$,

$$\mathbb{E}[\hat{Y}_l^r] \le \mathbb{P}\{\xi > l-1\} \, \mathbb{E}[Y_l^r \, 1\{T_l \ge n\} \mid \xi > l-1] = \left(\frac{1}{2}\right)^{l-1} \mathbb{E}[\hat{Y}_1^r].$$

Therefore,

$$\begin{aligned}
\mathbb{E}\left[(\hat{Y}_1 + \hat{Y}_2 + \cdots)^r \right] &= \lim_{l \to \infty} \mathbb{E}\left[(\hat{Y}_1 + \cdots + \hat{Y}_l)^r \right] \\
&\le \lim_{l \to \infty} \left[\mathbb{E}[\hat{Y}_1^r]^{1/r} + \cdots + \mathbb{E}[\hat{Y}_l^r]^{1/r} \right]^r \\
&= \mathbb{E}[\hat{Y}_1^r] \left[\sum_{l=1}^\infty \left(\frac{1}{2}\right)^{(l-1)/r} \right]^r = c \, \mathbb{E}[\hat{Y}_1^r].
\end{aligned}$$

□

2.5 LCLT — combinatorial approach

In this section, we only prove an LCLT with estimates for one-dimensional simple random walk, both discrete and continuous time. The extension to $p \in \mathcal{P}_d$ is straightforward using (1.2). The arguments are relatively elementary, but they do require a lot of calculation and estimation. Here is a basic outline:

- Establish the result for discrete time random walk by exact counting of paths. Along the way we will prove Stirling's formula.
- Prove an LCLT for Poisson random variables and use it to derive the result for one-dimensional continuous-time walks. (A result for d-dimensional continuous-time simple random walks follows immediately.)

We could continue this approach and prove an LCLT for multinomial random variables and use it to derive the result for discrete-time d-dimensional simple random walk, but we have chosen to omit this.

2.5.1 Stirling's formula and 1-d walks

Suppose S_n is a simple one-dimensional random walk starting at the origin. Determining the distribution of S_n reduces to an easy counting problem. In order for $X_1 + \cdots + X_{2n}$ to equal $2k$, exactly $n+k$ of the X_j must equal $+1$. Since all 2^{-2n} sequences of ± 1 are equally likely,

$$p_{2n}(2k) = \mathbb{P}\{S_{2n} = 2k\} = 2^{-2n} \binom{2n}{n+k} = 2^{-2n} \frac{(2n)!}{(n+k)!\,(n-k)!}. \tag{2.37}$$

We will use Stirling's formula, which we now derive, to estimate the factorial. In the proof, we will use some standard estimates about the logarithm, see Section 11.1.2.

Theorem 2.5.1 (Stirling's formula) *As $n \to \infty$,*

$$n! \sim \sqrt{2\pi}\, n^{n+(1/2)}\, e^{-n}.$$

In fact,

$$\frac{n!}{\sqrt{2\pi}\, n^{n+(1/2)}\, e^{-n}} = 1 + \frac{1}{12\,n} + O\left(\frac{1}{n^2}\right).$$

Proof Let $b_n = n^{n+(1/2)} e^{-n}/n!$. Then, (11.5) and Taylor's theorem imply

$$\begin{aligned}
\frac{b_{n+1}}{b_n} &= \frac{1}{e}\left(1+\frac{1}{n}\right)^n \left(1+\frac{1}{n}\right)^{1/2} \\
&= \left[1 - \frac{1}{2n} + \frac{11}{24n^2} + O\left(\frac{1}{n^3}\right)\right] \left[1 + \frac{1}{2n} - \frac{1}{8n^2} + O\left(\frac{1}{n^3}\right)\right] \\
&= 1 + \frac{1}{12n^2} + O\left(\frac{1}{n^3}\right).
\end{aligned}$$

Therefore,

$$\lim_{m\to\infty} \frac{b_m}{b_n} = \prod_{l=n}^{\infty} \left[1 + \frac{1}{12l^2} + O\left(\frac{1}{l^3}\right)\right] = 1 + \frac{1}{12n} + O\left(\frac{1}{n^2}\right).$$

The second equality is obtained by

$$\begin{aligned}
\log \prod_{l=n}^{\infty} \left[1 + \frac{1}{12l^2} + O\left(\frac{1}{l^3}\right)\right] &= \sum_{l=n}^{\infty} \log\left[1 + \frac{1}{12l^2} + O\left(\frac{1}{l^3}\right)\right] \\
&= \sum_{l=n}^{\infty} \frac{1}{12l^2} + \sum_{l=n}^{\infty} O\left(\frac{1}{l^3}\right) \\
&= \frac{1}{12n} + O\left(\frac{1}{n^2}\right).
\end{aligned}$$

This establishes that the limit

$$C = \lim_{m\to\infty} b_m$$

exists and

$$b_n = C\left[1 - \frac{1}{12n} + O\left(\frac{1}{n^2}\right)\right]. \tag{2.38}$$

There are a number of ways to determine the constant C. For example, if S_n denotes a one-dimensional simple random walk, then

$$\mathbb{P}\{|S_{2n}| \le \sqrt{2n}\,\log n\} = \sum_{|2k|\le\sqrt{2n}\,\log n} 4^{-n}\binom{2n}{n+k} = \sum_{|2k|\le\sqrt{2n}\,\log n} 4^{-n}\,\frac{(2n)!}{(n+k)!(n-k)!}.$$

Using (11.3), we see that as $n\to\infty$, if $|2k|\le\sqrt{2n}\,\log n$,

$$4^{-n}\,\frac{(2n)!}{(n+k)!(n-k)!} \sim \frac{\sqrt{2}}{C\sqrt{n}}\left(1+\frac{k}{n}\right)^{-(n+k)}\left(1-\frac{k}{n}\right)^{-(n-k)}$$

$$= \frac{\sqrt{2}}{C\sqrt{n}}\left(1-\frac{k^2/n}{n}\right)^{-n}\left(1+\frac{k^2/n}{k}\right)^{-k}\left(1-\frac{k^2/n}{k}\right)^{k} \sim \frac{\sqrt{2}}{C\sqrt{n}}\,e^{-k^2/n}.$$

Therefore,

$$\lim_{n\to\infty}\mathbb{P}\{|S_{2n}|\le\sqrt{2n}\,\log n\} = \lim_{n\to\infty}\sum_{|k|\le\sqrt{n/2}\,\log n}\frac{\sqrt{2}}{C\sqrt{n}}\,e^{-k^2/n} = \frac{\sqrt{2}}{C}\int_{-\infty}^{\infty} e^{-x^2}\,dx = \frac{\sqrt{2\pi}}{C}.$$

However, Chebyshev's inequality shows that

$$\mathbb{P}\{|S_{2n}|\ge\sqrt{2n}\,\log n\} \le \frac{\mathrm{Var}[S_{2n}]}{2n\,\log^2 n} = \frac{1}{\log^2 n} \longrightarrow 0.$$

Therefore, $C = \sqrt{2\pi}$. □

Comment. By adapting this proof, it is easy to see that one can find $r_1 = 1/12, r_2, r_3, \ldots$ such that for each positive integer k,

$$n! = \sqrt{2\pi}\, n^{n+(1/2)}\, e^{-n}\left[1 + \frac{r_1}{n} + \frac{r_2}{n^2} + \cdots + \frac{r_k}{n^k} + O\left(\frac{1}{n^{k+1}}\right)\right]. \tag{2.39}$$

We will now prove Theorem 2.1.1 and some difference estimates in the special case of simple random walk in one dimension by using (2.37) and Stirling's formula. As a warmup, we start with the probability of being at the origin.

Proposition 2.5.2 *For simple random walk in $\mathbb{Z}$, if n is a positive integer, then*

$$\mathbb{P}\{S_{2n}=0\} = \frac{1}{\sqrt{\pi n}}\left[1 - \frac{1}{8n} + O\left(\frac{1}{n^2}\right)\right].$$

Proof The probability is exactly

$$2^{-2n}\binom{2n}{n} = \frac{(2n)!}{4^n\,(n!)^2}.$$

By plugging into Stirling's formula, we see that the right hand side equals

$$\frac{1}{\sqrt{\pi n}}\,\frac{1+(24n)^{-1}+O(n^{-2})}{[1+(12n)^{-1}+O(n^{-2})]^2} = \frac{1}{\sqrt{\pi n}}\left[1-\frac{1}{8n}+O\left(\frac{1}{n^2}\right)\right].$$

$\square$

In the last proof, we just plugged into Stirling's formula and evaluated. We will now do the same thing to prove a version of the LCLT for one-dimensional simple random walk.

Proposition 2.5.3 *For simple random walk in $\mathbb{Z}$, if n is a positive integer and k is an integer with $|k|\le n$,*

$$p_{2n}(2k) = \mathbb{P}\{S_{2n}=2k\} = \frac{1}{\sqrt{\pi n}}\,e^{-k^2/n}\,\exp\left\{O\left(\frac{1}{n}+\frac{k^4}{n^3}\right)\right\}.$$

In particular, if $|k|\le n^{3/4}$, then

$$\mathbb{P}\{S_{2n}=2k\} = \frac{1}{\sqrt{\pi n}}\,e^{-k^2/n}\left[1+O\left(\frac{1}{n}+\frac{k^4}{n^3}\right)\right].$$

Comment. Note that for one-dimensional simple random walk,

$$2\,\overline{p}_{2n}(2k) = 2\,\frac{1}{\sqrt{(2\pi)\,(2n)}}\exp\left\{-\frac{(2k)^2}{2\,(2n)}\right\} = \frac{1}{\sqrt{\pi n}}\,e^{-k^2/n}.$$

Comment. While the theorem is stated for all $|k|\le n$, it is not a very strong statement when k is of order n. For example, for $n/2\le |k|\le n$, we can rewrite the conclusion as

$$p_{2n}(2k) = \frac{1}{\sqrt{\pi n}}\,e^{-k^2/n}\;e^{O(n)} = e^{O(n)},$$

which only tells us that there exists α such that

$$e^{-\alpha n}\le p_{2n}(2k)\le e^{\alpha n}.$$

In fact, $2\overline{p}_{2n}(2k)$ is not a very good approximation of $p_{2n}(2k)$ for large n. As an extreme example, note that

$$p_{2n}(2n) = 4^{-n},\qquad 2\,\overline{p}_{2n}(2n) = \frac{1}{\sqrt{\pi n}}\,e^{-n}.$$

Proof If $n/2\le |k|\le n$, the result is immediate using only the estimate $2^{-2n}\le \mathbb{P}\{S_{2n}=2k\}\le 1$. Hence, we may assume that $|k|\le n/2$. As noted before,

$$\mathbb{P}\{S_{2n}=2k\} = 2^{-2n}\binom{2n}{n+k} = \frac{(2n)!}{2^{2n}(n+k)!(n-k)!}.$$

If we restrict to $|k| \le n/2$, we can use Stirling's formula (Lemma 2.5.1) to see that

$$\mathbb{P}\{S_{2n} = 2k\} = \left[1 + O\left(\frac{1}{n}\right)\right] \frac{1}{\sqrt{\pi n}} \left(1 - \frac{k^2}{n^2}\right)^{-1/2} \left(1 - \frac{k^2}{n^2}\right)^{-n} \left(1 - \frac{2k}{n+k}\right)^{k}.$$

The last two terms approach exponential functions. We need to be careful with the error terms. Using (11.3) we get,

$$\left(1 - \frac{k^2}{n^2}\right)^{n} = e^{-k^2/n} \exp\left\{O\left(\frac{k^4}{n^3}\right)\right\}.$$

$$\begin{aligned}\left(1 - \frac{2k}{n+k}\right)^{k} &= e^{-2k^2/(n+k)} \exp\left\{-\frac{2k^3}{(n+k)^2} + O\left(\frac{k^4}{n^3}\right)\right\} \\ &= e^{-2k^2/(n+k)} \exp\left\{-\frac{2k^3}{n^2} + O\left(\frac{k^4}{n^3}\right)\right\},\end{aligned}$$

$$e^{-2k^2/(n+k)} = e^{-2k^2/n} \exp\left\{\frac{2k^3}{n^2} + O\left(\frac{k^4}{n^3}\right)\right\}.$$

Also, using $k^2/n^2 \le \max\{(1/n), (k^4/n^3)\}$, we can see that

$$\left(1 - \frac{k^2}{n^2}\right)^{-1/2} = \exp\left\{O\left(\frac{1}{n} + \frac{k^4}{n^3}\right)\right\}.$$

Combining all of this gives the theorem. □

Comment. We could also prove "difference estimates" by using the equalities

$$p_{2n}(2k+2) = \frac{n-k}{n+k+1}\, p_{2n}(2k),$$

$$p_{2(n+1)}(2k) = p_{2n}(2k)\, 4^{-1} \frac{(2n+1)(2n+2)}{(n+k+1)\,(n-k+1)}.$$

Corollary 2.5.4 *If S_n is simple random walk, then for all positive integers n and all $|k| < n$,*

$$\mathbb{P}\{S_{2n+1} = 2k+1\} = \frac{1}{\sqrt{\pi n}} \exp\left\{-\frac{(k+\frac{1}{2})^2}{n}\right\} \exp\left\{O\left(\frac{1}{n} + \frac{k^4}{n^3}\right)\right\}. \tag{2.40}$$

Proof Note that

$$\mathbb{P}\{S_{2n+1} = 2k+1\} = \frac{1}{2}\,\mathbb{P}\{S_{2n} = 2k\} + \frac{1}{2}\,\mathbb{P}\{S_{2n} = 2(k+1)\}.$$

Hence,

$$\mathbb{P}\{S_{2n+1} = 2k+1\} = \frac{1}{2\sqrt{\pi n}}[e^{-k^2/n} + e^{-(k+1)^2/n}] \exp\left\{O\left(\frac{1}{n} + \frac{k^4}{n^3}\right)\right\}.$$

But,

$$\exp\left\{-\frac{(k+\frac{1}{2})^2}{n}\right\} = e^{-k^2/n}\left[1-\frac{k}{n}+O\left(\frac{k^2}{n^2}\right)\right],$$

$$\exp\left\{-\frac{(k+1)^2}{n}\right\} = e^{-k^2/n}\left[1-\frac{2k}{n}+O\left(\frac{k^2}{n^2}\right)\right],$$

which implies

$$\frac{1}{2}\left[e^{-k^2/n}+e^{-(k+1)^2/n}\right] = \exp\left\{-\frac{(k+\frac{1}{2})^2}{n}\right\}\left[1+O\left(\frac{k^2}{n^2}\right)\right].$$

Using $k^2/n^2 \le \max\{(1/n),(k^4/n^3)\}$, we get (2.40). □

Comment. One might think that we should replace n in (2.40) with $n+(1/2)$. However,

$$\frac{1}{n+(1/2)} = \frac{1}{n}\left[1+O\left(\frac{1}{n}\right)\right].$$

Hence, the same statment with $n+(1/2)$ replacing n is also true.

2.5.2 LCLT for Poisson and continuous-time walks

The next proposition establishes the strong LCLT for Poisson random variables. This will be used for comparing discrete-time and continuous-time random walks with the same p. If N_t is a Poisson random variable with parameter t, then $\mathbb{E}[N_t]=t, \mathrm{Var}[N_t]=t$. The central limit theorem implies that as $t\to\infty$, the distribution of $(N_t-t)/\sqrt{t}$ approaches that of a standard normal. Hence, we might conjecture that

$$\begin{aligned}\mathbb{P}\{N_t=m\} &= \mathbb{P}\left\{\frac{m-t}{\sqrt{t}} \le \frac{N_t-t}{\sqrt{t}} < \frac{m+1-t}{\sqrt{t}}\right\}\\ &\approx \int_{(m-t)/\sqrt{t}}^{(m+1-t)/\sqrt{t}} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}\,dx \approx \frac{1}{\sqrt{2\pi t}}\, e^{-\frac{(m-t)^2}{2t}}.\end{aligned}$$

In the next proposition, we use a straightforward combinatorial argument to justify this approximation.

Proposition 2.5.5 *Suppose N_t is a Poisson random variable with parameter t, and m is an integer with $|m|\le t/2$. Then*

$$\mathbb{P}\{N_t=m\} = \frac{1}{\sqrt{2\pi t}}\, e^{-\frac{(m-t)^2}{2t}}\, \exp\left\{O\left(\frac{1}{\sqrt{t}}+\frac{|m-t|^3}{t^2}\right)\right\}.$$

Proof For notational ease, we will first consider the case where $t=n$ is an integer, and we let $m=n+k$. Let

$$q(n,k) = \mathbb{P}\{N_n=n+k\} = e^{-n}\,\frac{n^{n+k}}{(n+k)!},$$

and note the recursion formula

$$q(n,k) = \frac{n}{n+k}\, q(n,k-1).$$

Stirling's formula (Theorem 2.5.1) gives

$$q(n,0) = \frac{e^{-n}\, n^n}{n!} = \frac{1}{\sqrt{2\pi n}} \left[1 + O\left(\frac{1}{n}\right)\right]. \tag{2.41}$$

By the recursion formula, if $k \le n/2$,

$$q(n,k) = q(n,0) \left[\left(1+\frac{1}{n}\right)\left(1+\frac{2}{n}\right)\cdots\left(1+\frac{k}{n}\right)\right]^{-1},$$

and,

$$\log \prod_{j=1}^{k}\left(1+\frac{j}{n}\right) = \sum_{j=1}^{k} \log\left(1+\frac{j}{n}\right)$$

$$= \sum_{j=1}^{k}\left[\frac{j}{n} + O\left(\frac{j^2}{n^2}\right)\right] = \frac{k^2}{2n} + \frac{k}{2n} + O\left(\frac{k^3}{n^2}\right) = \frac{k^2}{2n} + O\left(\frac{1}{\sqrt{n}} + \frac{k^3}{n^2}\right).$$

The last equality uses the inequality

$$\frac{k}{n} \le \max\left\{\frac{1}{\sqrt{n}}, \frac{k^3}{n^2}\right\},$$

which will also be used in other estimates in this proof. Using (2.41), we get

$$\log q(n,k) = -\log\sqrt{2\pi n} - \frac{k^2}{2n} + O\left(\frac{1}{\sqrt{n}} + \frac{k^3}{n^2}\right),$$

and the result for $k \ge 0$ follows by exponentiating.

Similarly,

$$q(n,-k) = q(n,0)\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)$$

and

$$\begin{aligned}\log q(n,-k) &= -\log\sqrt{2\pi n} + \log\prod_{j=1}^{k-1}\left(1-\frac{j}{n}\right)\\ &= -\log\sqrt{2\pi n} - \frac{k^2}{2n} + O\left(\frac{1}{\sqrt{n}} + \frac{k^3}{n^2}\right).\end{aligned}$$

The proposition for integer n follows by exponentiating. For general t, let $n = \lfloor t \rfloor$ and note that

$$\begin{aligned}\mathbb{P}\{N_t = n+k\} &= \mathbb{P}\{N_n = n+k\}\, e^{-(t-n)}\left(1+\frac{t-n}{n}\right)^{n+k}\\ &= \mathbb{P}\{N_n = n+k\}\left(1+\frac{t-n}{n}\right)^{k}\left[1+O\left(\frac{1}{n}\right)\right]\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{P}\{N_n = n+k\}\left[1 + O\left(\frac{|k|+1}{n}\right)\right] \\
&= (2\pi n)^{-1/2}\, e^{-k^2/(2n)}\, \exp\left\{O\left(\frac{1}{\sqrt{n}} + \frac{k^3}{n^2}\right)\right\} \\
&= (2\pi t)^{-1/2}\, e^{-(k+n-t)^2/(2t)}\, \exp\left\{O\left(\frac{1}{\sqrt{t}} + \frac{|n+k-t|^3}{t^2}\right)\right\}.
\end{aligned}
$$

The last step uses the estimates

$$\frac{1}{\sqrt{t}} = \frac{1}{\sqrt{n}}\left[1 + O\left(\frac{1}{t}\right)\right], \qquad e^{-\frac{k^2}{2t}} = e^{-\frac{k^2}{2n}}\, \exp\left\{O\left(\frac{k^2}{t^2}\right)\right\}.$$

□

We will use this to prove a version of the local central limit theorem for one-dimensional, continuous-time simple random walk.

Theorem 2.5.6 *If $\tilde S_t$ is continuous-time one-dimensional simple random walk, then if $|x| \le t/2$,*

$$\tilde p_t(x) = \frac{1}{\sqrt{2\pi t}}\, e^{-\frac{x^2}{2t}}\, \exp\left\{O\left(\frac{1}{\sqrt{t}} + \frac{|x|^3}{t^2}\right)\right\}.$$

Proof We will assume that $x = 2k$ is even; the odd case is done similarly. We know that

$$\tilde p_t(2k) = \sum_{m=0}^{\infty} \mathbb{P}\{N_t = 2m\}\, p_{2m}(2k).$$

Standard exponential estimates, see (11.11), show that for every $\epsilon > 0$, there exist c, β such that $\mathbb{P}\{|N_t - t| \ge \epsilon t\} \le c\, e^{-\beta t}$. Hence,

$$
\begin{aligned}
\tilde p_t(2k) &= \sum_{m=0}^{\infty} \mathbb{P}\{N_t = 2m\}\, p_{2m}(2k) \\
&= O(e^{-\beta t}) + \sum \mathbb{P}\{N_t = 2m\}\, p_{2m}(2k), \qquad (2.42)
\end{aligned}
$$

where here and for the remainder of this proof, we write just $\sum$ to denote the sum over all integers m with $|t - 2m| \le \epsilon t$. We will show that there is an ϵ such that

$$\sum \mathbb{P}\{N_t = 2m\}\, p_{2m}(2k) = \frac{1}{\sqrt{2\pi t}}\, e^{-\frac{x^2}{2t}}\, \exp\left\{O\left(\frac{1}{\sqrt{t}} + \frac{|x|^3}{t^2}\right)\right\}.$$

A little thought shows that this and (2.42) imply the theorem.

By Proposition 2.5.3 we know that

$$p_{2m}(2k) = \mathbb{P}\{S_{2m} = 2k\} = \frac{1}{\sqrt{\pi m}}\, e^{-\frac{k^2}{m}}\, \exp\left\{O\left(\frac{1}{m} + \frac{k^4}{m^3}\right)\right\},$$

and by Proposition 2.5.5 we know that

$$\mathbb{P}\{N_t = 2m\} = \frac{1}{\sqrt{2\pi t}}\, e^{-\frac{(2m-t)^2}{2t}}\, \exp\left\{O\left(\frac{1}{\sqrt{t}} + \frac{|2m-t|^3}{t^2}\right)\right\}.$$

Also, we have

$$\frac{1}{2m} = \frac{1}{t}\,\left[1 + O\left(\frac{|2m-t|}{t}\right)\right], \qquad \frac{1}{\sqrt{2m}} = \frac{1}{\sqrt{t}}\,\left[1 + O\left(\frac{|2m-t|}{t}\right)\right],$$

which implies

$$e^{-\frac{k^2}{m}} = e^{-\frac{2k^2}{t}}\,\exp\left\{O\left(\frac{k^2\,|2m-t|}{t^2}\right)\right\}.$$

Combining all of this, we can see that the sum in (2.42) can be written as

$$\frac{1}{\sqrt{2\pi t}}\,e^{-\frac{x^2}{2t}}\exp\left\{O\left(\frac{1}{\sqrt{t}} + \frac{|x|^3}{t^2}\right)\right\}\sum\frac{2}{\sqrt{2\pi t}}\,e^{-\frac{(2m-t)^2}{2t}}\ \exp\left\{O\left(\frac{|2m-t|^3}{t^2}\right)\right\}.$$

We now choose ϵ so that $|O(|2m-t|^3/t^2)| \leq (2m-t)^2/(4t)$ for all $|2m-t| \leq \epsilon t$. We will now show that

$$\sum\frac{2}{\sqrt{2\pi t}}\,e^{-\frac{(2m-t)^2}{2t}}\ \exp\left\{O\left(\frac{|2m-t|^3}{t^2}\right)\right\} = 1 + O\left(\frac{1}{\sqrt{t}}\right),$$

which will complete the argument. Since

$$e^{-\frac{(2m-t)^2}{2t}}\ \exp\left\{O\left(\frac{|2m-t|^3}{t^2}\right)\right\} \leq e^{-\frac{(2m-t)^2}{4t}},$$

is easy to see that the sum over $|2m-t| > t^{2/3}$ decays faster than any power of t. For $|2m-t| \leq t^{2/3}$ we write

$$\exp\left\{O\left(\frac{|2m-t|^3}{t^2}\right)\right\} = 1 + O\left(\frac{|2m-t|^3}{t^2}\right).$$

The estimate

$$\begin{aligned}\sum_{|2m-t|\leq t^{2/3}}\frac{2}{\sqrt{2\pi t}}\,e^{-\frac{(2m-t)^2}{2t}} &= \sum_{|m|\leq t^{2/3}/2}\frac{2}{\sqrt{2\pi t}}\,e^{-2(m/\sqrt{t})^2}\\ &= O\left(\frac{1}{\sqrt{t}}\right) + 2\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-2y^2}\,dy = 1 + O\left(\frac{1}{\sqrt{t}}\right)\end{aligned}$$

is a standard approximation of an integral by a sum. Similarly,

$$\sum_{|2m-t|\leq t^{2/3}} O\left(\frac{|2m-t|^3}{t^2}\right)\,\frac{2}{\sqrt{2\pi t}}\,e^{-\frac{(2m-t)^2}{2t}} \leq \frac{c}{\sqrt{t}}\int_{-\infty}^{\infty}\frac{|y|^3}{\sqrt{2\pi}}e^{-2y^2}\,dy = O\left(\frac{1}{\sqrt{t}}\right).$$

□

Exercises

Exercise 2.1 Suppose $p \in \mathcal{P}'_d$, $\epsilon \in (0,1)$ and $\mathbb{E}[|X_1|^{2+\epsilon}] < \infty$. Show that the characteristic function has the expansion

$$\phi(\theta) = 1 - \frac{\theta\cdot\Gamma\theta}{2} + o(|\theta|^{2+\epsilon}), \quad \theta \to 0.$$

Show that the δ_n in (2.28) can be chosen so that $n^{\epsilon/2}\,\delta_n \to 0$.

Exercise 2.2 Show that if $p \in \mathcal{P}_d^*$, there exists a c such that for all $x \in \mathbb{Z}^d$ and all positive integers n,

$$|p_n(x) - p_n(0)| \leq c\,\frac{|x|}{n^{(d+1)/2}}.$$

(Hint: first show the estimate for aperiodic $p \in \mathcal{P}_d'$ with bounded support and then use (2.32). Alternatively, one can use Lemma 2.4.3 at time $n/2$, the Markov property, and (2.33).)

Exercise 2.3 Show that Lemma 2.3.2 holds for $p \in \mathcal{P}^*$.

Exercise 2.4 Suppose $p \in \mathcal{P}_d'$ with $\mathbb{E}[|X|^3] < \infty$. Show that there is a $c < \infty$ such that for all $|y| = 1$,

$$|p_n(0) - p_n(y)| \leq \frac{c}{n^{(d+2)/2}}.$$

Exercise 2.5 Suppose $p \in \mathcal{P}_d^*$. Let $A \subset \mathbb{Z}^d$ and

$$h(x) = \mathbb{P}^x\{S_n \in A \text{ i.o.}\}.$$

Show that if $h(x) > 0$ for some $x \in \mathbb{Z}^d$, then $h(x) = 1$ for all $x \in \mathbb{Z}^d$.

Exercise 2.6 Suppose S_n is a random walk with increment distribution $p \in \mathcal{P}_d$. Show that there exists a $b > 0$ such that

$$\sup_{n>1} \mathbb{E}\left[\exp\left\{\frac{b|S_n|^2}{n}\right\}\right] < \infty.$$

Exercise 2.7 Suppose $X_1, X_2, \ldots$ are independent, identically distributed random variables in $\mathbb{Z}^d$ with $\mathbb{P}\{X_1 = 0\} < 1$ and let $S_n = X_1 + \cdots + X_n$.

- Show that there exists an r such that for all n
$$\mathbb{P}\{|S_{rn^2}| \geq n\} \geq \frac{1}{2}.$$
- Show that there exist c, t such that for all $b > 0$,
$$\mathbb{P}\left\{\max_{1\leq j\leq n^2} |S_j| \leq bn\right\} \leq c\, e^{-t/b}.$$

Exercise 2.8 Find r_2, r_3 in (2.39).

Exercise 2.9 Let S_n denote one-dimensional simple random walk. In this exercise we will prove without using Stirling's formula that there exists a constant C such that

$$p_{2n}(0) = \frac{C}{\sqrt{n}}\left[1 - \frac{1}{8n} + O\left(\frac{1}{n^2}\right)\right].$$

a. Show that if $n \geq 1$,

$$p_{2(n+1)} = \left(1 + \frac{1}{2n}\right)\left(1 + \frac{1}{n}\right)^{-1} p_{2n}.$$

b. Let $b_n = \sqrt{n}\, p_{2n}(0)$. Show that $b_1 = 1/2$ and for $n \geq 1$,

$$\frac{b_{n+1}}{b_n} = 1 + \frac{1}{8n^2} + O\left(\frac{1}{n^3}\right).$$

c. Use this to show that $b_\infty = \lim b_n$ exists and is positive. Moreover,

$$b_n = b_\infty \left[1 - \frac{1}{8n} + O\left(\frac{1}{n^2}\right)\right].$$

Exercise 2.10 Show that if $p \in \mathcal{P}'_d$ with $\mathbb{E}[|X_1|^3] < \infty$, then

$$\nabla_j^2 p_n(x) = \nabla_j^2 \overline{p}_n(x) + O(n^{-(d+3)/2}).$$

Exercise 2.11 Suppose $q : \mathbb{Z}^d \to \mathbb{R}$ has finite support, and k is a positive integer such that for all $l \in \{1, \ldots, k-1\}$ and all $j_1, \ldots, j_l \in \{1, \ldots, d\}$,

$$\sum_{x=(x^1,\ldots,x^d)\in\mathbb{Z}^d} x^{j_1}\, x^{j_2}\, \ldots\, x^{j_l}\, q(x) = 0.$$

Then we call the operator

$$\Lambda f(x) := \sum_y f(x+y)\, q(y)$$

a difference operator of order (at least) k. The order of the operator is the largest k for which this is true. Suppose Λ is a difference operator of order $k \geq 1$.

- Suppose g is a C^∞ function on $\mathbb{R}^d$. Define g_ϵ on $\mathbb{Z}^d$ by $g_\epsilon(x) = g(\epsilon x)$. Show that
$$|\Lambda g_\epsilon(0)| = O(|\epsilon|^k), \quad \epsilon \to 0.$$
- Show that if $p \in \mathcal{P}'_d$ with $\mathbb{E}[|X_1|^3] < \infty$, then
$$\Lambda p_n(x) = \Lambda \overline{p}_n(x) + O(n^{-(d+1+k)/2}).$$
- Show that if $p \in \mathcal{P}'_d$ is symmetric with $\mathbb{E}[|X_1|^4] < \infty$, then
$$\Lambda p_n(x) = \Lambda \overline{p}_n(x) + O(n^{-(d+2+k)/2}).$$

3
Approximation by Brownian motion

3.1 Introduction

Suppose $S_n = X_1 + \cdots + X_n$ is a one-dimensional simple random. We make this into a (random) continuous function by linear interpolation,

$$S_t = S_n + (t - n)\,[S_{n+1} - S_n], \quad n \leq t \leq n + 1.$$

For fixed integer n, the LCLT describes the distribution of S_n. A corollary of LCLT is the usual central limit theorem that states that the distribution of $n^{-1/2}\,S_n$ converges to that of a standard normal random variable. A simple extension of this is the following: suppose $0 < t_1 < t_2 < \ldots < t_k = 1$. Then as $n \to \infty$ the distribution of

$$n^{-1/2}\,(S_{t_1 n}, S_{t_2 n}, \ldots, S_{t_k n})$$

converges to that of

$$(Y_1, Y_1 + Y_2, \ldots, Y_1 + Y_2 + \cdots Y_k),$$

where $Y_1, \ldots, Y_k$ are independent mean zero normal random varaibles with variables $t_1, t_2 - t_1, \ldots,$ $t_k - t_{k-1}$, respectively.

The *functional central limit theorem* (also called *the invariance principle* or *Donsker's theorem*) for random walk extends this result to the random function

$$W_t^{(n)} := n^{-1/2}\,S_{tn}. \tag{3.1}$$

The functional central limit theorem states roughly that as $n \to \infty$, the distribution of this random function converges to the distribution of a random function $t \mapsto B_t$. From what we know about the simple random walk, here are some properties that would be expected of the random function B_t:

- If $s < t$, the distribution of $B_t - B_s$ is $N(0, t - s)$.
- If $0 \leq t_0 < t_1 < \ldots < t_k$, then $B_{t_1} - B_{t_0}, \ldots, B_{t_k} - B_{t_{k-1}}$ are independent random variables.

These two properties follows almost immediately from the central limit theorem. The third property is not as obvious.

- The function $t \mapsto B_t$ is continuous.

Although this is not obvious, we can guess this from the heuristic argument:

$$\mathbb{E}[(B_{t+\Delta t} - B_t)^2] \approx \Delta t,$$

which indicates that $|B_{t+\Delta t} - B_t|$ should be of order $\sqrt{\Delta t}$. A process satisfying these assumptions will be called a *Brownian motion* (we will define it more precisely in the next section).

There are a number of ways to make rigorous the idea that $W^{(n)}$ approaches a Brownian motion in the limit. For example, if we restrict to $0 \leq t \leq 1$, then $W^{(n)}$ and B are random variables taking values in the metric space $C[0,1]$ with the supremum norm. There is a well understood theory of convergence in distribution of random variables taking values in metric spaces.

We will take a different approach using a method that is often called *strong approximation* of random walk by Brownian motion. We start by defining a Brownian motion B on a probability space and then define the random walk S_n as a function of the Brownian motion, i.e., for each realization of random function B_t we associate a particular random walk path. We will do this in a way so that the random walk S_n has the distribution of simple random walk. We will then do some estimates to show that there exist positive numbers c, a such that if $W_t^{(n)}$ is as defined in (3.1), then for all $r \leq n^{1/4}$,

$$\mathbb{P}\{\|B - W^{(n)}\| \geq r\, n^{-1/4}\, \sqrt{\log n}\,\} \leq c\, e^{-ra}, \tag{3.2}$$

where $\|\cdot\|$ denotes the supremum norm on $C[0,1]$. The convergence in distribution follows from the strong estimate (3.2).

Comment. There is a general approach here that is worth emphasizing. Suppose we have a discrete process and we want to show that it converges after some scaling to a continuous process. A good approach for proving such a result is to first study the conjectured limit process and then to show that the scaled discrete process is a small perturbation of the limit process.

We start by establishing (3.2) for one-dimensional simple random walk using *Skorokhod embedding*. We then extend this to continuous-time walks and all increment distributions $p \in \mathcal{P}$. The extension will not be difficult; the hard work is done in the one-dimensional case.

We will not handle the general case of $p \in \mathcal{P}'$ in this book. One can give strong approximations in this case to show that the random walk approaches Brownian motion. However, the rate of convergence depends on the moment assumptions. In particular, the estimate (3.2) will not hold assuming only mean zero and finite second moment.

3.2 Construction of Brownian motion

A *standard (one-dimensional) Brownian motion with respect to a filtration* $\mathcal{F}_t$ is a collection of random variables $B_t, t \geq 0$ satisfying the following:

(a) $B_0 = 0$;

(b) if $s < t$, then $B_t - B_s$ is an $\mathcal{F}_t$-measurable random variable, independent of $\mathcal{F}_s$, with a $N(0, t-s)$ distribution;

(c) with probability one, $t \mapsto B_t$ is a continuous function.

If the filtration is not given explicitly, then it is assumed to be the natural filtration, $\mathcal{F}_t = \sigma\{B_s : 0 \leq s \leq t\}$. In this section, we will construct a Brownian motion and derive an important estimate on the oscillations of the Brownian motion.

We will show how to construct a Brownian motion. There are technical difficulties involved in defining a collection of random variables $\{B_t\}$ indexed over an uncountable set. However, if we know *a priori* that the distribution should be supported on continuous functions, then we know that the random function $t \mapsto B_t$ should be determined by its value on a countable, dense subset of times. This observation leads us to a method of constructing Brownian motion: define the process on a countable, dense set of times and then extend the process to all times by continuity.

Suppose $(\Omega, \mathcal{F}, \mathbb{P})$ is any probability space that is large enough to contain a countable collection of independent $N(0,1)$ random variables which for ease we will index by

$$N_{n,k}, \quad n = 0, 1, \ldots; \quad k = 0, 1, \ldots$$

We will use these random variables to define a Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$. Let

$$D_n = \left\{ \frac{k}{2^n} : \quad k = 0, 1, \ldots \right\}, \qquad D = \bigcup_{n=0}^{\infty} D_n$$

denote the nonnegative dyadic rationals. Our strategy will be as follows:

- define B_t for t in D satisfying conditions (a) and (b);
- derive an estimate on the oscillation of $B_t, t \in D$, that implies that with probability one the paths are uniformly continuous on compact intervals;
- define B_t for other values of t by continuity.

The first step is straightforward using a basic property of normal random variables. Suppose X, Y are independent normal random variables, each mean 0 and variance $1/2$. Then $Z = X + Y$ is $N(0,1)$. Moreover, the conditional distribution of X given the value of Z is normal with mean $Z/2$ and variance $1/2$. This can be checked directly using the density of the normals. Alternatively, one can check that if Z, N are independent $N(0,1)$ random variables then

$$X := \frac{Z}{2} + \frac{N}{2}, \qquad Y := \frac{Z}{2} - \frac{N}{2}, \tag{3.3}$$

are independent $N(0, 1/2)$ random variables. To verify this, one only notes that (X, Y) has a joint normal distribution with $\mathbb{E}[X] = \mathbb{E}[Y] = 0, \mathbb{E}[X^2] = \mathbb{E}[Y^2] = 1/2, \mathbb{E}[XY] = 0$. (See Corollary 11.3.1.) This tells us that in order to define X, Y we can start with independent random variables N, Z and then use (3.3).

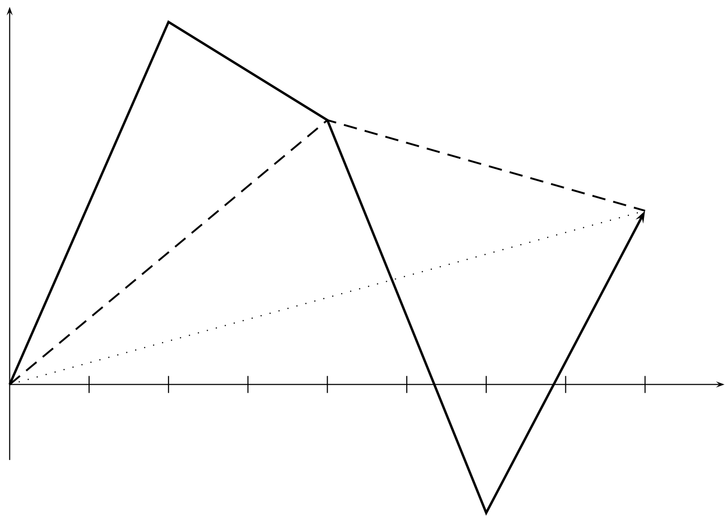

Figure 3.1: The dyadic construction

We start by defining B_t for $t \in D_0$ by $B_0 = 0$ and

$$B_j = N_{0,1} + \cdots + N_{0,j}.$$

We then continue recursively using (3.3). Suppose B_t has been defined for all $t \in D_n$. Then we define B_t for $t \in D_{n+1} \setminus D_n$ by

$$B_{\frac{2k+1}{2^{n+1}}} = B_{\frac{k}{2^n}} + \frac{1}{2}\left[B_{\frac{k+1}{2^n}} - B_{\frac{k}{2^n}}\right] + 2^{-(n+2)/2}\, N_{2k+1,n+1}.$$

By induction, one can check that for each n the collection of random variables $Z_{k,n} := B_{k/2^n} - B_{(k-1)/2^n}$ are independent, each with a $N(0, 2^{-n})$ distribution. Since this is true for each n, we can see that (a) and (b) hold (with the natural filtration) provided that we restrict to $t \in D$. The scaling property for normal random variables shows that for each integer n, the distribution of the random variables

$$2^{n/2}\, B_{t/2^n}, \quad t \in D,$$

has the same distribution as the random variables

$$B_t, \quad t \in D.$$

We define the *oscillation* of B_t (restricted to $t \in D$) by

$$\mathrm{osc}(B;\delta,T) = \sup\{|B_t - B_s| : s,t \in D; 0 \le s,t \le T; |s-t| \le \delta\}.$$

For fixed δ, T, this is an $\mathcal{F}_T$-measurable random variable. We write $\mathrm{osc}(B;\delta)$ for $\mathrm{osc}(B;\delta,1)$. Let

$$M_n = \max_{0 \le k < 2^n} \sup\left\{|B_{t+k2^{-n}} - B_{k2^{-n}}| : t \in D; 0 \le t \le 2^{-n}\right\}.$$

The random variable M_n is similar to $\mathrm{osc}(B;2^{-n})$ but is easier to analyze. Note that if $r \le 2^{-n}$,

$$\mathrm{osc}(B;r) \le \mathrm{osc}(B;2^{-n}) \le 3\, M_n. \tag{3.4}$$

To see this, suppose $\delta \leq 2^{-n}, 0 < s < t \leq s + \delta \leq 1$, and $|B_s - B_t| \geq \epsilon$. Then there exists a k such that either $k2^{-n} \leq s < t \leq (k+1)2^{-n}$ or $(k-1)2^{-n} \leq s \leq k2^{-n} < t \leq (k+1)2^{-n}$. In either case, the triangle inequality tells us that $M_n \geq \epsilon/3$. We will prove a proposition that bounds the probability of large values of $\mathrm{osc}(B; \delta, T)$. We start with a lemma which gives a similar bound for M_n.

Lemma 3.2.1 *For every integer n and every $\delta > 0$,*

$$\mathbb{P}\{M_n > \delta\, 2^{-n/2}\} \leq 4 \sqrt{\frac{2}{\pi}}\, \frac{2^n}{\delta}\, e^{-\delta^2/2}.$$

Proof Note that

$$\mathbb{P}\{M_n > \delta\, 2^{-n/2}\} \leq 2^n\, \mathbb{P}\left\{ \sup_{0 \leq t \leq 2^{-n}} |B_t| > \delta\, 2^{-n/2} \right\} = 2^n\, \mathbb{P}\left\{ \sup_{0 \leq t \leq 1} |B_t| > \delta \right\}.$$

Here the supremums are taken over $t \in D$. Also note that

$$\begin{aligned} \mathbb{P}\{ \sup\{|B_t| : 0 \leq t \leq 1, t \in D\} > \delta \} &= \lim_{n\to\infty} \mathbb{P}\{ \max\{|B_{k2^{-n}}| : k = 1, \ldots, 2^n\} > \delta \} \\ &\leq 2 \lim_{n\to\infty} \mathbb{P}\{ \max\{B_{k2^{-n}} : k = 1, \ldots, 2^n\} > \delta \}. \end{aligned}$$

The reflection principle (see Proposition 1.5.2 and the remark following) shows that

$$\begin{aligned} \mathbb{P}\{ \max\{B_{k2^{-n}} : k = 1, \ldots, 2^n\} > \delta \} &\leq 2\, \mathbb{P}\{B_1 > \delta\} \\ &= 2 \int_\delta^\infty \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}\, dx \\ &\leq 2 \int_\delta^\infty \frac{1}{\sqrt{2\pi}}\, e^{-x\delta/2}\, dx = 2\sqrt{\frac{2}{\pi}}\, \delta^{-1} e^{-\delta^2/2}. \end{aligned}$$

□

Proposition 3.2.2 *There exists a $c > 0$ such that for every $0 < \delta \leq 1$, $r \geq 1$, and positive integer T,*

$$\mathbb{P}\{\mathrm{osc}(B; \delta, T) > c\, r \sqrt{\delta \log(1/\delta)}\} \leq c\, T\, \delta^{r^2}.$$

Proof It suffices to prove the result for $T = 1$ since for general T we can estimate separately the oscillations over the $2T - 1$ intervals $[0, 1]$, $[1/2, 3/2]$, $[1, 2]$, $\ldots$, $[T - 1, T]$. Also, it suffices to prove the result for $\delta \leq 1/4$. Suppose that $2^{-n-1} \leq \delta \leq 2^{-n}$. Using (3.4), we see that

$$\mathbb{P}\{\mathrm{osc}(B; \delta) > c\, r \sqrt{\delta \log(1/\delta)}\} \leq \mathbb{P}\left\{ M_n > \frac{cr}{3\sqrt{2}} \sqrt{2^{-n} \log(1/\delta)} \right\}.$$

By Lemma 3.2.1, if c is chosen sufficiently large, the probability on the right-hand side is bounded by a constant times

$$\exp\left\{ -\frac{1}{4} \left(\frac{c^2 r^2}{18}\right) \log(1/\delta) \right\},$$

which for c large enough is bounded by a constant times δ^{r^2}. □

Corollary 3.2.3 *With probability one, for every integer $T < \infty$, the function $t \mapsto B_t, t \in D$ is uniformly continuous on $[0, T]$.*

Proof Uniform continuity on $[0, T]$ is equivalent to saying that $\operatorname{osc}(B; 2^{-n}, T) \longrightarrow 0$ as $n \to \infty$. The previous proposition implies that there is a c_1 such that

$$\mathbb{P}\{\operatorname{osc}(B; 2^{-n}, T) > c_1 \, 2^{-n/2} \sqrt{n}\} \leq c_1 \, T \, 2^{-n}.$$

In particular,

$$\sum_{n=1}^{\infty} \mathbb{P}\{\operatorname{osc}(B; 2^{-n}, T) > c_1 \, 2^{-n/2} \sqrt{n}\} < \infty,$$

which implies by Borel-Cantelli that with probability one $\operatorname{osc}(B; 2^{-n}, T) \leq c_1 \, T \, 2^{-n/2} \sqrt{n}$ for all n sufficiently large. □

Given the corollary, we can define B_t for $t \notin D$ by continuity, i.e.,

$$B_t = \lim_{t_n \to t} B_{t_n},$$

where $t_n \in D$ with $t_n \to t$. It is not difficult to show that this satisfies the definition of Brownian motion (we omit the details). Moreover, since B_t has continuous paths, we can write

$$\operatorname{osc}(B; \delta, T) = \sup\{|B_t - B_s| : 0 \leq s, t \leq T; |s - t| > \delta\}.$$

We restate the estimate and include a fact about scaling of Brownian motion. Note that if B_t is a standard Brownian motion and $a > 0$, then $Y_t := a^{-1/2} \, B_{at}$ is also a standard Brownian motion.

Theorem 3.2.4 (Modulus of continuity of Brownian motion) *There is a $c < \infty$ such that if B_t is a standard Brownian motion, $0 < \delta \leq 1$, $r \geq 1, T \geq 1$,*

$$\mathbb{P}\{\operatorname{osc}(B; \delta, T) > c \, r \, \sqrt{\delta \log(1/\delta)}\} \leq c \, T \, \delta^{r^2}.$$

Moreover, if $T > 0$, then $\operatorname{osc}(B; \delta, T)$ has the same distribution as $\sqrt{T} \operatorname{osc}(B, \delta/T)$. In particular, if $T \geq 1$,

$$\mathbb{P}\{\operatorname{osc}(B; 1, T) > c \, r \, \sqrt{\log T}\} = \mathbb{P}\{\operatorname{osc}(B; 1/T) > c \, r \, \sqrt{(1/T) \, \log T}\} \leq c \, T^{-r^2}. \tag{3.5}$$

3.3 Skorokhod embedding

We will now define a procedure that takes a Brownian motion path B_t and produces a random walk S_n. The idea is very straightforward. Start the Brownian motion and wait until it reaches $+1$ or -1. If it hits $+1$ first we let $S_1 = 1$; otherwise, we set $S_1 = -1$. Now we wait until the new increment of the Brownian motion reaches $+1$ or -1 and we use this value for the increment of the random walk.

To be more precise, let B_t be a standard one-dimensional Brownian motion, and let

$$\tau = \inf\{t \geq 0 : |B_t| = 1\}.$$

Symmetry tells us that $\mathbb{P}\{B_\tau = 1\} = \mathbb{P}\{B_\tau = -1\} = 1/2$.

Lemma 3.3.1 $\mathbb{E}[\tau] = 1$ *and there exists a* $b < \infty$ *such that* $\mathbb{E}[e^{b\tau}] < \infty$.

Proof Note that for integer n

$$\mathbb{P}\{\tau > n\} \leq \mathbb{P}\{\tau > n-1, |B_n - B_{n-1}| \leq 2\} = \mathbb{P}\{\tau > n-1\}\, \mathbb{P}\{|B_n - B_{n-1}| \leq 2\},$$

which implies for integer n,

$$\mathbb{P}\{\tau > n\} \leq \mathbb{P}\{|B_n - B_{n-1}| \leq 2\}^n = e^{-\rho n},$$

with $\rho > 0$. This implies that $\mathbb{E}[e^{b\tau}] < \infty$ for $b < \rho$. If $s < t$, then $\mathbb{E}[B_t^2 - t \mid \mathcal{F}_s] = B_s^2 - s$ (Exercise 3.1). This shows that $B_t^2 - t$ is a continuous martingale. Also,

$$\mathbb{E}[|B_t^2 - t|; \tau > t] \leq (t+1)\, \mathbb{P}\{\tau > t\} \longrightarrow 0.$$

Therefore, we can use the optional sampling theorem (Theorem 11.2.10) to conclude that $\mathbb{E}[B_\tau^2 - \tau] = 0$. Since $\mathbb{E}[B_\tau^2] = 1$, this implies that $\mathbb{E}[\tau] = 1$. □

More generally, let $\tau_0 = 0$ and

$$\tau_n = \inf\{t \geq \tau_{n-1} : |B_t - B_{\tau_{n-1}}| = 1\}.$$

Then $S_n := B_{\tau_n}$ is a simple one-dimensional random walk†. Let $T_n = \tau_n - \tau_{n-1}$. The random variables $T_1, T_2, \ldots$ are independent, identically distributed, with mean one satisfying $\mathbb{E}[e^{bT_j}] < \infty$ for some $b > 0$. As before, we define S_t for noninteger t by linear interpolation. Let

$$\Theta(B, S; n) = \max\{|B_t - S_t| : 0 \leq t \leq n\}.$$

In other words, $\Theta(B, S; n)$ is the distance between the continuous functions B and S in $C[0, n]$ using the usual supremum norm. If $j \leq t < j+1 \leq n$, then

$$|B_t - S_t| \leq |S_j - S_t| + |B_j - B_t| + |B_j - S_j| \leq 1 + \operatorname{osc}(B; 1, n) + |B_j - B_{\tau_j}|.$$

Hence for integer n,

$$\Theta(B, S; n) \leq 1 + \operatorname{osc}(B; 1, n) + \max\{|B_j - B_{\tau_j}| : j = 1, \ldots, n\}. \tag{3.6}$$

We can estimate the probabilities for the second term with (3.5). We will concentrate on the last term. Before doing the harder estimates, let us consider how large an error we should expect. Since $T_1, T_2, \ldots$ are i.i.d. random variables with mean 1 and finite variance, the central limit theorem says roughly that

$$|\tau_n - n| = \left| \sum_{j=1}^{n} [T_j - 1] \right| \approx \sqrt{n}.$$

Hence we would expect that

$$|B_n - B_{\tau_n}| \approx \sqrt{|\tau_n - n|} \approx n^{1/4}.$$

From this reasoning, we can see that we expect $\Theta(B, S; n)$ to be at least of order $n^{1/4}$. The next theorem shows that it is unlikely that the actual value is much greater than $n^{1/4}$.

† We actually need the *strong Markov property* for Brownian motion to justify this and the next assertion. This is not difficult to prove, but we will not do it in this book.

Theorem 3.3.2 *There exist* $0 < c, a < \infty$ *such that for all* $r \leq n^{1/4}$ *and all integers* $n \geq 3$

$$\mathbb{P}\{\Theta(B, S; n) > r\, n^{1/4}\, \sqrt{\log n}\} \leq c\, e^{-ar}.$$

Proof It suffices to prove the theorem for $r \geq 9$. Suppose $1 \leq r \leq n^{1/4}$. If $|B_n - B_{\tau_n}|$ is large, then either $|n - \tau_n|$ is large or the oscillation of B is large. Using (3.6), we see that the event $\{\Theta(B, S; n) \geq r\, n^{1/4}\, \sqrt{\log n}\}$ is contained in the union of the two events

$$\{\, \mathrm{osc}(B; r\sqrt{n}, 2n) \geq (r/3)\, n^{1/4}\, \sqrt{\log n}\, \},$$

$$\left\{ \max_{1 \leq j \leq n} |\tau_j - j| \geq r\, \sqrt{n} \right\}.$$

Indeed, if $\mathrm{osc}(B; r\sqrt{n}, 2n) \leq (r/3)\, n^{1/4}\, \sqrt{\log n}$ and $|\tau_j - j| \leq r\, \sqrt{n}$ for $j = 1, \ldots, n$, then the three terms on the right-hand side of (3.6) are each bounded by $(r/3)\, n^{1/4}\, \sqrt{\log n}$.

Theorem 3.2.4 gives for $1 \leq r \leq n^{1/4}$,

$$\begin{aligned}
\mathbb{P}\{\, \mathrm{osc}(B; r\sqrt{n}, 2n) > (r/3)\, n^{1/4}\, \sqrt{\log n}\, \} & \\
\leq\ & 3\, \mathbb{P}\{\, \mathrm{osc}(B; r\sqrt{n}, n) > (r/3)\, n^{1/4}\, \sqrt{\log n}\, \} \\
=\ & 3\, \mathbb{P}\{\, \mathrm{osc}(B; r\, n^{-1/2}) > (r/3)\, n^{-1/4}\, \sqrt{\log n}\} \\
\leq\ & 3\, \mathbb{P}\left\{ \mathrm{osc}(B; r\, n^{-1/2}) > (\sqrt{r}/3)\, \sqrt{r\, n^{-1/2}\, \log(n^{1/2}/r)} \right\} \\
\leq\ & c\, e^{-ar}
\end{aligned}$$

For the second event, consider the martingale

$$M_j = \tau_j - j.$$

Using (11.11) on M_j and $-M_j$, we get

$$\mathbb{P}\left\{ \max_{1 \leq j \leq n} |\tau_j - j| \geq r\, \sqrt{n} \right\} \leq c\, e^{-r}. \tag{3.7}$$

□

Extending the Skorokhod approximation to continuous time simple random walk $\tilde{S}_t$ is not difficult although in this case the path $t \mapsto \tilde{S}_t$ is not continuous. Let N_t be a Poisson process with parameter 1 defined on the same probability space and independent of the Brownian motion B. Then

$$\tilde{S}_t := S_{N_t}$$

has the distribution of the continuous-time simple random walk. Since $N_t - t$ is a martingale, and the Poisson distribution has exponential moments, another application of (11.11) shows that for $r \geq 1$,

$$\mathbb{P}\left\{ \max_{0 \leq s \leq t} |N_s - s| \geq r\, \sqrt{t} \right\} \leq c\, e^{-r}.$$

Let

$$\Theta(B, \tilde{S}; n) = \sup\{|B_t - \tilde{S}_t| : 0 \leq t \leq n\}.$$

Then the following is proved similarly.

Theorem 3.3.3 *There exist* $0 < c, a < \infty$ *such that for all* $1 \leq r \leq n^{1/4}$ *and all positive integers* n

$$\mathbb{P}\{\Theta(B, \tilde{S}; n) \geq r\, n^{1/4}\, \sqrt{\log n}\} \leq c\, e^{-ar}.$$

3.4 Higher dimensions

It is not difficult to extend Theorems 3.3.2 and 3.3.3 to $p \in \mathcal{P}_d$ for $d > 1$. A d-dimensional Brownian motion with covariance matrix Γ with respect to a filtration $\mathcal{F}_t$ is a collection of random variables $B_t, t \geq 0$ satisfying the following:

(a) $B_0 = 0$;

(b) if $s < t$, then $B_t - B_s$ is an $\mathcal{F}_t$-measurable random $\mathbb{R}^d$-valued variable, independent of $\mathcal{F}_s$, whose distribution is joint normal with mean zero and covariance matrix $(t - s)\,\Gamma$.

(c) with probability one, $t \mapsto B_t$ is a continuous function.

Lemma 3.4.1 *Suppose* $B^{(1)}, \ldots, B^{(l)}$ *are independent one-dimensional standard Brownian motions and* $v_1, \ldots, v_l \in \mathbb{R}^d$. *Then*

$$B_t := B_t^{(1)}\, v_1 + \cdots + B_t^{(l)}\, v_l$$

is a Brownian motion in $\mathbb{R}^d$ *with covariance matrix* $\Gamma = AA^T$ *where* $A = [v_1\, v_2\, \cdots\, v_l]$.

Proof Straightforward and left to the reader. □

In particular, a standard d-dimensional Brownian motion is of the form

$$B_t = (B_t^{(1)}, \ldots, B_t^{(d)})$$

where $B^{(1)}, \ldots, B^{(d)}$ are independent one-dimensional Brownian motions. Its covariance matrix is the identity.

The next theorem shows that one can define d-dimensional Brownian motions and d-dimensional random walks on the same probability space so that their paths are close to each other. Although the proof will use Skorokhod embedding, it is not true that the d-dimensional random walk is embedded into the d-dimensional Brownian motion. In fact, it is easy to see that it is impossible to have an embedded walk since if $d > 1$, with probability one a d-dimensional Brownian motion B_t visits the countable set $\mathbb{Z}^d$ after time 0 is zero.

Theorem 3.4.2 *Let* $p \in \mathcal{P}_d$ *with covariance matrix* Γ. *There exist* c, a *and a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *on which are defined a Brownian motion* B *with covariance matrix* Γ; *a discrete-time random walk* S *with increment distribution* p; *and a continuous-time random walk* $\tilde{S}$ *with increment distribution* p *such that for all positive integers* n *and all* $1 \leq r \leq n^{1/4}$,

$$\mathbb{P}\{\Theta(B, S; n) \geq r\, n^{1/4}\, \sqrt{\log n}\} \leq c\, e^{-ar},$$

$$\mathbb{P}\{\Theta(B, \tilde{S}; n) \geq r\, n^{1/4}\, \sqrt{\log n}\} \leq c\, e^{-ar}.$$

Proof Suppose $v_1, \ldots, v_l$ are the points such $p(v_j) = p(-v_j) = q_j/2$ and $p(z) = 0$ for all other $z \in \mathbb{Z}^d \setminus \{0\}$. Let $L_n = (L_n^1, \ldots, L_n^l)$ is a multinomial process with parameters $q_1, \ldots, q_l$, and let $B^1, \ldots, B^l$ be independent one-dimensional Brownian motions. Let $S^1, \ldots, S^l$ be the random walks derived from $B^1, \ldots, B^l$ by Skorokhod embedding. As was noted in (1.2),

$$S_n := S^1_{L_n^1}\, v_1 + \ldots + S^l_{L_n^l}\, v_l,$$

has the distribution of a random walk with increment distribution p. Also,

$$B_t := B_t^1\, v_1 + \cdots + B_t^l\, v_l,$$

is a Brownian motion with covariance matrix Γ. The proof now proceeds as in the previous cases. One fact that is used is that the L_n^j have a binomial distribution and hence we can get an exponential estimate

$$\mathbb{P}\left\{ \max_{1\le j\le n} |L_j^i - q^i\, j| \ge a\sqrt{n} \right\} \le c\, e^{-a}.$$

$\square$

3.5 An alternative formulation

Here we give a slightly different, but equivalent, form of the strong approximation from which we get (3.2). We will illustrate this in the case of one-dimensional simple random walk. Suppose B_t is a standard Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$). For positive integer n, let $B_t^{(n)}$ denote the Brownian motion

$$B_t^{(n)} = n^{-1/2}\, B_{nt}.$$

Let $S_t^{(n)}$ denote the simple random walk derived from $B^{(n)}$ using the Skorokhod embedding. Then we know that for all positive integers T,

$$\mathbb{P}\left\{ \max_{0\le t\le Tn} |S_t^{(n)} - B_t^{(n)}| \ge c\, r\, (Tn)^{1/4}\, \sqrt{\log(Tn)} \right\} \le c\, e^{-ar}.$$

If we let

$$W_t^{(n)} = n^{-1/2}\, S_{tn}^{(n)},$$

then this becomes

$$\mathbb{P}\left\{ \max_{0\le t\le T} |W_t^{(n)} - B_t| \ge c\, r\, T^{1/4}\, n^{-1/4}\, \sqrt{\log(Tn)} \right\} \le c\, e^{-ar}.$$

In particular, if $r = c_1 \log n$ where $c_1 = c_1(T)$ is chosen sufficiently large,

$$\mathbb{P}\left\{ \max_{0\le t\le T} |W_t^{(n)} - B_t| \ge c_1\, n^{-1/4}\, \log^{3/2} n \right\} \le c_1\, n^{-2}.$$

By the Borel-Cantelli lemma, with probability one

$$\max_{0\le t\le T} |W_t^{(n)} - B_t| \ge c_1\, n^{-1/4}\, \log^{3/2} n$$

for all n sufficiently large. In particular, with probability one $W^{(n)}$ converges to B in the metric space $C[0,T]$.

By using a multinomial process (in the discrete-time case) or a Poisson process (in the continuous-time) case, we can prove the following.

Theorem 3.5.1 *Suppose $p \in \mathcal{P}_d$ with covariance matrix Γ. There exist $c < \infty, a > 0$ and a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which are defined a d-dimensional Brownian motion B_t with covariance matrix Γ; an infinite sequence of discrete-time p-walks, $S^{(1)}, S^{(2)}, \ldots$; and an infinite sequence of continuous time p-walks $\tilde{S}^{(1)}, \tilde{S}^{(2)}, \ldots$ such that the following holds for every $r > 0, T \geq 1$. Let*

$$W_t^{(n)} = n^{-1/2}\, S_{nt}^{(n)}, \qquad \tilde{W}_t^{(n)} = n^{-1/2}\, \tilde{S}_{nt}^{(n)}.$$

Then,

$$\mathbb{P}\left\{\max_{0 \leq t \leq T} |W_t^{(n)} - B_t| \geq c\, r\, T^{1/4}\, n^{-1/4}\, \sqrt{\log(Tn)}\right\} \leq c\, e^{-ar}.$$

$$\mathbb{P}\left\{\max_{0 \leq t \leq T} |\tilde{W}_t^{(n)} - B_t| \geq c\, r\, T^{1/4}\, n^{-1/4}\, \sqrt{\log(Tn)}\right\} \leq c\, e^{-ar}.$$

In particular, with probability one $W^{(n)} \to B$ and $\tilde{W}^{(n)} \to B$ in the metric space $C^d[0,T]$.

Exercises

Exercise 3.1 Show that if B_t is a standard Brownian motion with respect to the filtration $\mathcal{F}_t$ and $s < t$, then $\mathbb{E}[B_t^2 - t \mid \mathcal{F}_s] = B_s^2 - s$.

Exercise 3.2 Let X be an integer-valued random variable with $\mathbb{P}\{X = 0\} = 0$ and $\mathbb{E}[X] = 0$.

(a) Show that there exists numbers $r_j \in (0, \infty]$,

$$r_1 \leq r_2 \leq \cdots, \qquad r_{-1} \leq r_{-2} \leq \cdots,$$

such that if B_t is a standard Brownian motion and

$$T = \inf\{t : B_t \in \mathbb{Z} \setminus \{0\}, t \leq r_{B_t}\},$$

then B_T has the same distribution as X.

(b) Show that if X has bounded support, then there exists a $b > 0$ with $\mathbb{E}[e^{bT}] < \infty$.

(c) Show that $\mathbb{E}[T] = \mathbb{E}[X^2]$.

(Hint: you may wish to consider first the cases where X is supported on $\{1, -1\}$, $\{1, 2, -1\}$, and $\{1, 2, -1, -2\}$, respectively.)

4
Green's Function

4.1 Recurrence and transience

A random walk S_n with increment distribution $p \in \mathcal{P}_d \cup \mathcal{P}_d^*$ is called *recurrent* if $\mathbb{P}\{S_n = 0 \text{ i.o.}\} = 1$. If the walk is not recurrent it is called *transient.* We will also say that p is recurrent or transient. It is easy to see using the Markov property that p is recurrent if and only if for each $x \in \mathbb{Z}^d$,

$$\mathbb{P}^x\{S_n = 0 \text{ for some } n \geq 1\} = 1,$$

and p is transient if and only if the *escape probability*, q, is positive, where q is defined by

$$q = \mathbb{P}\{S_n \neq 0 \text{ for all } n \geq 1\}.$$

Theorem 4.1.1 *If $p \in \mathcal{P}_d'$ with $d = 1, 2$, then p is recurrent. If $p \in \mathcal{P}_d^*$ with $d \geq 3$, then p is transient. For all p,*

$$q = \left[\sum_{n=0}^{\infty} p_n(0)\right]^{-1}, \tag{4.1}$$

where the left-hand equals zero if the sum is divergent.

Proof Let $Y = \sum_{n=0}^{\infty} 1\{S_n = 0\}$ denote the number of visits to the origin and note that

$$\mathbb{E}[Y] = \sum_{n=0}^{\infty} \mathbb{P}\{S_n = 0\} = \sum_{n=0}^{\infty} p_n(0).$$

If $p \in \mathcal{P}_d'$ with $d = 1, 2$, the LCLT (see Theorem 2.1.1 and Theorem 2.3.9) implies that $p_n(0) \sim c\, n^{-d/2}$ and the sum is infinite. If $p \in \mathcal{P}_d^*$ with $d \geq 3$, then (2.33) shows that $p_n(0) \leq c\, n^{-d/2}$ and hence $\mathbb{E}[Y] < \infty$. We can compute $\mathbb{E}(Y)$ in terms of q. Indeed, the Markov property shows that, $\mathbb{P}\{Y = j\} = (1-q)^{j-1}\, q$. Therefore, if $q > 0$,

$$\mathbb{E}[Y] = \sum_{j=0}^{\infty} j\, \mathbb{P}\{Y = j\} = \sum_{j=0}^{\infty} j\, (1-q)^{j-1}\, q = \frac{1}{q}.$$

$\square$

4.2 Green's generating function

If $p \in \mathcal{P} \cup \mathcal{P}^*$ and $x, y \in \mathbb{Z}^d$, we define the *Green's generating function* to be the power series in ξ:

$$G(x, y; \xi) = \sum_{n=0}^{\infty} \xi^n \, \mathbb{P}^x\{S_n = y\} = \sum_{n=0}^{\infty} \xi^n \, p_n(y - x).$$

Note that the sum is absolutely convergent for $|\xi| < 1$. We write just $G(y; \xi)$ for $G(0, y, \xi)$. If $p \in \mathcal{P}$, then $G(x; \xi) = G(-x; \xi)$.

The generating function is defined for complex ξ, but there is a very nice interpretation of the sum for positive $\xi \leq 1$. Suppose T is a random variable independent of the random walk S with a geometric distribution,

$$\mathbb{P}\{T = j\} = \xi^{j-1} \, (1 - \xi), \qquad j = 1, 2, \ldots,$$

i.e., $\mathbb{P}\{T > j\} = \xi^j$. We think of T as a "killing time" for the walk and we will refer to such T as a geometric random variable with *killing rate* $1 - \xi$. At each time j, if the walker has not already been killed, the process is killed with probability $1 - \xi$, where the killing is independent of the walk. If the random walk starts at the origin, then the expected number of visits to x before being killed is given by

$$\mathbb{E}\left[\sum_{j<T} 1\{S_j = x\}\right] = \mathbb{E}\left[\sum_{j=0}^{\infty} 1\{S_j = x; T > j\}\right]$$

$$= \sum_{j=0}^{\infty} \mathbb{P}\{S_j = x; T > j\} = \sum_{j=0}^{\infty} p_j(x) \, \xi^j = G(x; \xi).$$

Theorem 4.1.1 states that a random walk is transient if and only if $G(0; 1) < \infty$, in which case the escape probability is $G(0; 1)^{-1}$. For a transient random walk, we define the *Green's function* to be

$$G(x, y) = G(x, y; 1) = \sum_{n=0}^{\infty} p_n(y - x).$$

We write $G(x) = G(0, x)$; if $p \in \mathcal{P}$, then $G(x) = G(-x)$. The strong Markov property implies that

$$G(0, x) = \mathbb{P}\{S_n = x \text{ for some } n \geq 0\} \, G(0, 0). \tag{4.2}$$

Similarly, we define

$$\tilde{G}(x, y; \xi) = \int_0^{\infty} \xi^t \, p_t(x, y) \, dt.$$

For $\xi \in (0, 1)$ this is the expected amount of time spent at site y by a continuous-time random walk with increment distribution p before an independent "killing time" that has an exponential distribution with rate $-\log(1 - \xi)$. We will now show that if we set $\xi = 1$, we get the same Green's function as that induced by the discrete walk.

Proposition 4.2.1 *If $p \in \mathcal{P}_d^*$ is transient, then*

$$\int_0^{\infty} \tilde{p}_t(x) \, dt = G(x).$$

Proof Let S_n denote a discrete-time walk with distribution p, N_t an independent Poisson process with parameter 1, and let $\tilde{S}_t$ denote the continuous-time walk $\tilde{S}_t = S_{N_t}$. Let

$$Y_x = \sum_{n=0}^{\infty} 1\{S_n = x\}, \qquad \tilde{Y}_x = \int_0^{\infty} 1\{\tilde{S}_t = x\}\, dt,$$

denote the amount of time spent at x by S and $\tilde{S}$, respectively. Then $G(x) = \mathbb{E}[Y_x]$. If we let $T_n = \inf\{t : N_t = n\}$, then we can write

$$\tilde{Y}_x = \sum_{n=0}^{\infty} 1\{S_n = x\}\,(T_{n+1} - T_n).$$

Independence of S and N implies

$$\mathbb{E}[1\{S_n = x\}\,(T_{n+1} - T_n)] = \mathbb{P}\{S_n = x\}\,\mathbb{E}[T_{n+1} - T_n] = \mathbb{P}\{S_n = x\}.$$

Hence $\mathbb{E}[\tilde{Y}_x] = \mathbb{E}[Y_x]$. □

Remark. Suppose p is the increment distribution of a random walk in $\mathbb{Z}^d$. For $\epsilon > 0$, let p_ϵ denote the increment of the "lazy walker" given by

$$p_\epsilon(x) = \begin{cases} (1-\epsilon)\,p(x), & x \neq 0 \\ \epsilon + (1-\epsilon)\,p(0), & x = 0 \end{cases}$$

If p is irreducible and periodic on $\mathbb{Z}^d$, then for each $\epsilon > 0$, p_ϵ is irreducible and aperiodic. Let $\mathcal{L}, \phi$ denote the generator and characteristic function for p, respectively. Then the generator and characteristic fucntion for p_ϵ are

$$\mathcal{L}_\epsilon = (1-\epsilon)\,\mathcal{L}, \qquad \phi_\epsilon(\theta) = \epsilon + (1-\epsilon)\,\phi(\theta). \tag{4.3}$$

If p has mean zero and covariance matrix Γ, then p_ϵ has mean zero and covariance matrix

$$\Gamma_\epsilon = (1-\epsilon)\,\Gamma, \qquad \det\Gamma_\epsilon = (1-\epsilon)^d \det\Gamma. \tag{4.4}$$

If p is transient, and G, G_ϵ denote the Green's function for p, p_ϵ, respectively, then similarly to the last proposition we can see that

$$G_\epsilon(x) = \frac{1}{1-\epsilon}\,G(x). \tag{4.5}$$

In a number of our proofs in this book we will assume that the walk is aperiodic; results for periodic walks can then be derived using these relations.

If $n \geq 1$, let $f_n(x,y)$ denote the probability that a random walk starting at x first visits y at time n (not counting time $n = 0$), i.e.,

$$f_n(x,y) = \mathbb{P}^x\{S_n = y; S_1 \neq y, \ldots, S_{n-1} \neq y\} = \mathbb{P}^x\{\tau_y = n\},$$

where

$$\tau_y = \min\{j \geq 1 : S_j = y\}, \qquad \overline{\tau}_y = \min\{j \geq 0 : S_j = y\}.$$

Let $f_n(x) = f_n(0,x)$ and note that

$$\mathbb{P}^x\{\tau_y < \infty\} = \sum_{n=1}^{\infty} f_n(x,y) = \sum_{n=1}^{\infty} f_n(y-x) \leq 1.$$

Define the *first visit generating function* by

$$F(x,y;\xi) = F(y-x;\xi) = \sum_{n=1}^{\infty} \xi^n f_n(y-x).$$

If $\xi \in (0,1)$, then

$$F(x,y;\xi) = \mathbb{P}^x\{\tau_y < T_\xi\},$$

where T_ξ denotes an independent geometric random variable satisifying $\mathbb{P}\{T_\xi > n\} = \xi^n$.

Proposition 4.2.2 *If $n \geq 1$,*

$$p_n(y) = \sum_{j=1}^{n} f_j(y)\, p_{n-j}(0).$$

If $\xi \in \mathbb{C}$,

$$G(y;\xi) = \delta(y) + F(y;\xi)\, G(0;\xi), \tag{4.6}$$

where δ denotes the delta function. In particular, if $|F(0,\xi)| < 1$,

$$G(0;\xi) = \frac{1}{1 - F(0;\xi)}. \tag{4.7}$$

Proof The first equality follows from

$$\mathbb{P}^x\{S_n = y\} = \sum_{j=1}^{n} \mathbb{P}\{\tau_j = y; S_n - S_j = 0\} = \sum_{j=1}^{n} \mathbb{P}\{\tau_j = y\}\, p_{n-j}(0).$$

The second follows from the equality

$$\sum_{n=1}^{\infty} p_n(x)\, \xi^n = \left[\sum_{n=1}^{\infty} f_n(x)\, \xi^n\right] \left[\sum_{m=0}^{\infty} p_m(0)\, \xi^m\right],$$

which follows from the first equality. For $\xi \in (0,1]$, there is a probabilistic interpretation of (4.6). If $y \neq 0$, the expected number of visits to y (before time T_ξ) is the product of the probability of reaching y and the expected number of visits to y given that y is reached. If $y = 0$, we have to add an extra 1 to account for $p_0(y)$. □

Remark. If $\xi \in (0,1)$, the identity (4.7) can be considered as a generalization of (4.1). Note that

$$F(0;\xi) = \sum_{j=1}^{\infty} \mathbb{P}\{\tau_0 = j; T_\xi > j\} = \mathbb{P}\{\tau_0 < T_\xi\}$$

represents the probability that a random walk killed at rate $1-\xi$ returns to the origin before being killed. Hence, the probability that the walker does not return to the origin before being killed is

$$1 - F(0;\xi) = G(0;\xi)^{-1}. \tag{4.8}$$

The right-hand side is the reciprocal of the expected number of visits before killing. If p is transient, we can plug $\xi = 1$ into this expression and get (4.1).

Proposition 4.2.3 *Suppose $p \in \mathcal{P}_d \cup \mathcal{P}_d^*$ with characteristic function θ. Then if $x \in \mathbb{Z}^d, |\xi| < 1$,*

$$G(x;\xi) = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \frac{1}{1-\xi\phi(\theta)}\, e^{-ix\cdot\theta}\, d\theta.$$

If $d \geq 3$, this holds for $\xi = 1$, i.e.,

$$G(x) = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \frac{1}{1-\phi(\theta)}\, e^{-ix\cdot\theta}\, d\theta.$$

Proof All of the integrals in this proof will be over $[-\pi,\pi]^d$. The formal calculation, using Corollary 2.2.3, is

$$\begin{aligned} G(x;\xi) = \sum_{n=0}^{\infty} \xi^n\, p_n(x) &= \sum_{n=0}^{\infty} \xi^n \frac{1}{(2\pi)^d} \int \phi(\theta)^n\, e^{-ix\cdot\theta}\, d\theta \\ &= \frac{1}{(2\pi)^d} \int \left[\sum_{n=0}^{\infty} (\xi\,\phi(\theta))^n\right] e^{-ix\cdot\theta}\, d\theta \\ &= \frac{1}{(2\pi)^d} \int \frac{1}{1-\xi\,\phi(\theta)}\, e^{-ix\cdot\theta}\, d\theta. \end{aligned}$$

The interchange of the sum and the integral in the second equality is justified by the dominated convergence theorem as we now describe. For each N,

$$\left|\sum_{n=0}^{N} \xi^n \phi(\theta)^n e^{-ix\cdot\theta}\right| \leq \frac{1}{1-|\xi|\,|\phi(\theta)|}.$$

If $|\xi| < 1$, then the right-hand side is bounded by $1/[1-|\xi|]$. If $p \in \mathcal{P}_d^*$ and $\xi = 1$, then (2.13) shows that the right-hand side is bounded by $c\,|\theta|^{-2}$ for some c. If $d \geq 3$, $|\theta|^{-2}$ is integrable on $[-\pi,\pi]^d$. If $p \in \mathcal{P}_d$ is bipartite, we can use (4.3) and (4.5). $\square$

Some results are easier to prove for geometrically killed random walks than for walks restricted to a fixed number of steps. This is because stopping time arguments work more nicely for such walks. Suppose that S_n is a random walk, τ is a stopping time for the random walk, and T is an independent geometric random variable. Then on the event $\{T > \tau\}$ the distribution of $T-\tau$ given $S_n, n = 0,\ldots,\tau$ is the same as that of T. This "loss of memory" property for geometric and exponential random variables can be very useful. The next proposition gives an example of such a result, The result for fixed length random walks can be deduced from the geometrically killed walk result by using *Tauberian theorems.* Tauberian theorems are one of the major tools for deriving facts about a sequence from its generating functions. We will only use some simple Tauberian theorems; see Section 11.5.

Proposition 4.2.4 *Suppose $p \in \mathcal{P}_d \cup \mathcal{P}'_d$, $d = 1, 2$. Let*

$$q(n) = \mathbb{P}\{S_j \neq 0 : j = 1, \ldots, n\}.$$

Then as $n \to \infty$,

$$q(n) \sim \begin{cases} r\,\pi^{-1}\,n^{-1/2}, & d = 1 \\ r\,(\log n)^{-1}, & d = 2. \end{cases}$$

where $r = (2\pi)^{d/2}\,\sqrt{\det \Gamma}$.

Proof We will assume $p \in \mathcal{P}'_d$; it is not difficult to extend this to bipartite $p \in \mathcal{P}_d$. We will establish the corresponding facts about the generating functions for $q(n)$: as $\xi \to 1-$,

$$\sum_{n=0}^{\infty} \xi^n\,q(n) \sim \frac{r}{\mathbf{\Gamma}(1/2)}\,\frac{1}{\sqrt{1-\xi}}, \quad d = 1, \tag{4.9}$$

$$\sum_{n=0}^{\infty} \xi^n q(n) \sim \frac{r}{1-\xi}\left[\log\left(\frac{1}{1-\xi}\right)\right]^{-1}, \quad d = 2. \tag{4.10}$$

Since the sequence $q(n)$ is monotone in n, Propositions 11.5.2 and 11.5.3 imply the proposition (recall that $\mathbf{\Gamma}(1/2) = \sqrt{\pi}$).

Let T be a geometric random variable with killing rate $1-\xi$. Then (4.8) tells us that

$$\mathbb{P}\{S_j \neq 0 : j = 1, \ldots, T-1\} = G(0;\xi)^{-1}.$$

Also,

$$\mathbb{P}\{S_j \neq 0 : j = 1, \ldots, T-1\} = \sum_{n=0}^{\infty} \mathbb{P}\{T = n+1\}\,q(n) = (1-\xi)\sum_{n=0}^{\infty} \xi^n\,q(n).$$

Using (2.28) and Lemma 11.5.1, we can see that as $\xi \to 1-$,

$$G(0;\xi) = \sum_{n=0}^{\infty} \xi^n\,p_n(0) = \sum_{n=0}^{\infty} \xi^n\,\left[\frac{1}{r\,n^{d/2}} + o\left(\frac{1}{n^{d/2}}\right)\right] \sim \frac{1}{r}\,F\left(\frac{1}{1-\xi}\right),$$

where

$$F(s) = \begin{cases} \mathbf{\Gamma}(1/2)\,\sqrt{s}, & d = 1 \\ \log s, & d = 2 \end{cases}$$

This gives (4.9) and (4.10). $\square$

Corollary 4.2.5 *Suppose S_n is a random walk with increment distribution $p \in \mathcal{P}'_d$ and*

$$\tau = \tau_0 = \min\{j \geq 1 : S_j = 0\}.$$

Then $\mathbb{E}[\tau] = \infty$.

Proof If $d \geq 3$, then transience implies that $\mathbb{P}\{\tau = \infty\} > 0$. For $d = 1, 2$, the results comes from the previous proposition that tells us

$$\mathbb{P}\{\tau > n\} \geq \begin{cases} c\, n^{-1/2}, & d = 1, \\ c\,(\log n)^{-1}, & d = 2. \end{cases}$$

$\square$

4.3 Green's function, transient case

In this section, we will study the Green's function for $p \in \mathcal{P}_d, d \geq 3$. The Green's function $G(x,y) = G(y,x) = G(y-x)$ is given by

$$G(x) = \sum_{n=0}^{\infty} p_n(x) = \mathbb{E}\left[\sum_{n=0}^{\infty} 1\{S_n = x\}\right] = \mathbb{E}^x\left[\sum_{n=0}^{\infty} 1\{S_n = 0\}\right].$$

Note that

$$G(x) = 1\{x = 0\} + \sum_y p(x,y)\, \mathbb{E}^y\left[\sum_{n=0}^{\infty} 1\{S_n = 0\}\right] = \delta(x) + \sum_y p(x,y)\, G(y),$$

In other words,

$$\mathcal{L}G(x) = -\delta(x) = \begin{cases} -1, & x = 0, \\ 0, & x \neq 0. \end{cases}$$

Recall from (4.2) that

$$G(x) = \mathbb{P}\{\overline{\tau}_x < \infty\}\, G(0).$$

Comment. In the calculations above as well as throughout this section, we use the symmetry of the Green's function, $G(x,y) = G(y,x)$. For nonsymmetric random walks, one must be careful to distinguish between $G(x,y)$ and $G(y,x)$.

The next theorem gives the asymptotics of the Green's function as $|x| \to \infty$. Recall that $\mathcal{J}^*(x)^2 = d\,\mathcal{J}(x)^2 = x \cdot \Gamma^{-1} x$. Since Γ is nonsingular, $\mathcal{J}^*(x) \asymp \mathcal{J}(x) \asymp |x|$.

Theorem 4.3.1 *Suppose $p \in \mathcal{P}_d$ with $d \geq 3$. As $|x| \to \infty$,*

$$G(x) = \frac{C_d^*}{\mathcal{J}^*(x)^{d-2}} + O\left(\frac{1}{|x|^d}\right) = \frac{C_d}{\mathcal{J}(x)^{d-2}} + O\left(\frac{1}{|x|^d}\right),$$

where

$$C_d^* = \frac{C_d}{d^{(d/2)-1}} = \frac{\mathbf{\Gamma}(\frac{d-2}{2})}{2\,\pi^{d/2}\sqrt{\det \Gamma}}.$$

Here Γ denotes the covariance matrix and $\mathbf{\Gamma}$ denotes the Gamma function. In particular, for simple random walk,

$$G(x) = \frac{d^{d-1}\,\mathbf{\Gamma}(\frac{d-2}{2})}{2\,\pi^{d/2}}\,\frac{1}{|x|^{2-d}} + O\left(\frac{1}{|x|^d}\right).$$

The last statement of the theorem follows from the first statement using $\Gamma = d^{-1} I, \mathcal{J}(x) = |x|$ for simple random walk. It suffices to prove the first statememt for aperiodic p; the proof for bipartite p follows using (4.4) and (4.5). The proof of the theorem will consist of two estimates:

$$G(x) = \sum_{n=0}^{\infty} p_n(x) = O\left(\frac{1}{|x|^d}\right) + \sum_{n=1}^{\infty} \overline{p}_n(x), \tag{4.11}$$

and

$$\sum_{n=1}^{\infty} \overline{p}_n(x) = \frac{C_d^*}{\mathcal{J}^*(x)^{d-2}} + o\left(\frac{1}{|x|^d}\right).$$

The second estimate uses a simple lemma.

Lemma 4.3.2 *Let $b > 1$. Then as $r \to \infty$,*

$$\sum_{n=1}^{\infty} n^{-b}\, e^{-r/n} = \frac{\boldsymbol{\Gamma}(b-1)}{r^{b-1}} + O\left(\frac{1}{r^{b+1}}\right)$$

Proof The sum is a Riemann sum approximation of the integral

$$I_r := \int_0^\infty t^{-b}\, e^{-r/t}\, dt = \frac{1}{r^{b-1}} \int_0^\infty y^{b-2}\, e^{-y}\, dy = \frac{\boldsymbol{\Gamma}(b-1)}{r^{b-1}}. \tag{4.12}$$

If $f : (0,\infty) \to \mathbb{R}$ is a C^2 function and n is a positive integer, then Lemma 11.1.1 gives

$$\left| f(n) - \int_{n-(1/2)}^{n+(1/2)} f(s)\, ds \right| \leq \frac{1}{24} \sup\{|f''(t)| : |t-n| \leq 1/2\}.$$

Choosing $f(t) = t^{-b}\, e^{-r/t}$, we get

$$\left| n^{-b}\, e^{-n/r} - \int_{n-(1/2)}^{n+(1/2)} t^{-b}\, e^{-r/t}\, dt \right| \leq c\, \frac{1}{n^{b+2}} \left[1 + \frac{r^2}{n^2}\right] e^{-r/n}, \qquad n \geq \sqrt{r}.$$

(The restriction $n \geq \sqrt{r}$ is used to guarantee that $e^{-r/(n+(1/2))} \leq c\, e^{-r/n}$.) Therefore,

$$\begin{aligned} \sum_{n \geq \sqrt{r}} \left| n^{-b}\, e^{-n/r} - \int_{n-(1/2)}^{n+(1/2)} t^{-b}\, e^{-r/t}\, dt \right| &\leq c \sum_{n \geq \sqrt{r}} \frac{1}{n^{b+2}} \left[1 + \frac{r^2}{n^2}\right] e^{-r/n} \\ &\leq c \int_0^\infty t^{-(b+2)} \left(1 + \frac{r^2}{t^2}\right) e^{-r/t}\, dt \\ &\leq c\, r^{-(b+1)} \end{aligned}$$

The last step uses (4.12). It is easy to check that the sum over $n < \sqrt{r}$ and the integral over $t < \sqrt{r}$ decay faster than any power of r. $\square$

Proof of Theorem 4.3.1. Using Lemma 4.3.2 with $b = d/2, r = \mathcal{J}^*(x)^2/2$, we have

$$\sum_{n=1}^{\infty} \overline{p}_n(x) = \sum_{n=1}^{\infty} \frac{1}{(2\pi n)^{d/2}\sqrt{\det \Gamma}}\, e^{-\mathcal{J}^*(x)^2/(2n)} = \frac{\boldsymbol{\Gamma}(\frac{d-2}{2})}{2\,\pi^{d/2}\sqrt{\det\Gamma}}\, \frac{1}{\mathcal{J}^*(x)^{(d-2)}} + O\left(\frac{1}{|x|^{d+2}}\right).$$

Hence we only need to prove (4.11). A simple estimate shows that

$$\sum_{n<|x|} \overline{p}_n(x)$$

as a function of x decays faster than any power of x. Similarly, using Proposition 2.1.2,

$$\sum_{n<|x|} p_n(x) = o(|x|^{-d}). \tag{4.13}$$

Using (2.5), we can see that

$$\sum_{n>|x|^2} |p_n(x) - \overline{p}_n(x)| \le c \sum_{n>|x|^2} n^{-(d+2)/2} = O(|x|^{-d}).$$

Let $k = d+3$. For $|x| \le n \le |x|^2$, (2.3) implies that there is an r such that

$$|p_n(x) - \overline{p}_n(x)| \le c \left[\left(\frac{|x|}{\sqrt{n}}\right)^k e^{-r|x|^2/n} \frac{1}{n^{(d+2)/2}} + \frac{1}{n^{(d+k-1)/2}} \right]. \tag{4.14}$$

Note that

$$\sum_{n\ge|x|} n^{-(d+k-1)/2} = O(|x|^{-(d+k-3)/2}) = O(|x|^{-d}),$$

and

$$\sum_{n\ge|x|} \left(\frac{|x|}{\sqrt{n}}\right)^k e^{-r\,|x|^2/n} \frac{1}{n^{(d+2)/2}} \le c \int_0^\infty \left(\frac{|x|}{\sqrt{t}}\right)^k \frac{e^{-r|x|^2/t}}{(\sqrt{t})^{d+2}}\, dt \le c\,|x|^{-d}.$$

Remark. The error term in this theorem is very small. In order to prove that it is this small we used the sharp estimate (4.14) which uses the fact that the third moments of the increment distribution are zero. If $p \in \mathcal{P}_d'$ with bounded increments but with nonzero third moments, a similar asymptotic expansion exists for the Green's function except that the error term is $O(|x|^{-(d-1)})$. We have used bounded increments (or at least the existence of sufficiently large moments) in an important way in (4.13). Asymptotics can be proved under weaker moment assumptions; however, mean zero, finite variance is not sufficient to conclude that the Green's function is asymptotic to $c\,\mathcal{J}^*(x)^{2-d}$ for $d \ge 4$. See Exercise 4.4.

Comment. Often one does not use the full force of these asymptotics. An important thing to remember is that $G(x) \asymp |x|^{2-d}$. There are a number of ways to remember the exponent $2-d$. For example, the central limit theorem implies that the random walk should visit on the order of R^2 points in the ball of radius R. Since there are R^d points in this ball, the probability that a particular point is visited is of order R^{2-d}. In the case of standard d-dimensional Brownian motion, the Green's function is proportional to $|x|^{2-d}$. This is the unique (up to multiplicative constant) harmonic, radially symmetric function on $\mathbb{R}^d \setminus \{0\}$ that goes to zero as $|x| \to \infty$ (see Exercise 4.3).

Corollary 4.3.3 *If $p \in \mathcal{P}_d$, then*

$$\nabla_j G(x) = \nabla_j \frac{C_d}{\mathcal{J}^*(x)^{d-2}} + O(|x|^{-d}).$$

In particular, $\nabla_j G(x) = O(|x|^{-d+1})$. *Also,*

$$\nabla_j^2 G(x) = O(|x|^{-d}).$$

Remark. We could also prove this corollary with improved error terms by using the difference estimates for the LCLT such as Theorem 2.3.6, but we will not need the sharper results in this book. If $p \in \mathcal{P}_d'$ with bounded increments but nonzero third moments, we could also prove difference estimates for the Green's function using Theorem 2.3.6. The starting point is to write

$$\nabla_y G(x) = \sum_{n=0}^{\infty} \nabla_y \overline{p}_n(x) + \sum_{n=0}^{\infty} [\nabla_y p_n(x) - \nabla_y \overline{p}_n(x)].$$

4.3.1 Asymptotics under weaker assumptions

In this section we establish the asymptotics for G for some $p \in \mathcal{P}_d', d \geq 3$. We will follow the basic outline of the proof of Theorem 4.3.1. Let $\overline{G}(x) = C_d^*/\mathcal{J}^*(x)^{d-2}$ denote the dominant term in the asymptotics. From that proof we see that

$$G(x) = \overline{G}(x) + o(|x|^{-d}) + \sum_{n=0}^{\infty} [p_n(x) - \overline{p}_n(x)].$$

In the discussion below, we let $\alpha \in \{0, 1, 2\}$. If $\mathbb{E}[|X_1|^4] < \infty$ and the third moments vanish, we can set $\alpha = 2$. If this is not the case, but $\mathbb{E}[|X_1|^3] < \infty$, we can set $\alpha = 1$. Otherwise, we set $\alpha = 0$. By Theorems 2.3.5 and 2.3.9 we can see that there exists a sequence $\delta_n \to \infty$ such that

$$\sum_{n \geq |x|^2} |p_n(x) - \overline{p}_n(x)| \leq c \sum_{n \geq |x|^2} \frac{\delta_n + \alpha}{|x|^{(d+\alpha)/2}} = \begin{cases} o(|x|^{2-d}), & \alpha = 0 \\ O(|x|^{2-d-\alpha}), & \alpha = 1, 2. \end{cases}$$

This is the order of magnitude that we will try to show for the error term, so this estimate is finished. The sum that is more difficult to hand and which in some cases requires additional moment conditions is

$$\sum_{n < |x|^2} [p_n(x) - \overline{p}_n(x)].$$

Theorem 4.3.4 *Suppose* $p \in \mathcal{P}_3'$. *Then*

$$G(x) = \overline{G}(x) + o\left(\frac{1}{|x|}\right).$$

If $\mathbb{E}[|X_1|^3] < \infty$ *we can write*

$$G(x) = \overline{G}(x) + O\left(\frac{\log |x|}{|x|^2}\right).$$

If $\mathbb{E}[|X_1|^4] < \infty$ *and the third moments vanish, then*

$$G(x) = \overline{G}(x) + O\left(\frac{1}{|x|^3}\right).$$

Proof By Theorem 2.3.10, there exists $\delta_n \to 0$ such that

$$\sum_{n<|x|^2} |p_n(x) - \overline{p}_n(x)| \le c \sum_{n<|x|^2} \frac{\delta_n + \alpha}{|x|^2\, n^{(1+\alpha)/2}}.$$

□

The next theorem shows that if we assume enough moments of the distribution, then we get the asymptotics as in Theorem 4.3.1. Note that as $d \to \infty$, the number of moments assumed grows.

Theorem 4.3.5 *Suppose $p \in \mathcal{P}'_d, d \ge 3$.*

- *If $\mathbb{E}|X_1|^{d+1}] < \infty$, then*

$$G(x) = \overline{G}(x) + O(|x|^{1-d}).$$

- *If $\mathbb{E}|X_1|^{d+3}] < \infty$ and the third moments vanish,*

$$G(x) = \overline{G}(x) + O(|x|^{-d}).$$

Proof Let $\alpha = 1$ under the weaker assumption and $\alpha = 2$ under the stronger assumption, set $k = d + 2\alpha - 1$ so that $\mathbb{E}[|X_1|^k] < \infty$. As mentioned above, it suffices to show that

$$\sum_{n<|x|^2} [p_n(x) - \overline{p}_n(x)] = O(|x|^{2-d-\alpha}).$$

Let $\epsilon = 2(1+\alpha)/(1+2\alpha)$. As before,

$$\sum_{n<|x|^\epsilon} \overline{p}_n(x)$$

decays faster than any power of n. Using (2.36), we have

$$\sum_{n<|x|^\epsilon} p_n(x) \le \frac{c}{|x|^k} \sum_{n<|x|^\epsilon} n^{\frac{k-d}{2}} = O(|x|^{2-d-\alpha}).$$

(The value of ϵ was chosen as the largest value for which this holds.) For the range $|x|^\epsilon \le n < |x|^2$, we use the estimate from Theorem 2.3.8:

$$|p_n(x) - \overline{p}_n(x)| \le \frac{c}{n^{(d+\alpha)/2}} \left[|x/\sqrt{n}|^{k-1}\, e^{-r|x|^2/n} + n^{-(k-2-\alpha)/2} \right].$$

As before,

$$\sum_{n \ge |x|^\epsilon} \frac{|x/\sqrt{n}|^{k-1}}{n^{(d+\alpha)/2}}\, e^{-r|x|^2/n} \le c \int_0^\infty \frac{|x|^{k-1}}{(\sqrt{t})^{k+d+\alpha-1}}\, e^{-r|x|^2/t}\, dt = O(|x|^{2-d-\alpha}).$$

Also,

$$\sum_{n \ge |x|^\epsilon} \frac{1}{n^{\frac{d}{2}+\frac{k}{2}-1}} = O(|x|^{-\epsilon(d+\alpha-\frac{5}{2})}) \le O(|x|^{2-d-\alpha}),$$

provided that

$$\left(d + \alpha - \frac{5}{2}\right) \frac{2(1+\alpha)}{1+2\alpha} \ge d - 2 + \alpha,$$

which can be readily checked for $\alpha = 1, 2$ if $d \geq 3$.

$\square$

4.4 Potential kernel

4.4.1 Two dimensions

If $p \in \mathcal{P}_2'$, the *potential kernel* is the function

$$a(x) = \sum_{n=0}^{\infty}[p_n(0) - p_n(x)] = \lim_{N\to\infty}\left[\sum_{n=0}^{N} p_n(0) - \sum_{n=0}^{N} p_n(x)\right]. \tag{4.15}$$

Exercise 2.2 shows that $|p_n(0) - p_n(x)| \leq c\,|x|\,n^{-3/2}$, so the first sum converges absolutely. However, since $p_n(0) \asymp n^{-1}$, we cannot write

$$a(x) = \left[\sum_{n=0}^{\infty} p_n(0)\right] - \left[\sum_{n=0}^{\infty} p_n(x)\right]. \tag{4.16}$$

If $p \in \mathcal{P}_2$ is bipartite, the potential kernel for $x \in (\mathbb{Z}^2)_e$ is defined in the same way. If $x \in (\mathbb{Z}^2)_o$ we can define $a(x)$ by the second expression in (4.15). Many authors use the term Green's function for the potential kernel.

Comment. If $p \in \mathcal{P}_d^*$ is transient, then (4.16) is valid, and $a(x) = G(0) - G(x)$, where G is the Green's function for p. Since $|p_n(0) - p_n(x)| \leq c\,|x|\,n^{-3/2}$ for all $p \in \mathcal{P}_2^*$, the same argument shows that a exists for such p.

Proposition 4.4.1 *If $p \in \mathcal{P}_2'$, then $2\,a(x)$ is the expected number of visits to x by a random walk starting at x before its first visit to the origin.*

Proof We delay this until the next section; see (4.32). $\square$

Remark. Using Proposition 4.4.1, we can see that if p_ϵ is defined as in (4.3), and a_ϵ denotes the potential kernel for p_ϵ then

$$a_\epsilon(x) = \frac{1}{1-\epsilon}\,a(x). \tag{4.17}$$

This will allow us to assume aperiodicity in many of our proofs.

If we write

$$p_{2k+1}(x) = \sum_y p_{2k}(x-y)\,p(y),$$

we see that

$$a(x) = \sum_y a(x-y)\,p(y).$$

More generally, if $p \in \mathcal{P}_2$, we can see that $a(x)$ satisfies: $a(0) = 0$ and

$$\mathcal{L}a(x) = \delta_0(x) = \begin{cases} 1, & x = 0 \\ 0, & x \neq 0. \end{cases}$$

Proposition 4.4.2 *If $p \in \mathcal{P}_2^* \cup \mathcal{P}_2$, then*

$$a(x) = \frac{1}{(2\pi)^2} \int_{[-\pi,\pi]^2} \frac{1 - e^{-ix\cdot\theta}}{1 - \phi(\theta)} \, d\theta.$$

Proof By the remark above, it suffices to consider $p \in \mathcal{P}_2^*$. The formal calculation is

$$\begin{aligned} a(x) = \sum_{n=0}^{\infty} [p_n(0) - p_n(x)] &= \sum_{n=0}^{\infty} \frac{1}{(2\pi)^2} \int \phi(\theta)^n \, [1 - e^{-ix\cdot\theta}] \, d\theta \\ &= \frac{1}{(2\pi)^2} \int \left[\sum_{n=0}^{\infty} \phi(\theta)^n \right] [1 - e^{-ix\cdot\theta}] \, d\theta \\ &= \frac{1}{(2\pi)^2} \int \frac{1 - e^{-ix\cdot\theta}}{1 - \phi(\theta)} \, d\theta. \end{aligned}$$

All of the integrals are over $[-\pi, \pi]^2$. To justify the interchange of the sum and the integral i we use (2.13) to obtain the estimate

$$\left| \sum_{n=0}^{N} \phi(\theta)^n \, [1 - e^{-ix\cdot\theta}] \right| \leq \frac{|1 - e^{-ix\cdot\theta}|}{1 - |\phi(\theta)|} \leq \frac{c\,|x\theta|}{|\theta|^2} \leq \frac{c\,|x|}{|\theta|}.$$

Since $|\theta|^{-1}$ is an integrable function in $[-\pi, \pi]^2$, the dominated convergence theorem may be applied. □

Theorem 4.4.3 *If $p \in \mathcal{P}_2$, there exists a constant $C = C(p)$ such that as $|x| \to \infty$,*

$$a(x) = \frac{1}{\pi \sqrt{\det \Gamma}} \log[\mathcal{J}^*(x)] + C + O(|x|^{-2}).$$

For simple random walk,

$$a(x) = \frac{2}{\pi} \log |x| + \frac{2\gamma + \log 8}{\pi} + O(|x|^{-2}),$$

where γ is Euler's constant.

Proof We will assume that p is aperiodic; the bipartite case is done similarly. We write

$$a(x) = \sum_{n \leq \mathcal{J}^*(x)^2} p_n(0) - \sum_{n \leq \mathcal{J}^*(x)^2} p_n(x) + \sum_{n > \mathcal{J}^*(x)^2} [p_n(0) - p_n(x)].$$

We know from (2.23) that

$$p_n(0) = \frac{1}{2\pi \sqrt{\det \Gamma}} \frac{1}{n} + O\left(\frac{1}{n^2}\right).$$

We therefore get

$$\sum_{n \leq \mathcal{J}^*(x)^2} p_n(0) = 1 + O(|x|^{-2}) + \sum_{1 \leq n \leq \mathcal{J}^*(x)^2} \frac{1}{2\pi \sqrt{\det \Gamma}} \frac{1}{n} + \sum_{n=1}^{\infty} \left[p_n(0) - \frac{1}{2\pi \sqrt{\det \Gamma}} \frac{1}{n} \right],$$

where the last sum is absolutely convergent. Also,

$$\sum_{1\le n\le \mathcal{J}^*(x)^2} \frac{1}{n} = 2\,\log[\mathcal{J}^*(x)] + \gamma + O(|x|^{-2}),$$

where γ is Euler's constant (see Lemma 11.1.2). Hence,

$$\sum_{n\le \mathcal{J}^*(x)^2} p_n(0) = \frac{1}{\pi\,\sqrt{\det\Gamma}}\,\log[\mathcal{J}^*(x)] + c' + O(|x|^{-2})$$

for some constant c'.

Proposition 2.1.2 shows that

$$\sum_{n\le|x|} p_n(x)$$

decays faster than any power of $|x|$. Theorem 2.3.8 implies that there exists c, r such that for $n \le |x|^2$,

$$\left|p_n(x) - \frac{1}{2\pi\,n\,\sqrt{\det\Gamma}}\,e^{-\mathcal{J}^*(x)^2/2n}\right| \le c\,\left[\frac{|x|^5\,e^{-r|x|^2/n}}{n^2} + \frac{1}{n^3}\right]. \tag{4.18}$$

Therefore,

$$\sum_{n\le \mathcal{J}^*(x)^2} \left|p_n(x) - \frac{1}{2\pi\,n\,\sqrt{\det\Gamma}}\,e^{-\mathcal{J}^*(x)^2/2n}\right|$$

$$\le o(|x|^{-2}) + c \sum_{|x|<n\le \mathcal{J}^*(x)^2} \left[\frac{|x/\sqrt{n}|^5\,e^{-r|x|^2/n}}{n^2} + \frac{1}{n^3}\right] \le c\,|x|^{-2}.$$

The last estimate is done as in the final step of the proof of Theorem 4.3.1. Similarly, to the proof of Lemma 4.3.2 we can see that

$$\begin{aligned}\sum_{|x|\le n\le \mathcal{J}^*(x)^2} \frac{1}{n}\,e^{-\mathcal{J}^*(x)^2/2n} &= \int_0^{\mathcal{J}^*(x)^2} \frac{1}{t}\,e^{-\mathcal{J}^*(x)^2/2t}\,dt + O(|x|^{-2}) \\ &= \int_1^{\infty} \frac{1}{y}\,e^{-y/2}\,dy + O(|x|^{-2}).\end{aligned}$$

The integral contributes a constant. At this point we have shown that

$$\sum_{n\le \mathcal{J}^*(x)^2} [p_n(0) - p_n(x)] = \frac{1}{\pi\,\sqrt{\det\Gamma}}\,\log[\mathcal{J}^*(x)] + C' + O(|x|^{-2})$$

for some constant C'. For $n > \mathcal{J}^*(x)^2$, we use Theorem 2.3.8 and Lemma 4.3.2 again to conclude that

$$\sum_{n>\mathcal{J}^*(x)^2} [p_n(0) - p_n(x)] = c\,\int_0^1 \frac{1}{y}\,\left[1 - e^{-y/2}\right]\,dy + O(|x|^{-2}).$$

For simple random walk in two dimensions, it follows that

$$a(x) = \frac{2}{\pi} \log |x| + k + O(|x|^{-2}),$$

for some constant k. To determine k, we use

$$\phi(\theta^1, \theta^2) = 1 - \frac{1}{2}\,[\cos\theta^1 + \cos\theta^2].$$

Plugging this into Proposition 4.4.2 and doing the integral (details omitted), we get an exact expression for $x_n = (n, n)$ for integer $n > 0$

$$a(x_n) = \frac{4}{\pi} \left[1 + \frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{2n-1}\right].$$

However, we also know that

$$a(x_n) = \frac{2}{\pi} \log n + \frac{2}{\pi} \log \sqrt{2} + k + O(n^{-2}).$$

Therefore,

$$k = \lim_{n\to\infty} \left[-\frac{1}{\pi} \log 2 - \frac{2}{\pi} \log n + \frac{4}{\pi} \sum_{j=1}^{n} \frac{1}{2j-1} \right].$$

Using Lemma 11.1.2 we can see that as $n \to \infty$,

$$\sum_{j=1}^{n} \frac{1}{2j-1} = \sum_{j=1}^{2n} \frac{1}{j} - \sum_{j=1}^{n} \frac{1}{2j} = \frac{1}{2} \log n + \log 2 + \frac{1}{2}\gamma + o(1).$$

Therefore,

$$k = \frac{3}{\pi} \log 2 + \frac{2}{\pi}\gamma.$$

□

Comment. Roughly speaking, $a(x)$ is the difference between the expected number of visits to 0 and the expected number of visits to x by some large time N. Let us consider $N >> |x|^2$. By time $|x|^2$, the random walker has visited the origin about

$$\sum_{n<|x|^2} p_n(0) \sim \sum_{n<|x|^2} \frac{c}{n} \sim 2\,c \log |x|,$$

times where $c = (2\pi\sqrt{\Gamma})^{-1}$. It has visited x about $O(1)$ times. From time $|x|^2$ onward, $p_n(x)$ and $p_n(0)$ are roughly the same and the sum of the difference from then on is $O(1)$. This shows why we expect

$$a(x) = c\, \log |x| + O(1).$$

Note that $\log |x| = \log \mathcal{J}^*(x) + O(1)$.

Comment. Although we have included the exact value of the constant for simple random walk, we will never need to use this value.

Corollary 4.4.4 *If $p \in \mathcal{P}_2$,*

$$\nabla_j a(x) = \nabla_j \left[\frac{1}{\pi\sqrt{\det\Gamma}}\,\log[\mathcal{J}^*(x)]\right] + O(|x|^{-2}).$$

In particular, $\nabla_j a(x) = O(|x|^{-1})$. Also,

$$\nabla_j^2 a(x) = O(|x|^{-2}).$$

4.4.2 Asymptotics under weaker assumptions

We can prove asymptotics for the potential kernel under weaker assumptions. Let

$$\overline{a}(x) = [\pi\sqrt{\det\Gamma}]^{-1}\,\log[\mathcal{J}^*(x)]$$

denote the leading term in the asymptotics.

Theorem 4.4.5 *Suppose $p \in \mathcal{P}_2'$. Then*

$$a(x) = \overline{a}(x) + o(\log|x|).$$

If $\mathbb{E}[|X_1|^3] < \infty$, then there exists $C < \infty$ such that

$$a(x) = \overline{a}(x) + C + O(|x|^{-1}).$$

If $\mathbb{E}[|X_1|^6] < \infty$ and the third moments vanish, then

$$a(x) = \overline{a}(x) + C + O(|x|^{-2}).$$

Proof Let $\alpha = 0, 1, 2$ under the three possible assumptions, respectively. We start with $\alpha = 1, 2$ for which we can write

$$\sum_{n=0}^{\infty}[p_n(0) - p_n(x)] = \sum_{n=0}^{\infty}[\overline{p}_n(0) - \overline{p}_n(x)] + \sum_{n=0}^{\infty}[p_n(0) - \overline{p}_n(0)] + \sum_{n=0}^{\infty}[\overline{p}_n(x) - p_n(x)] \tag{4.19}$$

The estimate

$$\sum_{n=0}^{\infty}[\overline{p}_n(0) - \overline{p}_n(x)] = \overline{a}(x) + \tilde{C} + O(|x|^{-2})$$

is done as in Theorem 4.4.3. Since $|p_n(0) - \overline{p}_n(0)| \le c\,n^{-3/2}$, the second sum on the right-hand side of (4.19) converges, and we set

$$C = \tilde{C} + \sum_{n=0}^{\infty}[p_n(0) - \overline{p}_n(0)].$$

We write

$$\left|\sum_{n=0}^{\infty}[\overline{p}_n(x) - p_n(x)]\right| \le \sum_{n<|x|^2}|\overline{p}_n(x) - p_n(x)| + \sum_{n\ge|x|^2}|\overline{p}_n(x) - p_n(x)|.$$

By (2.5),

$$\sum_{n \geq |x|^2} |\overline{p}_n(x) - p_n(x)| \leq c \sum_{n \geq |x|^2} n^{-(2+\alpha)/2} = O(|x|^{-\alpha}).$$

For $\alpha = 1$, Theorem 2.3.10 gives

$$\sum_{n < |x|^2} |\overline{p}_n(x) - p_n(x)| \leq \sum_{n < |x|^2} \frac{c}{|x|^2 \, n^{1/2}} = o(|x|^{-1}).$$

If $\mathbb{E}[|X_1|^4] < \infty$ and the third moments vanish, a similar arugments shows that the sum on the left-hand side is bounded by $O(|x|^{-2} \log |x|)$ which is a little bigger than we want. However, if we also assume that $\mathbb{E}[|X_1|^6] < \infty$, then we get an estimate as in (4.18), and we can show as in Theorem 4.4.3 that this sum is $O(|x|^{-2})$.

If we only assume that $p \in \mathcal{P}_2'$, then we cannot write (4.19) because the second sum on the right-hand side might diverge. Instead, we write

$$\sum_{n=0}^{\infty} [p_n(0) - p_n(x)] =$$

$$\sum_{n \geq |x|^2} [p_n(0) - p_n(x)] + \sum_{n < |x|^2} [\overline{p}_n(0) - \overline{p}_n(x)] + \sum_{n < |x|^2} [p_n(0) - \overline{p}_n(0)] + \sum_{n < |x|^2} [\overline{p}_n(x) - p_n(x)].$$

As before,

$$\sum_{n < |x|^2} [\overline{p}_n(0) - \overline{p}_n(x)] = \overline{a}(x) + O(1).$$

Also, Exercise 2.2, Theorem 2.3.10, and (2.28) imply

$$\sum_{n \geq |x|^2} |p_n(0) - p_n(x)| \leq \sum_{n \geq |x|^2} \frac{c\,|x|}{n^{3/2}} = O(1),$$

$$\sum_{n < |x|^2} |\overline{p}_n(x) - p_n(x)| \leq \sum_{n < |x|^2} \frac{c}{|x|^2} = O(1),$$

$$\sum_{n < |x|^2} |\overline{p}_n(0) - p_n(0)| \leq \sum_{n < |x|^2} o\left(\frac{1}{n}\right) = o(\log |x|).$$

□

4.4.3 One dimension

If $p \in \mathcal{P}_1'$, the potential kernel is defined in the same way

$$a(x) = \lim_{N \to \infty} \left[\sum_{n=0}^{N} p_n(0) - \sum_{n=0}^{N} p_n(x) \right].$$

In this case, the convergence is a little more subtle. We will restrict ourselves to the third moment case, $\mathbb{E}[|X|^3] < \infty$. In this case, the proof of the next proposition shows that the sum converges absolutely.

Proposition 4.4.6 *Suppose $p \in \mathcal{P}_1'$ with $\mathbb{E}[|X|^3] < \infty$. Then there is a c such that for all x,*

$$\big|a(x)\,\sigma^2 - |x|\big| \leq c\,\log|x|.$$

If $\mathbb{E}[|X|^4] < \infty$ and $\mathbb{E}[X^3] = 0$, then there is a C such that ,

$$a(x) = \frac{|x|^2}{\sigma^2} + C + O(|x|^{-1}).$$

Proof Without loss of generality, assume $x > 0$. Let $\alpha = 1$ under the weaker assumption and $\alpha = 2$ under the stronger assumption. Theorem 2.3.6 gives

$$p_n(0) - p_n(x) = \overline{p}_n(0) - \overline{p}_n(x) + x\,O(n^{-(2+\alpha)/2}),$$

which shows that

$$\left|\sum_{n \geq x^2} [p_n(0) - p_n(x)] - [\overline{p}_n(0) - \overline{p}_n(x)]\right| \leq c\,x \sum_{n \geq x^2} n^{-(2+\alpha)/2} \leq c x^{1-\alpha}.$$

If $\alpha = 1$, Theorem 2.3.5 gives

$$\sum_{n < x^2} [p_n(0) - \overline{p}_n(0)] \leq c \sum_{n < x^2} n^{-1} = O(\log x).$$

If $\alpha = 2$, Theorem 2.3.5 gives $|p_n(0) - \overline{p}_n(0)| = O(n^{-3/2})$ and hence

$$\sum_{n < x^2} [p_n(0) - \overline{p}_n(0)] = C + O(|x|^{-1}), \qquad C := \sum_{n=0}^{\infty} [p_n(0) - \overline{p}_n(0)],$$

In both cases, Theorem 2.3.10 gives

$$\sum_{n < x^2} [p_n(x) - \overline{p}_n(x)] \leq \frac{c}{x^2} \sum_{n < x^2} n^{(1-\alpha)/2} \leq c\,x^{1-\alpha}.$$

Therefore,

$$\begin{aligned} a(x) &= e(x) + \sum_{n=0}^{\infty} [\overline{p}_n(0) - \overline{p}_n(x)] \\ &= e(x) + \sum_{n=1}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2 n}} \left[1 - e^{-\frac{x^2}{2\sigma^2 n}}\right], \end{aligned} \tag{4.20}$$

where $e(x) = O(\log x)$ if $\alpha = 1$ and $e(x) = C + O(x^{-1})$ if $\alpha = 2$. Standard estimates, which we omit, show that up to an error of order $O(x^{-1})$, the sum in (4.20) equals

$$\int_0^{\infty} \frac{1}{\sqrt{2\pi\sigma^2 t}} \left[1 - e^{-\frac{x^2}{2\sigma^2 t}}\right] dt = \frac{2x}{\sigma^2\sqrt{2\pi}} \int_0^{\infty} \frac{1}{u^2} \left(1 - e^{-u^2/2}\right) du = \frac{x}{\sigma^2}.$$

□

Theorem 4.4.7 *If $p \in \mathcal{P}_1$, and $x > 0$*

$$a(x) = \frac{x}{\sigma^2} + \mathbb{E}^x\left[a(S_T) - \frac{S_T}{\sigma^2}\right], \tag{4.21}$$

where $T = \min\{n : S_n \le 0\}$. *There exists* $\beta > 0$ *such that for* $x > 0$,

$$a(x) = \frac{x}{\sigma^2} + C + O(e^{-\beta x}), \quad x \to \infty, \tag{4.22}$$

where

$$C = \lim_{y\to\infty} \mathbb{E}^y \left[a(S_T) - \frac{S_T}{\sigma^2} \right].$$

In particular, for simple random walk, $a(x) = |x|$.

Proof Assume $y > x$, let $T_y = \min\{n : S_n \le 0 \text{ or } S_n \ge y\}$, and consider the bounded martingale $S_{n\wedge T_y}$. Then the optional sampling theorem implies that

$$x = \mathbb{E}^x[S_0] = \mathbb{E}^x[S_{T_y}] = \mathbb{E}^x[S_T; T \le T_y] + \mathbb{E}^x[S_{T_y} : T > T_y].$$

If we let $y \to \infty$, we see that

$$\lim_{y\to\infty} \mathbb{E}^x[S_{T_y}; T > T_y] = x - \mathbb{E}^x[S_T].$$

Also, since $\mathbb{E}[S_{T_y} \mid T_y < T] = y + O(1)$, we can see that

$$\lim_{n\to\infty} y\,\mathbb{P}^x\{T_y < T\} = x - \mathbb{E}^x[S_T].$$

We now consider the bounded martingale $M_n = a(S_{n\wedge T_y})$. Then the optional sampling theorem implies that

$$a(x) = \mathbb{E}^x[M_0] = \mathbb{E}^x[M_{T_y}] = \mathbb{E}^x[a(S_T); T \le T_y] + \mathbb{E}^x[a(S_{T_y}) : T > T_y].$$

As $y \to \infty$, $\mathbb{E}^x[a(S_T); T < T_y] \to \mathbb{E}^x[a(S_T)]$. Also, as $y \to \infty$,

$$\mathbb{E}^x[a(S_{T_y}) : T > T_y] \sim \mathbb{P}^x\{T_y < T\} \left[\frac{y}{\sigma^2} + O(1) \right] \sim \frac{x - \mathbb{E}^x[S_T]}{\sigma^2}.$$

This gives (4.21).

We will sketch the proof of (4.22); we leave it as an exercise (Exercise 4.10) to fill in the details. we will show that there exists a β such that if $0 < x < y < \infty$, then

$$\sum_{j=0}^{\infty} |\mathbb{P}^x\{S_T = -j\} - \mathbb{P}^y\{S_T = -j\}| = O(e^{-\beta x}). \tag{4.23}$$

Even though we have written this as an infinite sum, the terms are nonzero only for j less than the range of the walk. Let $\tau_z = \min\{n \ge 0 : S_n \le z\}$. Irreducibility and aperiodicity of the random walk can be used to see that there is an $\epsilon > 0$ such that for all $z > 0$, $\mathbb{P}^{z+1}\{S_{\tau_z} = z\} = \mathbb{P}^1\{\tau_0 = 0\} > \epsilon$. Let

$$f(r) = f_{-j}(r) = \sup_{x,y\ge r} |\mathbb{P}^x\{S_T = -j\} - \mathbb{P}^y\{S_T = -j\}|.$$

Then if R denotes the range of the walk, we can see that

$$f(r+1) \le (1-\epsilon)\, f(r-R).$$

Iteration of this inequality gives $f(kR) \le (1-\epsilon)^{k-1} f(R)$ and this gives (4.23). □

Comment. The potential kernel in one dimension is not as important as the potential kernel or Green's function in higher dimenions. For $d \geq 2$, we use the fact that the potential kernel or Green's function is harmonic on $\mathbb{Z}^d \setminus \{0\}$ and that we have very good estimates for the asymptotics. For $d = 1$, many of these same arguments can be done with the function $f(x) = x$ which is obviously harmonic.

4.5 Fundamental solutions

If $p \in \mathcal{P}$, the Green's function G for $d \geq 3$ or the potential kernel a for $d = 2$ is often called the *fundamental solution of the generator* $\mathcal{L}$ since

$$\mathcal{L}G(x) = -\delta(x), \qquad \mathcal{L}a(x) = \delta(x). \tag{4.24}$$

More generally, we write

$$\mathcal{L}_x G(x,y) = \mathcal{L}_x G(y,x) = -\delta(y-x), \quad \mathcal{L}_x a(x,y) = \mathcal{L}_x a(y,x) = \delta(y-x),$$

where $\mathcal{L}_x$ denotes $\mathcal{L}$ applied to the x variable.

Remark. Symmetry of walks in $\mathcal{P}$ is necessary to derive (4.24). If $p \in \mathcal{P}^*$ is transient, the Green's function G does not satisfy (4.24). Instead it satisfies $\mathcal{L}^R G(x) = -\delta_0(x)$ where $\mathcal{L}^R$ denotes the generator of the "backwards random walk" with increment distribution $p^R(x) = p(-x)$. The function $f(x) = G(-x)$ satisfies $\mathcal{L}f(x) = -\delta_0(x)$ and is therefore the fundamental solution of the generator. Similarly, if $p \in \mathcal{P}_2'$, the fundamental solution of the generator is $f(x) = a(-x)$.

Proposition 4.5.1 *Suppose $p \in \mathcal{P}_d$ with $d \geq 2$, and $f : \mathbb{Z}^d \to \mathbb{R}$ is a function satisfying $f(0) = 0$, $f(x) = o(|x|)$ as $x \to \infty$, and $\mathcal{L}f(x) = 0$ for $x \neq 0$. Then, there exists $b \in \mathbb{R}$ such that*

$$f(x) = b\,[G(x) - G(0)], \quad d \geq 3,$$

$$f(x) = b\,a(x), \quad d = 2.$$

Proof See Propositions 6.4.6 and 6.4.8. □

Remark. The assumption $f(x) = o(|x|)$ is clearly needed since the function $f(x^1, \ldots, x^d) = x^1$ is harmonic.

Assume $d \geq 3$. if $f : \mathbb{Z}^d \to \mathbb{R}$ is a function with finite support we define

$$Gf(x) = \sum_{y \in \mathbb{Z}^d} G(x,y)\, f(y) = \sum_{y \in \mathbb{Z}^d} G(y-x)\, f(y). \tag{4.25}$$

Note that if f is supported on A, then $\mathcal{L}Gf(x) = 0$ for $x \notin A$. Also if $x \in A$,

$$\mathcal{L}Gf(x) = \mathcal{L}_x \sum_{y \in \mathbb{Z}^d} G(x,y)\, f(y) = \sum_{y \in \mathbb{Z}^d} \mathcal{L}_x G(x,y)\, f(y) = -f(x). \tag{4.26}$$

In other words $-G = \mathcal{L}^{-1}$. For this reason the Green's function is often called the *inverse of the (negative of the) Laplacian.* Similarly, if $d = 2$, and f has finite support, we define

$$af(x) = \sum_{y \in \mathbb{Z}^d} a(x,y)\, f(y) = \sum_{y \in \mathbb{Z}^d} a(y-x)\, f(y). \tag{4.27}$$

In this case we get

$$\mathcal{L}af(x) = f(x),$$

i.e., $a = \mathcal{L}^{-1}$.

4.6 Green's function for a set

If $A \subset \mathbb{Z}^d$ and S is a random walk with increment distribution p, let

$$\tau_A = \min\{j \geq 1 : S_j \notin A\}, \qquad \overline{\tau}_A = \min\{j \geq 0 : S_j \notin A\}. \tag{4.28}$$

If $A = \mathbb{Z}^d \setminus \{x\}$, we write just $\tau_x, \overline{\tau}_x$, which is consistent with the definition of τ_x given earlier in this chapter. Note that $\tau_A, \overline{\tau}_A$ agree if $S_0 \in A$, but are different if $S_0 \notin A$. If p is transient or A is a proper subset of $\mathbb{Z}^d$ we define

$$G_A(x,y) = \mathbb{E}^x \left[\sum_{n=0}^{\overline{\tau}_A - 1} 1\{S_n = y\} \right] = \sum_{n=0}^{\infty} \mathbb{P}^x\{S_n = y; n < \overline{\tau}_A\}.$$

Lemma 4.6.1 *Suppose $p \in \mathcal{P}_d$ and A is a proper subset of $\mathbb{Z}^d$.*

- $G_A(x,y) = 0$ *unless* $x, y \in A$.
- $G_A(x,y) = G_A(y,x)$ *for all* x, y.
- *For $x \in A$, $\mathcal{L}_x G_A(x,y) = \delta(y-x)$. In particular if $f(y) = G_A(x,y)$, then f vanishes on ∂A and satisfies $\mathcal{L}f(y) = -\delta(y-x)$ on A.*
- *For each $y \in A$,*

$$G_A(y,y) = \frac{1}{\mathbb{P}^y\{\tau_A < \tau_y\}} < \infty.$$

- *If $x, y \in A$, then*

$$G_A(x,y) = \mathbb{P}^x\{\overline{\tau}_y < \tau_A\}\, G_A(y,y).$$

- $G_A(x,y) = G_{A-x}(0, y-x)$ *where* $A - x = \{z - x : z \in A\}$.

Proof Easy and left to the reader. The second assertion may be surprising at first, but symmetry of the random walk implies that for $x, y \in A$,

$$\mathbb{P}^x\{S_n = y; n < \tau_A\} = \mathbb{P}^y\{S_n = x; n < \tau_A\}.$$

Indeed if $z_o = x, z_1, z_2, \ldots, z_{n-1}, z_n = y \in A$, then

$$\mathbb{P}^x\{S_1 = z_1, S_2 = z_2, \ldots, S_n = y\} = \mathbb{P}^y\{S_1 = z_{n-1}, S_2 = z_{n-2}, \ldots, S_n = x\}.$$

$\square$

The next proposition gives an important relation between the Green's function for a set and the Green's function or the potential kernel.

Proposition 4.6.2 *Suppose $p \in \mathcal{P}_d$, $A \subset \mathbb{Z}^d$, $x, y \in \mathbb{Z}^d$.*

(a) If $d \geq 3$,

$$G_A(x,y) = G(x,y) - \mathbb{E}^x[G(S_{\overline{\tau}_A}, y); \overline{\tau}_A < \infty] = G(x,y) - \sum_z \mathbb{P}^x\{S_{\overline{\tau}_A} = z\}\, G(z,y).$$

(b) If $d = 1, 2$ and A is finite,

$$G_A(x,y) = \mathbb{E}^x[a(S_{\overline{\tau}_A}, y)] - a(x,y) = \left[\sum_z \mathbb{P}^x\{S_{\overline{\tau}_A} = z\}\, a(z,y)\right] - a(x,y). \tag{4.29}$$

Proof The result is trivial if $x \notin A$. We will assume $x \in A$ in which case $\tau_A = \overline{\tau}_A$.

If $d \geq 3$, let $Y_y = \sum_{n=0}^{\infty} 1\{S_n = y\}$ denote the total number of visits to the point y. Then

$$Y_y = \sum_{n=0}^{\tau_A - 1} 1\{S_n = y\} + \sum_{n=\tau_A}^{\infty} 1\{S_n = y\}.$$

If we assume $S_0 = x$ and take expectations of both sides, we get

$$G(x,y) = G_A(x,y) + \mathbb{E}^x[G(S_{\tau_A}, y)].$$

The $d = 1, 2$ case could be done using a similar approach, but it is easier to use a different argument. If $S_0 = x$ and g is any function, then it is easy to check that

$$M_n = g(S_n) - \sum_{j=0}^{n-1} \mathcal{L}g(S_j)$$

is a martingale. We apply this to $g(z) = a(z,y)$ for which $\mathcal{L}g(z) = \delta(z - y)$. Then,

$$a(x,y) = \mathbb{E}^x[M_0] = \mathbb{E}^x[M_{n \wedge \tau_A}] = \mathbb{E}^x[a(S_{n \wedge \tau_A}, y)] - \mathbb{E}^x\left[\sum_{j=0}^{(n \wedge \tau_A)-1} 1\{S_j = y\}\right].$$

Since A is finite, the dominated convergence theorem implies that

$$\lim_{n \to \infty} \mathbb{E}^x[a(S_{n \wedge \tau_A}, y)] = \mathbb{E}^x[a(S_{\tau_A}, y)]. \tag{4.30}$$

The monotone convergence theorem implies

$$\lim_{n \to \infty} \mathbb{E}^x\left[\sum_{j=0}^{(n \wedge \tau_A)-1} 1\{S_j = y\}\right] = \mathbb{E}^x\left[\sum_{j=0}^{\tau_A - 1} 1\{S_j = y\}\right] = G_A(x,y).$$

□

The finiteness assumption on A was used in (4.30). The next proposition generalizes this to all proper subsets A of $\mathbb{Z}^d, d = 1, 2$. Recall that $\mathcal{B}_n = \{x \in \mathbb{Z}^d : |x| < n\}$. Define a function F_A by

$$F_A(x) = \lim_{n \to \infty} \frac{\log n}{\pi \sqrt{\det \Gamma}}\, \mathbb{P}^x\{\tau_{\mathcal{B}_n} < \overline{\tau}_A\}, \qquad d = 2,$$

$$F_A(x) = \lim_{n \to \infty} \frac{n}{\sigma^2}\, \mathbb{P}^x\{\tau_{\mathcal{B}_n} < \overline{\tau}_A\}, \quad d = 1.$$

The existence of these limits is established in the next proposition. Note that $F_A \equiv 0$ on $\mathbb{Z}^d \setminus A$ since $\mathbb{P}^x\{\overline{\tau}_A = 0\} = 1$ for $x \in \mathbb{Z}^d \setminus A$. Although we will not prove it now, F_A is nontrivial if and only if A is finite.

Proposition 4.6.3 *Suppose $p \in \mathcal{P}_d, d = 1, 2$ and A is a proper subset of $\mathbb{Z}^d$. Then if $x, y \in \mathbb{Z}^2$,*

$$G_A(x, y) = \mathbb{E}^x[a(S_{\overline{\tau}_A}, y)] - a(x, y) + F_A(x).$$

Proof The result is trivial if $x \notin A$ so we will suppose that $x \in A$. Choose $n > |x|, |y|$ and let $A_n = A \cap \{|z| < n\}$. Using (4.29), we have

$$G_{A_n}(x, y) = \mathbb{E}^x[a(S_{\tau_{A_n}}, y)] - a(x, y).$$

Note also that

$$\mathbb{E}^x[a(S_{\tau_{A_n}}, y)] = \mathbb{E}^x[a(S_{\tau_A}, y); \tau_A \le \tau_{\mathcal{B}_n}] + \mathbb{E}^x[a(S_{\tau_{\mathcal{B}_n}}, y); \tau_A > \tau_{\mathcal{B}_n}].$$

The monotone convergence theorem implies that as $n \to \infty$,

$$G_{A_n}(x, y) \longrightarrow G_A(x, y), \qquad \mathbb{E}^x[a(S_{\tau_A}, y); \tau_A \le \tau_{\mathcal{B}_n}] \longrightarrow \mathbb{E}^x[a(S_{\tau_A}, y)].$$

Therefore,

$$\lim_{n\to\infty} \mathbb{E}^x[a(S_{\tau_{\mathcal{B}_n}}, y); \tau_A > \tau_{\mathcal{B}_n}] = G_A(x, y) + a(x, y) - \mathbb{E}^x[a(S_{\tau_A}, y)].$$

However, $n \le |S_{\tau_{\mathcal{B}_n}}| \le n + R$ where R denotes the range of the increment distribution. Hence Theorems 4.4.3 and 4.4.7 show that as $n \to \infty$,

$$\mathbb{E}^x[a(S_{\tau_{\mathcal{B}_n}}, y); \tau_A > \tau_{\mathcal{B}_n}] \sim \mathbb{P}^x\{\tau_A > \tau_{\mathcal{B}_n}\} \, \frac{\log n}{\pi \sqrt{\det \Gamma}}, \qquad d = 2,$$

$$\mathbb{E}^x[a(S_{\tau_{\mathcal{B}_n}}, y); \tau_A > \tau_{\mathcal{B}_n}] \sim \mathbb{P}^x\{\tau_A > \tau_{\mathcal{B}_n}\} \, \frac{n}{\sigma^2}, \qquad d = 1.$$

□

Remark. We proved that for $d = 1, 2$,

$$F_A(x) = G_A(x, y) + a(x, y) - \mathbb{E}^x[a(S_{\overline{\tau}_A}, y)]. \tag{4.31}$$

This holds for all y. If we choose $y \in \mathbb{Z}^d \setminus A$, then $G_A(x, y) = 0$, and hence we can write

$$F_A(x) = a(x, y) - \mathbb{E}^x[a(S_{\overline{\tau}_A}, y)].$$

Using this expression it is easy to see that

$$\mathcal{L} F_A(x) = 0, \qquad x \in A.$$

Also, if $\mathbb{Z}^d \setminus A$ is finite,

$$F_A(x) = a(x) + O_A(1), \qquad x \to \infty.$$

In the particular case $A = \mathbb{Z}^d \setminus \{0\}, y = 0$, we see that

$$F_{\mathbb{Z}^d \setminus \{0\}}(x) = a(x).$$

Applying (4.31) with $y = x$, we can see that

$$G_{\mathbb{Z}^d\setminus\{0\}}(x,x) = F_{\mathbb{Z}^d\setminus\{0\}}(x) + a(0,x) = 2\,a(x). \tag{4.32}$$

The next simple proposition relates Green's functions to "escape probabilities" from sets. The proof uses a *last-exit decomposition.* Note that the last time a random walk visits a set is a random time that is not a stopping time. If $A \subset A'$, the event $\{\overline{\tau}_{\mathbb{Z}^d\setminus A} < \overline{\tau}_{A'}\}$ is the event that the random walk visits A before leaving A'.

Proposition 4.6.4 (Last-Exit Decomposition) *Suppose $p \in \mathcal{P}_d$ and $A \subset \mathbb{Z}^d$. Then,*

- *If A' is a proper subset of $\mathbb{Z}^d$ with $A \subset A'$,*

$$\mathbb{P}^x\{\overline{\tau}_{\mathbb{Z}^d\setminus A} < \overline{\tau}_{A'}\} = \sum_{z\in A} G_{A'}(x,z)\,\mathbb{P}^z\{\tau_{\mathbb{Z}^d\setminus A} > \tau_{A'}\}.$$

- *If $\xi \in (0,1)$ and T_ξ is an independent geometric random variable with killing rate $1-\xi$, then*

$$\mathbb{P}^x\{\overline{\tau}_{\mathbb{Z}^d\setminus A} < T_\xi\} = \sum_{z\in A} G(x,z;\xi)\,\mathbb{P}^z\{\tau_{\mathbb{Z}^d\setminus A} \geq T_\xi\}.$$

- *If $d \geq 3$ and A is finite,*

$$\begin{aligned} \mathbb{P}^x\{S_j \in A \text{ for some } j \geq 0\} &= \mathbb{P}^x\{\overline{\tau}_{\mathbb{Z}^d\setminus A} < \infty\} \\ &= \sum_{z\in A} G(x,z)\,\mathbb{P}^z\{\tau_{\mathbb{Z}^d\setminus A} = \infty\}. \end{aligned}$$

Proof We will prove the first; the other two are left as Exercise 4.9. We assume $x \in A'$ (for otherwise the result is trivial). On the event $\{\overline{\tau}_{\mathbb{Z}^d\setminus A} < \overline{\tau}_{A'}\}$, let σ denote the largest $k < \overline{\tau}_{A'}$ such that $S_k \in A$. Then,

$$\begin{aligned} \mathbb{P}^x\{\overline{\tau}_{\mathbb{Z}^d\setminus A} < \overline{\tau}_{A'}\} &= \sum_{k=0}^{\infty}\sum_{z\in A} \mathbb{P}^x\{\sigma = k; S_\sigma = z\} \\ &= \sum_{z\in A}\sum_{k=0}^{\infty} \mathbb{P}^x\{S_k = z; k < \tau_{A'}; S_j \notin A, j = k+1,\ldots,\tau_{A'}\}. \end{aligned}$$

The Markov property implies that

$$\mathbb{P}^x\{S_j \notin A, j = k+1,\ldots,\tau_{A'} \mid S_k = z; k < \tau_{A'}\} = \mathbb{P}^z\{\tau_{A'} < \tau_{\mathbb{Z}^d\setminus A}\}.$$

Therefore,

$$\begin{aligned} \mathbb{P}^x\{\overline{\tau}_{\mathbb{Z}^d\setminus A} < \overline{\tau}_{A'}\} &= \sum_{z\in A}\sum_{k=0}^{\infty} \mathbb{P}^x\{S_k = z; k < \tau_{A'}\}\,\mathbb{P}^z\{\tau_{A'} < \tau_{\mathbb{Z}^d\setminus A}\} \\ &= \sum_{z\in A} G_{A'}(x,z)\,\mathbb{P}^z\{\tau_{A'} < \tau_{\mathbb{Z}^d\setminus A}\}. \end{aligned}$$

□

The next proposition uses a last-exit decomposition to describe the distribution of a random walk conditioned to not return to its starting point before a killing time. The killing time is either geometric or the first exit time from a set.

Proposition 4.6.5 *Suppose S_n is a p-walk with $p \in \mathcal{P}_d$; $0 \in A \subset \mathbb{Z}^d$; and $\xi \in (0,1)$. Let T_ξ be a geometric random variable independent of the random walk with killing rate $1-\xi$. Let*

$$\rho = \max\{j \geq 0 : j \leq \tau_A, S_j = 0\}, \quad \rho^* = \max\{j \geq 0; j < T_\xi, S_j = 0\}.$$

- *The distribution of $\{S_j : \rho \leq j \leq \tau_A\}$ is the same as the conditional distribution of $\{S_j : 0 \leq j \leq \tau_A\}$ given $\rho = 0$.*
- *The distribution of $\{S_j : \rho^* \leq j < T_\xi\}$ is the same as the conditional distribution of $\{S_j : 0 \leq j < T_\xi\}$ given $\rho^* = 0$.*

Proof The usual Markov property implies that for any positive integer j, any $x_1, x_2, \ldots, x_{k-1} \in A \setminus \{0\}$ and any $x_k \in Z^d \setminus A$,

$$\begin{aligned}
&\mathbb{P}\{\rho = j, \tau_A = j+k, S_{j+1} = x_1, \ldots, S_{j+k} = x_k\} \\
&\qquad = \mathbb{P}\{S_j = 0, \tau_A > j, S_{j+1} = x_1, \ldots, S_{j+k} = x_k\} \\
&\qquad = \mathbb{P}\{S_j = 0, \tau_A > j\}\, \mathbb{P}\{S_1 = x_1, \ldots, S_k = x_k\}.
\end{aligned}$$

The other equality is done similarly. □

Exercises

Exercise 4.1 Suppose $p \in \mathcal{P}_d$ and S_n is a p-walk. Suppose $A \subset \mathbb{Z}^d$ and that $\mathbb{P}^x\{\overline{\tau}_A = \infty\} > 0$ for some $x \in A$. Show that for every $\epsilon > 0$, there is a y with $\mathbb{P}^y\{\overline{\tau}_A = \infty\} > 1 - \epsilon$.

Exercise 4.2 Suppose $d = 1$. Show that the only function satisfying the conditions of Proposition 4.5.1 is the zero function.

Exercise 4.3 Find all radially symmetric functions f in $\mathbb{R}^d \setminus \{0\}$ satisfying $\Delta f(x) = 0$ for all $x \in \mathbb{R}^d \setminus \{0\}$.

Exercise 4.4 For each positive integer k find positive integer d and $p \in \mathcal{P}'_d$ such that $\mathbb{E}[|X_1|^k] < \infty$ and

$$\limsup_{|x| \to \infty} |x|^{d-2}\, G(x) = \infty.$$

(Hint: Consider a sequence of points $z_1, z_2, \ldots$ going to infinity and define $\mathbb{P}\{X_1 = z_j\} = q_j$. Note that $G(z_j) \geq q_j$. Make a good choice of $z_1, z_2, \ldots$ and $q_1, q_2, \ldots$)

Exercise 4.5 Suppose $X_1, X_2, \ldots$ are independent, identically distributed random variables in $\mathbb{Z}$ with mean zero. Let $S_n = X_1 + \cdots + X_n$ denote the corresponding random walk and let

$$G_n(x) = \sum_{j=0}^{n} \mathbb{P}\{S_j = x\}$$

be the expected number of visits to x in the first n steps of the walk.

(a) Show that $G_n(x) \le G_n(0)$ for all n.

(b) Use the law of large numbers to conclude that for all $\epsilon > 0$ there is an N_ϵ such that for $n \ge N_\epsilon$,

$$\sum_{|x| \le \epsilon n} G_n(x) \ge \frac{n}{2}.$$

(c) Show that

$$G(0) = \lim_{n\to\infty} G_n(0) = \infty$$

and conclude that the random walk is recurrent.

Exercise 4.6 Suppose $A \subsetneq \mathbb{Z}^d$ and $x, y \in A$. Show that

$$G_A(x,y) = \lim_{n\to\infty} G_{A_n}(x,y),$$

where $A_n = \{z \in A : |z| < n\}$.

Exercise 4.7 Let S_n denote simple random walk in $\mathbb{Z}^2$ starting at the origin and let $\rho = \min\{j \ge 1 : S_j = 0 \text{ or } \mathbf{e}_1\}$. Show that $\mathbb{P}\{S_\rho = 0\} = 1/2$.

Exercise 4.8 Suppose $p \in \mathcal{P}_1$ and let $A = \{1, 2, \ldots\}$. Show that

$$F_A(x) = \frac{x - \mathbb{E}^x[S_T]}{\sigma^2},$$

where $T = \min\{j \ge 0 : S_j \le 0\}$ and F_A is as in (4.31).

Exercise 4.9 Finish the details in Proposition 4.6.4.

Exercise 4.10 Finish the details in Theorem 4.4.7.

5
One-dimensional walks

5.1 Gambler's ruin estimate

One of the basic estimates for one-dimensional random walks with zero mean and finite variance goes under the name of *gambler's ruin estimate.* We will not restrict to integer-valued random walks. For this section we assume that $X_1, X_2, \ldots$ are independent, identically distributed random variables with $\mathbb{E}[X_1] = 0, \mathbb{E}[X_1^2] = \sigma^2 > 0$. We let $S_n = S_0 + X_1 + \cdots + X_n$ be the corresponding random walk. If $r > 0$, we let

$$\eta_r = \min\{n \geq 0 : S_n \leq 0 \text{ or } S_n \geq r\},$$

$$\eta = \eta_\infty = \min\{n \geq 0 : S_n \leq 0\}.$$

For simple random walk, the gambler's ruin estimate is particularly nice (and is exact).

Proposition 5.1.1 *If S_n is one-dimensional simple random walk and $j < k$ are positive integers, then*

$$\mathbb{P}^j\{S_{\eta_k} = k\} = \frac{j}{k}.$$

Proof Since $M_n := S_{n \wedge \eta_k}$ is a bounded martngale, the optional sampling theorem implies that

$$j = \mathbb{E}^j[M_0] = \mathbb{E}^j[M_{\eta_k}] = k\,\mathbb{P}^j\{S_{\eta_k} = k\}.$$

$\square$

Proposition 5.1.2 *If S_n is one-dimensional simple random walk, then for positive integer n,*

$$\mathbb{P}^1\{\eta > 2n\} = \mathbb{P}^1\{S_{2n} > 0\} - \mathbb{P}^1\{S_{2n} < 0\} = \mathbb{P}\{S_{2n} = 0\} = \frac{1}{\sqrt{\pi\, n}} + O\left(\frac{1}{n^{3/2}}\right).$$

Proof Symmetry and the Markov property tell us that each $k < 2n$ and each positive integer x,

$$\mathbb{P}^1\{\eta = k; S_{2n} = x\} = \mathbb{P}^1\{\eta = k\}\, p_{2n-k}(x) = \mathbb{P}^1\{\eta = k; S_{2n} = -x\}.$$

Therefore,

$$\mathbb{P}^1\{\eta \leq 2n; S_{2n} = x\} = \mathbb{P}^1\{\eta \leq 2n; S_{2n} = -x\}.$$

But, $\mathbb{P}^1\{\eta > 2n; S_{2n} = -x\} = 0$. Therefore,

$$\begin{aligned}
\mathbb{P}^1\{\eta > 2n\} &= \sum_{x>0} \mathbb{P}\{\eta > 2n; S_{2n} = x\} \\
&= \sum_{x>0} [p_{2n}(1,x) - p_{2n}(1,-x)] \\
&= p_{2n}(1,1) + \sum_{x>0} [p_{2n}(1,x+2) - p_{2n}(1,-x)] \\
&= p_{2n}(0,0) = 4^{-n}\binom{2n}{n} = \frac{1}{\sqrt{\pi\, n}} + O\left(\frac{1}{n^{3/2}}\right).
\end{aligned}$$

□

The basic argument in the proof of Proposition 5.1.1 extends to more general walks. However, there is a complication arising from the fact that we do not know the exact value of S_{η_k}. We start by showing that the optional sampling theorem step is valid.

Lemma 5.1.3 *If $X_1, X_2, \ldots$ are i.i.d. random variables in $\mathbb{R}$ with $\mathbb{E}(X_j) = 0$, then for every $r > 0$ and every $x \in \mathbb{R}$,*

$$\mathbb{E}^x[S_{\eta_r}] = x. \tag{5.1}$$

Proof We start by showing that $\mathbb{E}^x[|S_{\eta_r}|] < \infty$. There exists an integer m and a $\delta > 0$ such that

$$\mathbb{P}\{|X_1 + \cdots + X_m| > r\} \geq \delta.$$

Therefore for all x and all positive integers j,

$$\mathbb{P}^x\{\eta_r > jm\} \leq (1-\delta)^m.$$

In particular, $\mathbb{E}^x[\eta_r] < \infty$. By the Markov property,

$$\mathbb{P}^x\{|S_{\eta_r}| \geq r + y; \eta_r = k\} \leq \mathbb{P}^x\{\eta_r > k-1; |X_k| \geq y\} = \mathbb{P}^x\{\eta_r > k-1\}\,\mathbb{P}\{|X_k| \geq y\}.$$

Summing over k gives

$$\mathbb{P}^x\{|S_{\eta_r}| \geq r + y\} \leq \mathbb{E}^x[\eta_r]\,\mathbb{P}\{|X_k| \geq y\}.$$

Hence

$$\begin{aligned}
\mathbb{E}^x\left[|S_{\eta_r}|\right] = \int_0^\infty \mathbb{P}^x\{|S_{\eta_r}| \geq y\}\, dy &\leq \mathbb{E}^x[\eta_r]\left[r + \int_0^\infty \mathbb{P}\{|X_k| \geq y\} dy\right] \\
&= \mathbb{E}^x[\eta_r]\ (\ r + \mathbb{E}\left[|X_j|\right]\ ) < \infty.
\end{aligned}$$

Given this, the martingale $M_n := S_{n\wedge\eta_r}$ is dominated by the integrable random variable $r + |S_{\eta_r}|$. Hence it is a uniformly integrable martingale, and (5.1) follows from the optional sampling theorem (Theorem 11.2.3). □

Proposition 5.1.4 *For every $\epsilon > 0$ and $K < \infty$, there exist $0 < c_1 < c_2 < \infty$ such that if $\mathbb{P}\{|X_1| > K\} = 0$ and $\mathbb{P}\{X_1 \geq \epsilon\} \geq \epsilon$, then for all $0 < x < r$,*

$$c_1\,\frac{x+1}{r} \leq \mathbb{P}^x\{S_{\eta_r} \geq r\} \leq c_2\,\frac{x+1}{r}.$$

Proof It suffices to show that for $K \leq x \leq r/2$,

$$c_1 \frac{x}{r} \leq \mathbb{P}^x\{S_{\eta_r} \geq r\} \leq c_2 \frac{x}{r}.$$

We know that $\mathbb{E}^x[S_{\eta_r}] = x$. If $S_{\eta_r} \geq r$, then $r \leq S_{\eta_r} \leq r + K$. If $S_{\eta_r} \leq 0$, then $-K \leq S_{\eta_r} \leq 0$. This allows us to estimate the expectation in terms of $\mathbb{P}\{S_{\eta_r} \geq r\}$. Indeed,

$$\mathbb{E}^x[S_{\eta_r}] \leq \mathbb{E}^x[S_{\eta_r}; S_{\eta_r} \geq r] \leq \mathbb{P}^x\{S_{\eta_r} \geq r\}\,(r + K) \leq 4r\,\mathbb{P}^x\{S_{\eta_r} \geq r\},$$

$$\mathbb{E}^x[S_{\eta_r}] \geq \mathbb{E}^x[S_{\eta_r}; S_{\eta_r} \geq r] - K \geq r\,\mathbb{P}^x\{S_{\eta_r} \geq r\} - x.$$

□

If $p \in \mathcal{P}_d, d \geq 2$, then p induces an infinite family of one-dimensional non-lattice random walks $S_n \cdot \theta$ where $|\theta| = 1$. It is useful to gambler's ruin estimates for these walks that are uniform over all θ. One simple geometric fact is that there is an $\epsilon > 0$, independent of θ such that $\mathbb{P}\{S_1 \cdot \theta \geq \epsilon\} \geq \epsilon$, see Exercise 1.7. Also, $c_1 \leq \mathbb{E}[|S_1 \cdot \theta|^2] \leq c_2$. The invariance principle implies that there is a $c' > 0$ such that for all n sufficiently large

$$\mathbb{P}\left\{\max_{0 \leq j \leq n} |S_j| < \sqrt{n}\right\} \geq c'.$$

Using this and the geometric fact, we can see that there is a c such that for every $|\theta| = 1$ amd positive integer n,

$$\mathbb{P}\left\{\min_{0 \leq j \leq n} S_j \cdot \theta > -\sqrt{n}\right\} > c.$$

In section 5.1.1 we prove a gambler's ruin estimate for walks assuming only mean zero and finite variance. Enough care is taken in the estimate so that the constants obtained will be uniform over the all the walks $S_n \cdot \theta$.

5.1.1 General case

We prove the gambler's ruin estimate assuming only mean zero and finite variance. While we will not attempt to get the best values for the constants, we do show that the constants can be chosen uniformly over a wide class of distributions. In this section we fix $K < \infty$, $\delta, b > 0$ and $0 < \rho < 1$, and we let $\mathcal{A} = \mathcal{A}(K, \delta, b, \rho)$ the collection of distributions on X_1 with $\mathbb{E}[X_1] < \infty$,

$$\mathbb{E}[X_1^2] = \sigma^2 \leq K^2,$$

$$\mathbb{P}\{X_1 \geq 1\} \geq \delta,$$

$$\inf_n \mathbb{P}\{S_1, \ldots, S_{n^2} > -n\} \geq b,$$

$$\rho \leq \inf_{n > 0} \mathbb{P}\{S_{n^2} \leq -n\}.$$

It is easy to check that for any mean zero, finite variance random walk S_n we can find a $t > 0$ and some K, δ, b, ρ such that the estimates above hold for tS_n. Also, if $p \in \mathcal{P}_d$, the comments at the end of the previous section imply that we can find t, K, δ, b, ρ such that the estimates hold for $tS_n \cdot \theta$ for all $|\theta| = 1$.

Theorem 5.1.5 (Gambler's ruin) *If $X_1, X_2, \ldots$ are i.i.d. random variables with mean zero and finite variance, there exist $0 < c_1 < c_2 < \infty$ (depending on K, δ, b, ρ but not otherwise on the distribution of X_1) such that for all $0 < x < r$,*

$$c_1 \frac{x}{r} \le \mathbb{P}^x\{\eta > r^2\} \le c_2 \frac{x+1}{r},$$

$$c_1 \frac{x}{r} \le \mathbb{P}^x\{S_{\eta_r} \ge r\} \le c_2 \frac{x+1}{r}.$$

We will prove this in a number of steps. We start with the upper bound. Let

$$\eta_r^* = \min\{n > 0 : S_n \le 0 \text{ or } S_n \ge r\}, \quad \eta^* = \eta_\infty^* = \min\{n > 0 : S_n \le 0\}.$$

Note that η_r^* differs from η_r in that the minimum is taken over $n > 0$ rather than $n \ge 0$.

Lemma 5.1.6

$$\mathbb{P}\{\eta^* > n\} \le \frac{4K}{\delta\sqrt{n}}, \qquad \mathbb{P}\{\eta_n^* < \eta^*\} \le \frac{4K}{b\delta\, n}.$$

Proof Let $q_n = \mathbb{P}\{\eta^* > n\} = \mathbb{P}\{S_1, \ldots, S_n > 0\}$. Then

$$\mathbb{P}\{S_1, \ldots, S_n \ge 1\} \ge \delta q_{n-1} \ge \delta q_n.$$

Let $J_{k,n}$ be the event

$$J_{k,n} = \{S_{k+1}, \ldots, S_n \ge S_k + 1\}.$$

We will also use $J_{k,n}$ to denote the indicator function of the event. Let $m_n = \min\{S_j : 0 \le j \le n\}$, $M_n = \max\{S_j : 0 \le j \le n\}$. For each real $x \in [m_n, M_n]$, there is at most one integer k such that $S_k \le x$ and $S_j > x, k < j \le n$. On the event $J_{k,n}$ there is an interval of length at least one such that for x in the interval, $S_k \le x$ and $S_j > x, k < j \le n$. In other words, there are $\sum_k J_{k,n}$ nonoverlapping intervals contained in $[m_n, M_n]$ each of length at least one. Therefore,

$$\sum_{k=0}^{n-1} J_{k,n} \le M_n - m_n.$$

But, $\mathbb{P}(J_{k,n}) \ge \delta q_{n-k} \ge \delta q_n$. Therefore,

$$n\delta q_n \le \mathbb{E}[M_n - m_n] \le 2\,\mathbb{E}[\max\{|S_j| : j \le n\}].$$

Martingale maximal inequalities (Theorem 11.2.5) give

$$\mathbb{P}\{\max\{|S_j| : j \le n\} \ge t\} \le \frac{\mathbb{E}[S_n^2]}{t^2} \le \frac{K^2 n}{t^2}.$$

Therefore,

$$\begin{aligned} \frac{n\delta q_n}{2} \le \mathbb{E}[\max\{|S_j| : j \le n\}] &= \int_0^\infty \mathbb{P}\{\max\{|S_j| : j \le n\} \ge t\}\, dt \\ &\le K\sqrt{n} + \int_{K\sqrt{n}}^\infty K^2\, n\, t^{-2}\, dt = 2K\sqrt{n}. \end{aligned}$$

This gives the first inequality. By using the strong Markov property, we can see that

$$\mathbb{P}\{\eta^* > n^2 \mid \eta_n^* < \eta^*\} \geq \mathbb{P}\{S_j - S_{\eta_n^*} > -n, 1 \leq j \leq n^2 \mid \eta_n^* < \eta^*\} \geq b.$$

Hence,

$$b\,\mathbb{P}\{\eta_n^* < \eta^*\} \leq \mathbb{P}\{\eta^* > n^2\}, \tag{5.2}$$

which gives the second inequality. □

Lemma 5.1.7 (Overshoot lemma I) *For all $x > 0$,*

$$\mathbb{P}^x\{|S_\eta| \geq m\} \leq \frac{1}{\rho}\,\mathbb{E}[X_1^2; |X_1| \geq m]. \tag{5.3}$$

Moreover if $\alpha > 0$ and $\mathbb{E}[|X_1|^{2+\alpha}] < \infty$, then

$$\mathbb{E}\left[|S_\eta|^\alpha\right] \leq \frac{\alpha}{\rho}\,\mathbb{E}[|X_1|^{2+\alpha}].$$

Comment. Since $\mathbb{E}^x[\eta] = \infty$, we cannot use the proof from Lemma 5.1.3.

Proof Fix $\epsilon > 0$. For nonnegative integers k, let

$$Y_k = \sum_{n=0}^{\eta} 1\{k\epsilon < S_n \leq (k+1)\epsilon\}$$

be the number of times the random walk visits $(k\epsilon, (k+1)\epsilon]$ before hitting $(-\infty, 0]$, and let

$$g(x,k) = \mathbb{E}^x[Y_k] = \sum_{n=0}^{\infty} \mathbb{P}^x\{k\epsilon < S_n \leq (k+1)\epsilon; \eta > n\}.$$

Note that if $m, x > 0$,

$$\begin{aligned}
\mathbb{P}^x\{|S_\eta| \geq m\} &= \sum_{n=0}^{\infty} \mathbb{P}^x\{|S_\eta| \geq m; \eta = n+1\} \\
&= \sum_{n=0}^{\infty}\sum_{k=0}^{\infty} \mathbb{P}^x\{|S_\eta| \geq m; \eta = n+1, k\epsilon < S_n \leq (k+1)\epsilon\} \\
&\leq \sum_{k=0}^{\infty}\sum_{n=0}^{\infty} \mathbb{P}^x\{\eta > n; k\epsilon < S_n \leq (k+1)\epsilon; |S_{n+1} - S_n| \geq m + k\epsilon\} \\
&= \sum_{k=0}^{\infty} g(x,k)\,\mathbb{P}\{|X_1| \geq m + k\epsilon\} \\
&= \sum_{k=0}^{\infty} g(x,k) \sum_{l=k}^{\infty} \mathbb{P}\{m + l\epsilon \leq |X_1| < m + (l+1)\epsilon\} \\
&= \sum_{l=0}^{\infty} \mathbb{P}\{m + l\epsilon \leq |X_1| < m + (l+1)\epsilon\} \sum_{k=0}^{l} g(x,k).
\end{aligned}$$

We claim that for all x, y,

$$\sum_{0 \leq k\epsilon < y} g(x,k) \leq \frac{y^2}{\rho}. \tag{5.4}$$

To see this, let

$$H_y = \max_x \sum_{0 \leq k\epsilon < y} g(x,k).$$

Then for any $x \leq y$,

$$\sum_{0 \leq k\epsilon < y} g(x,k) \leq y^2 + \mathbb{P}^x\{\eta \geq y^2\}\mathbb{E}\left[\sum_{n \geq y^2} 1\{S_n \leq y; n < \eta\} \;\mid \eta \geq y^2\right] \leq y^2 + (1-\rho)\, H_y.$$

By taking the supremum over x we get $H_y \leq y^2 + (1-\rho)H_y$ which gives (5.4). We therefore have

$$\begin{aligned}
\mathbb{P}^x\{|S_\eta| \geq m\} &\leq \frac{1}{\rho}\sum_{l=0}^{\infty} \mathbb{P}\{m + l\epsilon \leq |X_1| < m + (l+1)\epsilon\}(l\epsilon + \epsilon)^2 \\
&\leq \frac{1}{\rho}\,\mathbb{E}[(X_1+\epsilon)^2; |X_1| \geq m].
\end{aligned}$$

Letting $\epsilon \to 0$, we get (5.3).

To get the second estimate, let F denote the distribution function of $|X_1|$. Then

$$\begin{aligned}
\mathbb{E}[|S_\eta|^\alpha] &= \alpha \int_0^\infty t^{\alpha-1}\,\mathbb{P}\{|S_\eta| \geq t\}\,dt \\
&\leq \frac{\alpha}{\rho}\int_0^\infty t^{\alpha-1}\,\mathbb{E}[X_1^2; |X_1| \geq t]\,dt \\
&\leq \frac{\alpha}{\rho}\int_0^\infty \mathbb{E}[X_1^{1+\alpha}; |X_1| \geq t]\,dt \\
&= \frac{\alpha}{\rho}\int_0^\infty \int_t^\infty x^{1+\alpha}\,dF(x)\,dt \\
&= \frac{\alpha}{\rho}\int_0^\infty \left[\int_0^x dt\right] x^{1+\alpha}\,dF(x) = \frac{\alpha}{\rho}\,\mathbb{E}[|X_1|^{2+\alpha}].
\end{aligned}$$

□

Comment. The lemma states that the overshoot random variable has two fewer moments than the increment distribution. When the starting point is close to the origin, one might expect that the overshoot would be smaller since there are fewer chances for the last step before entering $(-\infty, 0]$ to be much larger than a typical step. The next lemma confirms this intuition and shows that one gains one moment if one starts near the origin.

Lemma 5.1.8 (Overshoot lemma II) *Let*

$$c' = \frac{32\,K}{b\delta}.$$

Then for all $0 < x \leq 1$,

$$\mathbb{P}^x\{|S_\eta| \geq m\} \leq \frac{c'}{\rho}\, \mathbb{E}[|X_1|; |X_1| \geq m].$$

Moreover if $\alpha > 0$ *and* $\mathbb{E}[|X_1|^{1+\alpha}] < \infty$, *then*

$$\mathbb{E}[|S_\eta|^\alpha] \leq \frac{\alpha\, c'}{\rho}\, \mathbb{E}[|X_1|^{1+\alpha}].$$

Proof The proof proceeds exactly as in Theorem 5.1.7 up to (5.4) which we replace with a stronger estimate that is valid for $0 < x \leq 1$:

$$\sum_{0 \leq k\epsilon < y} g(x,k) \leq \frac{c'\, y}{\rho}. \tag{5.5}$$

To derive this estimate we note that

$$\sum_{2^{j-1} \leq k\epsilon < 2^j} g(x,k)$$

equals the product of the probability of reaching level 2^{j-1} before dropping below zero and the expected number of visits in this range given that event. The first probability is less than $[4K/(b\delta 2^{j-1}]$ and the conditional expectation, as estimated in (5.4), is less than $2^{2j}/rho$. Therefore,

$$\sum_{0 \leq k\epsilon < 2^j} g(x,k) \leq \frac{1}{\rho} \sum_{l=1}^{j} \left(\frac{4K}{b\delta 2^{l-1}}\right) 2^{2l} \leq \frac{1}{\rho}\left(\frac{16K}{b\delta}\right) 2^j.$$

For general y we write $2^{j-1} < y \leq 2^j$ and this gives (5.5).

Given this, the same argument gives

$$\mathbb{P}^x\{|S_\eta| \geq m\} \leq \frac{c'}{\rho}\, \mathbb{E}[|X_1|; |X_1| \geq m],$$

and

$$\begin{aligned}
\mathbb{E}[|S_\eta|^\alpha] &= \alpha \int_0^\infty t^{\alpha-1}\, \mathbb{P}\{|S_\eta| \geq t\}\, dt \\
&\leq \frac{\alpha\, c'}{\rho} \int_0^\infty t^{\alpha-1}\, \mathbb{E}[|X_1|; |X_1| \geq t]\, dt \\
&\leq \frac{\alpha\, c'}{\rho} \int_0^\infty \mathbb{E}[X_1^\alpha; |X_1| \geq t]\, dt = \frac{\alpha\, c'}{\rho}\, \mathbb{E}[|X_1|^{1+\alpha}].
\end{aligned}$$

□

Remark. The inequalities (5.4) and (5.5) are important. There is a $c < \infty$ such that for all y, $\mathbb{E}^y[\eta_n^*] < cn^2$. However,

$$\mathbb{E}^x[\eta_n^*] \leq cn, \qquad 0 < x \leq 1. \tag{5.6}$$

Lemma 5.1.9

$$\mathbb{P}\{\eta_n^* < \eta^*\} \geq \frac{c^*}{n},$$

where

$$c^* = \frac{\rho}{2\,(\rho + 2c' K^2)},$$

and ρ, c' *are as in Lemma 5.1.8. Also,*

$$\mathbb{P}\{\eta^* \geq n\} \geq b\,\mathbb{P}\{\eta^*_{\sqrt{n}} < \eta^*\} \geq \frac{b\,c^*}{\sqrt{n}}.$$

Proof The last assertion follows immediately from the first and the strong Markov property as in (5.2). To prove the first it suffices to prove that

$$\mathbb{P}^1\{\eta^*_n < \eta^*\} \geq \frac{c^*}{n}.$$

Using (5.6), we have

$$\begin{aligned}
\mathbb{P}^1\{|S_{\eta^*_n}| \geq s + n\} &\leq \sum_{l=0}^{\infty} \mathbb{P}^1\{\eta^*_n = l+1; |S_{\eta^*_n}| \geq s\} \\
&\leq \sum_{l=0}^{\infty} \mathbb{P}^1\{\eta^*_n > l; |X_1| \geq s\} \\
&\leq \mathbb{P}\{|X_1| \geq s\}\,\mathbb{E}^1[\eta^*_n] \\
&\leq \frac{c'}{\rho}\,n\,\mathbb{P}\{|X_1| \geq s\}.
\end{aligned}$$

In particular, if $t > 0$,

$$\begin{aligned}
\mathbb{E}^1\left[|S_{\eta^*_n}|; |S_{\eta^*_n}| \geq (1+t)\,n\right] &= \int_{tn}^{\infty} \mathbb{P}^1\{|S_{\eta^*_n}| \geq s + n\}\,ds \\
&\leq \frac{c'\,n}{\rho} \int_{tn}^{\infty} \mathbb{P}\{|X_1| \geq s\}\,ds \\
&= \frac{c'\,n}{\rho}\mathbb{E}[|X_1|; |X_1| \geq tn] \leq \frac{c'\,K^2}{\rho\,t}.
\end{aligned}$$

Consider the martingale $M_n = S_{n \wedge \eta^*_n}$. Since

$$1 = \mathbb{E}^1[M_0] = \mathbb{E}^1[M_\infty] \leq \mathbb{E}^1[S_{\eta^*_n}; S_{\eta^*_n} \geq n],$$

if we let $t_0 = 2c' K^2/\rho$, then

$$\mathbb{E}^1[|S_{\eta^*_n}|; |S_{\eta^*_n}| \geq (1+t_0)\,n] \leq \frac{1}{2},$$

and hence

$$\mathbb{E}^1[S_{\eta^*_n}; n \leq S_{\eta^*_n} \leq (1+t_0)n] \geq \frac{1}{2},$$

which implies

$$\mathbb{P}^1\{\eta^*_n < \eta^*\} \geq \mathbb{P}^1\{n \leq S_{\eta^*_n} \leq (1+t_0)n\} \geq \frac{1}{2(1+t_0)n}.$$

$\square$

Proof (of Theorem 5.1.5) Lemmas 5.1.6 and 5.1.9 prove the result for $0 \leq x \leq 1$. The result is easy if $x \geq r/2$ so we will assume $1 \leq x \leq r/2$. Clearly the function $x \mapsto \mathbb{P}^x\{S_{\eta_r} \geq r\}$ is nondecreasing in x. Therefore,

$$\mathbb{P}\{S_{\eta_r} \geq r\} = \mathbb{P}\{S_{\eta_x} \geq x\}\, \mathbb{P}\{S_{\eta_r} \geq r \mid S_{\eta_x} \geq x\} \geq \mathbb{P}\{S_{\eta_x} \geq x\}\, \mathbb{P}^x\{S_{\eta_r} \geq r\}.$$

Hence by Lemmas 5.1.6 and 5.1.9,

$$\mathbb{P}^x\{S_{\eta_r} \geq r\} \leq \frac{\mathbb{P}\{S_{\eta_r} \geq r\}}{\mathbb{P}\{S_{\eta_x} \geq x\}} \leq \frac{4K}{c^* b\delta}\, \frac{x}{r}.$$

For the other direction, we first claim that there is a c such that $\mathbb{E}^x[\eta_r] \leq c_2\, xr$. We have already shown that $\mathbb{E}^y[\eta_r] \leq cr$ for $0 < y \leq 1$. But the strong Markov property can be used to see that

$$\mathbb{E}^x[\eta_r] \leq \mathbb{E}^1[\eta_r] + \mathbb{E}^{x-1}[\eta_r].$$

Hence we have the estimate by induction. Once we have this, we can show as in the previous lemma that

$$\mathbb{E}^x[|S_{\eta_r^*}|; |S_{\eta_r^*}| \geq (1+t)\, r\} \leq \frac{c_2\, K^2\, x}{t},$$

and hence if $t_0 = 2c_2K^2$,

$$\mathbb{E}^x[S_{\eta_r^*}; S_{\eta_r^*} \geq (1+t_0)\, r\} \leq \frac{x}{2},$$

$$\mathbb{E}^x[|S_{\eta_r^*}|; n \leq S_{\eta_r^*} \leq (1+t_0)\, r\} \geq \frac{x}{2},$$

$$\mathbb{P}^x\{r \leq S_{\eta_r^*} \leq (1+t_0)\, r\} \geq \frac{x}{2(1+t_0)r}.$$

As we have already shown, this implies

$$\mathbb{P}^x\{\eta^* \geq r^2\} \geq b\, \frac{x}{2(1+t_0)r}.$$

□

5.2 One-dimensional killed walks

A *symmetric defective increment distribution (on $\mathbb{Z}$)* is a set of nonnegative numbers $\{p_k : k \in \mathbb{Z}\}$ with $\sum p_k < 1$ and $p_{-k} = p_k$ for all k. Given a symmetric defective increment distribution, we have the corresponding symmetric random walk with killing S_j. This is a Markov chain with state space $\mathbb{Z} \cup \{\infty\}$, where ∞ is an absorbing state, and

$$\mathbb{P}\{S_{j+1} = k + l \mid S_j = k\} = p_l, \qquad \mathbb{P}\{S_{j+1} = \infty \mid S_j = k\} = p_\infty.$$

Here $p_\infty = 1 - \sum p_k$. We let

$$T = \min\{j : S_j = \infty\}$$

denote the *killing time* for the random walk. Note that $\mathbb{P}\{T = j\} = p_\infty\, (1 - p_\infty)^{j-1}, j \in \{1, 2, \ldots\}$.

Examples.

- Suppose $p(j)$ is the increment distribution of a symmetric one-dimensional random walk and $s \in [0,1)$. Then $p_j = s\, p(j)$ is a defective increment distribution corresponding to the random walk with killing rate $1-s$. Conversely, if p_j is a symmetric defective increment distribution, and $p(j) = p_j/(1-p_\infty)$, then $p(j)$ is an increment distribution of a symmetric one-dimensional random walk (not necessarily aperiodic or irreducible). If we then kill this walk at rate $1-p_\infty$ we get back p_j.
- Suppose $\overline{S}_j$ is a symmetric random walk in $\mathbb{Z}^d, d \geq 2$ which we write $\overline{S}_j = (Y_j, Z_j)$ where Y_j is a random walk in $\mathbb{Z}$ and Z_j is a random walk in $\mathbb{Z}^{d-1}$. Suppose the random walk is killed at rate $1-s$ and let $\hat{T}$ denote the killing time. Let
$$\tau = \min\{j \geq 1 : Z_j = 0\}. \tag{5.7}$$
Then set
$$p_k = \mathbb{P}\{Y_\tau = k; \tau < \hat{T}\}.$$
Note that
$$\begin{aligned} p_k &= \sum_{j=1}^{\infty} \mathbb{P}\{\tau = j; Y_j = k; j < \hat{T}\} \\ &= \sum_{j=1}^{\infty} s^j\, \mathbb{P}\{\tau = j; Y_j = k\} = \mathbb{E}[s^{\hat{T}}; Y_{\hat{T}} = k; \hat{T} < \infty]. \end{aligned}$$
If Z is a transient random walk, then $\mathbb{P}\{\tau < \infty\} < 1$ and we can let $s = 1$.
- Suppose $\overline{S}_j = (Y_j, Z_j)$ is as in the previous example and suppose $A \subset \mathbb{Z}^{d-1} \setminus \{0\}$. Let
$$\sigma_A = \min\{j : Z_j \in A\},$$
$$p_k = \mathbb{P}\{Y_\tau = k; \tau < \sigma_A\}.$$
If $\mathbb{P}\{Z_j \in A \text{ for some } j\} > 0$, then $\{p_k\}$ is a defective increment distribution.

Given a symmetric defective increment distribution $\{p_k\}$ define the events
$$V_+ = \{S_j > 0 : j = 1, \ldots, T-1\}, \quad \overline{V}_+ = \{S_j \geq 0 : j = 1, \ldots, T-1\},$$
$$V_- = \{S_j < 0 : j = 1, \ldots, T-1\}, \quad \overline{V}_- = \{S_j \leq 0 : j = 1, \ldots, T-1\}.$$
Symmetry implies that $\mathbb{P}(V_+) = \mathbb{P}(V_-), \mathbb{P}(\overline{V}_+) = \mathbb{P}(\overline{V}_-)$. Note that
$$\mathbb{P}(V_+ \cap \overline{V}_-) = \mathbb{P}(\overline{V}_+ \cap V_-) = \mathbb{P}\{T = 1\} = p_\infty. \tag{5.8}$$

Define the defective increment distribution $p_{k,-}$, which is supported on $k = 0, -1, -2, \ldots$, by saying that $p_{k,-}$ is the probability that the first visit to $\{\cdots, -2, -1, 0\}$ after time 0 occurs at position k and this occurs before the killing time T, i.e.,
$$p_{k,-} = \sum_{j=1}^{\infty} \mathbb{P}\{S_j = k \;;\; S_l > 0, l = 1, \ldots, j-1 \;;\; j < T\}.$$

Define $p_{k,+}$ similarly and note that $p_{k,+} = p_{k,-}$ Note that

$$\mathbb{P}(\overline{V}_+) = \mathbb{P}(V_+) + p_{0,-}\,\mathbb{P}(\overline{V}_+),$$

and hence

$$\mathbb{P}(V_+) = (1 - p_{0,-})\,\mathbb{P}(\overline{V}_+). \tag{5.9}$$

In the next proposition we prove a nonintuitive fact.

Proposition 5.2.1 *The events $\overline{V}_+$ and V_- are independent. In particular,*

$$\mathbb{P}(V_-) = \mathbb{P}(V_+) = (1 - p_{0,+})\,\mathbb{P}(\overline{V}_+) = \sqrt{p_\infty\,(1 - p_{0,+})}. \tag{5.10}$$

Proof Independence is equivalent to the statement $\mathbb{P}(V_- \cap \overline{V}_+) = \mathbb{P}(V_-)\,\mathbb{P}(\overline{V}_+)$. We will prove the equivalent statement $\mathbb{P}(V_- \cap \overline{V}^c_+) = \mathbb{P}(V_-)\,\mathbb{P}(\overline{V}^c_+)$. Note that $V_- \cap \overline{V}^c_+$ is the event that $T > 1$ but no point in $\{0, 1, 2\ldots\}$ is visited in times $\{1, \ldots, T-1\}$. In particular, at least one point in $\{\ldots, -2, -1\}$ is visited before time T.

Let

$$\rho = \max\{k \in \mathbb{Z} : S_j = k \text{ for some } j = 1, \ldots, T-1\},$$

$$\eta = \min\{j \geq 1 : S_j > 0\}, \qquad \xi_k = \max\{j \geq 0 : S_j = k; j < T\}.$$

In other words, ρ is the rightmost point visited after time zero, and ξ_k is the last time that k is visited. Then,

$$\mathbb{P}(V_- \cap \overline{V}^c_+) = \sum_{k=1}^{\infty} \mathbb{P}\{\rho = -k\} = \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \mathbb{P}\{\rho = -k; \xi_{-k} = j\}.$$

Note that the event $\{\rho = -k; \xi_{-k} = j\}$ is the same as the event

$$\{S_j = -k\ ;\ j < T\ ;\ S_l \leq -k, l = 1, \ldots, j-1\ ;\ S_l < -k, l = j+1, \ldots, T-1\}.$$

Since,

$$\mathbb{P}\{S_l < -k, l = j+1, \ldots, T-1 \mid S_j = -k\ ;\ j < T\ ;\ S_l \leq -k, l = 1, \ldots, j-1\} = \mathbb{P}(V_-),$$

we get

$$\mathbb{P}\{\rho = -k; \xi_{-k} = j\} = \mathbb{P}\{S_j = -k\ ;\ j < T\ ;\ S_l \leq -k, l = 1, \ldots, j-1\}\,\mathbb{P}(V_-),$$

We now do a path reversal using the fact that the probability of the path $[x_0 = 0, x_1, \ldots, x_j]$ is the same as the probability of $[x_j - x_j, x_{j-1} - x_j, \ldots, x_0 - x_j]$. This uses the symmetry of the random walk. Using this we can see that

$$\mathbb{P}\{S_j = -k\ ;\ j < T\ ;\ S_l \leq -k, l = 1, \ldots, j-1\} = \mathbb{P}\{\eta = j; j < T; S_j = k\}.$$

But,

$$\sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \mathbb{P}\{\eta = j; j < T; S_j = k\} = \mathbb{P}\{\eta < T\} = \mathbb{P}(\overline{V}^c_-) = \mathbb{P}(\overline{V}^c_+).$$

This establishes independence. The equality (5.10) follows from

$$p_\infty = \mathbb{P}(V_- \cap \overline{V}_+) = \mathbb{P}(V_-)\,\mathbb{P}(\overline{V}_+) = \mathbb{P}(V_-)\,\mathbb{P}(V_+)\,(1-p_{0,+}) = \frac{\mathbb{P}(V_+)^2}{(1-p_{0,+})}.$$

□

5.3 Hitting a half-line

In this section we will give an application of Proposition 5.2.1 to walks in $\mathbb{Z}^d$. Suppose $d \geq 2$ and S_n is a random walk with increment distribution $p \in \mathcal{P}_d$. Let us write $S_n = (Y_n, Z_n)$ where Y_n is a one-dimensional walk and Z_n is $(d-1)$-dimensional walk. Let Γ denote the covariance matrix for S_n and let Γ^* denote the covariance matrix for the $(d-1)$-dimensional walk Z_n. Let $T = \min\{j > 0 : Z_j = 0\}$ be the first time that the random walk returns to the line $\{(j,x) : \mathbb{Z} \times \mathbb{Z}^{d-1} : x = 0\}$. Let T_+ be the corresponding quantity for the half-line

$$T_+ = \min\left\{n > 0 : S_n \in \{(j,x) : \mathbb{Z} \times \mathbb{Z}^{d-1} : j > 0, x = 0\}\right\},$$

and let

$$p_{0,+} = \mathbb{P}\{Y_T = 0\}.$$

Proposition 5.3.1 *If $p \in \mathcal{P}_d, d = 2, 3$, there is an C such that as $n \to \infty$,*

$$\mathbb{P}\{T_+ > n\} \sim \begin{cases} C\,n^{-1/4}, & d = 2, \\ C\,(\log n)^{-1/2}, & d = 3. \end{cases}$$

Proof Let $\sigma = \sigma_\xi$ denotes a geometric random variable, independent of the random walk, with killing rate $1-\xi$, i.e., $\mathbb{P}\{\sigma > k\} = \xi^k$. Let $q_n = \mathbb{P}\{T_+ > n\}$. Then,

$$\mathbb{P}\{T_+ > \sigma\} = \sum_{n=1}^{\infty} \mathbb{P}\{\sigma = n; T_+ > n\} = \sum_{n=1}^{\infty} (1-\xi)\,\xi^{n-1}\,q_n.$$

By Propositions 11.5.2 and 11.5.3, it suffices to show that $\mathbb{P}\{T_+ > \sigma\} \sim c\,(1-\xi)^{1/4}$ if $d = 2$ and $\mathbb{P}\{T_+ > \sigma\} \sim c\,[\log(1-\xi)]^{-1/2}$ if $d = 3$. Proposition 5.2.1 tells us that

$$q(\xi) = \sqrt{\frac{p_\infty(\xi)}{1 - p_{0,+}(\xi)}},$$

where $p_\infty(\xi) = \mathbb{P}\{T > \xi\}$ and $1 - p_{0,+}(\xi) = \mathbb{P}\{T \leq \xi; Y_T \neq 0\}$. Clearly, as $\xi \to 1-$, $1 - p_{0,+}(\xi) \to 1 - p_{0,+}$. By using (4.9) and (4.10) on the random walk Z_n, we can see that

$$\mathbb{P}\{T > \xi\} \sim c\,(1-\xi)^{1/2}, \quad d = 2,$$

$$\mathbb{P}\{T > \xi\} \sim c\,\log\left(\frac{1}{1-\xi}\right), \quad d = 3.$$

□

Comment. From the proof one can see that the constant C can be determined in terms of Γ^* and $p_{0,+}$. We do not need the exact value and the proof is a little easier to follow if we do not try to keep track of this constant.

Comment. The proof uses the surprising fact that the events "avoid the positive x^1-axis" and "avoid the negative x^1- axis" are independent up to a multiplicative constant. This idea does not extend to other sets, for example the event "avoid the positive x^1-axis" and "avoid the positive x^2-axis" are not independent up to a multiplicative constant in two dimensions. However, they are in three dimensions, which is a nontrivial fact.

In Section 6.7, we will need some estimates for two-dimensional random walks avoiding a half-line. For the remainder of this section, let $d = 2$ and let $S_n = (Y_n, Z_n)$ be the random walk. Let

$$\zeta_r = \min\{n > 0 : Y_n \leq -r\},$$

$$\rho_r = \min\{n > 0 : S_n \notin (-r, r) \times (-r, r)\},$$

Proposition 5.3.2 *If $p \in \mathcal{P}_2$, then*

$$\mathbb{P}\{T_+ > \rho_r\} \asymp r^{-1/2}. \tag{5.11}$$

Moreover, for all z,

$$\mathbb{P}^z\{\xi_r < T_+\} \leq c\,|z|^{1/2}\,r^{-1/2} \tag{5.12}$$

Also, there is a $c < \infty$ such that if $A_k = \{j\mathbf{e}_1 : j = -k, -k+1, \ldots\}$, and $n \geq k$,

$$\mathbb{P}\{T_{A_k} > \xi_r\} \leq c\,k^{-1/2}\,r^{-1/2}. \tag{5.13}$$

Proof Using the functional central limit theorem, it is not difficult to show that there is a c such that for r sufficiently large, $\mathbb{P}\{\zeta_r = \rho_r < r^2\} \geq c$. However, the strong Markov property shows that

$$\mathbb{P}\{\zeta_r = \rho_r < r^2 \mid T < \rho_r\} \geq \mathbb{P}\{\zeta < r^2\}.$$

Therefore, $\mathbb{P}\{\zeta_r = \rho_r < r^2 \mid T_+ > \rho_r\} \geq c$ and

$$\mathbb{P}\{\rho_r < T_+\} \geq \mathbb{P}\{\zeta_r = \rho_r; T_+ > \rho_r\} \geq c\,\mathbb{P}\{T > \rho_r\} \geq c\,r^{-1/2}.$$

The functional central limit theorem again can be used to show that

$$\mathbb{P}\{T_+ > n^2 \mid \zeta_r < T_+\} \geq c$$

and hence,

$$\mathbb{P}\{\rho_r < T_+\} \leq c\,\mathbb{P}\{\rho_r < T_+, T_+ > n^2\} \leq \mathbb{P}\{T_+ > n/2\} \leq c\,n^{-1/2}.$$

This gives (5.11).

Let $A = \{j\mathbf{e}_1 : j = 0, 1, \ldots\}$. It suffices to prove (5.12) for $|z| \leq r/4$. Let $\sigma_r = \min\{j : |S_j - S_0| \geq r\}$. Let $h_r(z) = \mathbb{P}^z\{S[0, \sigma_r] \cap A = \emptyset\}$. Note that

$$h_{r+|z|}(z) \leq \mathbb{P}^z\{T_+ > \xi_r\} \leq h_{r-|z|}(z).$$

Since $h_r(z) \leq h_r(z-(k,0))$ if $k \geq 0$, the maximum of h_r on $(-4|z|, 4|z|) \times (-4|z|, 4|z|)$ is taken on on the set $(-4|z|, -2|z|) \times (-4|z|, 4|z|)$. Therefore,

$$\mathbb{P}^z\{T_+ > \xi_r\} \leq h_{3r/4}(z) \leq \sup h_{3r/4}(w) \leq \mathbb{P}^w\{T_+ > \xi_{r/2}\}.$$

There the supremum is taken over $w \in (-4|z|, -2|z|) \times (-4|z|, 4|z|)$. The Harnack inequality shows that for $w, w' \in (-6|z|, -2|z|) \times (-4|z|, 4|z|)$,

$$\mathbb{P}^{w'}\{T_+ > \xi_{r/2}\} \leq c\,\mathbb{P}^w\{T_+ > \xi_{r/2}\}.$$

But,

$$\begin{aligned} c_1 r^{-1/2} &\geq \mathbb{P}\{T_+ \geq \xi_r\} \\ &\geq \mathbb{P}\{\rho_{2|z|} = \zeta_{2|z|} < T_+\}\,\mathbb{P}\{T_+ \geq \xi_r \mid \rho_{2|z|} = \zeta_{2|z|} < T_+\} \\ &\geq c\mathbb{P}\{\zeta_{2|z|} < T_+\}\,\mathbb{P}^w\{T_+ > \zeta_{r/2}\} \\ &\geq c\,|z|^{-1/2}\,\mathbb{P}^w\{T_+ > \zeta_{r/2}\}. \end{aligned}$$

Therefore, $\mathbb{P}^w\{T_+ > \zeta_{r/2}\} \leq c_1 |z|^{1/2}\, r^{-1/2}$ which gives (5.12).

To get (5.13), we write

$$\mathbb{P}\{\xi_r < T_{A_k}\} = \mathbb{P}\{\xi_k < T_{A_k}\}\mathbb{P}\{\xi_r < T_{A_k} \mid \xi_k < T_{A_k}\} \leq [ck^{-1}]\,[ck^{1/2}\, r^{-1/2}].$$

$\square$

6

Potential Theory

6.1 Introduction

There is a close relationship between random walks with increment distribution p and functions that are harmonic with respect to the generator $\mathcal{L} = \mathcal{L}_p$.

We start by setting some notation. We fix $p \in \mathcal{P}$. If $A \subsetneqq \mathbb{Z}^d$, we let

$$\partial A = \{x \in \mathbb{Z}^d \setminus A : p(y,x) > 0 \text{ for some } y \in A\}$$

denote the *(outer) boundary* of A and we let $\overline{A} = A \cup \partial A$. Our defintion of $\partial A, \overline{A}$ depends on the choice of p; we hope this will not be confusing. In the case of simple random walk,

$$\partial A = \{x \in \mathbb{Z}^d \setminus A : |y - x| = 1 \text{ for some } y \in A\}.$$

Since p has finite range, if A is finite, then $\partial A, \overline{A}$ are finite. The *inner boundary* of $A \subset \mathbb{Z}^d$ is defined by

$$\partial_i A = \partial(\mathbb{Z}^d \setminus A) = \{x \in A : p(x,y) > 0 \text{ for some } y \notin A\}.$$

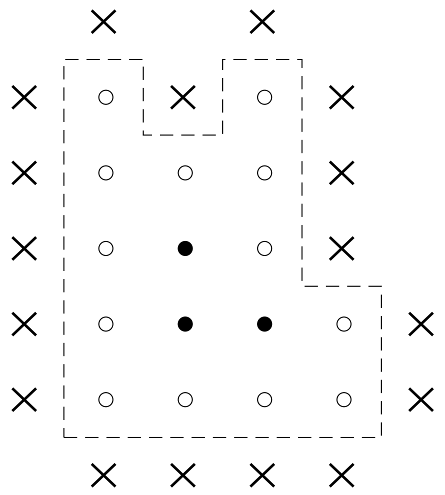

Figure 6.1: A set A: the points in $A \setminus \partial_i A$, $\partial_i A$ and ∂A are marked by $\bullet$, $\circ$ and $\times$, respectively

A function $f : \overline{A} \to \mathbb{R}$ is *harmonic (with respect to p) or p-harmonic in A* if

$$\mathcal{L}f(y) := \sum_x p(x)\,[f(y+x) - f(y)] = 0$$

for every $y \in A$. Note that we cannot define $\mathcal{L}f(y)$ for all $y \in A$ unless f is defined on $\overline{A}$.

We say that A is *connected (with respect to p)* if for every $x, y \in A$, there is a finite sequence $x = z_0, z_1, \ldots, z_k = y$ of points in A with $p(z_{j+1} - z_j) > 0, j = 0, \ldots, k$.

Comment. We will be proving a number of results about functions on subsets of $\mathbb{Z}^d$. These results have continuous analogues. The set A corresponds to an open set $D \subset \mathbb{R}^d$, the outer boundary ∂A corresponds to the usual topological boundary ∂D, and $\overline{A}$ corresponds to the closure $\overline{D} = D \cup \partial D$. The term domain is often used for open, connected subsets of $\mathbb{R}^d$. Finiteness assumptions on A correspond to boundedness assumptions on D.

Proposition 6.1.1 *Suppose S_n is a random walk with increment distribution $p \in \mathcal{P}_d$ starting at $x \in \mathbb{Z}^d$. Suppose $f : \mathbb{Z}^d \to \mathbb{R}$. Then*

$$M_n := f(S_n) - \sum_{j=0}^{n-1} \mathcal{L}f(S_j)$$

is a martingale. In particular, if f is harmonic on $A \subset \mathbb{Z}^d$, then $Y_n := f(S_{n \wedge \overline{\tau}_A})$ is a martingale, where $\overline{\tau}_A$ is as defined in (4.28).

Proof Immediate from the definition. □

Proposition 6.1.2 *Suppose $p \in \mathcal{P}_d$ and $f : \mathbb{Z}^d \to \mathbb{R}$ is bounded and harmonic on $\mathbb{Z}^d$. Then f is constant.*

Proof We may assume p is aperiodic; if not consider $\hat{p} = (1/2)\,p + (1/2)\delta_0$ and note that f is p-harmonic if and only if it is $\hat{p}$-harmonic. Let $x, y \in \mathbb{Z}^d$. By Lemma 2.4.3 we can define random walks $S, \hat{S}$ on the same probability space so that S is a random walk starting at x; $\hat{S}$ is a random walk starting at y; and

$$\mathbb{P}\{S_n \neq \hat{S}_n\} \leq c\,|x-y|\,n^{-1/2},$$

In particular,

$$|\mathbb{E}[f(S_n)] - \mathbb{E}[f(\hat{S}_n)]| \leq c\,|x-y|\,n^{-1/2}\,\|f\|_\infty \to 0.$$

Proposition 6.1.1 implies that $f(x) = \mathbb{E}[f(S_n)], f(y) = \mathbb{E}[f(\hat{S}_n)]$. □

Comment. The fact that all bounded harmonic functions are constant is closely related to the fact that a random walk eventually forgets its starting point. Lemma 2.4.3 gives a precise formulation of this forgetfulness property. The last proposition is not true for simple random walk on a regular tree.

6.2 Dirichlet problem

The standard *Dirichlet problem* for harmonic functions is to find a harmonic function on a region with specified values on the boundary.

Theorem 6.2.1 (Dirichlet problem I) *Suppose $p \in \mathcal{P}_d$ and $A \subset \mathbb{Z}^d$ satisfying $\mathbb{P}^x\{\overline{\tau}_A < \infty\} = 1$ for all $x \in A$. Suppose $F : \partial A \to \mathbb{R}$ is a bounded function. Then there is a unique bounded function $f : \overline{A} \to \mathbb{R}$ satisfying*

$$\mathcal{L}f(x) = 0, \quad x \in A, \tag{6.1}$$

$$f(x) = F(x), \quad x \in \partial A. \tag{6.2}$$

It is given by

$$f(x) = \mathbb{E}^x[F(S_{\overline{\tau}_A})]. \tag{6.3}$$

Proof A simple application of the Markov property shows that f defined by (6.3) satisfies (6.1) and (6.2). Now suppose f is a bounded function satisfying (6.1) and (6.2). Then $M_n := f(S_{n\wedge\overline{\tau}_A})$ is a bounded martingale. Hence, the optional sampling theorem (Theorem 11.2.3) implies that

$$f(x) = \mathbb{E}^x[M_0] = \mathbb{E}^x[M_{\overline{\tau}_A}] = \mathbb{E}^x[F(S_{\overline{\tau}_A})].$$

□

Remark.

- If A is finite, then ∂A is also finite and all functions on $\overline{A}$ are bounded. Hence for each F on ∂A, there is a unique function satisfying (6.1) and (6.2). In that case we could have proved this using linear algebra since (6.1) and (6.2) give $\#(A)$ linear equations in $\#(A)$ unknowns. However, algebraic methods do not yield the nice probabilistic form (6.3).
- If A is infinite, there may well be more than one solution to the Dirichlet problem if we allow unbounded solutions. For example, if $d = 1$, p is simple random walk, $A = \{1, 2, 3, \ldots\}$, and $F(0) = 0$, then there is an infinite number of solutions of the form $f_b(x) = bx$. If $b \neq 0$, f_b is unbounded.
- Under the conditions of the theorem, it follows that any function f on $\overline{A}$ that is harmonic on A satisfies the *maximum principle*:

$$\sup_{x\in\overline{A}} |f(x)| = \sup_{x\in\partial A} |f(x)|.$$

- If $d = 1, 2$ and A is a proper subset of $\mathbb{Z}^d$, then we know by recurrence that $\mathbb{P}^x\{\overline{\tau}_A < \infty\} = 1$ for all $x \in A$.
- If $d \geq 3$ and $\mathbb{Z}^d \setminus A$ is finite, then there are points $x \in A$ with $\mathbb{P}^x\{\overline{\tau}_A = \infty\} > 0$. The function

$$f(x) = \mathbb{P}^x\{\overline{\tau}_A = \infty\}$$

  is a bounded function satisfying (6.1) and (6.2) with $F \equiv 0$ on ∂A. Hence, the condition $\mathbb{P}^x\{\overline{\tau}_A < \infty\} = 1$ is needed to guarantee uniqueness. However, as Proposition 6.2.2 below shows, all solutions with $F \equiv 0$ on ∂A are multiples of f.

Remark. This theorem has a well-known continuous analogue. Suppose $f : \{|z| \in \mathbb{R}^d : |z| \leq 1\} \to \mathbb{R}$ is a continuous function with $\Delta f(x) = 0$ for $|x| < 1$. Then

$$f(x) = \mathbb{E}^x[f(B_T)],$$

where B is a standard d-dimensional Brownian motion and T is the first time t that $|B_t| = 1$. If $|x| < 1$, the distribution of B_T given $B_0 = x$ has a density with respect to surface measure on $\{|z| = 1\}$. This density $h(x,z) = c\,(1-|x|^2)/|x-z|^d$ is called the *Poisson kernel* and we can write

$$f(x) = c \int_{|z|=1} f(z)\, \frac{1-|x|^2}{|x-z|^d}\, ds(z), \tag{6.4}$$

where s dentoes surface measure. To verify that this is correct, one can check directly that f as defined above is harmonic in the ball and satisfies the boundary condition on the sphere. Two facts follow almost immediately from this integral formula:

- **Derivative estimates.** For every k, there is a $c = c(k) < \infty$ such that if f is harmonic in the unit ball and D denotes a kth order derivative, then $|Df(0)| \leq c_k \, \|f\|_\infty$.
- **Harnack inequality.** For every $r < 1$, there is a $c = c_r < \infty$ such that if f is a positive harmonic function on the unit ball, then $f(x) \leq c\, f(y)$ for $|x|, |y| \leq r$.

An important aspect of these estimates is the fact that the constants do not depend on f. We will prove the analogous results for random walk in Section 6.3.

Proposition 6.2.2 (Dirichlet problem II) *Suppose $p \in \mathcal{P}_d$ and $A \subsetneq \mathbb{Z}^d$. Suppose $F : \partial A \to \mathbb{R}$ is a bounded function. Then the only bounded functions $f : \overline{A} \to \mathbb{R}$ satisfying (6.1) and (6.2) are of the form*

$$f(x) = \mathbb{E}^x[F(S_{\overline{\tau}_A}); \overline{\tau}_A < \infty] + b\, \mathbb{P}^x\{\overline{\tau}_A = \infty\}, \tag{6.5}$$

for some $b \in \mathbb{R}$.

Proof We may assume that p is aperiodic. We also assume that $\mathbb{P}^x\{\overline{\tau}_A = \infty\} > 0$ for some $x \in A$; if not, then Theorem 6.2.1 applies. Assume that f is a bounded function satisfying (6.1) and (6.2). Since $M_n := f(S_{n \wedge \overline{\tau}_A})$ is a martingale, we know that

$$\begin{aligned} f(x) = \mathbb{E}^x[M_0] = \mathbb{E}^x[M_n] = \mathbb{E}^x[f(S_{n\wedge\overline{\tau}_A})] \\ = \mathbb{E}^x[f(S_n)] - \mathbb{E}^x[f(S_n); \overline{\tau}_A < n] + \mathbb{E}^x[F(S_{\overline{\tau}_A}); \overline{\tau}_A < n]. \end{aligned}$$

Using Lemma 2.4.3, we can see that for all x, y,

$$\lim_{n\to\infty} |\mathbb{E}^x[f(S_n)] - \mathbb{E}^y[f(S_n)]| = 0.$$

Therefore,

$$|f(x) - f(y)| \leq 2\, \|f\|_\infty \, [\mathbb{P}^x\{\overline{\tau}_A < \infty\} + \mathbb{P}^y\{\overline{\tau}_A < \infty\}].$$

Let $U_\epsilon = \{z \in \mathbb{Z}^d : \mathbb{P}^z\{\overline{\tau}_A = \infty\} \geq 1 - \epsilon\}$. Since $\mathbb{P}^x\{\overline{\tau}_A = \infty\} > 0$ for some x, one can easily check (Exercise 4.1) that U_ϵ is non-empty for each $\epsilon \in (0,1)$. Then,

$$|f(x) - f(y)| \leq 4\,\epsilon\, \|f\|_\infty, \qquad x, y \in U_\epsilon.$$

Hence, there is a b (which we will call $f(\infty)$) such that

$$|f(x) - b| \leq 4\,\epsilon\,\|f\|_\infty, \qquad x \in U_\epsilon.$$

Let ρ_ϵ be the minimum of $\overline{\tau}_A$ and the smallest n such that $S_n \in U_\epsilon$. Then for every $x \in \mathbb{Z}^d$, the optional sampling theorem implies

$$f(x) = \mathbb{E}^x[f(S_{\rho_\epsilon})] = \mathbb{E}^x[F(S_{\overline{\tau}_A}); \overline{\tau}_A \leq \rho_\epsilon] + \mathbb{E}^x[f(S_{\rho_\epsilon}); \overline{\tau}_A > \rho_\epsilon].$$

(Here we use the fact that $\mathbb{P}^x\{\tau_A \wedge \rho_\epsilon < \infty\} = 1$ which can be verified easily.) By the dominated convergence theorem,

$$\lim_{\epsilon \to 0} \mathbb{E}^x[F(S_{\overline{\tau}_A}); \overline{\tau}_A \leq \rho_\epsilon] = \mathbb{E}^x[F(S_{\overline{\tau}_A}); \overline{\tau}_A < \infty].$$

Also,

$$|\mathbb{E}^x[f(S_{\rho_\epsilon}); \overline{\tau}_A > \rho_\epsilon] - b\,\mathbb{P}^x\{\tau_A > \rho_\epsilon\}| \leq 4\epsilon\|f\|_\infty\,\mathbb{P}^x\{\tau_A > \rho_\epsilon\},$$

and since $\rho_\epsilon \to \infty$,

$$\lim_{\epsilon \to 0} \mathbb{E}^x[f(S_{\rho_\epsilon}); \overline{\tau}_A > \rho_\epsilon] = b\,\mathbb{P}^x\{\tau_A = \infty\}.$$

This gives (6.5). $\square$

Remark. We can think of (6.5) as a generalization of (6.3) where we have added a boundary point at infinity. The constant b in the last proposition is the boundary value at infinity and can be written as $F(\infty)$. The fact that there is a single boundary value at infinity is closely related to Proposition 6.1.2.

Definition. If $p \in \mathcal{P}_d$ and $A \subset \mathbb{Z}^d$, then the *Poisson kernel* is the function $H : \overline{A} \times \partial A \to [0,1]$ defined by

$$H_A(x,y) = \mathbb{P}^x\{\overline{\tau}_A < \infty; S_{\overline{\tau}_A} = y\}.$$

As a slight abuse of notation we will also write

$$H_A(x,\infty) = \mathbb{P}^x\{\overline{\tau}_A = \infty\}.$$

Note that

$$\sum_{y \in \partial A} H_A(x,y) = \mathbb{P}^x\{\overline{\tau}_A < \infty\}.$$

For fixed $y \in \partial A$, $f(x) = H_A(x,y)$ is a function on $\overline{A}$ that is harmonic on A and equals $\delta(\cdot - y)$ on ∂A. If p is recurrent, there is a unique such function. If p is transient, f is the unique such function that tends to 0 as x tends to infinity. We can write (6.3) as

$$f(x) = \mathbb{E}^x[F(S_{\overline{\tau}_A})] = \sum_{y \in \partial A} H_A(x,y)\,F(y), \tag{6.6}$$

and (6.5) as

$$f(x) = \mathbb{E}^x[F(S_{\overline{\tau}_A}); \overline{\tau}_A < \infty] + b\,\mathbb{P}^x\{\overline{\tau}_A = \infty\} = \sum_{y \in \partial A \cup \{\infty\}} H_A(x,y)\,F(y),$$

where $F(\infty) = b$. The expression (6.6) is a random walk analogue of (6.4).

Proposition 6.2.3 *Suppose $p \in \mathcal{P}_d$ and $A \subsetneqq \mathbb{Z}^d$. Let $g : A \to \mathbb{R}$ be a function with finite support. Then, the function*

$$f(x) = \sum_{y \in A} G_A(x, y)\, g(y) = \mathbb{E}^x \left[\sum_{j=0}^{\overline{\tau}_A - 1} g(S_j) \right],$$

is the unique bounded function on $\overline{A}$ that vanishes on ∂A and satisfies

$$\mathcal{L} f(x) = -g(x), \quad x \in A. \tag{6.7}$$

Proof Since g has finite support,

$$|f(x)| \leq \sum_{y \in A} G_A(x, y)\, |g(y)| < \infty,$$

and hence f is bounded. We have already noted in Lemma 4.6.1 that f satisfies (6.7). Now suppose f is a bounded function vanishing on ∂A satisfying (6.7). Then, Proposition 6.1.1 implies that

$$M_n := f(S_{n \wedge \overline{\tau}_A}) + \sum_{j=0}^{n \wedge \overline{\tau}_A - 1} g(S_j),$$

is a martingale. Note that $|M_n| \leq \|f\|_\infty + Y$ where

$$Y = \sum_{j=0}^{\overline{\tau}_A - 1} |g(S_j)|,$$

and that

$$\mathbb{E}^x[Y] = \sum_y G_A(x, y)\, |g(y)| < \infty.$$

Hence M_n is dominated by an integrable random variable and we can use the optional sampling theorem (Theorem 11.2.3) to conclude that

$$f(x) = \mathbb{E}^x[M_0] = \mathbb{E}^x[M_{\overline{\tau}_A}] = \mathbb{E}^x \left[\sum_{j=0}^{\overline{\tau}_A - 1} g(S_j) \right].$$

$\square$

Remark. Suppose $A \subset \mathbb{Z}^d$ is finite with $\#(A) = m$. Then $G_A = [G_A(x, y)]_{x,y \in A}$ is an $m \times m$ symmetric matrix with nonnegative entries. Let $\mathcal{L}^A = [\mathcal{L}^A(x, y)]_{x,y \in A}$ be the $m \times m$ symmetric matrix defined by

$$\mathcal{L}^A(x, y) = p(x, y), \;\; x \neq y; \qquad \mathcal{L}^A(x, x) = p(x, x) - 1.$$

If $g : A \to \mathbb{R}$ and $x \in A$, then $\mathcal{L}^A g(x)$ is the same as $\mathcal{L} g(x)$ where g is extended to $\overline{A}$ by setting $g \equiv 0$ on ∂A. The last proposition can be rephrased as $\mathcal{L}^A [G_A g] = -g$, or in other words, $G_A = -(\mathcal{L}^A)^{-1}$.

Corollary 6.2.4 *Suppose $p \in \mathcal{P}_d$ and $A \subset \mathbb{Z}^d$ is finite. Let $g : A \to \mathbb{R}, F : \partial A \to \mathbb{R}$ be given. Then, the function*

$$f(x) = \mathbb{E}^x[F(S_{\overline{\tau}_A})] + \mathbb{E}^x\left[\sum_{j=0}^{\overline{\tau}_A - 1} g(S_j)\right] = \sum_{z \in \partial A} H_A(x,z)\, F(z) + \sum_{y \in A} G_A(x,y)\, g(y), \tag{6.8}$$

is the unique function on $\overline{A}$ that satisfies

$$\mathcal{L}f(x) = -g(x), \quad x \in A.$$

$$f(x) = F(x), \quad x \in \partial A.$$

In other words, for any $f : \overline{A} \to \mathbb{R}, x \in \overline{A}$,

$$f(x) = \mathbb{E}^x[f(S_{\overline{\tau}_A})] - \mathbb{E}^x\left[\sum_{j=0}^{\overline{\tau}_A - 1} \mathcal{L}f(S_j)\right]. \tag{6.9}$$

Proof Use the fact that $h(x) := f(x) - \mathbb{E}^x[F(S_{\overline{\tau}_A})]$ satisfies the conditions of the previous proposition. □

Proposition 6.2.5 *Let $\rho_n = \overline{\tau}_{\mathcal{B}_n} = \inf\{j \geq 0 : |S_j| \geq n\}$. Then if $p \in \mathcal{P}_d$ with range* R *and $|x| < n$,*

$$[n^2 - |x|^2] \leq (\mathrm{tr}\Gamma)\, \mathbb{E}^x[\rho_n] \leq [(n + \mathrm{R})^2 - |x|^2].$$

Proof In Exercise 1.4 it was shown that $M_j =: |S_{j \wedge \rho_n}|^2 - (\mathrm{tr}\Gamma)(j \wedge \rho_n)$ is a martingale. Also, $\mathbb{E}^x[\rho_n] < \infty$ for each x, so M_j is dominated by the integrable random variable $(n + \mathrm{R})^2 + (\mathrm{tr}\Gamma)\, \rho_n$. Hence,

$$|x|^2 = \mathbb{E}^x[M_0] = \mathbb{E}^x[M_{\rho_n}] = \mathbb{E}^x[|S_{\rho_n}|^2] - (\mathrm{tr}\Gamma)\, \mathbb{E}^x[\rho_n].$$

However, $n \leq |S_{\rho_n}| < (n + R)$. □

Proposition 6.2.6 *Suppose $d \geq 3, p \in \mathcal{P}_d$ and $\rho_n = \overline{\tau}_{\mathcal{C}_n} = \min\{j : S_j \notin \mathcal{C}_n\} = \min\{j : \mathcal{J}(S_j) \geq n\}$. Then for $x \in \mathcal{C}_n$,*

$$\mathbb{E}^x\left[\sum_{j=0}^{\rho_n - 1} \frac{1}{1 \vee \mathcal{J}(S_j)^2}\right] = \frac{2d}{d-2}\,[\log n - \log \mathcal{J}(x)] + O(|x|^{-1}).$$

Proof Let $f(x) = \log_+[\mathcal{J}(S_j)^2]$, where $\log_+ y = 0 \vee \log y$. Then plugging into (6.9) gives

$$\mathbb{E}^x\left[\sum_{j=0}^{\overline{\tau}_A - 1} \mathcal{L}f(S_j)\right] = 2\,[\log n - \log \mathcal{J}(x)] + O(|x|^{-1}).$$

Using (1.4) and (1.5), we see that

$$\mathcal{L}f(y) = \frac{d-2}{d\mathcal{J}(y)^2} + O(|y|^{-4}),$$

and it is easy to check that

$$\mathbb{E}^x\left[\sum_{j=0}^{\overline{\tau}_A-1}\frac{1}{1\vee|S_j|^4}\right]\le O(|x|^{-1}).$$

□

6.3 Difference estimates and Harnack's inequality

In the next two sections we will prove useful results about random walk and harmonic functions. The main tools in the proofs are the optional sampling theorem and the estimates for the Green's function and the potential kernel. The basic idea in many of the proofs is to define a martingale in terms of the Green's function or potential kernel and then to stop it at a region at which that function is approximately constant. We recall that

$$\mathcal{B}_n=\{z\in\mathbb{Z}^d:|z|<n\},\quad \mathcal{C}_n=\{z\in\mathbb{Z}^d:\mathcal{J}(z)<n\}.$$

Also, there is a $\delta>0$ such that

$$\mathcal{C}_{\delta n}\subset\mathcal{B}_n\subset\mathcal{C}_{n/\delta}.$$

We set

$$\xi_n=\tau_{\mathcal{C}_n}=\min\{j\ge 1:S_j\notin\mathcal{C}_n\},\qquad \xi_n^*=\tau_{\mathcal{B}_n}=\min\{j\ge 1:S_j\notin\mathcal{B}_n\}.$$

As the next proposition points out, the Green's function and potential kernel are almost constant on $\partial\mathcal{C}_n$. We recall that Theorems 4.3.1 and 4.4.3 imply that as $x\to\infty$,

$$G(x)=\frac{C_d}{\mathcal{J}(x)^{d-2}}+O\left(\frac{1}{|x|^d}\right),\qquad d\ge 3, \tag{6.10}$$

$$a(x)=C_2\,\log\mathcal{J}(x)+\gamma_2+O\left(\frac{1}{|x|^2}\right). \tag{6.11}$$

Here $C_2=[\pi\sqrt{\det\Gamma}]^{-1}$ and $\gamma_2=C+C_2\,\log\sqrt{2}$ where C is as in Theorem 4.4.3.

Proposition 6.3.1 *If $p\in\mathcal{P}_d, d\ge 3$ then for $x\in\partial\mathcal{C}_n\cup\partial_i\mathcal{C}_n$,*

$$G(x)=\frac{C_d}{n^{d-2}}+O(n^{1-d}),\quad d\ge 2,$$

$$a(x)=C_2\,\log n+\gamma_2+O(n^{-1}),\quad d=2,$$

where C_d, C_2, γ_2 are as (6.10) and (6.11).

Proof This follows immediately from (6.10) and (6.11) and the estimate

$$\mathcal{J}(x)=n+O(1),\quad x\in\partial\mathcal{C}_n\cup\partial_i\mathcal{C}_n.$$

Note that the error term $O(n^{1-d})$ comes from the estimates

$$[n+O(1)]^{2-d}=n^{2-d}+O(n^{1-d}),\qquad \log[n+O(1)]=\log n+O(n^{-1}).$$

□

Comment. Many of the arguments in this section use $\mathcal{C}_n$ rather than $\mathcal{B}_n$ because we can then use Proposition 6.3.1. We recall that for simple random walk $\mathcal{B}_n = \mathcal{C}_n$.

Comment. Proposition 6.3.1 requires the walk to have bounded increments. If the walk does not have bounded increments, then many of the arguments in this chapter still hold. However, one needs to worry about "overshoot" estimates, i.e., giving upper bounds for the probability that the first visit of a random walk to the complement of $\mathcal{C}_n$ is far from $\mathcal{C}_n$. These kinds of estimate can be done in a spirit similar to Lemmas 5.1.7 and 5.1.8, but they complicate the arguments. For this reason, we restrict our attention to walks with bounded increments.

Comment. Proposition 6.3.1 gives the estimates for the Green's function or potential kernel on $\partial\mathcal{C}_n$. In order to prove these estimates, it suffices for the error terms in (6.10) and (6.11) to be $O(|x|^{1-d})$ rather than $O(|x|^{-d})$. For this reason, many of the ideas of this section extend to random walks with bounded increments that are not necessarily symmetric (see Theorems 4.3.5 and 4.4.5). However, in this case we would need to deal with the Green's functions for the reversed walk as well as the Green's functions for the forward walk, and this complicates the notation in the arguments. For this reason, we restrict our attention to symmetric walks.

Proposition 6.3.2 *If $p \in \mathcal{P}_d$,*

$$G_{\mathcal{C}_n}(0,0) = G(0,0) - \frac{C_d}{n^{d-2}} + O(n^{1-d}), \qquad d \geq 3,$$

$$G_{\mathcal{C}_n}(0,0) = C_2 \log n + \gamma_2 + O(n^{-1}), \quad d = 2. \tag{6.12}$$

where C_d, γ_2 are as defined in Proposition 6.3.1.

Proof Applying Lemma 4.6.2 at $x = y = 0$ gives

$$G_{\mathcal{C}_n}(0,0) = G(0,0) - \mathbb{E}[G(S_{\tau_{\mathcal{C}_n}}, 0)], \quad d \geq 3,$$

$$G_{\mathcal{C}_n}(0,0) = \mathbb{E}[a(S_{\tau_{\mathcal{C}_n}}, 0)], \quad d = 2.$$

We now apply Proposition 6.3.1. □

Comment. It follows from Proposition 6.3.2 that

$$G_{\mathcal{B}_n}(0,0) = G(0,0) - \frac{C_d}{n^{d-2}} + O(n^{1-d}), \quad d \geq 3,$$

$$a_{\mathcal{B}_n}(0,0) = C_2 \log n + O(1), \quad d = 2.$$

One can show that $a_{\mathcal{B}_n}(0,0) = C_2 \log n + \hat{\gamma}_2 + O(n^{-1})$ for a different constant $\hat{\gamma}_2$, but we will not need this and will not bother doing the argument.

We will now prove difference estimates and a Harnack inequality for harmonic functions. There are different possible approaches to proving these results. One would be to use the result for Brownian motion and approximate. We will use a different approach where we start with the known difference estimates for the Green's function G and the potential kernel a and work from there. We begin by proving a difference estimate for G_A. We then use this to prove a result on probabilities that is closely related to the gambler's ruin estimate for one-dimensional walks.

Lemma 6.3.3 *If $p \in \mathcal{P}_d, d \geq 2$, then for every $\epsilon > 0$, $r < \infty$, there is a c such that if $\mathcal{B}_{\epsilon n} \subset A \subsetneqq \mathbb{Z}^d$, then for every $|x| > \epsilon n$ and every $|y| \leq r$,*

$$|G_A(0,x) - G_A(y,x)| \leq \frac{c}{n^{d-1}}.$$

$$|2\, G_A(0,x) - G_A(y,x) - G_A(-y,x)| \leq \frac{c}{n^d}.$$

Proof It suffices to prove the result for finite A for we can approximate any A by finite sets (see Exercise 4.6). Assume that $x \in A$, for otherwise the result is trivial. By symmetry $G_A(0,x) = G_A(x,0), G_A(y,x) = G_A(x,y)$. By Lemma 4.6.2,

$$G_A(x,0) - G_A(x,y) = G(x,0) - G(x,y) - \sum_{z \in \partial A} H_A(x,z)\, [G(z,0) - G(z,y)], \quad d \geq 3,$$

$$G_A(x,y) - G_A(x,0) = a(x,0) - a(x,y) - \sum_{z \in \partial A} H_A(x,z)\, [a(z,0) - a(z,y)], \quad d = 2.$$

There are similar expressions for the second differences. The difference estimates for the Green's function and the potential kernel (Corollaries 4.3.3 and 4.4.4) give

$$|G(z) - G(z+y)| \leq c_\epsilon\, n^{1-d}, \quad |2G(z) - G(z+y) - G(z-y)| \leq c_\epsilon\, n^{-d}$$

for $d \geq 3$ and

$$|a(z) - a(z+y)| \leq c_\epsilon\, n^{-1}, \quad |2a(z) - a(z+y) - a(z-y)| \leq c_\epsilon\, n^{-2}$$

for $d = 2$, provided that $|y| \leq r$ and $|z| \geq (\epsilon/2)\, n$. □

The next lemma is very closely related to the one-dimensional gambler's ruin estimate. This lemma is particularly useful for x on or near the boundary of $\mathcal{C}_n$. For x in $\mathcal{C}_n \setminus \mathcal{C}_{n/2}$ that are away from the boundary, there are sharper estimates. See Propositions 6.4.1 and 6.4.2.

Lemma 6.3.4 *Suppose $p \in \mathcal{P}_d, d \geq 2$. There exist c_1, c_2 such that for all n sufficiently large and all $x \in \overline{\mathcal{C}_n} \setminus \mathcal{C}_{n/2}$,*

$$\mathbb{P}^x\{S_{\tau_{\mathcal{C}_n \setminus \mathcal{C}_{n/2}}} \in \mathcal{C}_{n/2}\} \geq c_1\, n^{-1}, \tag{6.13}$$

and if $x \in \partial \mathcal{C}_n$,

$$\mathbb{P}^x\{S_{\tau_{\mathcal{C}_n \setminus \mathcal{C}_{n/2}}} \in \mathcal{C}_{n/2}\} \leq c_2\, n^{-1}, \tag{6.14}$$

Proof We will do the proof for $d \geq 3$; the proof for $d = 2$ is identical replacing the Green's function with the potential kernel. It follows from (6.10) that there exist r, c such that for all n sufficiently large and all $y \in \mathcal{C}_{n-r}, z \in \partial \mathcal{C}_n$,

$$G(y) - G(z) \geq\, cn^{1-d}. \tag{6.15}$$

By choosing n sufficiently large, we can assure that $\partial_i \mathcal{C}_{n/2} \cap \mathcal{C}_{n/4} = \emptyset$.

Suppose that $x \in \mathcal{C}_{n-r}$ and let $T = \tau_{\mathcal{C}_n \setminus \mathcal{C}_{n/2}}$. Applying the optional sampling theorem to the bounded martingale $G(S_{j \wedge T})$, we see that

$$G(x) = \mathbb{E}^x[G(S_T)] \leq \mathbb{E}^x[G(S_T); S_T \in \mathcal{C}_{n/2}] + \max_{z \in \partial \mathcal{C}_n} G(z).$$

Therefore, (6.15) implies that

$$\mathbb{E}^x[G(S_T); S_T \in \mathcal{C}_{n/2}] \geq c\, n^{1-d}.$$

For n sufficiently large, $G(S_T) \notin \mathcal{C}_{n/4}$ and hence (6.10) gives

$$\mathbb{E}^x[G(S_T); S_T \in \mathcal{C}_{n/2}] \leq c\, n^{2-d}\, \mathbb{P}^x\{\tau_{\mathcal{C}_n \setminus \mathcal{C}_{n/2}} < \tau_{\mathcal{C}_n}\}.$$

This establishes (6.13) for $x \in \mathcal{C}_{n-r}$.

To prove (6.13) for other x we note the following fact that holds for any $p \in \mathcal{P}_d$: there is an $\epsilon > 0$ such that for all $|x| \geq r$, there is a y with $p(y) \geq \epsilon$ and $\mathcal{J}(x+y) \leq \mathcal{J}(x) - \epsilon$. It follows that there is a $\delta > 0$ such that for all n sufficiently large and all $x \in \mathcal{C}_n$, there is probability at least δ that a random walk starting at x reaches $\mathcal{C}_{n-r}$ before leaving $\mathcal{C}_n$.

It suffices to prove (6.14) for $x \in \mathcal{C}_n \setminus \mathcal{C}_{n-r}$. For such x,

$$G(x) = C_d\, n^{2-d} + O(n^{1-d}).$$

Also,

$$\mathbb{E}^x[G(S_T) \mid S_T \in \mathcal{C}_{n/2}] = C_d\, 2^{d-2}\, n^{2-d} + O(n^{1-d}).$$

$$\mathbb{E}^x[G(S_T) \mid S_T \in \partial\mathcal{C}_n] = C_d\, n^{2-d} + O(n^{1-d}).$$

The optional sampling theorem gives

$$G(x) = \mathbb{E}^x[G(S_T)] =$$

$$\mathbb{P}^x\{S_T \in \mathcal{C}_{n/2}\}\, \mathbb{E}^x[G(S_T) \mid S_T \in \mathcal{C}_{n/2}] + \mathbb{P}^x\{S_T \in \partial\mathcal{C}_n\}\, \mathbb{E}^x[G(S_T) \mid S_T \in \partial\mathcal{C}_n].$$

The left-hand side equals $C_d\, n^{2-d} + O(n^{1-d})$ and the right-hand side equals

$$C_d\, n^{2-d} + O(n^{1-d}) + C_d\, [2^{d-2} - 1]\, n^{2-d}\, \mathbb{P}^x\{S_T \in \mathcal{C}_{n/2}\}.$$

Therefore $\mathbb{P}^x\{S_T \in \mathcal{C}_{n/2}\} = O(n^{-1})$. □

Proposition 6.3.5 *If $p \in \mathcal{P}_d, d \geq 2$ and $x \in \mathcal{C}_n$,*

$$G_{\mathcal{C}_n}(0,x) = C_d\, \left[\mathcal{J}(x)^{2-d} - n^{2-d}\right] + O(|x|^{1-d}), \quad d \geq 3,$$

$$G_{\mathcal{C}_n}(0,x) = C_2\, [\log n - \log \mathcal{J}(x)] + O(|x|^{-1}), \quad d = 2.$$

In particular, for every $0 < \epsilon < 1/2$, there exist c_1, c_2 such that for all n sufficiently large,

$$c_1\, n^{2-d} \leq G_{\mathcal{C}_n}(y,x) \leq c_2\, n^{2-d}, \qquad y \in \mathcal{C}_{\epsilon n},\;\; x \in \partial_i \mathcal{C}_{2\epsilon n} \cup \partial\mathcal{C}_{2\epsilon n}.$$

Proof Symmetry and Lemma 4.6.2 tell us that

$$G_{\mathcal{C}_n}(0,x) = G_{\mathcal{C}_n}(x,0) = G(x,0) - \mathbb{E}^x[G(S_{\tau_{\mathcal{C}_n}})], \quad d \geq 3,$$

$$G_{\mathcal{C}_n}(0,x) = G_{\mathcal{C}_n}(x,0) = \mathbb{E}^x[a(S_{\tau_{\mathcal{C}_n}})] - a(x), \quad d = 2. \tag{6.16}$$

Also, (6.10) and (6.11) give

$$G(x) = C_d\, \mathcal{J}(x)^{2-d} + O(|x|^{-d}), \quad d \geq 3,$$

$$a(x) = C_2 \log[\mathcal{J}(x)] + \gamma_2 + O(|x|^{-2}), \quad d = 2,$$

and Proposition 6.3.1 implies that

$$\mathbb{E}^x[G(S_{\tau_{\mathcal{C}_n}})] = \frac{C_d}{n^{d-2}} + O(n^{1-d}), \quad d \geq 3,$$

$$\mathbb{E}^x[a(S_{\tau_{\mathcal{C}_n}})] = C_2 \log n + \gamma_2 + O(n^{-1}), \quad d = 2.$$

Since $|x| \leq c\, n$, we can write $O(|x|^{-d}) + O(n^{1-d}) \leq O(|x|^{1-d})$. To get the final assertion we use the estimate

$$G_{\mathcal{C}_{(1-\epsilon)n}}(0, x-y) \leq G_{\mathcal{C}_n}(y,x) \leq G_{\mathcal{C}_{(1+\epsilon)n}}(0, x-y).$$

□

We now focus on $H_{\mathcal{C}_n}$, the distribution of the first visit of a random walker to the complement of $\mathcal{C}_n$. Our first lemma uses the last-exit decomposition.

Lemma 6.3.6 *If $p \in \mathcal{P}_d$, $x \in B \subset A \subsetneq \mathbb{Z}^d$, $y \in \partial A$,*

$$H_A(x,y) = \sum_{z \in B} G_A(x,z)\, \mathbb{P}^z\{S_{\tau_{A \setminus B}} = y\} = \sum_{z \in B} G_A(z,x)\, \mathbb{P}^y\{S_{\tau_{A\setminus B}} = z\}.$$

In particular,

$$H_A(x,y) = \sum_{z \in A} G_A(x,z)\, p(z,y) = \sum_{z \in \partial_i A} G_A(x,z)\, p(z,y).$$

Proof In the first display the first equality follows immediately from a Proposition 4.6.4, and the second equality uses the symmetry of p. The second display is the particular case $B = A$. □

Lemma 6.3.7 *If $p \in \mathcal{P}_d$, there exist c_1, c_2 such that for all n sufficiently large and all $x \in \mathcal{C}_{n/4}, y \in \partial \mathcal{C}_n$,*

$$\frac{c_1}{n^{d-1}} \leq H_{\mathcal{C}_n}(x,y) \leq \frac{c_2}{n^{d-1}}.$$

Comment. We think of $\partial\mathcal{C}_n$ as a $(d-1)$-dimensional subset of $\mathbb{Z}^d$ that contains on the order of n^{d-1} points. This lemma states that the hitting measure is mutually absolutely continuous with respect to the uniform measure on $\partial\mathcal{C}_n$ (with a constant independent of n).

Proof By the previous lemma,

$$H_{\mathcal{C}_n}(x,y) = \sum_{z \in \mathcal{C}_{n/2}} G_{\mathcal{C}_n}(z,x)\, \mathbb{P}^y\{S_{\tau_{\mathcal{C}_n \setminus \mathcal{C}_{n/2}}} = z\}.$$

Using Proposition 6.3.5 we see that for $z \in \partial_i \mathcal{C}_{n/2}, x \in \mathcal{C}_{n/4}$, $G_{\mathcal{C}_n}(z,x) \asymp n^{2-d}$. Also, Lemma 6.3.4 implies that

$$\sum_{z \in \mathcal{C}_{n/2}} \mathbb{P}^y\{S_{\tau_{\mathcal{C}_n \setminus \mathcal{C}_{n/2}}} = z\} \asymp n^{-1}.$$

□

Theorem 6.3.8 (Difference estimates) *If $p \in \mathcal{P}_d$ and $r < \infty$, there exists c such that the following holds for every n sufficiently large.*

(a) If $g : \overline{\mathcal{B}}_n \to \mathbb{R}$ is harmonic in $\mathcal{B}_n$ and $|y| \le r$,

$$|\nabla_y g(0)| \le c\, \|g\|_\infty\, n^{-1}, \tag{6.17}$$

$$|\nabla_y^2 g(0)| \le c\, \|g\|_\infty\, n^{-2}. \tag{6.18}$$

(b) If $f : \overline{\mathcal{B}}_n \to [0,\infty)$ is harmonic in $\mathcal{B}_n$ and $|y| \le r$, then

$$|\nabla_y f(0)| \le c\, f(0)\, n^{-1}, \tag{6.19}$$

$$|\nabla_y^2 f(0)| \le c\, f(0)\, n^{-2}. \tag{6.20}$$

Proof Choose $\epsilon > 0$ such that $\mathcal{C}_{2\epsilon n} \subset \mathcal{B}_n$. Choose n sufficiently large so that $\mathcal{B}_r \subset \mathcal{C}_{(\epsilon/2)n}$ and $\partial_i \mathcal{C}_{2\epsilon n} \cap \mathcal{C}_{\epsilon n} = \emptyset$. Let $H(x,z) = H_{\mathcal{C}_{2\epsilon n}}(x,z)$. Then for $|x| \le r$,

$$g(x) = \sum_{z \in \partial \mathcal{C}_{2\epsilon n}} H(x,z)\, g(z),$$

and similarly for f. Hence to prove the theorem, it suffices to establish (6.19) and (6.20) for $f(x) = H(x,z)$ (with c independent of n, z). Let $\rho = \rho_{n,\epsilon} = \tau_{\mathcal{C}_{2\epsilon n} \setminus \mathcal{C}_{\epsilon n}}, G(y,x) = G_{\mathcal{C}_{2\epsilon n}}(y,x)$. By Lemma 6.3.6, if $x \in \mathcal{C}_{(\epsilon/2)n}$,

$$f(x) = \sum_{w \in \partial_i \mathcal{C}_{\epsilon n}} G_{\mathcal{C}_{\epsilon n}}(w,x)\, \mathbb{P}^z\{S_\rho = w\}.$$

Lemma 6.3.7 shows that $f(x) \asymp n^{1-d}$. There is a $\delta > 0$ such that for n sufficiently large, $|w| \ge \delta n$ for $w \in \partial_i \mathcal{C}_{\epsilon n}$. The estimates (6.19) and (6.20) now follow from Lemma 6.3.3. Note also that we have established

$$f(z) \le c\, f(w), \qquad z, w \in \mathcal{C}_{(\epsilon/2)n}. \tag{6.21}$$

$\square$

Theorem 6.3.9 (Harnack inequality) *Suppose $p \in \mathcal{P}_d$, $U \subset \mathbb{R}^d$ is open and connected, and K is a compact subset of U. Then there exist $c = c(K,U,p) < \infty$ and positive integer $N = N(K,U,p)$ such that if $n \ge N$,*

$$U_n = \{x \in \mathbb{Z}^d : n^{-1} x \in U\}, \quad K_n = \{x \in \mathbb{Z}^d : n^{-1} x \in K\},$$

and $f : \overline{U_n} \to [0,\infty)$ is harmonic in U_n, then

$$f(x) \le c\, f(y), \quad x, y \in K_n.$$

Comment. This is the discrete analogue of the Harnack principle for positive harmonic functions in $\mathbb{R}^d$. Suppose $K \subset U \subset \mathbb{R}^d$ where K is compact and U is open. Then there exists $c(K,U) < \infty$ such that if $f : U \to (0,\infty)$ is harmonic, then

$$f(x) \le c(K,U)\, f(y), \quad x, y \in K.$$

Proof Without loss of generality we will assume that U is bounded. In (6.21) we showed that there exists $\delta > 0, c_0 < \infty$ such that

$$f(x) \leq c_0\, f(y) \quad \text{if} \quad |x-y| \leq \delta \operatorname{dist}(x, \partial U_n). \tag{6.22}$$

Let us call two points z, w in U *adjacent* if $|z-w| < (\delta/4)\, \max\{\operatorname{dist}(z,\partial U), \operatorname{dist}(w, \partial U)\}$. Let ρ denote the graph distance associated to this adjacency, i.e., $\rho(z,w)$ is the minimum k such that there exists a sequence $z = z_0, z_1, \ldots, z_k = w$ of points in U such that z_j is adjacent to z_{j-1} for $j = 1, \ldots, k$. Fix $z \in U$, and let $V_k = \{w \in U : \rho(z,w) \leq k\}, V_{n,k} = \{x \in \mathbb{Z}^d : n^{-1} x \in V_k\}$. For $k \geq 1$, V_k is open, and connectedness of U implies that $\cup V_k = U$. For n sufficiently large, if $x, y \in V_{n,k}$, there is a sequence of points $x = x_0, x_1, \ldots, x_k = y$ in $V_{n,k}$ such that $|x_j - x_{j-1}| < (\delta/2)\, \max\{\operatorname{dist}(x_j, \partial U), \operatorname{dist}(x_{j-1}, \partial U)\}$. Repeated application of (6.22) gives $f(x) \leq c_0^k\, f(y)$. Compactness of K implies that $K \subset V_k$ for some finite k, and hence $K_n \subset U_{n,k}$. $\square$

6.4 Further estimates

In this section we will collect some more facts about random walks in $\mathcal{P}_d$ restricted to the set $\mathcal{C}_n$. The first three propositions are similar to Lemma 6.3.4.

Proposition 6.4.1 *If $p \in \mathcal{P}_2$, $m < n$, $T = \tau_{\mathcal{C}_n \setminus \mathcal{C}_m}$, then for $x \in \mathcal{C}_n \setminus \mathcal{C}_m$,*

$$\mathbb{P}^x\{S_T \in \partial \mathcal{C}_n\} = \frac{\log \mathcal{J}(x) - \log m + O(m^{-1})}{\log n - \log m}.$$

Proof Let $q = \mathbb{P}^x\{S_T \in \partial\mathcal{C}_n\}$. The optional sampling theorem applied to the bounded martingale $M_j = a(S_{j\wedge T})$ gives

$$a(x) = \mathbb{E}^x[a(S_T)] = (1-q)\, \mathbb{E}^x[a(S_T) \mid S_T \in \partial_i \mathcal{C}_m] + q\, \mathbb{E}^x[a(S_T) \mid S_T \in \partial \mathcal{C}_n].$$

From (6.11) we know that

$$a(x) = C_2\, \log \mathcal{J}(x) + \gamma_2 + O(|x|^{-2}),$$

$$\mathbb{E}^x[a(S_T) \mid S_T \in \partial_i \mathcal{C}_m] = C_2\, \log m + \gamma_2 + O(m^{-1}),$$

$$\mathbb{E}^x[a(S_T) \mid S_T \in \partial \mathcal{C}_n] = C_2\, \log n + \gamma_2 + O(n^{-1}).$$

The last two equalities use the fact that $\mathcal{J}(z) = m + O(1), z \in \partial_i \mathcal{C}_m$ and $\mathcal{J}(w) = n + O(1), w \in \partial \mathcal{C}_n$. Solving for q gives the result. $\square$

Proposition 6.4.2 *If $p \in \mathcal{P}_d, d \geq 3$, $T = \tau_{\mathbb{Z}^d \setminus \mathcal{C}_m}$, then for $x \in \mathbb{Z}^d \setminus \mathcal{C}_m$,*

$$\mathbb{P}^x\{T < \infty\} = \left(\frac{m}{\mathcal{J}(x)}\right)^{d-2} \left[1 + O(m^{-1})\right].$$

Proof Since $G(y)$ is a bounded harmonic function on $\tau_{\mathbb{Z}^d \setminus \mathcal{C}_m}$ with $G(\infty) = 0$, (6.5) gives

$$G(x) = \mathbb{E}^x[G(S_T); T < \infty] = \mathbb{P}^x\{T < \infty\}\, \mathbb{E}^x[G(S_T) \mid T < \infty].$$

But (6.10) gives

$$G(x) = C_d\, \mathcal{J}(x)^{2-d}\, [1 + O(|x|^{-2})],$$

$$\mathbb{E}^x[G(S_T) \mid T < \infty] = C_d\, m^{2-d}\, [1 + O(m^{-1})].$$

□

Proposition 6.4.3 *If $p \in \mathcal{P}_2$, $n > 0$, and $T = \tau_{\mathcal{C}_n \setminus \{0\}}$, then for $x \in \mathcal{C}_n$,*

$$\mathbb{P}^x\{S_T = 0\} = \left[1 - \frac{\log \mathcal{J}(x)}{\log n}\right] \left[1 + O\left(\frac{1}{\log n}\right)\right].$$

Proof Recall that $\mathbb{P}^x\{S_t = 0\} = G_{\mathcal{C}_n}(x, 0)/G_{\mathcal{C}_n}(0, 0)$. The estimate then follows immediately from Propositions 6.3.2 and 6.3.5. □

Suppose $m \leq n/2, x \in \mathcal{C}_m, z \in \partial\mathcal{C}_n$. By applying Theorem 6.3.8 $O(m)$ times we can see that (for n sufficiently large)

$$H_{\mathcal{C}_n}(x, z) = \mathbb{P}^x\{S_{\xi_n} = z\} = H_{\mathcal{C}_n}(0, z)\, \left[1 + O\left(\frac{m}{n}\right)\right]. \tag{6.23}$$

We will use this in the next two propositions to estimate some conditional probabilities.

Proposition 6.4.4 *Suppose $p \in \mathcal{P}_d, d \geq 3$, $m < n/4$, and $\mathcal{C}_n \setminus \mathcal{C}_m \subset A \subset \mathcal{C}_n$. Suppose $x \in \mathcal{C}_{2m}$ with $\mathbb{P}^x\{S_{\tau_A} \in \partial\mathcal{C}_n\} > 0$ and $z \in \partial\mathcal{C}_n$. Then for n sufficiently large,*

$$\mathbb{P}^x\{S_{\tau_A} = z \mid S_{\tau_A} \in \partial\mathcal{C}_n\} = H_{\mathcal{C}_n}(0, z)\, \left[1 + O\left(\frac{m}{n}\right)\right]. \tag{6.24}$$

Proof It is easy to check that it suffices to verify (6.24) for $x \in \partial\mathcal{C}_{2m}$. Note that (6.23) gives

$$\mathbb{P}^x\{S_{\xi_n} = z\} = H_{\mathcal{C}_n}(0, z)\, \left[1 + O\left(\frac{m}{n}\right)\right],$$

and since $\partial A \setminus \partial\mathcal{C}_n \subset \mathcal{C}_m$,

$$\mathbb{P}^x\{S_{\xi_n} = z \mid S_{\tau_A} \notin \partial\mathcal{C}_n\} = H_{\mathcal{C}_n}(0, z)\, \left[1 + O\left(\frac{m}{n}\right)\right].$$

This gives

$$\mathbb{P}^x\{S_{\xi_n} = z; S_{\tau_A} \notin \partial\mathcal{C}_n\} = \mathbb{P}\{S_{\tau_A} \notin \partial\mathcal{C}_n\}\, H_{\mathcal{C}_n}(0, z)\, \left[1 + O\left(\frac{m}{n}\right)\right],$$

and hence

$$\mathbb{P}^x\{S_{\xi_n} = z; S_{\tau_A} \in \partial\mathcal{C}_n\} = \mathbb{P}\{S_{\tau_A} \in \partial\mathcal{C}_n\}\, H_{\mathcal{C}_n}(0, z) + H_{\mathcal{C}_n}(0, z)\, O\left(\frac{m}{n}\right).$$

Using Proposition 6.4.2, we can see there is a c such that

$$\mathbb{P}^x\{S_{\tau_A} \in \partial\mathcal{C}_n\} \geq \mathbb{P}^x\{S_j \notin \mathcal{C}_m \text{ for all } j\} \geq c, \quad x \in \partial\mathcal{C}_{2m},$$

which allows use to write the preceding expression as

$$\mathbb{P}^x\{S_{\xi_n} = z; S_{\tau_A} \in \partial\mathcal{C}_n\} = \mathbb{P}\{S_{\tau_A} \in \partial\mathcal{C}_n\}\, H_{\mathcal{C}_n}(0, z)\, \left[1 + O\left(\frac{m}{n}\right)\right].$$

□

For $d = 2$ we get a similar result but with a slightly larger error term.

Proposition 6.4.5 *Suppose $p \in \mathcal{P}_2$, $m < n/4$, and $\mathcal{C}_n \setminus \mathcal{C}_m \subset A \subset \mathcal{C}_n$. Suppose $x \in \mathcal{C}_{2m}$ with $\mathbb{P}^x\{S_{\tau_A} \in \partial\mathcal{C}_n\} > 0$ and $z \in \partial\mathcal{C}_n$. Then, for n sufficiently large,*

$$\mathbb{P}^x\{S_{\tau_A} = z \mid S_{\tau_A} \in \partial\mathcal{C}_n\} = H_{\mathcal{C}_n}(0, z)\ \left[1 + O\left(\frac{m\ \log(n/m)}{n}\right)\right]. \tag{6.25}$$

Proof The proof is essentially the same, except for the last step where Proposition 6.4.1 gives us

$$\mathbb{P}^x\{S_{\tau_A} \in \partial\mathcal{C}_n\} \geq \frac{c}{\log(n/m)}, \quad x \in \mathcal{C}_{2m},$$

so that

$$H_{\mathcal{C}_n}(0, z)\, O\left(\frac{m}{n}\right)$$

can be written as

$$\mathbb{P}^x\{S_{\tau_A} \in \partial\mathcal{C}_n\}\, H_{\mathcal{C}_n}(0, z)\, O\left(\frac{m\ \log(n/m)}{n}\right).$$

□

The next proposition is a stronger version of Proposition 6.2.2. Here we show that the boundedness assumption of that proposition can be replaced with an assumption of sublinearity.

Proposition 6.4.6 *Suppose $p \in \mathcal{P}_d, d \geq 3$ and $A \subset \mathbb{Z}^d$ with $\mathbb{Z}^d \setminus A$ finite. Suppose $f : \mathbb{Z}^d \to \mathbb{R}$ is harmonic on A and satisfies $f(x) = o(|x|)$ as $x \to \infty$. Then there exists $b \in \mathbb{R}$ such that for all x,*

$$f(x) = \mathbb{E}^x[f(S_{\overline{\tau}_A}); \overline{\tau}_A < \infty] + b\, \mathbb{P}^x\{\overline{\tau}_A = \infty\}.$$

Proof Without loss of generality, we may assume that $0 \notin A$. Also, we may assume that $f \equiv 0$ on $\mathbb{Z}^d \setminus A$; otherwise, we can consider

$$\hat{f}(x) = f(x) - \mathbb{E}^x[f(S_{\overline{\tau}_A}); \overline{\tau}_A < \infty].$$

The assumptions imply that there is a sequence of real numbers ϵ_n decreasing to 0 such that $|f(x)| \leq \epsilon_n\, n$ for all $x \in \overline{\mathcal{C}_n}$ and hence

$$|f(x) - f(y)| \leq\ 2\, \epsilon_n\, n, \quad x, y \in \partial\mathcal{C}_n.$$

Since $\mathcal{L}f \equiv 0$ on A, (6.8) gives

$$0 = f(0) = \mathbb{E}[f(S_{\xi_n})] - \sum_{y \in \mathbb{Z}^d \setminus A} G_{\mathcal{C}_n}(0, y)\, \mathcal{L}f(y), \tag{6.26}$$

which implies that

$$\lim_{n\to\infty} \mathbb{E}[f(S_{\xi_n})] = b := \sum_{y \in \mathbb{Z}^d \setminus A} G(0, y)\, \mathcal{L}f(y).$$

If $x \in A \cap \mathcal{C}_n$, the optional sampling theorem implies that

$$f(x) = \mathbb{E}^x[f(S_{\tau_A \wedge \xi_n})] = \mathbb{E}^x[f(S_{\xi_n}); \tau_A > \xi_n] = \mathbb{P}^x\{\tau_A > \xi_n\}\, \mathbb{E}^x[f(S_{\xi_n}) \mid \tau_A > \xi_n].$$

We can write

$$\mathbb{E}^x[f(S_{\xi_n}) \mid \tau_A > \xi_n] - \mathbb{E}[f(S_{\xi_n})] = \sum_{z \in \partial \mathcal{C}_n} f(z)\, [\mathbb{P}^x\{S_{\xi_n} = z \mid \tau_A > \xi_n\} - H_{\mathcal{C}_n}(0,z)].$$

For n large, (6.24) then implies that

$$|\mathbb{E}^x[f(S_{\xi_n}) \mid \tau_A > \xi_n] - \mathbb{E}[f(S_{\xi_n})]| \leq c\, \frac{|x|}{n} \sup_{y,z \in \partial \mathcal{C}_n} |f(z) - f(y)| \leq c\, |x|\, \epsilon_n. \tag{6.27}$$

Therefore,

$$\begin{aligned} f(x) &= \lim_{n\to\infty} \mathbb{P}^x\{\tau_A > \xi_n\}\, \mathbb{E}^x[f(S_{\xi_n}) \mid \tau_A > \xi_n] \\ &= \mathbb{P}^x\{\tau_A = \infty\} \lim_{n\to\infty} \mathbb{E}[f(S_{\xi_n})] = b\, \mathbb{P}^x\{\tau_A = \infty\}. \end{aligned}$$

□

Proposition 6.4.7 *Suppose $p \in \mathcal{P}_2$ and A is a finite subset of $\mathbb{Z}^2$ containing the origin. Let $T = T_A = \overline{\tau}_{\mathbb{Z}^2 \setminus A} = \min\{j \geq 0 : S_j \in A\}$. Then for each $x \in \mathbb{Z}^2$ the limit*

$$g_A(x) := \lim_{n\to\infty} C_2\, (\log n)\, \mathbb{P}^x\{\xi_n < T\} \tag{6.28}$$

exists. Moreover, if $y \in A$,

$$g_A(x) = a(x-y) - \mathbb{E}^x[a(S_T - y)]. \tag{6.29}$$

Proof If $y \in A$ and $x \in \mathcal{C}_n \setminus A$, the optional sampling theorem applied to the bounded martingale $M_j = a(S_{j\wedge T\wedge \xi_n} - y)$ implies

$$\begin{aligned} a(x-y) = \mathbb{E}^x[a(S_{T\wedge\xi_n} - y)] &= \mathbb{P}^x\{\xi_n < T\}\, \mathbb{E}^x[a(S_{\xi_n} - y) \mid \xi_n < T] + \mathbb{E}^x[a(S_T - y)] \\ &\quad - \mathbb{P}^x\{\xi_n < T\}\, \mathbb{E}^x[a(S_T - y) \mid \xi_n < T]. \end{aligned}$$

As $n \to \infty$,

$$\mathbb{E}^x[a(S_{\xi_n} - y) \mid \xi_n < T] \sim C_2 \log n.$$

Letting $n \to \infty$, we get the result. □

Remark. As mentioned before, it follows that the right-hand side of (6.29) is the same for all $y \in A$. Also, since there exists δ such that $\mathcal{C}_{\delta n} \subset \mathcal{B}_n \subset \mathcal{C}_{n/\delta}$ we can replace (6.28) with

$$g_A(x) := \lim_{n\to\infty} C_2\, (\log n)\, \mathbb{P}^x\{\xi_n^* < T\}.$$

The astute reader will note now that we already proved this Proposition in Proposition 4.6.3.

Proposition 6.4.8 *Suppose $p \in \mathcal{P}_2$ and A is a finite subset of $\mathbb{Z}^2$. Suppose $f : \mathbb{Z}^2 \to \mathbb{R}$ is harmonic on $\mathbb{Z}^2 \setminus A$; vanishes on A; and satisfies $f(x) = o(|x|)$ as $|x| \to \infty$. Then $f = b\, g_A$ for some $b \in \mathbb{R}$.*

Proof Without loss of generality, assume $0 \in A$ and let $T = T_A$ be as in the previous proposition. Using (6.8) and (6.12), we get

$$\mathbb{E}[f(S_{\mathcal{C}_n})] = \sum_{y \in A} G_{\mathcal{B}_n}(0,y)\, \mathcal{L}f(y) = C_2 \log n \sum_{y\in A} \mathcal{L}f(y) + O(1).$$

(Here and below the error terms may depend on A.) Similarly to (6.27) we use (6.25) to see that

$$|\mathbb{E}^x[f(S_{T\wedge\xi_n}) \mid \xi_n < T] - \mathbb{E}[f(S_{\xi_n})]| \leq c\,\frac{|x|\,\log n}{n} \sup_{y,z\in\partial C_n} |f(y)-f(z)| \leq c\,|x|\,\epsilon_n\,\log n.$$

$$\begin{aligned} f(x) = \mathbb{E}^x[f(S_{T\wedge\xi_n})] &= \mathbb{P}^x\{\xi_n < T\}\,\mathbb{E}[^x[f(S_{T_A\wedge\xi_n}) \mid \xi_n < T] \\ &= \mathbb{P}^x\{\xi_n < T\}\,\mathbb{E}[f(S_{\xi_n})] + |x|\,o(1) \\ &= b\,g_A(x) + o(1), \end{aligned}$$

where

$$b = \sum_{y\in A} \mathcal{L}f(y).$$

□

6.5 Capacity, transient case

If A is a finite subset of $\mathbb{Z}^d$, we let

$$T_A = \tau_{\mathbb{Z}^d\setminus A}, \quad \overline{T}_A = \overline{\tau}_{\mathbb{Z}^d\setminus A},$$

$$\mathrm{rad}(A) = \sup\{|x| : x \in A\}.$$

If $p \in \mathcal{P}_d, d \geq 3$, let

$$\mathrm{Es}_A(x) = \mathbb{P}^x\{T_A = \infty\}, \qquad g_A(x) = \mathbb{P}^x\{\overline{T}_A = \infty\}.$$

Note that $\mathrm{Es}_A(x) = 0$ if $x \in A\setminus\partial_i A$. If $d \geq 3$, then g_A is the unique function on $\mathbb{Z}^d$ that is zero on A; harmonic on $\mathbb{Z}^d\setminus A$; and satisfies $g_A(x) \sim 1$ as $|x| \to \infty$. Moreover, if $x \in A$,

$$\mathcal{L}g_A(x) = \sum_y p(y)\,g_A(x+y) = \mathrm{Es}_A(x).$$

Definition. If $d \geq 3$, the *capacity* of a finite set A is given by

$$\mathrm{cap}(A) = \sum_{x\in A} \mathrm{Es}_A(x) = \sum_{z\in\partial_i A} \mathrm{Es}_A(z) = \sum_{x\in A} \mathcal{L}g_A(x) = \sum_{z\in\partial_i A} \mathcal{L}g_A(z).$$

Comment. This definition may seem unmotivated. The next proposition shows that as $z \to \infty$, $n^{2-d}\,\mathrm{cap}(A)$ is asymptotic to a multiple of the probability that a random walk starting at z ever hits A.

Proposition 6.5.1 *If $p \in \mathcal{P}_d, d \geq 3$ and $A \subset \mathbb{Z}^d$ is finite, then*

$$\mathbb{P}^x\{T_A < \infty\} = \frac{C_d\,\mathrm{cap}(A)}{\mathcal{J}(x)^{d-2}}\left[1 + O\left(\frac{\mathrm{rad}(A)}{|x|}\right)\right], \qquad x \geq 2\,\mathrm{rad}(A).$$

Proof There is a δ such that $\mathcal{B}_n \subset \mathcal{C}_{n/\delta}$ for all n. We will first prove the result for $x \notin \mathcal{C}_{2\mathrm{rad}(A)/\delta}$. By the last-exit decomposition, Proposition 4.6.4,

$$\mathbb{P}^x\{\overline{T}_A < \infty\} = \sum_{y \in A} G(x,y)\,\mathrm{Es}_A(y).$$

For $y \in A$, $\mathcal{J}(x-y) = \mathcal{J}(x) + O(|y|)$. Therefore,

$$G(x,y) = C_d\,\mathcal{J}(x)^{2-d} + O\left(\frac{|y|}{|x|^{d-1}}\right) = \frac{C_d}{\mathcal{J}(x)^{d-2}}\left[1 + O\left(\frac{\mathrm{rad}(A)}{|x|}\right)\right].$$

This gives the result for $x \notin \mathcal{C}_{2\mathrm{rad}(A)/\delta}$. We can extend this to $x \geq 2\mathrm{rad}(A)$ by using the Harnack inequality (Theorem 6.3.9) on the set

$$\{z : 2\,\mathrm{rad}(A) \leq |z|; \mathcal{J}(z) \geq (3/\delta)\,\mathrm{rad}(A)\}.$$

Note that for x in this set, $\mathrm{rad}(A)/|x|$ is of order 1, so we only need to show that there is a c such that

$$\mathbb{P}^x\{\overline{T}_A < \infty\} \leq c\,\mathbb{P}^z\{\overline{T}_A < \infty\}$$

for x, z in this set. $\square$

Proposition 6.5.2 *If $p \in \mathcal{P}_d, d \geq 3$,*

$$\mathrm{cap}(\mathcal{C}_n) = C_d^{-1}\,n^{d-2} + O(n^{d-1}).$$

Proof By Proposition 4.6.4,

$$1 = \mathbb{P}\{\overline{T}_{\mathcal{C}_n} < \infty\} = \sum_{y \in \partial_i \mathcal{C}_n} G(0,y)\,\mathrm{Es}_{\mathcal{C}_n}(y),$$

But for $y \in \partial_i \mathcal{C}_n$, Proposition 6.3.1 gives

$$G(0,y) = C_d\,n^{2-d}\,[1 + O(n^{-1})].$$

Hence,

$$1 = C_d\,n^{2-d}\,\mathrm{cap}(\mathcal{C}_n)\,\left[1 + O(n^{-1})\right].$$

$\square$

Let

$$T_{A,n} = T_A \wedge \xi_n = \inf\{j \geq 1 : S_j \in A \text{ or } S_j \notin \mathcal{C}_n\}.$$

If $x \in A \subset \mathcal{C}_n$

$$\mathbb{P}^x\{T_A > \xi_n\} = \sum_{y \in \partial \mathcal{C}_n} \mathbb{P}^x\{S_{T_{A,n}} = y\} = \sum_{y \in \partial \mathcal{C}_n} \mathbb{P}^y\{S_{T_{A,n}} = x\}.$$

The last equation uses symmetry of the walk. Therefore,

$$\sum_{x \in A} \mathbb{P}^x\{T_A > \xi_n\} = \sum_{x \in A}\sum_{y \in \partial \mathcal{C}_n} \mathbb{P}^y\{S_{T_{A,n}} = x\} = \sum_{y \in \partial \mathcal{C}_n} \mathbb{P}^y\{T_A < \xi_n\}.$$

Also,

$$\operatorname{cap}(A) = \lim_{n\to\infty} \sum_{x\in A} \mathbb{P}^x\{T_A > \xi_n\} = \lim_{n\to\infty} \sum_{y\in\partial\mathcal{C}_n} \mathbb{P}^y\{T_A < \xi_n\}. \tag{6.30}$$

Comment. The last equation shows the relationship between the probability that a random walker on a set A "escapes" A to the probability that a random walker far away from A ever hits A. Basically every path from A to infinity can also be considered as a path from infinity to A by reversal. Another version of this relationship is given in Proposition 6.5.4.

Proposition 6.5.3 *If $p \in \mathcal{P}_d, d \geq 3,$ and A, B are finite subsets of $\mathbb{Z}^d$, then*

$$\operatorname{cap}(A\cup B) \leq \operatorname{cap}(A) + \operatorname{cap}(B) - \operatorname{cap}(A\cap B).$$

Proof Choose n such that $A \cup B \subset \mathcal{C}_n$. Then for $y \in \mathcal{C}_n$,

$$\begin{aligned} \mathbb{P}^y\{T_{A\cup B} < \xi_n\} &= \mathbb{P}^x\{T_A < \xi_n \text{ or } T_B < \xi_n\} \\ &= \mathbb{P}^y\{T_A < \xi_n\} + \mathbb{P}^y\{T_B < \xi_n\} - \mathbb{P}^y\{T_A < \xi_n, T_B < \xi_n\} \\ &\leq \mathbb{P}^y\{T_A < \xi_n\} + \mathbb{P}^y\{T_B < \xi_n\} - \mathbb{P}^y\{T_{A\cap B} < \xi_n\}. \end{aligned}$$

The proposition the follows from (6.30). □

Definition. If $p \in \mathcal{P}_d, d \geq 3$, and $A \subset \mathbb{Z}^d$ is finite, the *harmonic measure of A (from infinity)* is defined by

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

Note that hm_A is a probability measure supported on $\partial_i A$. As the next proposition shows, it can be considered as the hitting measure of A by a random walk "started at infinity conditioned to hit A".

Proposition 6.5.4 *If $p \in \mathcal{P}_d, d \geq 3,$ and $A \subset \mathbb{Z}^d$ is finite, then for $x \in A$,*

$$\operatorname{hm}_A(x) = \lim_{|y|\to\infty} \mathbb{P}^y\{S_{T_A} = x \mid T_A < \infty\}.$$

In fact, if $A \subset \mathcal{C}_{n/2}$ and $y \notin \mathcal{C}_n$, then

$$\mathbb{P}^y\{S_{T_A} = x \mid T_A < \infty\} = \operatorname{hm}_A(x) \left[1 + O\left(\frac{\operatorname{rad}(A)}{|y|}\right)\right]. \tag{6.31}$$

Proof If $A \subset \mathcal{C}_n$ and $y \notin \mathcal{C}_n$, the last-exit decomposition (Proposition 4.6.4) gives

$$\mathbb{P}^y\{S_{T_A} = x\} = \sum_{z\in\partial\mathcal{C}_n} G_{\mathbb{Z}^d\setminus A}(y,z)\, \mathbb{P}^z\{S_{T_{A,n}} = x\},$$

where, as before, $T_{A,n} = T_A \wedge \xi_n$. By symmetry and (6.24),

$$\begin{aligned} \mathbb{P}^z\{S_{T_{A,n}} = x\} = \mathbb{P}^x\{S_{T_{A,n}} = z\} &= \mathbb{P}^x\{\xi_n < T_A\}\, H_{\mathcal{C}_n}(0,z) \left[1 + O\left(\frac{\operatorname{rad}(A)}{n}\right)\right] \\ &= \operatorname{Es}_A(x)\, H_{\mathcal{C}_n}(0,z) \left[1 + O\left(\frac{\operatorname{rad}(A)}{n}\right)\right]. \end{aligned}$$

The last equality uses

$$\mathrm{Es}_A(x) = \mathbb{P}^x\{T_A = \infty\} = \mathbb{P}^x\{T_A > \xi_n\}\left[1 + O\left(\frac{\mathrm{rad}(A)^{d-2}}{n^{d-2}}\right)\right],$$

which follows from Proposition 6.4.2. Therefore,

$$\mathbb{P}^y\{S_{T_A} = x\} = \mathrm{Es}_A(x)\left[1 + O\left(\frac{\mathrm{rad}(A)}{n}\right)\right]\sum_{z\in\partial\mathcal{C}_n} G_{\mathbb{Z}^d\setminus A}(y,z)\, H_{\mathcal{C}_n}(0,z),$$

and by summing over x,

$$\mathbb{P}^y\{T_A < \infty\} = \mathrm{cap}(A)\left[1 + O\left(\frac{\mathrm{rad}(A)}{n}\right)\right]\sum_{z\in\partial\mathcal{C}_n} G_{\mathbb{Z}^d\setminus A}(y,z)\, H_{\mathcal{C}_n}(0,z).$$

By dividing the last two expressions, we get (6.31). □

Proposition 6.5.5 *If $p \in \mathcal{P}_d, d \geq 3$, and $A \subset \mathbb{Z}^d$ is finite, then*

$$\mathrm{cap}(A) = \sup \sum_{x\in A} f(x), \tag{6.32}$$

where the sum is over all functions $f \geq 0$ supported on A such that

$$Gf(y) := \sum_{x\in\mathbb{Z}^d} G(y,x)\, f(x) \leq 1$$

for all $y \in \mathbb{Z}^d$.

Proof Let $\hat{f}(x) = \mathrm{Es}_A(x)$. Note that Proposition 4.6.4 implies that for $y \in \mathbb{Z}^d$,

$$1 \geq \mathbb{P}^y\{\overline{T}_A < \infty\} = \sum_{x\in A} G(y,x)\, \mathrm{Es}_A(x).$$

Hence $G\hat{f} \leq 1$ and the supremum in (6.32) is at least as large as $\mathrm{cap}(A)$. Note also that $G\hat{f}$ is the unique bounded function on $\mathbb{Z}^d$ that is harmonic on $\mathbb{Z}^d \setminus A$; equals 1 on A; and approaches 0 at infinity. Suppose $f \geq 0$ with $Gf(y) \leq 1$ for all $y \in \mathbb{Z}^d$. Then Gf is the unique bounded function on $\mathbb{Z}^d$ that is harmonic on $\mathbb{Z}^d \setminus A$; equals $Gf \leq 1$ on A; and approaches zero at infinity. By the maximum principle, $Gf(y) \leq G\hat{f}(y)$ for all y. In particular, $G(\hat{f} - f)$ is harmonic on $\mathbb{Z}^d \setminus A$; is nonnegative on A; and approaches zero at infinity. We need to show that

$$\sum_{x\in A} f(x) \leq \sum_{x\in A} \hat{f}(x).$$

If $x, y \in A$, let

$$K_A(x,y) = \mathbb{P}^x\{S_{T_A} = y\}.$$

Note that $K_A(x,y) = K_A(y,x)$ and

$$\sum_{y\in A} K_A(x,y) = 1 - \mathrm{Es}_A(x).$$

If h is a bounded function on $\mathbb{Z}^d$ that is harmonic on $\mathbb{Z}^d \setminus A$ and has $h(\infty) = 0$, then $h(z) = \mathbb{E}[h(S_{\overline{T}_A}); \overline{T}_A < \infty]$. Using this one can easily check that for $x \in A$,

$$\mathcal{L}h(x) = \left[\sum_{y \in A} K_A(x,y)\, h(y) \right] - h(x).$$

Also, if $h \geq 0$,

$$\sum_{x \in A} \sum_{y \in A} K_A(x,y)\, h(y) = \sum_{y \in A} h(y) \sum_{x \in A} K_A(y,x) = \sum_{y \in A} h(y)\, [1 - \mathrm{Es}_A(y)] \leq \sum_{y \in A} h(y),$$

which implies

$$\sum_{x \in \mathbb{Z}^d} \mathcal{L}h(x) = \sum_{x \in A} \mathcal{L}h(x) \leq 0.$$

Then, using (4.26),

$$\sum_{x \in A} f(x) = -\sum_{x \in A} \mathcal{L}[Gf](x) \leq -\sum_{x \in A} \mathcal{L}[Gf](x) - \sum_{x \in A} \mathcal{L}[G(\hat{f} - f)](x) = \sum_{x \in A} \mathrm{Es}_A(x).$$

$\square$

Our defintion of capacity depends on the random walk p. The next proposition shows that capacities for different p's in the same dimension are comparable.

Proposition 6.5.6 *Suppose $p, q \in \mathcal{P}_d, d \geq 3$ and let $\mathrm{cap}_p, \mathrm{cap}_q$ denote the corresponding capacities. Then there is a $\delta = \delta(p,q) > 0$ such that for all finite $A \subset \mathbb{Z}^d$,*

$$\delta\, \mathrm{cap}_p(A) \leq \mathrm{cap}_q(A) \leq \delta^{-1}\, \mathrm{cap}_p(A).$$

Proof It follows from Theorem 4.3.1 that there exists δ such that

$$\delta\, G_p(x,y) \leq G_q(x,y) \leq \delta^{-1}\, G_p(x,y),$$

for all x, y. The proposition then follows from Proposition 6.5.5. $\square$

Definition. If $p \in \mathcal{P}_d, d \geq 3$, and $A \subset \mathbb{Z}^d$, then A is *transient* if

$$\mathbb{P}\{S_n \in A \text{ i.o.}\} = 0.$$

Otherwise, the set is called *recurrent.*

Lemma 6.5.7 *If $p \in \mathcal{P}_d, d \geq 3$, then a subset A of $\mathbb{Z}^d$ is recurrent if and only if for every $x \in \mathbb{Z}^d$,*

$$\mathbb{P}^x\{S_n \in A \ i.o.\} = 1.$$

Proof Let $F(y) = \mathbb{P}^y\{S_n \in A \text{ i.o.}\}$. Suppose there exists $y \in A$ and $\epsilon > 0$ such that $F(y) \geq \epsilon$. Then,

$$\mathbb{E}^y[F(S_n)] \geq \epsilon.$$

Then by coupling (Lemma 2.4.3) we can see that for all x, $F(x) = \mathbb{E}^x[F(S_n)] = \epsilon + |x|\, O(n^{-1/2})$, and hence $F(x) \geq \epsilon$. In particular, for every x there is an N_x such that

$$\mathbb{P}^x\{S_n \in A \text{ for some } n \leq N_x\} \geq \epsilon/2.$$

By iterating this we can see for all x,

$$\mathbb{P}^x\{S_n \in A \text{ for some } n < \infty\} = 1,$$

and the lemma follows easily. □

Clearly, all finite sets are transient; in fact, finite unions of transient sets are transient. If A is a subset such that

$$\sum_{x \in A} G(x) < \infty, \tag{6.33}$$

then A is transient. To see this, let S_n be a random walk starting at the origin and let V denote the number of visits to A,

$$V_A = \sum_{j=0}^{\infty} 1\{S_n \in A\}.$$

Then (6.33) implies that $\mathbb{E}[V_A] < \infty$ which implies that $\mathbb{P}\{V_A < \infty\} = 1$. In Exercise 6.2, it is shown that the converse is not true, i.e., there exist transient sets A with $\mathbb{E}[V_A] = \infty$.

Lemma 6.5.8 *Suppose $p \in \mathcal{P}_d, d \geq 3$, and $A \subset \mathbb{Z}^d$. Then A is transient if and only if*

$$\sum_{k=1}^{\infty} \mathbb{P}\{T^k < \infty\} < \infty, \tag{6.34}$$

where $T^k = T_{A_k}$ and $A_k = A \cap (\mathcal{C}_{2^k} \setminus \mathcal{C}_{2^{k-1}})$.

Proof Let E_k be the event $\{T^k < \infty\}$. Since the random walk is transient, A is transient if and only if $\mathbb{P}\{E_k \text{ i.o.}\} = 0$. Hence the Borel-Cantelli Lemma implies that any A satisfying (6.34) is transient.

Suppose

$$\sum_{k=1}^{\infty} \mathbb{P}\{T^k < \infty\} = \infty.$$

Then either the sum over even k or the sum over odd k is infinite. We will assume the former; the argument if the latter holds is almost identical. Let $B_{k,+} = (A_k \setminus A_{k-1}) \cap \{(z^1, \ldots, z^d) : z^1 \geq 0\}$ and $B_{k,-} = (A_k \setminus A_{k-1}) \cap \{(z^1, \ldots, z^d) : z^1 \leq 0\}$. Since $\mathbb{P}\{T^{2k} < \infty\} \leq \mathbb{P}\{T_{B_{2k,+}} < \infty\} + \mathbb{P}\{T_{B_{2k,-}} < \infty\}$, we know that either

$$\sum_{k=1}^{\infty} \mathbb{P}\{T_{B_{2k,+}} < \infty\} = \infty, \tag{6.35}$$

or the same equality with $B_{2k,-}$ replacing $B_{2k,+}$. We will assume (6.35) holds and write $\sigma_k = T_{B_{2k,+}}$.

An application of the Harnack inequality (we leave the details as Exercise 6.7) shows that there is a c such that for all $j \neq k$,

$$\mathbb{P}\{\sigma_j < \infty \mid \sigma_j \wedge \sigma_k = \sigma_k < \infty\} \leq c\,\mathbb{P}\{\sigma_j < \infty\}.$$

This implies

$$\mathbb{P}\{\sigma_j < \infty, \sigma_k < \infty\} \leq 2c\,\mathbb{P}\{\sigma_j < \infty\}\,\mathbb{P}\{\sigma_k < \infty\}.$$

Using this and a special form of the Borel-Cantelli Lemma (Corollary 11.6.2) we can see that

$$\mathbb{P}\{\sigma_j < \infty \text{ i.o.}\} > 0,$$

which implies that A is not transient. □

Corollary 6.5.9 (Wiener's test) *Suppose $p \in \mathcal{P}_d, d \geq 3$, and $A \subset \mathbb{Z}^d$. Then A is transient if and only if*

$$\sum_{k=1}^{\infty} 2^{(2-d)k}\,\mathrm{cap}(A_k) < \infty \tag{6.36}$$

where $A_k = A \cap (\mathcal{C}_{2^k} \setminus \mathcal{C}_{2^{k-1}})$. In particular, if A is transient for some $p \in \mathcal{P}_d$, then it is transient for all $p \in \mathcal{P}_d$.

Proof By Proposition 6.5.1, we can see that $\mathbb{P}\{T^k < \infty\} \asymp 2^{(2-d)k}\,\mathrm{cap}(A_k)$. □

Theorem 6.5.10 *Suppose $d \geq 3$, $p \in \mathcal{P}_d$, and S_n is a p-walk. Let A be the set of points visited by the random walk,*

$$A = S[0,\infty) = \{S_n : n = 0, 1, \ldots\}.$$

If $d = 3, 4$, then with probability one A is a recurrent set. If $d \geq 5$, then with probability one A is a transient set.

Proof Since a set is transient if and only if all its translates are transient, we see that for each n, A is recurrent if and only if the set

$$\{S_m - S_n : m = n, n+1, \ldots\}$$

is recurrent. Therefore, the event $\{A \text{ is recurrent}\}$ is a tail event, and the Kolmogorov 0-1 law implies that it has probability 0 or 1.

Let Y denote the random variable that equals the expected number of visits to A by an independent random walker $\tilde{S}_n$ starting at the origin. In other words,

$$Y = \sum_{x \in A} G(x) = \sum_{x \in \mathbb{Z}^d} 1\{x \in A\}\,G(x).$$

Then,

$$\mathbb{E}(Y) = \sum_{x \in \mathbb{Z}^d} \mathbb{P}\{x \in A\}\,G(x) = G(0)^{-1} \sum_{x \in \mathbb{Z}^d} G(x)^2.$$

Since $G(x) \asymp |x|^{2-d}$, we have $G(x)^2 \asymp |x|^{4-2d}$. By examining the sum, we see that $\mathbb{E}(Y) = \infty$ for

$d = 3, 4$ and $\mathbb{E}(Y) < \infty$ for $d \geq 5$. If $d \geq 5$, this gives $Y < \infty$ with probability one which implies that A is transient with probability one.

We now focus on $d = 4$ (it is easy to see that if the result holds for $d = 4$ then it also holds for $d = 3$). It suffices to show that $\mathbb{P}\{A \text{ is recurrent}\} > 0$. Let S^1, S^2 be independent random walks with increment distribution p starting at the origin, and let

$$\sigma_k^j = \inf\{n : S_n^j \not\in \mathcal{C}_{2^k}\}.$$

Let

$$V_k^j = [\mathcal{C}_{2^k} \setminus \mathcal{C}_{2^{k-1}}] \cap S^j[0, \sigma_{k+1}^j) = \{x \in \mathcal{C}_{2^k} \setminus \mathcal{C}_{2^{k-1}} : S_n^j = x \text{ for some } n \leq \sigma_{k+1}^j\}.$$

Let E_k be the event $\{V_k^1 \cap V_k^2 \neq \emptyset\}$. We will show that $\mathbb{P}\{E_k \text{ i.o.}\} > 0$ which will imply that with positive probability, $\{S_n^1 : n = 0, 1, \ldots\}$ is recurrent. Using Corollary 11.6.2, we can see that to prove this it suffices to show that

$$\sum_{k=1}^{\infty} \mathbb{P}(E_{3k}) = \infty, \tag{6.37}$$

and that there exists a constant $c < \infty$ such that for $m < k$,

$$\mathbb{P}(E_{3m} \cap E_{3k}) \leq c\, \mathbb{P}(E_{3m})\, \mathbb{P}(E_{3k}). \tag{6.38}$$

The event E_{3m} depends only on the values of S_n^j with $\sigma_{3m-1}^j \leq n \leq \sigma_{3m+1}^j$. Hence, the Harnack inequality implies $\mathbb{P}(E_{3k} \mid E_{3m}) \leq c\, \mathbb{P}(E_{3k})$ and hence (6.38) holds. To prove (6.37), let $J^j(k, x)$ denote the indicator function of the event that $S_n^j = x$ for some $n \leq \sigma_k^j$. Then,

$$Z_k := \#(V_k^1 \cap V_k^2) = \sum_{x \in \mathcal{C}_{2^k} \setminus \mathcal{C}_{2^{k-1}}} J^1(k, x)\, J^2(k, x).$$

There exist c_1, c_2 such that if $x, y \in \mathcal{C}_{2^k} \setminus \mathcal{C}_{2^{k-1}}$,

$$\mathbb{E}[J^j(k, x)] \geq c_1\, (2^k)^{-4}, \qquad \mathbb{E}[J^j(k, x)\, J^j(k, y)] \leq c_2\, (2^k)^{-2}\, [1 \vee |x - y|]^{2-d}.$$

(The latter inequality is obtained by noting that the probability that a random walker hits both x and y given that it hits at least one of them is bounded above by the probability that a random walker starting at the origin visits $y - x$.) Therefore,

$$\mathbb{E}[Z_k] = \sum_{x \in \mathcal{C}_{2^{3k}} \setminus \mathcal{C}_{2^{3k-1}}} \mathbb{E}[J^1(3k, x)]\, \mathbb{E}[J^2(3k, x)] \geq c \sum_{x \in \mathcal{C}_{2^{3k}} \setminus \mathcal{C}_{2^{3k-1}}} (2^{3k})^{-2} \geq c,$$

$$\begin{aligned} \mathbb{E}[Z_k^2] &= \sum_{x, y \in \mathcal{C}_{2^{3k}} \setminus \mathcal{C}_{2^{3k-1}}} \mathbb{E}[J^1(3k, x)\, J^1(3k, y)]\, \mathbb{E}[J^2(3k, x)\, J^2(3k, y)] \\ &\leq c \sum_{x, y \in \mathcal{C}_{2^{3k}} \setminus \mathcal{C}_{2^{3k-1}}} (2^{3k})^{-4}\, \frac{1}{[1 \vee |x - y|^2]^2} \leq ck. \end{aligned}$$

Using the second moment estimate, Lemma 11.6.1, this implies that $\mathbb{P}\{Z_k > 0\} \geq c/k$ which implies (6.37). □

Comment. The central limit theorem implies that the number of points in $\mathcal{B}_n$ visited by a random walk is of order n^2. Roughly speaking, we can say that a random walk path is a "two-dimensional" set. Asking whether or not this is recurrent is asking whether or not two random two-dimensional sets intersect. Using the example of planes in $\mathbb{R}^d$, we can see that the critical dimension is four.

6.6 Capacity in two dimensions

The theory of capacity in two dimensions is somewhat similar to that for $d \geq 3$, but there are significant differences caused by the recurrence of the random walk. We recall some facts; see Propositions 6.4.7 and 6.4.8. If $p \in \mathcal{P}_2$ and $0 \in A \subset \mathbb{Z}^2$ is finite, let

$$g_A(x) = a(x) - \mathbb{E}^x[a(S_{\overline{T}_A})] = \lim_{n\to\infty} C_2 \,(\log n)\, \mathbb{P}^x\{\xi_n < \overline{T}_A\}. \tag{6.39}$$

The function g_A is the unique function on $\mathbb{Z}^2$ that vanishes on A; is harmonic on $\mathbb{Z}^2 \setminus A$; and satisfies $g_A(x) \sim C_2 \log \mathcal{J}(x) \sim C_2 \log |x|$ as $x \to \infty$. If $y \in A$, we can also write

$$g_A(x) = a(x-y) - \mathbb{E}^x[a(S_{\overline{T}_A} - y)].$$

(This must give the same value by uniqueness.) For the rest of this section we will assume that $0 \in A$, but this is only for convenience. Note that $a(x) - g_A(x)$ is the unique bounded function on $\mathbb{Z}^2$ that is harmonic on $\mathbb{Z}^2 \setminus A$ and has boundary value a on A. We define the *harmonic measure of A (from infinity)* by

$$\mathrm{hm}_A(x) = \lim_{|y|\to\infty} \mathbb{P}^y\{S_{T_A} = x\}. \tag{6.40}$$

Since $\mathbb{P}^y\{T_A < \infty\} = 1$, this is the same as $\mathbb{P}^y\{S_{T_A} = x \mid T_A < \infty\}$ and hence agrees with the definition of harmonic measure for $d \geq 3$. It is not immediately obvious that the limit exists, but the next proposition establishes this fact.

Proposition 6.6.1 *Suppose $p \in \mathcal{P}_2$ and $0 \in A \subset \mathbb{Z}^2$ is finite. Then the limit in (6.40) exists and equals $\mathcal{L} g_A(x)$.*

Proof We fix A and let $r_A = \mathrm{rad}(A)$. Let n be sufficiently large so that $A \subset \mathcal{C}_{n/4}$. Using (6.25) on the set $\mathbb{Z}^2 \setminus A$, we see that if $x \in \partial_i A, y \in \partial \mathcal{C}_n$,

$$\mathbb{P}^y\{S_{T_A \wedge \xi_n} = x\} = \mathbb{P}^x\{S_{T_A \wedge \xi_n} = y\} = \mathbb{P}^x\{\xi_n < T_A\}\, H_{\mathcal{C}_n}(0,y) \left[1 + O\left(\frac{r_A \log n}{n}\right)\right].$$

If $z \in \mathbb{Z}^2 \setminus \mathcal{C}_n$, the last-exit decomposition (Proposition 4.6.4) gives

$$\mathbb{P}^z\{S_{T_A} = x\} = \sum_{y \in \partial \mathcal{C}_n} G_{\mathbb{Z}^2\setminus A}(z,y)\, \mathbb{P}^y\{S_{T_A \wedge \xi_n} = x\}.$$

Therefore,

$$\mathbb{P}^z\{S_{T_A} = x\} = \mathbb{P}^x\{\xi_n < T_A\}\, J(n,z) \left[1 + O\left(\frac{r_A \log n}{n}\right)\right], \tag{6.41}$$

where

$$J(n,z) = \sum_{y \in \partial \mathcal{C}_n} H_{\mathcal{C}_n}(0,y)\, G_{\mathbb{Z}^2\setminus A}(z,y).$$

If $x \in A$, the definition of $\mathcal{L}$, the optional sampling theorem, and the asymptotic expansion of g_A imply

$$\begin{aligned}
\mathcal{L}g_A(x) = \mathbb{E}^x[g_A(S_1)] &= \mathbb{E}^x[g_A(S_{T_A \wedge \xi_n})] \\
&= \mathbb{E}^x[g_A(S_{\xi_n}); \xi_n < T_A] \\
&= \mathbb{P}^x\{\xi_n < T_A\} \; [C_2 \, \log n + O_A(1)] \,. \qquad (6.42)
\end{aligned}$$

In particular,

$$\mathcal{L}g_A(x) = \lim_{n \to \infty} C_2 \,(\log n)\, \mathbb{P}^x\{\xi_n < T_A\}, \qquad x \in A. \qquad (6.43)$$

(This is the $d = 2$ analogue of the relation $\mathcal{L}g_A(x) = \mathrm{Es}_A(x)$ for $d \geq 3$.)

Note that

$$\begin{aligned}
\sum_{x \in \partial_i A} \mathbb{P}^x\{\xi_n < T_A\} &= \sum_{x \in \partial_i A} \sum_{y \in \partial \mathcal{C}_n} \mathbb{P}^x\{S_{\xi_n \wedge T_A} = y\} \\
&= \sum_{y \in \partial \mathcal{C}_n} \sum_{x \in \partial_i A} \mathbb{P}^y\{S_{\xi_n \wedge T_A} = x\} = \sum_{y \in \partial \mathcal{C}_n} \mathbb{P}^y\{T_A < \xi_n\}.
\end{aligned}$$

Proposition 6.4.3 shows that if $x \in A$, then the probability that a random walk starting at x reaches $\partial \mathcal{C}_n$ before visiting the origin is bounded above by $c \, \log r_A / \log n$. Therefore,

$$\mathbb{P}^y\{T_A < \xi_n\} = \mathbb{P}^x\{T_{\{0\}} < \xi_n\} \left[1 + O\left(\frac{\log r_A}{\log n}\right)\right].$$

Therefore,

$$\begin{aligned}
\sum_{x \in \partial_i A} \mathbb{P}^x\{\xi_n < T_A\} &= \sum_{y \in \partial \mathcal{C}_n} \mathbb{P}^y\{S_{\xi_n \wedge T_{\{0\}}} = 0\} \left[1 + O\left(\frac{\log r_A}{\log n}\right)\right] \\
&= \mathbb{P}\{\xi_n < T_{\{0\}}\} \left[1 + O\left(\frac{\log r_A}{\log n}\right)\right] \\
&= [C_2 \, \log n]^{-1} \left[1 + O\left(\frac{\log r_A}{\log n}\right)\right].
\end{aligned}$$

Combining this with (6.43) gives

$$\sum_{x \in A} \mathcal{L}g_A(x) = \sum_{x \in \partial_i A} \mathcal{L}g_A(x) = 1.$$

This is a difference between the $d = 2$ and $d \geq 3$ cases. If $d \geq 3$, the sum above grows as A gets large and gives the capacity of the set. For $d = 2$, this quantity is the same for all A.

Using (6.41) and $\sum_{x \in A} \mathbb{P}^z\{S_{T_A} = x\} = 1$, we see that

$$J(n,z) \sum_{x \in A} \mathbb{P}^x\{\xi_n < T_A\} = 1 + O\left(\frac{r_A \, \log n}{n}\right),$$

which by (6.42) implies that

$$J(n,z) = C_2 \, \log n \left[1 + O\left(\frac{r_A \, \log n}{n}\right)\right].$$

$\square$

We define the *capacity* of A by

$$\mathrm{cap}(A) = \lim_{y\to\infty} [a(y) - g_A(y)] = \sum_{x\in A} \mathrm{hm}_A(x)\, a(x-z),$$

where $z \in A$. The last proposition establishes the limit if $z = 0 \in A$, and for other z see (6.29). We have the expansion

$$g_A(x) = C_2 \log \mathcal{J}(x) + \gamma_2 - \mathrm{cap}(A) + o_A(1), \qquad |x| \to \infty.$$

The capacity is translation invariant, $\mathrm{cap}(A-y) = \mathrm{cap}(A)$. Note that singleton sets have capacity zero.

Proposition 6.6.2 *Suppose $p \in \mathcal{P}_2$.*
(a) If $0 \in A \subset B \subset \mathbb{Z}^d$ are finite, then $g_A(x) \geq g_B(x)$ for all x. In particular, $\mathrm{cap}(A) \leq \mathrm{cap}(B)$.
(b) If $A, B \subset \mathbb{Z}^d$ are finite subsets containing the origin, then for all x

$$g_{A\cup B}(x) \geq g_A(x) + g_B(x) - g_{A\cap B}(x). \tag{6.44}$$

In particular,

$$\mathrm{cap}(A \cup B) \leq \mathrm{cap}(A) + \mathrm{cap}(B) - \mathrm{cap}(A \cap B).$$

Proof The inequality $g_A(x) \geq g_B(x)$ follows immediately from (6.39). The inequality (6.44) follows from (6.39) and the observation

$$\begin{aligned}
\mathbb{P}^x\{\overline{T}_{A\cup B} < \xi_n\} &= \mathbb{P}^x\{\overline{T}_A < \xi_n \text{ or } \overline{T}_B < \xi_n\} \\
&= \mathbb{P}^x\{\overline{T}_A < \xi_n\} + \mathbb{P}^x\{\overline{T}_B < \xi_n\} - \mathbb{P}^x\{\overline{T}_A < \xi_n; \overline{T}_B < \xi_n\} \\
&\leq \mathbb{P}^x\{\overline{T}_A < \xi_n\} + \mathbb{P}^x\{\overline{T}_B < \xi_n\} - \mathbb{P}^x\{\overline{T}_{A\cap B} < \xi_n\},
\end{aligned}$$

which implies

$$\mathbb{P}^x\{\overline{T}_{A\cup B} > \xi_n\} \geq \mathbb{P}^x\{\overline{T}_A > \xi_n\} + \mathbb{P}^x\{\overline{T}_B > \xi_n\} - \mathbb{P}^x\{\overline{T}_{A\cap B} > \xi_n\}.$$

$\square$

We will now derive an analogue of Proposition 6.5.5. If A is a finite set, let a_A denote the $\#(A) \times \#(A)$ symmetric matrix with entries $a(x,y)$. Let a_A also denote the operator

$$a_A\, f(x) = \sum_{y\in A} a(x,y)\, f(y)$$

which is defined for all functions $f : A \to \mathbb{R}$ and all $x \in \mathbb{Z}^2$. Note that $x \mapsto a_A f(x)$ is harmonic on $\mathbb{Z}^2 \setminus A$.

Proposition 6.6.3 *Suppose $p \in \mathcal{P}_2$ and $0 \in A \subset \mathbb{Z}^2$ is finite. Then*

$$\mathrm{cap}(A) = \left[\sup \sum_{y\in A} f(y)\right]^{-1},$$

where the supremum is over all nonnegative functions f on A satisfying $a_A f(x) \leq 1$ for all $x \in A$.

If $A = \{0\}$ is a singleton set, the proposition is trivial since $a_A f(0) = 0$ for all f and hence the supremum is infinity. A natural first guess for other A (which turns out to be correct) is that the supremum is obtained by a function f satisfying $a_A f(x) = 1$ for all $x \in A$. If $\{a_A(x,y)\}_{x,y\in A}$ is invertible, there is a unique such function that can be written as $f = a_A^{-1}\mathbf{1}$ (where $\mathbf{1}$ denotes the vector of all 1s). The main ingredient in the proof of Proposition 6.6.3 is the next lemma that shows this inverse is well defined assuming A has at least two points.

Lemma 6.6.4 *Suppose $p \in \mathcal{P}_2$ and $0 \in A \subset \mathbb{Z}^2$ is finite with at least two points. Then a_A^{-1} exists and*

$$a_A^{-1}(x,y) = \mathbb{P}^x\{S_{T_A} = y\} - \delta(y-x) + \frac{\mathcal{L}g_A(x)\, \mathcal{L}g_A(y)}{\mathrm{cap}(A)}, \quad x, y \in A.$$

Proof We will first show that for all $x \in \mathbb{Z}^2$.

$$\sum_{z\in A} a(x,z)\, \mathcal{L}g_A(z) = \mathrm{cap}(A) + g_A(x). \tag{6.45}$$

To prove this, we will need the following fact (see Exercise 6.4):

$$\lim_{n\to\infty} [G_{\mathcal{C}_n}(0,0) - G_{\mathcal{C}_n}(x,y)] = a(x,y).$$

Consider the function

$$h(x) = \sum_{z\in A} a(x,z)\, \mathcal{L}g_A(z).$$

We first claim that h is constant on A. By a last-exit decomposition (Proposition 4.6.4), if $x, y \in A$,

$$1 = \mathbb{P}^x\{\overline{T}_A < \xi_n\} = \sum_{z\in A} G_{\mathcal{C}_n}(x,z)\, \mathbb{P}^z\{\xi_n < T_A\} = \sum_{z\in A} G_{\mathcal{C}_n}(y,z)\, \mathbb{P}^z\{\xi_n < T_A\}.$$

Hence,

$$(C_2 \log n) \sum_{z\in A} [G_{\mathcal{C}_n}(0,0) - G_{\mathcal{C}_n}(x,z)] \mathbb{P}^z\{\xi_n < T_A\} =$$

$$(C_2 \log n) \sum_{z\in A} [G_{\mathcal{C}_n}(0,0) - G_{\mathcal{C}_n}(y,z)] \mathbb{P}^z\{\xi_n < T_A\}].$$

Letting $n \to \infty$, and recalling that $C_2 \log n\, \mathbb{P}^z\{\xi_n < T_A\} \to \mathcal{L}g_A(z)$, we get the claim. Theorem 4.4.3 implies that

$$\lim_{x\to\infty} [a(x) - h(x)] = 0.$$

Hence, $a(x) - h(x)$ is a bounded function that is harmonic in $\mathbb{Z}^2 \setminus A$ and takes the value $a - h_A$ on A, where h_A denotes the constant value of h on A. This implies that $a(x) - h(x) = a(x) - g_A(x) - h_A$. Therefore,

$$h_A = \lim_{x\to\infty} [a(x) - g_A(x)] = \mathrm{cap}(A).$$

This establishes (6.45).

An application of the optional sampling theorem gives

$$G_{\mathcal{C}_n}(x,z) = \delta(z-x) + \mathbb{E}^x[G_{\mathcal{C}_n}(S_1,z)] = \delta(z-x) + \sum_{y\in A} \mathbb{P}^x\{S_{T_A\wedge\xi_n} = y\}\, G_{\mathcal{C}_n}(y,z).$$

Hence,

$$G_{\mathcal{C}_n}(0,0) - G_{\mathcal{C}_n}(x,z) = -\delta(z-x)$$
$$+ G_{\mathcal{C}_n}(0,0)\,\mathbb{P}^x\{\xi_n < T_A\} + \sum_{y\in A} \mathbb{P}^x\{S_{\tau_A\wedge\xi_n} = y\}\,[G_{\mathcal{C}_n}(0,0) - G_{\mathcal{C}_n}(y,z)].$$

Letting $n\to\infty$ and using (6.12), this gives

$$\delta(z-x) = -a(x,z) + \mathcal{L}g_A(x) + \sum_{y\in A}\mathbb{P}^x\{S_{T_A} = y\}\,a(y,z).$$

If $x,z\in A$, we can use (6.45) to write the previous equation as

$$\delta(z-x) = \sum_{y\in A}\left[\mathbb{P}^x\{S_{T_A} = y\} - \delta(y-x) + \frac{\mathcal{L}g_A(x)\,\mathcal{L}g_A(y)}{\mathrm{cap}(A)}\right] a(y,z),$$

provided that $\mathrm{cap}(A) > 0$. $\square$

Proof of Proposition 6.6.3 Let $\hat f(x) = \mathcal{L}g_A(x)/\mathrm{cap}(A)$. In (6.45) we showed that

$$\sum_{y\in A} a(x,y)\,\hat f(y) = 1, \qquad x\in A.$$

Suppose f satisfies the conditions in the statement of the proposition, and let $h = a\hat f - af$ which is nonnegative in A. Then, using Lemma 6.6.4,

$$\begin{aligned}\sum_{x\in A}\left[\sum_{y\in A} a_A^{-1}(x,y)\,h(y)\right] &\ge \sum_{x\in A}\sum_{y\in A}\mathbb{P}^x\{S_{\tau_A} = y\}\,h(y) - \sum_{x\in A} h(x)\\ &= \sum_{y\in A} h(y)\sum_{x\in A}\mathbb{P}^y\{S_{\tau_A} = x\} - \left[\sum_{y\in A} h(y)\right] = 0.\end{aligned}$$

In other words,

$$\sum_{y\in A}\hat f(y) \ge \sum_{y\in A} f(y).$$

Proposition 6.6.5 *If $p\in\mathcal{P}_2$,*

$$\mathrm{cap}(\mathcal{C}_n) = C_2\log n + \gamma_2 + O(n^{-1}),$$

Proof By definition,

$$g_{\mathcal{C}_n}(x) = C_2\,\log\mathcal{J}(x) + \gamma_2 - \mathrm{cap}(\mathcal{C}_n) + o(1), \quad x\to\infty.$$

But for $x \notin \mathcal{C}_n$,

$$g_{\mathcal{C}_n}(x) = a(x) - \mathbb{E}^x[a(S_{T_{\mathcal{C}_n}})] = C_2 \log \mathcal{J}(x) + \gamma_2 + O(|x|^{-2}) - [C_2 \log n + \gamma_2 + O(n^{-1})].$$

$\square$

Lemma 6.6.6 *if $p \in \mathcal{P}_2$, and $A \subset B \subset \mathbb{Z}^2$ are finite, then*

$$\mathrm{cap}(A) = \mathrm{cap}(B) - \sum_{y \in B} \mathrm{hm}_B(y)\, g_A(y).$$

Proof $g_A - g_B$ is a bounded function that is harmonic on $\mathbb{Z}^2 \setminus B$ with boundary value g_A on B. Therefore,

$$\begin{aligned} \mathrm{cap}(A) - \mathrm{cap}(B) &= \lim_{x \to \infty} [g_A(x) - g_B(x)] \\ &= \lim_{x \to \infty} \mathbb{E}^x[g_A(S_{\overline{T}_B}) - g_B(S_{\overline{T}_B})] = \sum_{y \in B} \mathrm{hm}_B(y)\, g_A(y). \end{aligned}$$

$\square$

Comment. Proposition 6.6.5 tells us that the capacity of an ellipse of diameter n is $C_2 \log n + O(1)$. The next lemma shows that this is also true for any connected set of diameter n. In particular, the capacities of the ball of raidus n and a line of radius n are asymptotic as $n \to \infty$. This is not true for capacities in $d \geq 3$.

Lemma 6.6.7 *If $p \in \mathcal{P}_2$, there exist c_1, c_2 such that if A is a finite subset of $\mathbb{Z}^2$ with* $\mathrm{rad}(A) < n$ *and satisfying*

$$\#\{x \in A : k - 1 \leq |x| < k\} \geq 1,$$

for $k = 1, \ldots, n$, then the following holds.

(a) If $x \in \partial \mathcal{C}_{2n}$,

$$\mathbb{P}^x\{T_A < \xi_{4n}\} \geq c_1.$$

(b) $|\mathrm{cap}(A) - C_2 \log n| \leq c_2$.

(c) If $x \in \partial \mathcal{C}_{2n}$, $m \geq 4n$, and $A_n = A \cap \mathcal{C}_n$, then

$$c_1 \leq \mathbb{P}^x\{T_{A_n} > \xi_m\} \log(m/n) \leq c_2. \tag{6.46}$$

Proof (a) Let δ be such that $\mathcal{B}_{\delta n} \subset \mathcal{C}_n$, and let B denote a subset of A contained in $\mathcal{B}_n$ such that

$$\#\{x \in B : k - 1 \leq |x| < k\} = 1$$

for each positive integer $k < n$. We will prove the estimate for B which will clearly imply the estimate for A. Let $V = V_{n,B}$ denote the number of visits to B before leaving $\mathcal{C}_{4n}$,

$$V = \sum_{j=0}^{\xi_{4n}-1} 1\{S_j \in B\} = \sum_{j=0}^{\infty} \sum_{z \in B} 1\{S_j = z; j < \xi_{4n}\}.$$

The strong Markov property implies that if $x \in \partial \mathcal{C}_{2n}$,

$$\mathbb{E}^x[V] = \mathbb{P}^x\{T_B < \xi_{4n}\}\, \mathbb{E}^x[V \mid T_B < \xi_{4n}] \leq \mathbb{P}^x\{T_B < \xi_{4n}\} \max_{z \in B} \mathbb{E}^z[V].$$

Hence, we need only find a c_1 such that $\mathbb{E}^x[V] \geq c_1 \, \mathbb{E}^z[V]$ for all $x \in \partial \mathcal{C}_{2n}, z \in B$. Note that $\#(B) = \delta n + O(1)$. By Exercise 6.10, we can see that $G_{\mathcal{C}_{4n}}(x,z) \geq c$ for $x, z \in \mathcal{C}_{2n}$. Therefore $\mathbb{E}^x[V] \geq c\, n$. If $z \in B$, there are at most $2k+1$ points w in B satisfying $|z-w| \leq 2k+1$. Using Proposition 6.3.5, we see that

$$G_{\mathcal{C}_{4n}}(z,w) \leq C_2 \, [\log n - \log |z-w| + O(1)].$$

Therefore,

$$\mathbb{E}^z[V] \leq \sum_{k=1}^{n} 2\, C_2 \, [\log n - \log k + O(1)] \leq c\, n.$$

The last inequality uses the estimate

$$\sum_{k=1}^{n} \log k = O(\log n) + \int_1^n \log x \, dx = n \, \log n - n + O(\log n).$$

(b) There exists a δ such that $\mathcal{B}_n \subset \mathcal{C}_{n/\delta}$ for all n and hence

$$\mathrm{cap}(A) \leq \mathrm{cap}(\mathcal{B}_n) \leq \mathrm{cap}(\mathcal{C}_{n/\delta}) \leq C_2 \, \log n + O(1).$$

Hence, we only need to give a lower bound on $\mathrm{cap}(A)$. By the previous lemma it suffices to give a bound $g_A(x) \leq c$ for $x \in \partial \mathcal{C}_{4n}$. For $m > 4n$, let

$$r_m = r_{m,n,A} = \max_{y \in \mathcal{C}_{2n}} \mathbb{P}^y\{\xi_m < T_A\},$$

$$r_m^* = r_{m,n,A} = \max_{y \in \mathcal{C}_{4n}} \mathbb{P}^y\{\xi_m < T_A\}.$$

Using part (a) and the strong Markov property, we see that there is a $\rho < 1$ such that $r_m \leq \rho \, r_m^*$. Also, if $y \in \mathcal{C}_{4n}$

$$\begin{aligned} \mathbb{P}^y\{\xi_m < T_A\} &= \mathbb{P}^y\{\xi_m < T_{\mathcal{C}_{2n}}\} + \mathbb{P}^y\{\xi_m > T_{\mathcal{C}_{2n}}\} \, \mathbb{P}^y\{\xi_m < T_A \mid \xi_m > T_{\mathcal{C}_{2n}}\} \\ &\leq \mathbb{P}^y\{\xi_m < T_{\mathcal{C}_{2n}}\} + \rho \, r_m^*. \end{aligned}$$

Proposition 6.4.1 tells us that there is a c_3 such that for $y \in \mathcal{C}_{4n}$,

$$\mathbb{P}^y\{\xi_m < T_{\mathcal{C}_{2n}}\} \leq \frac{c_3}{\log m - \log n + O(1)}.$$

Therefore,

$$g_A(y) = \lim_{m \to \infty} C_2 \, (\log m) \, \mathbb{P}^y\{\xi_m < T_A\} \leq \frac{C_2 \, c_3}{1-\rho}.$$

(c) The lower bound for (6.46) follows from Proposition 6.4.1 and the observation

$$\mathbb{P}^x\{T_{A_n} > \xi_m\} \geq \mathbb{P}^x\{T_{\mathcal{C}_n} > \xi_m\}.$$

For the upper bound let

$$u = u_n = \max_{x \in \partial \mathcal{C}_{2n}} \mathbb{P}^x\{T_{A_n} > \xi_m\}.$$

Consider a random walk starting at $y \in \partial \mathcal{C}_{2n}$ and consider $T_{\mathcal{C}_n} \wedge \xi_m$. Clearly,

$$\mathbb{P}^y\{T_{A_n} > \xi_m\} = \mathbb{P}^y\{\xi_m < T_{\mathcal{C}_n}\} + \mathbb{P}^y\{\xi_m < T_{\mathcal{C}_n}; \xi_m > T_{A_n}\}.$$

By Proposition 6.4.1, for all $y \in \mathcal{C}_{2n}$

$$\mathbb{P}^y\{\xi_m < T_{\mathcal{C}_n}\} \le \frac{c}{\log(m/n)}.$$

Let $\sigma = \sigma_n = \min\{j \ge T_{\mathcal{C}_n} : S_j \in \partial\mathcal{C}_{2n}\}$. Then, by the Markov property,

$$\mathbb{P}^y\{\xi_m < T_{\mathcal{C}_n}; \xi_m > T_{A_n}\} \le u\, \mathbb{P}^y\{S[0,\sigma] \cap A_n = \emptyset\}.$$

Part (a) shows that there is a $\rho < 1$ such that $\mathbb{P}^y\{S[0,\sigma] \cap A_n = \emptyset\} \le \rho$ and hence, we get

$$\mathbb{P}^y\{T_{A_n} > \xi_m\} \le \frac{c}{\log(m/n)} + \rho\, u.$$

Since this holds, for all $y \in \partial\mathcal{C}_{2n}$, this implies

$$u \le \frac{c}{\log(m/n)} + \rho\, u,$$

which gives us the upper bound. □

Comment. A major example of a set satisfying the condition of the theorem is a connected (with respect to simple random walk) subset of $\mathbb{Z}^2$ of radius between $n-1$ and n. In the case of simple random walk, there is another proof of part (a) that basically makes the observation that there is a positive probability that the simple random walk starting on $\partial\mathcal{C}_{2n}$ makes a closed loop about the origin contained in $\mathcal{C}_n$. One can justify this rigorously, and bound the probability from below independently of n, by using an approximation by Brownian motion. If the random walk makes a closed loop, then it must intersect any connected set. Unfortunately, it is not easy to modify this argument for random walks that take non-nearest neighbor steps.

6.7 Beurling estimates

We will prove results for simple random walk and then discuss at the end how one can generalize the results for $p \in \mathcal{P}$. Recall that a set $A \subset \mathbb{Z}^d$ is connected (for simple random walk) if any two points in A can be connected by a nearest neighbor path of points in A. The next theorem is a very useful estimate, especially in two dimensions.

Theorem 6.7.1 (Beurling estimate) *If $p \in \mathcal{P}_d$, $d = 2, 3$, there is a c such that if A is an infinte connected subset of $\mathbb{Z}^d$ containing the origin,*

$$\mathbb{P}\{\xi_n < T_A\} \le \frac{c}{n^{1/2}}, \quad d = 2, \tag{6.47}$$

$$\mathbb{P}\{\xi_n < T_A\} \le \frac{c}{(\log n)^{1/2}}, \quad d = 3. \tag{6.48}$$

We will prove this theorem first for simple random walk and then we will discuss the minor modifications needed to establish it for all $p \in \mathcal{P}$. We will give a proof for $d = 2$. This proof could be modified for $d = 3$, but we choose to give a different proof in this case in order to demonstrate a useful technique. In both cases, It turns out to be useful to consider a different class of sets than the connected sets.

Definition. Let $\mathcal{A}_d$ denote the infinite subsets of $\mathbb{Z}^d$ with the property that for each positive integer j,

$$\#\{z \in A : (j-1) \leq |z| < j\} = 1.$$

One important example of a set in $\mathcal{A}_d$ is the half-infinite line

$$L = \{j\mathbf{e}_1 : j = 0, 1, \ldots\}.$$

We state two immediate facts about $\mathcal{A}_d$.

- If A' is an infinite connected subset of $\mathbb{Z}^d$ containing the origin, then there exists a (not necessarily connected) $A \in \mathcal{A}_d$ with $A \subset A'$.
- If $z \in A \in \mathcal{A}_d$, then for every real $r > 0$.

$$\#\{w \in A : |z-w| \leq r\} \subset \{w \in A : |z| - r \leq |w| \leq |z| + r\} \leq 2r + 2. \tag{6.49}$$

To prove Theorem 6.7.1, we prove the following stronger result. Since we will start by restricting to simple random walk, we first state the theorem in this case.

Theorem 6.7.2 *For simple random walk in $\mathbb{Z}^d$, $d = 2, 3$, there is a c such that if $A \in \mathcal{A}_d$, then*

$$\mathbb{P}\{\xi_n < T_A\} \leq \frac{c}{n^{1/2}}, \quad d = 2, \tag{6.50}$$

$$\mathbb{P}\{\xi_n < T_A\} \leq \frac{c}{(\log n)^{1/2}}, \quad d = 3. \tag{6.51}$$

Proof (of (6.50)) We fix n and let $V = \{y_1, \ldots, y_n\}$ where y_j denotes the unique point in A with $j \leq |y_j| < j+1$. We let $K = \{x_1, \ldots, x_n\}$ where $x_j = j\mathbf{e}_1$. Let $\mathcal{B} = \mathcal{B}_{n^3}$, $G = G_{\mathcal{B}_{n^3}}, \xi = \xi_{n^3}$. By (6.46), there exist c_1, c_2 such that for $z \in \partial\mathcal{B}_{2n}$,

$$\frac{c_1}{\log n} \leq \mathbb{P}^z\{\xi < T_A\}, \mathbb{P}^z\{\xi < T_K\} \leq \frac{c_2}{\log n}.$$

Let

$$v(z) = \mathbb{P}^z\{\xi < T_V\}, \quad q(z) = \mathbb{P}^z\{\xi < T_K\}.$$

It suffices to prove that $v(0) \leq c\, n^{-1/2}(\log n)^{-1}$. We will use the following two estimates, see (5.13) and ????: there is a c_1 such that for $j = 1, \ldots, n$,

$$q(x_j) \leq c\, n^{-1/2} \left[j^{-1/2} + (n-j+1)^{-1/2}\right] [\log n]^{-1}; \tag{6.52}$$

If $|x|, |y| \leq n$, then

$$G_{n^3-n}(0,0) \leq G_{n^3}(x,x) \leq G_{n^3+n}(0,0),$$

and hence

$$G(x,x) = \frac{2}{\pi} \log n^3 + \gamma + O\left(\frac{1}{n^2}\right),$$

$$G(x,y) = \frac{2}{\pi} \log n^3 + \gamma - a(x,y) + O\left(\frac{1}{n^2}\right). \tag{6.53}$$

Using Proposition 4.6.4, we can write

$$\begin{aligned}
v(z) - q(z) &= \mathbb{P}\{\xi < T_V\} - \mathbb{P}\{\xi < T_K\} \\
&= \mathbb{P}\{\xi > T_K\} - \mathbb{P}\{\xi > T_V\} \\
&= \sum_{j=1}^{n} G(0, x_j)\, q(x_j) - \sum_{j=1}^{n} G(0, y_j)\, v(y_j) \\
&= \sum_{j=1}^{n} [G(0, x_j) - G(0, y_j)]\, q(x_j) + \sum_{j=1}^{n} G(0, y_j)\, [q(x_j) - v(y_j)].
\end{aligned}$$

Using (6.52) and (6.53), we get

$$(\log n) \left| \sum_{j=1}^{n} [G(0, x_j) - G(0, y_j)]\, q(x_j) \right| \le O(n^{-1}) + c \sum_{j=1}^{n} |a(x_j) - a(y_j)|\, j^{-1/2}\, n^{-1/2}.$$

But $|x_j| = j, |y_j| = j + O(1)$, and hence we get that

$$(\log n) \left| \sum_{j=1}^{n} [G(0, x_j) - G(0, y_j)]\, q(x_j) \right| \le O(n^{-1}) + c \sum_{j=1}^{n} \frac{1}{j^{3/2}\, n^{-1/2}} \le \frac{c}{n^{1/2}}.$$

Given (6.46), it now suffices to show that

$$\sum_{j=1}^{n} G(0, y_j)\, [q(x_j) - v(y_j)] \le \frac{c}{n^{1/2} \log n}. \tag{6.54}$$

Note that we are not taking absolute values here. Consider the function

$$F(z) = \sum_{j=1}^{n} G(z, y_j)\, [q(x_j) - v(y_j)],$$

and note that F is harmonic on $\mathcal{B} \setminus V$. Since $F \equiv 0$ on $\partial\mathcal{B}$, either $F \le 0$ everywhere (in which case (6.54) is trivial) or it takes its maximum on V. Therefore, it suffices to find a c such that for $k = 1, \ldots, n$,

$$\sum_{j=1}^{n} G(y_k, y_j)\, [q(x_j) - v(y_j)] \le c\, n^{-1/2}.$$

By using Proposition 4.6.4 again, we get

$$\sum_{j=1}^{n} G(y_k, y_j)\, v(y_j) = \mathbb{P}^{y_k}\{\overline{T}_V \le \xi\} = 1 = \mathbb{P}^{x_k}\{\overline{T}_K \le \xi\} = \sum_{j=1}^{n} G(x_k, x_j)\, q(x_j).$$

Therefore, it suffices to show that

$$\sum_{j=1}^{n} [G(y_k, y_j) - G(x_k, x_j)]\, q(x_j) \le c\, n^{-1/2}.$$

Note that $|x_k - x_j| = |k-j|$ and $|y_k - y_j| \geq |k-j| - 1$. Hence, using (6.53),

$$G(y_k, y_j) - G(x_k, x_j) \leq \frac{c}{|k-j|+1}$$

and

$$\sum_{j=1}^{n} [G(y_k, y_j) - G(x_k, x_j)]\, q(x_j) \leq c \sum_{j=1}^{n} \frac{1}{(|k-j|+1)\, j^{1/2}\, \log n},$$

and it is straightforward to show that the right-hand side is $O(n^{-1/2}\,(\log n)^{-1})$.

$\square$

Before proving the $d = 3$ case, we state some simple lemmas.

Lemma 6.7.3 *If $z \in A \in \mathcal{A}_3$ and $r > 0$, then*

$$\sum_{w \in A,\, |w-z| \leq r} G(z, w) \leq 2\, C_3\, \log r + O(1).$$

Proof We know that

$$\left| G(z, w) - \frac{C_3}{r} \right| \leq \frac{c}{r^2}, \quad r \leq |z - w| < r + 1.$$

The result follows from (6.49). $\square$

Lemma 6.7.4 *Suppose $A \in \mathcal{A}_3$ and $z \in \mathcal{B}_n$. Then,*

$$\mathbb{P}^z\{\xi_n > T_A\} \geq \frac{1}{2\, C_3\, \log n} \left[\sum_{w \in A} G_{\mathcal{B}_n}(z, w) \right] \left[1 + O\left(\frac{1}{\log n} \right) \right].$$

Proof Let Y be the total number of visits to A before leaving $\mathcal{B}_n$. Then,

$$\mathbb{P}^z\{\xi_n > T_A\} = \mathbb{P}^z\{Y \geq 1\} = \frac{\mathbb{E}^z[Y]}{\mathbb{E}^z[Y \mid Y \geq 1]}.$$

The previous lemma tells us that

$$\mathbb{E}^z[Y \mid Y \geq 1] \leq 2\, C_3\, \log n + O(1).$$

$\square$

Lemma 6.7.5 *Suppose $A \in \mathcal{A}_3$. Then there exists a constant $C(A)$ such that as $n \to \infty$,*

$$\sum_{w \in A} G_{\mathcal{B}_n}(0, w) = C_3\, \log n + C(A) + O\left(\frac{1}{n} \right).$$

In particular, if $n < m$,

$$\sum_{w \in A} [G_{\mathcal{B}_m}(0, w) - G_{\mathcal{B}_n}(0, w)] = C_3\, \log\left(\frac{m}{n} \right) + O\left(\frac{1}{n} \right). \tag{6.55}$$

Proof Let $A = \{w_1, w_2, \ldots, \}$ where $j-1 \leq |w_j| < j$. For $w \in \mathcal{B}_n$,

$$G_{\mathcal{B}_n}(0,w) = G_{\mathcal{B}_n}(w,0) = G(w) - \mathbb{E}^w[G(S_{\mathcal{B}_n})] = G(0,w) - \frac{C_3}{n} + O\left(\frac{1}{n^2}\right).$$

Therefore,

$$\sum_{j=1}^n G_{\mathcal{B}_n}(0,w_n) = C_3 + O\left(\frac{1}{n}\right) + \sum_{j=1}^n G(w_n).$$

But,

$$G(w_n) = \frac{C_3}{j} + O\left(\frac{1}{j^2}\right),$$

and hence there is a constant c such that

$$G_{\mathcal{B}_n}(0,w) = c + O\left(\frac{1}{n}\right) + \sum_{j=1}^n \frac{C_3}{n}.$$

Since

$$\sum_{j=1}^n \frac{1}{n} = \log n + \gamma + O\left(\frac{1}{n}\right),$$

we get the first assertion and the second follows by subtracting. □

Proof of (6.51). Let

$$F(n) = \sup_{A \in \mathcal{A}_3} \mathbb{P}\{\xi_{e^n} < T_A\}.$$

Then it suffices to show that $F(n) \leq c\, n^{-1/2}$. In order to establish this it suffices to find a sequence of positive numbers δ_n with $\sum \delta_n < \infty$ such that for all n sufficiently large,

$$F(n+1) \leq F(n)\left[1 - \frac{1}{2n} + \delta_n\right].$$

Indeed, the last condition implies that

$$\log F(n+1) \leq -\left[\sum_{j=1}^n \frac{1}{2n}\right] + O(1) = -\frac{1}{2}\log n + O(1).$$

(see Lemma 11.1.3). Let $\xi^n = \xi_{e^n}$. Note that

$$\begin{aligned} \mathbb{P}\{\xi^{n+1} < T_A\} &= \mathbb{P}\{\xi^n < T_A\}\mathbb{P}\{S[\xi^n, \xi^{n+1}] \cap A = \emptyset \mid \xi^n < T_A\} \\ &\leq F(n)\,\mathbb{P}\{S[\xi^n, \xi^{n+1}] \cap A = \emptyset \mid \xi^n < T_A\}. \end{aligned}$$

Roughly speaking, the proof consists of two estimates:

$$\mathbb{P}\{S[\xi^n, \xi^{n+1}] \cap A = \emptyset\} \leq 1 - \frac{1}{2n} + \delta_n,$$

and

$$\mathbb{P}\{S[\xi^n, \xi^{n+1}] \cap A = \emptyset \mid \xi^n < T_A\} \sim \mathbb{P}\{S[\xi^n, \xi^{n+1}] \cap A = \emptyset\}.$$

The approximate indepedence may be surprising, but it comes from the rate of decay of F. Suppose it were true that

$$\mathbb{P}\{\xi^n < T_A\} \sim c\,(\log n)^{-1/2}$$

(which happens, for example, if $A = L$). Then this implies that $\mathbb{P}\{\xi^n < T_A\} \sim \mathbb{P}\{\xi^{n-\log n} < T_A\}$, or, in other words,

$$\mathbb{P}\{\xi^n < T_A \mid \xi^{n-\log n} < T_A\} \longrightarrow 1.$$

Hence conditioning on the event $\{\xi^n < T_A\}$ is approximately the same thing as conditioning on the event $\{\xi^{n-\log n} < T_A\}$. However, the latter event deals only with the random walk up to time $\xi^{n-\log n}$ and hence we can expect that the conditional distribution of S_{ξ^n} given this event is very close to the unconditional distribution.

We now make this more precise. Note that if $x, t > 0$, then

$$\frac{1}{(x+t)^{1/2}} \geq \frac{1}{x^{1/2}} - \frac{t}{2x^{3/2}} = \frac{1}{x^{1/2}}\left[1 - \frac{t}{2x}\right]. \tag{6.56}$$

We will show the following: there exists an $N < \infty$ and a sequence of positive numbers δ_n with $\sum \delta_n < \infty$ such that the following holds for $n \geq N$. Suppose $A \in \mathcal{A}_3$ satisfies

$$\mathbb{P}\{\xi^n < T_A \mid \xi^{n-\log n} < T_A\} \geq 1 - \frac{\log n}{n}. \tag{6.57}$$

Then,

$$\mathbb{P}\{\xi^{n+1} < T_A \mid \xi^n < T_A\} \leq (1+\delta_n)\left(1 - \frac{1}{2n}\right). \tag{6.58}$$

To see that this suffices, define

$$b_n = n^{1/2} \max_{A \in \mathcal{A}_3} \mathbb{P}\{\xi^n < T_A\}, \quad \overline{b}_n = \max\{b_1, \ldots, b_n\}.$$

Choose integer N sufficiently large so that

$$(n+1)^{-1/2} \geq (n-1-\log n)^{-1/2}\left(1 - \frac{\log n}{n}\right), \quad n \geq N;$$

(6.56) shows that such an N exists. Let $A \in \mathcal{A}_3$. If (6.57) does not hold, Let k be the largest integer less than $n - \log n$. Then,

$$\begin{aligned}
\mathbb{P}\{\xi^{n+1} < T_A\} &\leq \mathbb{P}\{\xi^n < T_A\} \\
&\leq \left(1 - \frac{\log n}{n}\right)\mathbb{P}\{\xi^{n-\log n} < T_A\} \\
&\leq \left(1 - \frac{\log n}{n}\right)\mathbb{P}\{\xi^k < T_A\} \\
&\leq b_k\,(n - \log n - 1)^{-1/2}\left(1 - \frac{\log n}{n}\right) \\
&\leq \overline{b}_n\,(n+1)^{-1/2}.
\end{aligned}$$

If (6.57) holds, then so does (6.58) and

$$\mathbb{P}\{\xi^{n+1} < T_A\} \le (1+\delta_n)\left(1 - \frac{1}{2n}\right)\mathbb{P}\{\xi^n < T_A\}$$

$$\le (1+\delta_n)\,\overline{b}_n\, n^{-1/2}\left(1 - \frac{1}{2n}\right) \le (1+\delta_n)\,\overline{b}_n\,(n+1)^{-1/2}.$$

Therefore, $\overline{b}_{n+1} \le (1+\delta_n)\,\overline{b}_n$, and for all $n \ge N$,

$$\overline{b}_{n+1} \le \overline{b}_N \prod_{j=N}^{n}(1+\delta_j) \le \overline{b}_N \prod_{j=N}^{\infty}(1+\delta_j) < c\, b_N.$$

We now prove (6.58). Note that

$$\mathbb{P}\{T_A < \xi^{n+1} \mid \xi^n < T_A\} = \sum_{w\in\partial\mathcal{B}_{e^n}} \mathbb{P}\{S_{\xi^n} = w \mid \xi^n < T_A\}\,\mathbb{P}^w\{T_A < \xi^{n+1}\}.$$

Lemma 6.7.4 tells us that

$$\mathbb{P}^w\{T_A < \xi^{n+1}\} \ge \frac{1}{2C_3 n}\left[1 + O\left(\frac{1}{n}\right)\right]\sum_{z\in A} G_{\mathcal{B}_{e^{n+1}}}(w,z).$$

Hence, if suffices to prove

$$\sum_{w\in\partial\mathcal{B}_{e^n}} \mathbb{P}\{S_{\xi^n} = w \mid \xi^n < T_A\}\sum_{z\in A} G_{\mathcal{B}_{e^{n+1}}}(w,z) \ge C_3 + s_n,$$

where $\sum n^{-1}s_n < \infty$. If we replaced $\mathbb{P}\{S_{\xi^n} = w \mid \xi_n < T_A\}$ with $\mathbb{P}\{S_{\xi^n} = w\}$, then this would follow from (6.55). Consider three probability measures on $\partial\mathcal{B}_n$:

$$q_1(z) = q_1(z,n,A) = \mathbb{P}\{S_{\xi^n} = w \mid \xi^n < T_A\},$$

$$q_2(z) = q_2(z,n) = \mathbb{P}\{S_{\xi^n} = w\},$$

$$q_3(z) = q_3(z,n,A) = \mathbb{P}\{S_{\xi^n} = w \mid \xi^{n-\log n} < T_A\}.$$

Hence it suffices to show that

$$\sum_{w\in\partial\mathcal{B}_{e^n}} [q_2(w) - q_1(w)] \sum_{z\in A\cap\mathcal{B}_{n+1}} G(w,z) \le c\left(\frac{\log n}{n}\right)^{1/2}. \qquad (6.59)$$

Since the event $\{\xi^{n-\log n} < T_A\}$ only concerns the random walk up to the first time it reaches $\mathcal{B}_{n-\log n}$, (6.24) shows that

$$q_3(w) = q_2(w)\left[1 + O\left(\frac{1}{n}\right)\right].$$

But, if (6.57) holds, we get

$$q_1(w) \le q_3(w)\left[1 + O\left(\frac{\log n}{n}\right)\right] \le q_2(w)\left[1 + O\left(\frac{\log n}{n}\right)\right],$$

which implies (since q_1, q_2 are probability measures supported on $\partial \mathcal{C}_{e^n}$),

$$\sum_{w \in \partial \mathcal{B}_{e^n}} |q_2(w) - q_1(w)| \le O\left(\frac{\log n}{n}\right). \tag{6.60}$$

Recall also that

$$q_2(w) - q_1(w) \le q_2(w) \le c\, e^{-2n}. \tag{6.61}$$

Suppose $z \in A$, and consider

$$\begin{aligned} \sum_{w \in \partial \mathcal{B}_{e^n}} [q_2(w) - q_1(w)]\, G(w,z) &\le c \sum_{w \in \partial \mathcal{B}_{e^n}} \frac{1}{|z-w|+1} \\ &\le c \sup_m \sum_{w \in \partial \mathcal{B}_{e^n}} m(z) \frac{1}{|z-w|+1}, \end{aligned}$$

where the supremum is over all positive measures on $\partial \mathcal{C}_{e^n}$ satisfying

$$\sum_w m(w) \le \frac{\log n}{n}, \qquad m(w) \le e^{-2n}.$$

The maximizing measure gives measure e^{-2n} to the $O(e^n \sqrt{\log n/n})$ points in $\mathcal{B}_{e^n}$ that are closest to z. The number of such points that are distance r to $r+1$ of z is bounded above by cr (and is zero if $r < |z-w|$. Therefore the sum above is bounded above by

$$c \sum_{|z-w| \le r \le e^n \sqrt{\log n/n}} e^{-2n} \le c\, e^{-n} \sqrt{\log n/n}.$$

Since $\#(A \cap \mathcal{B}_{e^{n+1}}) \sim e^{n+1}$, summing over z gives (6.59).

Let

$$L_{j,\pm} = \{\pm j \mathbf{e}_i : j = 0, 1, \ldots\}.$$

Theorem 6.7.6 *Suppose $p \in \mathcal{P}_d, x \in \mathbb{Z}^2$, $J \le 6$ is a positive integer and $A = A_1 \cup \cdots \cup A_J$, where $A_1, \ldots, A_J$ are distinct sets chosen from $L_{1,+}, L_{1,-}, \ldots, L_{3,-}$. Then (except for the case $x = 0, J-6$) there exists a β (depending on x and the choice of $A_1, \ldots, A_J$, such that as $n \to \infty$,*

$$\mathbb{P}^x\{\xi_n < T_A\} \sim c_0, (\log n)^{-J/2}.$$

Moreover $\beta > 0$ unless the left-hand side equals zero for some n.

Comment. This result implies that the events $\{\xi_n < T_{A_j}\}$ are "independent up to a multiplicative constant", i.e.,

$$\mathbb{P}^x\{\xi_n < T_{A_1}, \ldots, \xi_n < T_{A_J}\} = c_1 \prod_{j=1}^J \mathbb{P}^x\{\xi_n < T_{A_j}\}.$$

This relation does not hold for $d = 2$.

Proof We will only do the proof for simple random walk and $J \leq 5$ (if $J = 6$, then $\mathbb{P}\{S_1 \in A\} = 1$.) Let $F(n) = \mathbb{P}\{\xi^n < T_A\}$; it suffices to find a c_0 such that $F(n) \sim c_0/n^{J/2}$. To establish this it suffices to show that

$$\mathbb{P}\{\xi^{n+1} > T_A \mid \xi^n < T_A\} = \frac{J}{n} + O\left(\frac{\log n}{n^2}\right), \tag{6.62}$$

see Exercise 6.9. Let $E_{n,j}$ denote the event

$$E_{n,j} = \{S[\xi^n, \xi^{n+1}] \cap A_j \neq \emptyset\},$$

and $E_n = E_{n,1} \cup \cdots \cup E_{n,J}$. Then to prove (6.62) it suffices to prove that

$$\mathbb{P}(E_n) = \frac{J}{2n} + O\left(\frac{\log n}{n^2}\right), \tag{6.63}$$

and

$$\mathbb{P}(E_n \mid \xi^n < T_A) = \mathbb{P}(E_n)\left[1 + O\left(\frac{\log n}{n}\right)\right]. \tag{6.64}$$

Let

$$V_n = \sum_{j=0}^{\xi^{n+1}-1} 1\{S_j \in A\},$$

denote the number of visits to A before leaving $\mathcal{B}_n$. By (6.55),

$$\mathbb{E}[V_n - V_{n-1}] = J\, C_3 + O(e^{-n}).$$

Let $\sigma_n = \min\{j \geq \xi^n : S_j \in A\}$. Note that

$$\mathbb{E}[V_n - V_{n-1}; |S_{\sigma_n}| < e^n/n^2] \leq \left[\sup_{|x| \geq e^n} \mathbb{P}^x\{S[0,\infty) \cap \mathcal{B}_{e^n/n^3} \neq \emptyset\}\right]\left[\sup_{x \in \mathcal{B}_{e^n/n^2}} \mathbb{E}^x[V_n]\right] \leq \frac{c}{n^2}.$$

Similarly,

$$\mathbb{E}[V_n - V_{n-1}; |S_{\sigma_n}| > e^{n+1}(1 - n^{-3})] \leq \frac{c}{n^2}.$$

Therefore,

$$\mathbb{E}[V_n - V_{n-1}; e^n/n^2 \leq |S_{\sigma_n \wedge \xi^{n+1}}| \leq e^{n+1}(1 - n^{-3})] = J\, C_3 + O\left(frac1n^2\right).$$

However, for $x \in \{z \in A : e^n/n^2 \leq |z| \leq e^{n+1}(1 - n^{-3})\}$,

$$\mathbb{E}^x[V_n] = 2C_3 n + O(\log n).$$

Therefore,

$$\mathbb{P}\{e^n/n^2 \leq |S_{\sigma_n \wedge \xi^{n+1}}| \leq e^{n+1}(1 - n^{-3})\} = \frac{J}{2n}\left[1 + O\left(\frac{\log n}{n}\right)\right],$$

and

$$\mathbb{P}\{\sigma_n < \xi^{n+1}\} = O\left(\frac{1}{n^2}\right) + \mathbb{P}\{e^n/n^2 \leq |S_{\sigma_n \wedge \xi^{n+1}}| \leq e^{n+1}(1 - n^{-3})\}.$$

This establishes (6.63).

The other estimate is done similarly to the preceding proof. We first note (see Exercise 6.8) that there is a c such that if a random walk starts on $\mathcal{B}_{e^m}$, then there is a positive probability that it distance ϵe^m away from A before it moves distance $(1/4)e^m$. Hence (using, e.g., approximation by Brownian motion), we see that for all $x \in \mathcal{B}_{e^m}$,

$$\mathbb{P}^x\{\mathrm{dist}(S_{\xi^{m+1}}, A) \geq \epsilon e^m \mid \xi^{m+1} < T_A\} \geq \epsilon.$$

This implies that $\mathbb{P}\{\xi^{m+2} < T_A\} \geq c\mathbb{P}\{\xi^{m+1} < T_A\}$, and this implies (see Exercise 6.8 again), that for all n and all $z \in \mathcal{B}_n$,

$$\mathbb{P}\{S_{\xi^n} = z \mid \xi^n < T_A\} \leq cn^{-2} \leq c\mathbb{P}\{S_{\xi^n} = z\}.$$

Hence,

$$\mathbb{P}\{\xi^n > T_A \mid \xi^{n-\log n} < T_A\} \leq c\,\mathbb{P}\{S(\xi^{n-\log n}, \xi^n] \cap A \neq \emptyset\}.$$

The latter probability is bounded by $c \log n/n$, and hence as before,

$$\mathbb{P}\{S_{\xi_n} = z \mid \xi^{n-\log n} < T_A\} \leq \mathbb{P}\{S_{\xi_n} = z\} \left[1 + O\left(\frac{\log n}{n}\right)\right].$$

This implies,

$$\mathbb{P}\{S(\xi^n, \xi^{n+1}] \cap A \neq \emptyset \mid \xi^n < T_A\} \leq \mathbb{P}\{S(\xi^n, \xi^{n+1}] \cap A\} \left[1 + O\left(\frac{\log n}{n}\right)\right].$$

The lower bound

$$\mathbb{P}\{S(\xi^n, \xi^{n+1}] \cap A \neq \emptyset \mid \xi^n < T_A\} \geq \mathbb{P}\{S(\xi^n, \xi^{n+1}] \cap A\} \left[1 + O\left(\frac{\log n}{n}\right)\right]$$

is done as in the previous proposition. This establishes (6.64). □

6.8 Eigenvalue of a set

Suppose $p \in \mathcal{P}_d$ and $A \subset \mathbb{Z}^d$ is finite and connected with $\#(A) = m$. The *(first) eigenvalue* of A is defined to be the number $\alpha_A = e^{-\lambda_A}$ such that as $n \to \infty$,

$$\mathbb{P}^x\{\tau_A > n\} \approx \alpha_A^n = e^{-\lambda_A n}.$$

Let P^A denote the $m \times m$ matrix $[p(x,y)]_{x,y \in A}$ and, as before, let $\mathcal{L}^A = P^A - I$. Note that $(P^A)^n$ is the matrix $[p_n^A(x,y)]$ where $p_n^A(x,y) = \mathbb{P}^x\{S_n = y; n < \tau_A\}$. We will say that $p \in \mathcal{P}_d$ is *aperiodic restricted to* A if there exists an n such that $(P^A)^n$ has all entries strictly positive; otherwise, we say that p is *bipartite restricted to* A. In order for p to be aperiodic restricted to A, p must be aperiodic. However, it is possible for p to be aperiodic but for p to be bipartite restricted to A (Exercise 6.13).

Proposition 6.8.1 *If $p \in \mathcal{P}_d$, $A \subset \mathbb{Z}^d$ is finite and connected, and p restricted to A is aperiodic, then there exist numbers $0 < \beta = \beta_A < \alpha = \alpha_A < 1$ such that if $x, y \in A$,*

$$p_n^A(x,y) = \alpha^n\, g_A(x)\, g_A(y) + O_A(\beta^n). \tag{6.65}$$

Here $g_A : A \to \mathbb{R}$ is the unique positive function satisfying

$$\sum_{x\in A} g_A(x)^2 = 1, \qquad P^A g_A(x) = \alpha_A\, g_A(x),\; x \in A.$$

In particular,

$$\mathbb{P}^x\{\tau_A > n\} = \tilde{g}_A(x)\,\alpha^n + O_A(\beta^n),$$

where

$$\tilde{g}_A(x) = g_A(x) \sum_{y\in A} g_A(y),$$

We write O_A to indicate that the implicit constant in the error (as well as β) depends on A.

Proof This is a general fact about irreducible Markov chains, see Proposition 11.4.3. In the notation of that proposition $v = w = g$. Note that

$$\mathbb{P}^x\{\tau_A > n\} = \sum_{y\in A} p_n^A(x,y).$$

. □

Corollary 6.8.2 *If $p \in \mathcal{P}_d$, $A \subset \mathbb{Z}^d$ is finite and connected, and p is bipartite restricted to A, then there exist numbers $0 < \beta = \beta_A < \alpha = \alpha_A < 1$ such that if $x, y \in A$ for all n sufficiently large,*

$$p_n^A(x,y) + p_{n+1}^A(x,y) = 2\,\alpha^n\, g_A(x)\, g_A(y) + O_A(\beta^n).$$

Here $g_A : A \to \mathbb{R}$ is the unique positive function satisfying

$$\sum_{x\in A} g_A(x)^2 = 1, \qquad P^A g_A(x) = \alpha\, g_A(x),\; x \in A.$$

Proof This can be proved similarly using Markov chains. We omit the proof. □

Proposition 6.8.3 *Suppose $p \in \mathcal{P}_d$; $\epsilon \in (0,1)$, and $p_\epsilon = \epsilon\,\delta_0 + (1-\epsilon)\,p$ is the corresponding lazy walker. Suppose A is a finite, connected subset of $\mathbb{Z}^d$ and let $\alpha, \alpha_\epsilon, g, g_\epsilon$ be the eigenvalues and eigenfunctions for A using p, p_ϵ, respectively. Then $1 - \alpha_\epsilon = (1-\epsilon)\,(1-\alpha)$ and $g_\epsilon = g$.*

Proof Let P^A, P_ϵ^A be the corresponding matrices. Then $P_\epsilon^A = (1-\epsilon)\,P^A + \epsilon\, I$ and hence

$$P_\epsilon^A\, g_A = [(1-\epsilon)\,\alpha + \epsilon]\, g_A.$$

□

Remark. A standard problem is to estimate λ_A or α_A as A gets large and $\alpha_A \to 1, \lambda_A \to 0$. In these cases it usually suffices to consider the eigenvalue of the lazy walker with $\epsilon = 1/2$. Indeed let $\tilde{\lambda}_A$ be the eigenvalue for the lazy walker. Since,

$$\lambda_A = 1 - \alpha_A + O((1-\alpha_A)^2), \quad \alpha_A \to 1-.$$

we get

$$\tilde{\lambda}_A = \frac{1}{2}\,\lambda_A + O(\lambda_A^2), \qquad \lambda_A \to 0.$$

Proposition 6.8.1 gives no bounds for the β. The optimal β is the norm of the second eigenvalue which measures the rate of convergence to the eigenvector g_A. In general, it is hard to estimate β and for very bad sets β can be very close to α. We will show that in the case of the nice set $\mathcal{C}_n$ there is an upper bound for β independent of n. We fix $p \in \mathcal{P}_d$ with $p(x,x) > 0$ and let $e^{-\lambda_m} = \alpha_{\mathcal{C}_m}, g_m = g_{\mathcal{C}_m}$, and $p_n^m(x,y) = p_n^{\mathcal{C}_m}(x,y)$. For $x \in \mathcal{C}_m$ we let

$$\rho_m(x) = \frac{\operatorname{dist}(x, \partial\mathcal{C}_m) + 1}{m},$$

and we set $\rho_m \equiv 0$ on $\mathbb{Z}^d \setminus \mathcal{C}_m$.

Lemma 6.8.4 *There exist c_1, c_2 such that for all m sufficiently large and all $x, y \in \mathcal{C}_m$,*

$$c_1\, \rho_m(x)\, \rho_m(y) \le m^d\, p_{m^2}^m(x,y) \le c_2\, \rho_m(x)\, \rho_m(y).$$

Also, there exist c_3, c_4 such that for every $n \ge m^2$, and all $x, y \in \mathcal{C}_m$,

$$c_3\, \rho_m(y)\, m^{-d} \le \mathbb{P}^x\{S_n = y \mid \tau_{\mathcal{C}_m} > n\} \le c_4\, \rho_m(y)\, m^{-d}.$$

Proof For notational ease, we will restrict to the case where m is even. (If m is odd, essentially the same proof works except $m^2/4$ must be replaced with $\lfloor m^2/4 \rfloor$, etc.) Then,

$$p_{m^2}^m(x,y) = \sum_{z,w} p_{m^2/4}^m(x,z)\, p_{m^2/2}^m(z,w)\, p_{m^2/4}^m(w,y).$$

The local central limit theorem implies that there is a c such that for all z, w, $p_{m^2/2}^m(z,w) \le p_{m^2/2}(z,w) \le c\, m^{-d}$. Therefore,

$$p_{m^2}^m(x,y) \le m^{-d}\, \mathbb{P}^x\{\tau_{\mathcal{C}_m} > m^2/4\}\; \mathbb{P}^y\{\tau_{\mathcal{C}_m} > m^2/4\}.$$

Gambler's ruin (see Theorem 5.1.5 and the remark immediately preceding the theorem) implies that $\mathbb{P}^x\{\tau_{\mathcal{C}_m} > m^2/4\} \le c\, \rho_m(x)$. This gives the upper bound.

For the lower bound, we first note that there is an $\epsilon > 0$ such that

$$p_{\lfloor \epsilon m^2 \rfloor}^m(z,w) \ge \epsilon\, m^{-d}, \qquad |z|, |w| \le \epsilon\, m.$$

Indeed, the Markov property implies that

$$p_{\lfloor \epsilon m^2 \rfloor}^m(z,w) \ge p_{\lfloor \epsilon m^2 \rfloor}(z,w) - \max\{p_k(\tilde{z}, w) : k \le \lfloor \epsilon m^2 \rfloor, \tilde{z} \in \mathbb{Z}^d \setminus \mathcal{C}_m\},$$

and the local central limit theorem establishes the estimate. Using this estimate and the invariance principle, one case see that for every $\epsilon > 0$, there is a c such that for $z, w \in \mathcal{C}_{(1-\epsilon)m}$,

$$p_{m^2/2}(z,w) \ge c\, n^{-d}.$$

SOME ESTIMATE LATER

The last assertion follows for $n = m^2$ by noting that

$$\mathbb{P}^x\{S_{m^2} = y \mid \tau_{\mathcal{C}_m} > m^2\} = \frac{p_{m^2}^m(x,y)}{\sum_z p_{m^2}^m(x,z)}$$

and

$$\sum_z p^m_{m^2}(x,z) \asymp \rho_m(x)\, m^{-d} \sum_z \rho_m(z) \asymp \rho_m(x).$$

For $n > m^2$, we can do similarly by conditioning on the walk at time $n - m^2$. $\square$

Given this, we can use standard facts about Markov chains, see Proposition 11.4.4.

Exercises

Exercise 6.1 Show that Proposition 6.1.2 holds for $p \in \mathcal{P}^*$.

Exercise 6.2 In this exercise we construct a transient subset A of $\mathbb{Z}^3$ with

$$\sum_{y\in A} G(0,y) = \infty. \tag{6.66}$$

Here G denotes the Green's function for simple random walk. Our set will be of the form

$$A = \bigcup_{k=1}^{\infty} A_k, \qquad A_k = \{z \in \mathbb{Z}^3 : |z - 2^k\, \mathbf{e}_1| \le \epsilon_k\, 2^k\}.$$

for some $\epsilon_k \to 0$.

(a) Show that (6.66) holds if and only if $\sum_{k=1}^{\infty} \epsilon_k^3\, 2^{2k} = \infty$.
(b) Show that A is transient if and only if $\sum_{k=1}^{\infty} \epsilon_k < \infty$.
(c) Find a transient A satsifying (6.66).

Exercise 6.3 Suppose $p \in \mathcal{P}_d, d \ge 3$. Show that there exists a sequence $K_n \to \infty$ such that if $A \subset \mathbb{Z}^d$ is a finite set with at least n points, then $\mathrm{cap}(A) \ge K_n$.

Exercise 6.4 Show that if $p \in \mathcal{P}_2$ and $r > 0$,

$$\lim_{n\to\infty} [G_{\mathcal{C}_{n+r}}(0,0) - G_{\mathcal{C}_n}(0,0)] = 0.$$

Use this and (6.16) to conclude that for all x, y,

$$\lim_{n\to\infty} [G_{\mathcal{C}_n}(0,0) - G_{\mathcal{C}_n}(x,y)] = a(x,y).$$

Exercise 6.5 Suppose $p \in \mathcal{P}_d$ and $A \subset \mathbb{Z}^d$ is finite. Define

$$Q_A(f,g) = \sum_{x,y\in\overline{A}} p(x,y)\,[f(y) - f(x)]\,[g(y) - g(x)].$$

and $Q_A(f) = Q_A(f,f)$. Let $F : \partial A \to \mathbb{R}$ be given. Show that the infimum of $Q_A(f)$ restricted to functions $f : \overline{A} \to \mathbb{R}$ with $f \equiv F$ on ∂A is obtained by the unique harmonic function with boundary value F.

Exercise 6.6

Write the two-dimensional integer lattice in complex form, $\mathbb{Z}^2 = \mathbb{Z} + i\mathbb{Z}$ and let A be the upper half plane $A = \{j + ik \in \mathbb{Z}^2 : k > 0\}$. Show that for simple random walk

$$G_A(x,y) = a(\overline{x}, y) - a(x,y), \quad x, y \in A,$$

$$H_A(x,j) = \frac{1}{4}\,[a(x, j-i) - a(x, j+i)] + \delta(x-j), \quad x \in \overline{A}, j \in \mathbb{Z}.$$

where $\overline{j+ik} = j - ik$ denotes complex conjugate. Find

$$\lim_{k \to \infty} k\, H_A(ik, j).$$

Exercise 6.7 Supply the details of the Harnack inequality argument in Lemma 6.5.8.

Exercise 6.8 Suppose $p \in \mathcal{P}_d$.

a. Show that there is a $c < \infty$ such that if $x \in A \subset \mathcal{C}_n$ and $z \in \mathcal{C}_{2n}$,

$$\mathbb{P}^x\{S_{\xi_{2n}} = z \mid \xi_{2n} < T_A\} \le c\, n^{1-d}\, \frac{\mathbb{P}^x\{\xi_n < T_A\}}{\mathbb{P}^x\{\xi_{2n} < T_A\}}.$$

b. Let A be the line $\{j\mathbf{e}_1 : j \in \mathbb{Z}\}$. Show that there is an $\epsilon > 0$ such that for all n sufficiently large,

$$\mathbb{P}\{\mathrm{dist}(S_{\xi_n}, A) \ge \epsilon n \mid \xi_n < T_A\} \ge \epsilon.$$

(Hint: you can use the gambler's ruin estimate to estimate $\mathbb{P}^x\{\xi_{n/2} < T_A\}/\mathbb{P}^x\{\xi_n < T_A\}$.)

Exercise 6.9 Suppose $\alpha > 0$ and $F(n)$ is a sequence of positive numbers. Define δ_n by

$$F(n+1) = F(n)\, \left[1 - \frac{\alpha}{n} + \delta_n\right].$$

Show that if $\sum |\delta_n| < \infty$, then there exists a $c > 0$ such that as $n \to \infty$,

$$F(n) = c\, n^{-\alpha}\, [1 + O(r_n)],$$

where $r_n = \sum_{m \ge n} |\delta_n|$.

Exercise 6.10 Show that for each $p \in \mathcal{P}_2$ and each $r \in (0,1)$, there is a c such that for all n sufficiently large,

$$G_{\mathcal{C}_n}(x,y) \ge c, \quad x, y \in \mathcal{C}_{rn},$$

$$G_{\mathcal{B}_n}(x,y) \ge c, \quad x, y \in \mathcal{B}_{rn}.$$

Exercise 6.11 Suppose $p \in \mathcal{P}_2$ and let $A = \{x_1, x_2\}$ be a two-point set.

a. Prove that $\mathrm{hm}_A(x_1) = 1/2$.

b. Show that there is a $c < \infty$ such that if $A \subset \mathcal{C}_n$, then for $y \in \mathbb{Z}^2 \setminus \mathcal{C}_{2n}$,

$$\left| \mathbb{P}^y\{S_{T_A} = x_1\} - \frac{1}{2} \right| \le \frac{c}{\log n}.$$

(Hint: Suppose $\mathbb{P}^y\{S_{T_A} = x_j\} \ge 1/2$ and let V be the set of z such that $\mathbb{P}^z\{S_{T_A} = x_j\} \le 1/2$. Let $\sigma = \min\{j : S_j \in V\}$. Then it suffices to prove that $\mathbb{P}^y\{T_A < \sigma\} \le c/\log n$.)

c. Show that there is a $c < \infty$ such that if $A = \mathbb{Z}^2 \setminus \{x\}$ with $x \neq 0$, then

$$\left| G_A(0,0) - \frac{4}{\pi} \log |x| \right| \leq c.$$

Exercise 6.12 Suppose $p \in \mathcal{P}_2$. Show that there exist $c_1, c_2 > 0$ such that the following holds.

a. If n is sufficiently large, A is a set as in Lemma 6.6.7, and $A_n = A \cap \{|z| \geq n/2\}$, then for $x \in \partial\mathcal{B}_{n/2}$,

$$\mathbb{P}^x\{T_A < \xi_n^*\} \geq c.$$

b. If $x \in \partial\mathcal{B}_{n/2}$,

$$G_{\mathbb{Z}^2\setminus A}(x,0) \leq c.$$

c. If A' is a set with $\mathcal{B}_{n/2} \subset A' \subset \mathbb{Z}^2 \setminus A_n$,

$$\left| G_{A'}(0,0) - \frac{2}{\pi} \log n \right| \leq c.$$

Exercise 6.13 Give an example of an aperiodic $p \in \mathcal{P}_d$ and a finite connected (with respect to p) set A for which p is bipartite restricted to A.

7
Dyadic coupling

7.1 Introduction

In this chapter, we will study the *dyadic or KMT coupling* which is a coupling of Brownian motion and random walk for which the paths are signficantly closer to each other than than in the Skorokhod embedding. Recall that if (S_n, B_n) are coupled by the Skorokhod embedding, then typically one expects $|S_n - B_n|$ to be of order $n^{1/4}$. In the dyadic coupling, $|S_n - B_n|$ will be of order $\log n$. We mainly restrict our consideration to one dimension, although we discuss some higher dimensional versions in Section 7.6.

Suppose $p \in \mathcal{P}_1'$ and

$$S_n = X_1 + \cdots + X_n$$

is a p-walk. Suppose that there exists $b > 0$ such that

$$\mathbb{E}[X_1^2] = \sigma^2, \quad \mathbb{E}[e^{b|X_1|}] < \infty. \tag{7.1}$$

Then by Theorem 2.3.11, there exist N, c, ϵ such that if we define $\delta(n,x)$ by

$$p_n(x) := \mathbb{P}\{S_n = x\} = \frac{1}{\sqrt{2\pi\sigma^2 n}}\, e^{-\frac{x^2}{2\sigma^2 n}}\, \exp\{\delta(n,x)\},$$

then for all $n \geq N$ and $|x| \leq \epsilon n$,

$$|\delta(n,x)| \leq c\left[\frac{1}{\sqrt{n}} + \frac{|x|^3}{n^2}\right]. \tag{7.2}$$

Theorem 7.1.1 *Suppose $p \in \mathcal{P}_d'$ satisfies (7.1) and (7.2). Then one can define on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a Brownian motion B_t with variance parameter σ^2 and a random walk with increment distribution p such that the following holds. For each $\alpha < \infty$, there is a c_α such that*

$$\mathbb{P}\left\{\max_{1\leq j\leq n} |S_j - B_j| \geq c_\alpha \log n\right\} \leq c_\alpha\, n^{-\alpha}. \tag{7.3}$$

Remark. From the theorem, it is easy to conclude the corresponding result for bipartite or continuous-time walks with $p \in \mathcal{P}_1$. In particular, the result holds for discrete-time and continuous-time simple random walk.

Comment. We will describe the dyadic coupling formally in Section 7.4, but we will give a basic idea here. Suppose that $n = 2^m$. One starts by defining S_{2^m} as closely to B_{2^m} as possible. Using the local central limit theorem, we can do this in a way so that with very high probability $|S_{2^m} - B_{2^m}|$ is of order 1. We then define $S_{2^{m-1}}$ using the values of $B_{2^m}, B_{2^{m-1}}$, and again get an error of order 1. We keep subdividing intervals using binary splitting, and every time we construct the value of S at the middle point of a new interval. If at each subdivision we get an error of order 1, the total error should be at most of order m, the number of subdivisions needed. (Typically it might be less because of cancellation.)

Comment. The assumption $\mathbb{E}[e^{b|X_1|}] < \infty$ for some $b > 0$ is necessary for (7.3) to hold at $j = 1$. Suppose $p \in \mathcal{P}_1'$ such that for each n there is a coupling with

$$\mathbb{P}\{|S_1 - B_1| \geq \hat{c} \log n\} \leq \hat{c}\, n^{-1}.$$

It is not difficult to show that as $n \to \infty$, $\mathbb{P}\{|B_1| \geq \hat{c} \log n\} = o(n^{-1})$, and hence

$$\mathbb{P}\{|S_1| \geq 2\hat{c} \log n\} \leq \mathbb{P}\{|S_1 - B_1| \geq \hat{c} \log n\} + \mathbb{P}\{|B_1| \geq \hat{c} \log n\} \leq 2\, \hat{c}\, n^{-1}$$

for n sufficiently large. If we let $x = 2\hat{c} \log n$, this becomes

$$\mathbb{P}\{|X_1| \geq x\} \leq 2\hat{c}\, e^{-x/(2\hat{c})},$$

for all x sufficiently large which implies $\mathbb{E}[e^{b|X_1|}] < \infty$ for $b < (2\hat{c})^{-1}$.

The coupling will be defined in Section 7.4. Estimates will be needed to show that the coupling satisfies (7.3). The proof is the essentially the same for all values of σ^2. For ease of notation we will assume that $\sigma^2 = 1$. It also suffices to prove the result for $n = 2^m$ and we will assume this in Sections 7.4 and 7.5.

For the remainder of this chapter, we fix b, ϵ, c_0, N and assume that p is an increment distribution satisfying

$$\mathbb{E}[e^{b|X_1|}] < \infty, \tag{7.4}$$

and if $\delta(n, x)$ is defined by

$$p_n(x) = \frac{1}{\sqrt{2\pi n}}\, e^{-\frac{x^2}{2n}} \exp\{\delta(n, x)\},$$

then

$$|\delta(n, x)| \leq c_0 \left[\frac{1}{\sqrt{n}} + \frac{|x|^3}{n^2}\right], \quad n \geq N,\ |x| \leq \epsilon n. \tag{7.5}$$

7.2 Some estimates

In this section we collect a few lemmas about random walk that will be used in establishing (7.3). The reader may wish to skip this section at first reading and come back to the estimates as they are needed.

Lemma 7.2.1 *Suppose S_n is a random walk with increment distribution p satisfying (7.4) and (7.5). Define $\delta_n^*(n, x, y)$ by*

$$\mathbb{P}\{S_n = x \mid S_{2n} = y\} = \frac{1}{\sqrt{\pi n}} \exp\left\{-\frac{(x - (y/2))^2}{n}\right\} \exp\{\delta_*(n, x, y)\}.$$

Then if $n \geq N$, $|x|,|y| \leq \epsilon n/2$,

$$|\delta_*(n,x,y)| \leq 9\, c_0 \left[\frac{1}{\sqrt{n}} + \frac{|x|^3}{n^2} + \frac{|y|^3}{n^2}\right].$$

Comment. Without the conditioning, S_n is approximately normal with mean zero and variance n. Conditioned on the event $S_{2n} = y$, S_n is approximately normal with mean $y/2$ and variance $n/2$. Note that specifying the value at time $2n$ reduces the variance of S_n.

Proof Note that

$$\mathbb{P}\{S_n = x \mid S_{2n} = y\} = \frac{\mathbb{P}\{S_n = x, S_{2n} - S_n = y - x\}}{\mathbb{P}\{S_{2n} = y\}} = \frac{p_n(x)\, p_n(y-x)}{p_{2n}(y)}.$$

Since $|x|,|y|,|x-y| \leq \epsilon n$, we can apply (7.5). Note that

$$|\delta_*(n,x,y)| \leq |\delta(n,x)| + |\delta(n,y-x)| + |\delta(2n,y)|.$$

We use the simple estimate $|y-x|^3 \leq 8(|x|^3 + |y|^3)$. □

Lemma 7.2.2 *If $x_1, x_2, \ldots, x_n \in \mathbb{R}$, then*

$$\sum_{j=1}^{n} \frac{(x_1 + \cdots + x_j)^2}{2^j} \leq 2\, \frac{\sqrt{2}+1}{\sqrt{2}-1} \left[\sum_{j=1}^{n} \frac{x_j^2}{2^j}\right]. \tag{7.6}$$

Proof By homogeneity we may assume that $\sum 2^{-j}\, x_j^2 = 1$. Let $y_j = 2^{-j/2}\, x_j, \mathbf{y} = (y_1, \ldots, y_n)$. Then

$$\begin{aligned}
\sum_{i=1}^{n} \frac{(x_1 + \cdots + x_i)^2}{2^i} &= \sum_{i=1}^{n} \sum_{1 \leq j,k \leq i} \frac{x_j x_k}{2^i} \\
&= \sum_{j=1}^{n} \sum_{k=1}^{n} x_j x_k \sum_{i=j \vee k}^{n} 2^{-i} \\
&\leq 2 \sum_{j=1}^{n} \sum_{k=1}^{n} 2^{-(j \vee k)}\, x_j x_k \\
&= 2 \sum_{j=1}^{n} \sum_{k=1}^{n} 2^{-|k-j|/2}\, y_j y_k \\
&= 2 \langle A\mathbf{y}, \mathbf{y} \rangle \leq 2\lambda\, \|\mathbf{y}\|^2 = 2\lambda,
\end{aligned}$$

where $A = A_n$ is the $n \times n$ symmetric matrix with entries $a(j,k) = 2^{-|k-j|/2}$ and $\lambda = \lambda_n$ denotes the largest eigenvalue of A. Since λ is bounded by the maximum of the row sums,

$$\lambda \leq 1 + 2 \sum_{j=1}^{\infty} 2^{-j/2} = \frac{\sqrt{2}+1}{\sqrt{2}-1}.$$

□

Comment. We will use the fact that the left-hand side of (7.6) is bounded by a constant times the term in brackets on the right-hand side. The exact constant is not important.

Lemma 7.2.3 *Suppose S_n is a random walk with increment distribution satisfying (7.4) and (7.5). Then for every α there exists a $c = c(\alpha)$ such that*

$$\mathbb{P}\left\{\sum_{\log_2 n < j \leq n} \frac{S_{2^j}^2}{2^j} \geq c\, n\right\} \leq c\, e^{-\alpha n}. \tag{7.7}$$

Comment. Consider the random variables $U_j = S_{2^j}^2/2^j$ and note that $\mathbb{E}[U_j] = 1$. Suppose that $U_1, U_2, \ldots$ were independent. If it were also true that there exist t, c such that $\mathbb{E}[e^{tU_j}] \leq c$ for all j, then (7.7) would be a standard large deviation estimate similar to Theorem 11.2.5. To handle the lack of independence, we consider the independent random variables $[S_{2^j} - S_{2^{j-1}}]^2/2^j$ and use (7.6). If the increment distribution is bounded then we also get $\mathbb{E}[e^{tU_j}] \leq c$ for some t, see Exercise 2.6. However, if the range is infinite this expectation may be infinite for all $t > 0$, see Exercise 7.3. To overcome this difficulty, we use a striaghtforward truncation argument.

Proof We fix $\alpha > 0$ and allow constants in this proof to depend on α. Using (7.4), we see that there is a β such that

$$\mathbb{P}\{|S_n| \geq n\} \leq e^{-\beta n}.$$

Hence, we can find c_1 such that

$$\sum_{\log_2 n < j \leq n} [\, \mathbb{P}\{|S_{2^j}| \geq c_1 2^j\} + \mathbb{P}\{|S_{2^j} - S_{2^{j-1}}| \geq c_1 2^j\}\,] = O(e^{-\alpha n}).$$

Fix this c_1, and let j_0 be the smallest integer greater than $\log_2 n$. Let $Y_j = 0$ for $j < j_0$; $Y_{j_0} = S_{2^{j_0}}$; and for $j > j_0$, let $Y_j = S_{2^j} - S_{2^{j-1}}$. Then, except for an event of probability $O(e^{-\alpha n})$, $|Y_j| \leq c_1 2^j$ for $j \geq j_0$ and hence

$$\mathbb{P}\left\{\sum_{j=j_0}^{n} \frac{Y_j^2}{2^j} \neq \sum_{j=j_0}^{n} \frac{Y_j^2}{2^j}\, 1\{|Y_j| \leq c_1 2^j\}\right\} \leq O(e^{-\alpha n}).$$

Note that

$$\sum_{\log_2 n < j \leq n} \frac{S_{2^j}^2}{2^j} = \sum_{j=j_0}^{n} \frac{S_{2^j}^2}{2^j} = \sum_{j=1}^{n} \frac{(Y_1 + \cdots + Y_j)^2}{2^j} \leq c \sum_{j=j_0}^{n} \frac{Y_j^2}{2^j}.$$

The last step uses (7.6). Therefore it suffices to prove that

$$\mathbb{P}\left\{\sum_{j=1}^{n} \frac{Y_j^2}{2^j}\, 1\{|Y_j| \leq c_1 2^j\} \geq cn\right\} \leq e^{-\alpha n}.$$

The estimates (7.4) and (7.5) imply that there is a $t > 0$ such that for each n,

$$\mathbb{E}\left[\exp\left\{\frac{t\, S_n^2}{n}\right\};\, |S_n| \leq c_1 n\right] \leq e$$

(see Exercise 7.2). Therefore,

$$\mathbb{E}\left[\exp\left\{t\sum_{j=1}^{n}\frac{Y_j^2}{2^j}\,1\{|Y_j|\le c_1 2^j\}\right\}\right]\le e^n,$$

which implies

$$\mathbb{P}\left\{\sum_{j=1}^{n}\frac{Y_j^2}{2^j}\,1\{|Y_j|\le c_1 2^j\}\ge t^{-1}\,(\alpha+1)\,n\right\}\le e^{-\alpha n}.$$

$\square$

7.3 Quantile coupling

In this section we consider the simpler problem of coupling S_n and B_n for a fixed n. The following definition is more general than we need. We will only use the quantile coupling when F is supported on $\mathbb{Z}$ or $(1/2)\mathbb{Z}$.

Definition. Suppose F is the distribution function of a discrete random variable supported on the locally finite set

$$\cdots < a_{-1} < a_0 < a_1 < \cdots,$$

and Z is a random variable with a continuous, strictly increasing distribution function G. Let r_k be defined by $G(r_k) = F(a_k)$, i.e., if $F(a_k) > F(a_k-)$,

$$G(r_k) - G(r_{k-1}) = F(a_k) - F(a_k-).$$

Let f be the step function

$$f(z) = a_k \text{ if } r_{k-1} < z \le r_k,$$

and let X be the random variable $f(Z)$. We call X the *quantile coupling* of F with Z and f the *quantile coupling function* of F and G.

Note that the event $\{X = a_k\}$ is the same as the event $\{r_{k-1} < Z \le r_k\}$. Hence,

$$\mathbb{P}\{X = a_k\} = \mathbb{P}\{r_{k-1} < Z \le r_k\} = G(r_k) - G(r_{k-1}) = F(a_k) - F(a_k-),$$

and X has distribution function F. Also, if

$$G(a_k - t) \le F(a_{k-1}) < F(a_k) \le G(a_k + t), \tag{7.8}$$

then on the event $\{X = a_k\}$, $|X - Z| = |a_k - Z| \le t$. Hence, if we wish to prove that $|X - Z| \le t$ on the event $\{X = a_k\}$, it suffices to establish (7.8).

We will need results concerning the closeness of the quantile coupling for the random walk and normal random variables. Let Φ denote the standard normal distribution function, and let Φ_{σ^2} denote the distribution function of a mean zero normal random variable with variance σ^2.

Proposition 7.3.1 *For every ϵ, b, c_0, N there exist c, δ such that if S_n is a random walk with increment distribution p satisfying (7.4) and (7.5) the following holds for $n \ge N$. Let F_n denote* *the distribution function of* S_n*, and suppose* Z *has distribution function* Φ_n*. Let* (X, Z) *be the quantile coupling of* F_n *with* Z*, then*

$$|X - Z| \leq c \left[1 + \frac{X^2}{n}\right], \quad |X| \leq \delta n.$$

Proposition 7.3.2 *For every* ϵ, b, c_0, N *there exist* c, δ *such that the following holds for* $n \geq N$*. Suppose* S_n *is a random walk with increment distribution* p *satisfying (7.4) and (7.5). Suppose* $|y| \leq \delta n$ *with* $\mathbb{P}\{S_{2n} = y\} > 0$*. Let* $F_{n,y}$ *denote the conditional distribution function of* $S_n - [S_{2n}/2]$ *given* $S_{2n} = y$*, and suppose* Z *has distribution function* $\Phi_{n/2}$*. Let* (X, Z) *be the quantile coupling of* $F_{n,y}$ *with* Z*, Then,*

$$|X - Z| \leq c \left[1 + \frac{X^2}{n} + \frac{y^2}{n}\right], \quad |X|, |y| \leq \delta n.$$

Using (7.8), we can see that to prove these propositions it suffices to prove the following estimate on the distribution functions.

Lemma 7.3.3 *For every* ϵ, b, c_0, N *there exist* c, δ *such that if* S_n *is a random walk with increment distribution* p *satisfying (7.4) and (7.5) the following holds for* $n \geq N$*. Let* $F_n, F_{n,y}$ *be as in the propositions above. Then for* $|x|, |y| \leq \delta n$*.*

$$\Phi_n \left(x - c\left[1 + \frac{x^2}{n}\right]\right) \leq F_n(x-1) \leq F_n(x) \leq \Phi_n\left(x + c\left[1 + \frac{x^2}{n}\right]\right). \tag{7.9}$$

$$\Phi_{n/2} \left(x - c\left[1 + \frac{x^2}{n} + \frac{y^2}{n}\right]\right) \leq F_{n,y}(x-1) \leq F_{n,y}(x) \leq \Phi_{n/2}\left(x + c\left[1 + \frac{x^2}{n} + \frac{y^2}{n}\right]\right).$$

Proof This is a straightforward estimate of sums and integrals and we leave the details as an exercise (Exercise 7.4). It suffices to prove the result for integer $x \geq 0$. Constants in this proof are allowed to depend on ϵ, b, c_0 and we assume $n \geq N$. If F is a distribution function, we write $\overline{F} = 1 - F$. Note that (7.4) and (7.5) imply that we can write

$$p_n(x) = \int_{x-1}^{x} \frac{1}{\sqrt{2\pi n}}\, e^{-t^2/(2n)} \exp\left\{O\left(\frac{1}{\sqrt{n}} + \frac{t^3}{n^2}\right)\right\} dt, \quad |x| \leq \epsilon n,$$

and hence for some a and all $|x| \leq \epsilon n$,

$$\overline{F}_n(x) = O(e^{-an}) + \int_{x}^{\epsilon n} \frac{1}{\sqrt{2\pi n}}\, e^{-t^2/(2n)} \exp\left\{O\left(\frac{1}{\sqrt{n}} + \frac{t^3}{n^2}\right)\right\} dt.$$

From this we can conclude that for $|x| \leq \epsilon n$,

$$\overline{F}_n(x) = \overline{\Phi}_n(x) \exp\left\{O\left(\frac{1}{\sqrt{n}} + \frac{x^3}{n^2}\right)\right\}, \tag{7.10}$$

and from this we can conclude (7.9). The second inequality is done similarly by using Lemma 7.2.1 to derive

$$\overline{F}_{n,y}(x) = \overline{\Phi}_{n/2}(x) \exp\left\{O\left(\frac{1}{\sqrt{n}} + \frac{x^3}{n^2} + \frac{y^3}{n^2}\right)\right\},$$

for $|x|, |y| \leq \delta n$.

$\square$

Comment. To derive Propositions 7.3.1 and Proposition 7.3.2 we use only estimates on the distribution functions $F_n, F_{n,y}$ and not pointwise estimates (local central limit theorem). However, the pointwise estimate (7.5) is used in the proof of Lemma 7.2.1 which is what we use to estimate $F_{n,y}$.

7.4 The dyadic coupling

In this section we will define the dyadic coupling. Throughout this section we fix $n = 2^m$. We will start with a Brownian motion B_t and will define the random variables $S_1, S_2, \ldots, S_{2^m}$ as functions of the random variables $B_1, B_2, \ldots, B_{2^m}$.

In Chapter 3, we constructed a Brownian motion from a collection of indepedent normal random variables by a dyadic construction. Here we will start with the Brownian motion, B_t, and obtain the independent normals. We will only use the random variables $B_1, B_2, \ldots, B_{2^m}$. Define $\Gamma_{k,j}$ by

$$\Gamma_{k,j} = B_{k2^{m-j}} - B_{(k-1)2^{m-j}}, \qquad j = 0, 1, \ldots, m; \;\; k = 1, 2, \ldots, 2^{m-j}.$$

For each j, $\{\Gamma_{k,j} : k = 1, 2, \ldots, 2^{m-j}\}$ are independent normal random variables with mean zero and variance 2^j. Let $Z_{1,0} = B_{2^m}$ and define

$$Z_{2k+1,j}, \quad j = 1, \ldots, m, \;\; k = 0, 1, \ldots, 2^{m-1} - 1,$$

recursively by

$$\Gamma_{2k+1,j} = \frac{1}{2}\Gamma_{k,j-1} + Z_{2k+1,j},$$

and hence

$$\Gamma_{2k+2,j} = \frac{1}{2}\Gamma_{k,j-1} - Z_{2k+1,j}.$$

One can check (see Corollary 11.3.1) that the random variables $\{Z_{2k+1,j} : j = 0, \ldots, 2^m, \; k = 0, 1, \ldots, 2^{m-1} - 1\}$ are independent, mean zero, normal random variables with $\mathbb{E}[Z_{1,0}^2] = 2^m$ and $\mathbb{E}[Z_{2k+1,j}^2] = 2^{m-j-1}$ for $j \geq 1$.

Let $f_m(\cdot)$ denote the quantile coupling function for the distribution functions of S_{2^m} and B_{2^m}. If $y \in \mathbb{Z}$, let $f_j(\cdot; y)$ denote the quantile coupling function for the conditional distribution of

$$S_{2^j} - \frac{1}{2} S_{2^{j+1}}$$

given $S_{2^{j+1}} = y$ and a normal random variable with mean zero and variance 2^{j-1}. This is well defined if $\mathbb{P}\{S_{2^{j+1}} = y\} > 0$ which is all that we need. Note that the range of $f_j(\cdot; y)$ is contained in $(1/2)\mathbb{Z}$. This conditional distribution is symmetric about the origin, see Exercise 7.1. This implies that $f_j(-z; y) = -f_j(z; y)$. We can now define the coupling.

- Let $S_{2^m} = f_m(B_{2^n})$.
- Suppose the values of $S_{l2^{m-j-1}}, l = 1, \ldots, 2^{j+1}$ are known. Let

$$\Delta_{k,j} = S_{k2^{m-j}} - S_{(k-1)2^{m-j}}.$$

Then we let

$$S_{(2k+1)2^{m-j}} = \frac{1}{2}\left[S_{(k-1)2^{m-j+1}} + S_{k2^{m-j+1}}\right] + f_{m-j}(Z_{2k+1,j}, \Delta_{k,j-1}),$$

so that

$$\Delta_{2k+1,j} = \frac{1}{2}\Delta_{k+1,j-1} + f_{m-j}(Z_{2k+1,j}, \Delta_{k+1,j-1}),$$

$$\Delta_{2k+2,j} = \frac{1}{2}\Delta_{k+1,j-1} - f_{m-j}(Z_{2k+1,j}, \Delta_{k+1,j-1}).$$

It follows immediately from the defintion that $S_1, S_2, \ldots, S_{2^m}$ have the distribution of the random walk with increment p. Also Exercise 7.1 shows that $\Delta_{2k+1,j}$ and $\Delta_{2k+2,j}$ have the same condtional distribution given $\Delta_{k+1,j-1}$.

We give some important properties of the coupling.

- The random variables $\Gamma_{2k+1,j}\Delta_{2k+1,j}, \Gamma_{2k+2,j}, \Delta_{2k+2,j}$ are deterministic functions of $\Gamma_{k+1,j-1}$, $\Delta_{k+1,j-1}$, $Z_{2k+1,j}$. Given $(\Gamma_{k+1,j-1}, \Delta_{k+1,j-1})$, the conditional distributions of $(\Gamma_{2k+1,j}, \Delta_{2k+1,j})$ and $(\Gamma_{2k+2,j}, \Delta_{2k+2,j})$ are the same. Also, the condtional distribution of $(\Gamma_{2k+1,j}, \Delta_{2k+1,j})$ and $(\Gamma_{2k+2,j}, \Delta_{2k+2,j})$ given $(\Gamma_{k+1,j-1}, \Delta_{k+1,j-1})$ is the same as the conditional distribution given

$$\left\{(\Gamma_{i,l}, \Delta_{i,l}) : l = 0, \ldots, j-1, i = 0, \ldots, 2^{m-l}\right\}.$$

- Let

$$R_{k,j} := \sum_{l=0}^{j} |\Delta_{k_l,l} - \Gamma_{k_l,l}|,$$

where k_l is the integer satisfying

$$(k_l - 1)\, 2^{m-l} < k2^{m-j} \le k_l 2^{m-l}.$$

By the last comment and induction, we can see that the random variables $R_{1,j}, \ldots, R_{2^{m-j},j}$ are identically distributed. (They are not independent.)

- Note that

$$S_{(2k+1)2^{m-j}} - B_{(2k+1)2^{m-j}} = [S_{k2^{m-j+1}} - B_{k2^{m-j+1}}] + [\Delta_{2k+1,j} - \Gamma_{2k+1,j}].$$

By iterating, we can see that

$$|S_{k2^{m-j}} - B_{k2^{m-j}}| \le R_{k,j}. \tag{7.11}$$

We also see by induction that the random variables

$$S_{(2k+1)2^{m-j}} - B_{(2k+1)2^{m-j}}, \quad m = 0, 1, \ldots, 2^{j-1} - 1,$$

are identically distributed.

- For $k = 1$,

$$\begin{aligned}\Delta_{1,j} - \Gamma_{1,j} &= S_{2^{m-j}} - B_{2^{m-j}} \\ &= \frac{1}{2}\left[S_{2^{m-j-1}} - B_{2^{m-j-1}}\right] + f_{m-j}(Z_{1,j}; S_{2^{m-j-1}}) - Z_{1,j}.\end{aligned}$$

In particular,

$$R_{1,j} \leq 2 \left[|S_{2^m} - B_{2^m}| + \sum_{l=1}^{j} |f_{m-j}(Z_{1,l}; S_{2^{m-l-1}}) - Z_{1,l}| \right]. \tag{7.12}$$

7.5 Proof of Theorem 7.1.1

Recall that $n = 2^m$. It suffices to show that for each α there is a c_α such that for each integer j,

$$\mathbb{P}\left\{|S_i - B_i| \geq c_\alpha \log n\right\} \leq c_\alpha\, n^{-\alpha}. \tag{7.13}$$

Indeed if the above holds, then

$$\mathbb{P}\left\{\max_{1 \leq i \leq n} |S_i - B_i| \geq c_\alpha \log n\right\} \leq \sum_{i=1}^{n} \mathbb{P}\left\{|S_i - B_i| \geq c_\alpha \log n\right\} \leq c_\alpha\, n^{-\alpha+1}.$$

We claim in fact, that it suffices to find a sequence $0 = i_0 < i_1 < \cdots < i_l = n$ such that $|i_k - i_{k-1}| \leq c_\alpha \log n$ and such that (7.13) holds for these indices. Indeed, if we prove this and $|j - i_k| \leq c_\alpha \log n$, then exponential estimates show that there is a c'_α such that

$$\mathbb{P}\{|S_j - S_{i_k}| \geq c'_\alpha \log n\} + \mathbb{P}\{|B_j - B_{i_k}| \geq c'_\alpha \log n\} \leq c'_\alpha\, n^{-\alpha},$$

and hence the triangle inequality gives (7.13) (with a different constant).

For the remainder of this section we fix α and allow constants to depend on α. By the the previous paragraph and (7.11), it suffices to find a c such that for $\log_2 m + c \leq j \leq m$, and $k = 1, \ldots, 2^{m-j}$,

$$\mathbb{P}\left\{R_{k,j} \geq c_\alpha\, m\right\} \leq c_\alpha\, e^{-\alpha m},$$

and as pointed out, it suffices to prove this for $k = 1$, i.e., for $j = \log_2 m + c, \ldots, m$,

$$\mathbb{P}\left\{R_{1,j} \geq c_\alpha\, m\right\} \leq c_\alpha\, e^{-\alpha m}. \tag{7.14}$$

This is the estimate we will derive.

Let δ be the minimum of the two values in Propositions 7.3.1 and 7.3.2, and recall that there is a $\beta = \beta(\delta)$ such that

$$\mathbb{P}\{|S_{2^j}| \geq \delta 2^j\} \leq \exp\{-\beta 2^j\}$$

In particular, we can find a c_3 such that

$$\sum_{\log_2 m + c_3 \leq j \leq m} \mathbb{P}\{|S_{2^j}| \geq \delta 2^j\} \leq O(e^{-\alpha m}).$$

Proposition 7.3.1 tells us that on the event $\{S_{2^m}| \leq \delta 2^m$,

$$|S_{2^m} - B_{2^m}| \leq c \left[1 + \frac{S_{2^m}^2}{2^m}\right].$$

Proposition 7.3.2 tells us that on the event $\{|S_{2^j}|, |S_{2^{j+1}}| \leq \delta 2^j\}$,

$$|Z_{1,j+1} - J_{1,j+1}| \leq c \left[1 + \frac{S_{2^{j+1}}^2 + S_{2^j}^2}{2^j}\right].$$

Hence, by (7.12), we see that for $\log_2 m + c_3 \leq i \leq m$,

$$|S_{2^i} - B_{2^i}| \leq R_{1,i} \leq c \left[m + \sum_{\log_2 m \leq j \leq m} \frac{S_{2^j}^2}{2^j} \right].$$

We now use (7.7) to conclude (7.14) for $j \geq \log_2 m + c_3$.

7.6 Higher dimensions

We will not try to extend the arguments to higher dimensions in this book, but we do mention two results that come out immediately.

Theorem 7.6.1 *One can define on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a Brownian motion B_t in $\mathbb{R}^2$ with covariance matrix $(1/2)\, I$ and a simple random walk in $\mathbb{Z}^2$. such that the following holds. For each $\alpha < \infty$, there is a c_α such that*

$$\mathbb{P}\left\{ \max_{1 \leq j \leq n} |S_j - B_j| \geq c_\alpha \log n \right\} \leq c_\alpha\, n^{-\alpha}.$$

Proof We use the trick from Exercise 1.6. Let $(S_{n,1}, B_{n,1}), (S_{n,2}, B_{n,2})$ be independent dyadic couplings of one-dimensional simple random walk and Brownian motion. Let

$$S_n = \left(\frac{S_{n,1} + S_{n,2}}{2}, \frac{S_{n,1} - S_{n,2}}{2} \right),$$

$$B_n = \left(\frac{B_{n,1} + B_{n,2}}{2}, \frac{B_{n,1} - B_{n,2}}{2} \right).$$

$\square$

Theorem 7.6.2 *If $p \in \mathcal{P}_d$, one can define on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a Brownian motion B_t in $\mathbb{R}^d$ with covariance matrix Γ and a continuous-time random walk $\tilde{S}_t$ with increment distribution p such that the following holds. For each $\alpha < \infty$, there is a c_α such that*

$$\mathbb{P}\left\{ \max_{1 \leq j \leq n} |\tilde{S}_j - B_j| \geq c_\alpha \log n \right\} \leq c_\alpha\, n^{-\alpha}. \tag{7.15}$$

Proof Recall from (1.3) that we can write any such $\tilde{S}_t$ as

$$\tilde{S}_t = \tilde{S}^1_{q_1 t}\, x_1 + \cdots + \tilde{S}^l_{q_l t}\, x_l,$$

where $q_1, \ldots, q_l > 0; x_1, \ldots, x_l \in \mathbb{Z}^d$; and $\tilde{S}^1, \ldots, \tilde{S}^l$ are independent one-dimensional simple continuous-time random walks. Choose l independent couplings as in Theorem 7.1.1,

$$(S^1_t, B^1_t), (S^2_t, B^2_t), \ldots, (S^l_t, B^l_t),$$

where $B^1, \ldots, B^l$ are standard Brownian motions. Let

$$B_t = B^1_{q_1 t}\, x_1 + \cdots + B^l_{q_l t}\, x_l.$$

This satisfies (7.15). $\square$

7.7 Coupling first visits

Proposition 7.7.1 *Suppose $p \in \mathcal{P}_d$. Then one can define on the same probability a (discrete-time) random walk S_n with increment distribution p; a continuous-time random walk $\tilde{S}_t$ with increment distribution p; and a Brownian motion B_t with covariance matrix Γ such that for each $n, r > 0$,*

$$\mathbb{P}\{|S_{\xi_n} - B_{\xi'_n}| \geq r \log n\} = \mathbb{P}\{|\tilde{S}_{\tilde{\xi}_n} - B_{\xi'_n}| \geq r \log n\} \leq \frac{c}{r}.$$

Here,

$$\xi_n = \min\{j : \mathcal{J}(S_j) \geq n\}, \quad \tilde{\xi}_n = \min\{t : \mathcal{J}(\tilde{S}_t) \geq n\}, \quad \xi'_n = \min\{t : \mathcal{J}(B_t) = n\}.$$

Proof It suffices to prove the result for $\tilde{S}_t, B_t$, for then we can define S_j to be the discrete-time walk obtained from $\tilde{S}_t$ "when it moves". By Theorem 7.6.2 we can define $\tilde{S}, B$ on the same probability space such that except for an event of probability $O(n^{-4})$,

$$|\tilde{S}_t - B_t| \leq c_1 \log n, \quad 0 \leq t \leq n^3.$$

We claim that $\mathbb{P}\{\tilde{\xi}_n > n^3\}$ decay exponentially in n. Indeed, the central limit theorem shows that there is a $c > 0$ such that for for n sufficiently large and $|x| < n$, $\mathbb{P}^x\{\tilde{\xi}_n \leq n^2\} \geq c$. Iterating this gives $\mathbb{P}^x\{\tilde{\xi}_n > n^3\} \leq (1-c)^n$. Similarly, $\mathbb{P}\{\xi'_n > n^3\}$ decays exponentially. Therefore, except on an event of probability $O(n^{-4})$,

$$|\tilde{S}_t - B_t| \leq c_1 \log n, \quad 0 \leq t \leq \max\{\tilde{\xi}_n, \xi'_n\}. \tag{7.16}$$

Let

$$\sigma_n = \min\{t : \operatorname{dist}(\tilde{S}_t, \mathbb{Z}^d \setminus \mathcal{C}_n) \leq 2\, c_1 \log n\}.$$

Using the gambler's ruin estimate, we see that

$$\mathbb{P}\{|S_{\sigma_n} - S_j| \geq r \log n \text{ for some } \sigma_n \leq j \leq \tilde{\xi}_n\} \leq \frac{c_2}{r}.$$

If we define σ'_n similarly for Brownian motion, we get

$$\mathbb{P}\{|B_{\sigma'_n} - B_t| \geq r \log n \text{ for some } \sigma'_n \leq t \leq \xi'_n\} \leq \frac{c_2}{r}.$$

We now let

$$\rho_n = \min\{t : \operatorname{dist}(\tilde{S}_t, \partial\mathcal{C}_n) \leq c_1 \log n \text{ or } \operatorname{dist}(B_t, \partial\mathcal{C}_n) \leq c_1 \log n\},$$

and note that on the event that (7.16) holds,

$$\sigma_n \wedge \sigma'_n \leq \rho_n \leq \tilde{\xi}_n \wedge \xi'_n,$$

□

Definition. A finite subset A of $\mathbb{Z}^d$ is *simply connected* if A and $\mathbb{Z}^d \setminus A$ are connected. If $x \in \mathbb{Z}^d$, let $\mathcal{S}_x$ denote the closed cube in $\mathbb{R}^d$ of side length one, centered at x, with sides perpendicular to the coordinate axes. If $A \subset \mathbb{Z}^d$, let D_A be the domain defined as the interior of $\cup_{x \in A} \mathcal{S}_x$.. The *inradius* of A is defined by

$$\operatorname{inrad}(A) = \min\{|y| : y \in \mathbb{Z}^d \setminus A\}.$$

Proposition 7.7.2 *Suppose $p \in \mathcal{P}_2$. Then one can define on the same probability a (discrete-time) random walk S_n with increment distribution p and a Brownian motion B_t with covariance matrix Γ such that the following holds. If A is a finite, simply connected set containing the origin and*

$$\rho_A = \inf\{t : B_t \notin D_A\},$$

then each if $r > 0$,

$$\mathbb{P}\{|S_{\tau_A} - B_{\rho_A}| \geq r \log[\mathrm{inrad}(A)]\} \leq \frac{c}{\sqrt{r}}.$$

Proof Similar to the last proposition except that the gambler's ruin estimate is replaced with the Beurling estimate. □

Exercises

Exercise 7.1 Suppose $S_n = X_1 + \cdots + X_n$ where $X_1, X_2, \ldots$ are independent, identically distributed random variables. Suppose $\mathbb{P}\{S_{2n} = 2y\} > 0$ for some $y \in \mathbb{R}$. Show that the conditional distribution of

$$S_n - y$$

conditioned on $\{S_{2n} = 2y\}$ is symmetric about the origin.

Exercise 7.2 Suppose S_n is a random walk in $\mathbb{Z}$ whose increment distribution satisfies (7.4) and (7.5) and let $C < \infty$. Show that there exists a $t = t(b, \epsilon, c_0, C) > 0$ such that for all n,

$$\mathbb{E}\left[\exp\left\{\frac{t\, S_n^2}{n}\right\}; |S_n| \leq Cn\right] \leq e.$$

Exercise 7.3 Suppose $\tilde{S}_t$ is continuous-time simple random walk in $\mathbb{Z}$.

(i) Show that there is a $c > 0$ such that for all integers n,

$$\mathbb{P}\{\tilde{S}_n = n^2\} \geq c^{-1} \exp\{-cn^2 \log n\}.$$

(ii) Show that if $t > 0$,

$$\mathbb{E}\left[\exp\left\{\frac{t\, \tilde{S}_n^2}{n}\right\}\right] = \infty,$$

Exercise 7.4 Let Φ be the normal distribution function, and let $\overline{\Phi} = 1 - \Phi$.

(i) Show that as $x \to \infty$,

$$\bar{\Phi}(x) \sim \frac{e^{-x^2/2}}{\sqrt{2\pi}} \int_0^\infty e^{-xt}\, dt = \frac{1}{x\sqrt{2\pi}}\, e^{-x^2/2}.$$

(ii) Show that there is a c such that for all $0 \leq t \leq x < \infty$,

$$\overline{\Phi}(x - t) - \overline{\Phi}(x) \geq c\, x\, t\, \overline{\Phi}(x).$$

(iii) Show that there is a c such that for all $x > 0$,

$$\overline{\Phi}(2x) \leq c\, e^{-x^2}\, \overline{\Phi}(x).$$

(iv) Show that there exists an $a > 0$ such that if $x \geq 1$ and $0 \leq t \leq x$,

$$\overline{\Phi}(x + t) \leq e^{-axt}\, \overline{\Phi}(x) \leq e^{axt}\overline{\Phi}(x) \leq \overline{\Phi}(x - t).$$

(v) Prove (7.10).

(vi) Prove (7.9).

8
Addtional topics on Simple Random Walk

8.1 Poisson kernel for a cube

8.2 Approximating continuous harmonic functions

Discrete harmonic functions appropriately scaled become continuous harmonic functions. We will discuss some quantitative versions of this principle in this chapter. We let $\mathcal{U} = \mathcal{U}_d = \{x \in \mathbb{R}^d : |x| < 1\}$ denote the unit ball in $\mathbb{R}^d$.

Proposition 8.2.1 *There exists c such that the following is true for all positive integers n, m. Suppose $f : (n+m)\mathcal{U} \to \mathbb{R}$ is a harmonic function with $\|f\|_\infty \leq 1$. Then there is a function $\hat{f}$ on $\overline{\mathcal{B}}_n$ with $\mathcal{L}\hat{f}(x) = 0, x \in \mathcal{B}_n$ and such that*

$$|f(x) - \hat{f}(x)| \leq \frac{c}{m^2}, \quad x \in \mathcal{B}_n. \tag{8.1}$$

In fact, we can choose

$$\hat{f}(x) = \mathbb{E}^x[f(S_{\xi_n})].$$

Proof Note that f is defined on $(n+1)\mathcal{U}_d$ and hence $\hat{f}$ is well defined. By definition we know that $\mathcal{L}\hat{f}(x) = 0, x \in \mathcal{B}_n$. We need to prove (8.1). By (6.9), we know that

$$f(x) = \mathbb{E}^x\left[f(S_{\xi_n}) - \sum_{j=0}^{\xi_n - 1} \mathcal{L}f(S_j)\right] = \hat{f}(x) - \phi(x),$$

where

$$\phi(x) = \sum_{z \in \mathcal{B}_n} G_{\mathcal{B}_n}(x, z)\, \mathcal{L}f(z).$$

We have already noted that there is a c such that all 4th order derivatives of f at x are bounded above by $c\,(n+m-|x|)^{-4}$. Therefore,

$$|\phi(x)| \leq c \sum_{k=0}^{n-1} \sum_{k \leq |z| < k+1} \frac{G(x,z)}{(n+m-k)^4}.$$

We claim that there is a c_1 such that for all x,

$$\sum_{n-l\leq k\leq n-1} G_{\mathcal{B}_n}(x,z) \leq c\, l^2.$$

Indeed, the proof of this is essentially the same as the proof of (5.4). Once we have this, a standard summation by parts gives that $|\phi(x)| \leq c/m^2$. □

8.3 Diffusion Limited Aggregation (DLA)

8.4 Interior Diffusion Limited Aggregation (IDLA)

8.5 The cone exponent

Definition. Suppose D is a nonempty, open, connected subset of the unit sphere $\{u \in \mathbb{R}^d : |u| = 1\}$ whose closure is a proper subset of the sphere. Then the *cone* associated to D is the open, connected set

$$V = V_D = \{ru : u \in D, r > 0\}.$$

We will call V a *standard cone* if D is of the form $\{u : |u| = 1, |u - u_0| < \epsilon\}$ for some $|u_0| = 1$, $\epsilon > 0$.

Note that if V is a cone and $r > 0$, then $rV = V$. In this section, we study the probability that a simple random walk stays in V up to a specified time. It will be easier to first consider the analogous probability for Brownian motion. It is easier to study Brownian motion, because Brownian motion satisfies a scaling relation.

Let B_t denote a standard Brownian motion and let

$$T_n = \inf\{t \geq 0 : |B_t| = n\}, \quad T^n = T_{e^n},$$

$$\sigma_r = \inf\{t \geq 0 : |B_t - B_0| = r\}.$$

The *cone exponent* for V, $\theta = \theta(V)$, is defined by the relation

$$\mathbb{P}^x\{B[0, T_n] \subset V\} \approx n^{-\theta}, \quad n \to \infty, x \in V,$$

assuming such a θ exists. The scaling relation for Brownian motion implies that for $r > 0$,

$$\mathbb{P}^{rx}\{B[0, T_{rn}] \subset V\} = \mathbb{P}^x\{B[0, T_n] \subset V\}.$$

If V is a cone, we let $V_n = \{x \in V : |x| \leq n\}, V^n = V_{e^n}$, and if $x \in V^0$, let μ_x be the conditional distribution of B_{T^1} conditioned on the event $\{B[0, T^0] \subset V\}$.

We will say that the cone V is *nice* if it satisfies the following conditions:

- There exist c_1, α_1 such that for all $z \in V$,

$$\mathbb{P}^z\{B[0, T_1] \subset V\} \geq c_1\, [\mathrm{dist}(z, \partial V)]^{\alpha_1}. \tag{8.2}$$

- For $\delta > 0$, let

$$V_-(\delta) = \{z \in V : \mathrm{dist}(z, \partial V) \geq \delta\},$$

$$V_+(\delta) = V \cup \{z \in \mathbb{R}^d \setminus V : \mathrm{dist}(z, \partial V) \leq \delta\}.$$

There exist c_2, α_2 such that if $z \notin V_-(\delta)$, then

$$\mathbb{P}^z\{B[0,\sigma_r] \subset V_+(\delta)\} \le c_2\,(\delta/r)^{\alpha_2}. \tag{8.3}$$

In particular, there exists a c_3 such that if $z \notin V_-(\delta)$, then

$$\mathbb{P}^z\{S[0,\sigma_{c_3\delta}] \subset V_+(\delta)\} \le 1/2. \tag{8.4}$$

Remark.

- Suppose there exists $c' > 0$ such that all $z \in \partial V_-(\delta)$, the probability that a Brownian reaches $V_-(2\delta)$ without leaving V is at least c'. Then (8.2) holds. Indeed, if we let

$$j_n = \inf \mathbb{P}^z\{B[0,T_1] \subset V\},$$

  where the infinmum is over all z with $\mathrm{dist}(z,\partial V) \ge 2^{-n}$, then $j_{n+1} \ge c'\, j_n$.
- Similarly, suppose there exist c', K such that for all $z \in V \setminus V_-(\delta)$,

$$\mathbb{P}^z\{S[0,\sigma_{K\delta}] \subset V_+(\delta)\} \le 1 - c'.$$

  Then by iterating we get (8.3).

The next proposition gives a useful criterion for showing existence of θ. If V is a cone, we let $V_n = \{x \in V : |x| \le n\}, V^n = V_{e^n}$, and if $x \in V^0$, let μ_x be the conditional distribution of B_{T^0} conditioned on the event $\{B[0,T^0] \subset V\}$.

Proposition 8.5.1 *Suppose V is a cone. Suppose there exists a probability measure μ with $\mu(D) = 1$ such that if*

$$\epsilon_n = \sup\{\|\mu_x - \mu\| : x \in V, |x| \le e^{-n}\},$$

then

$$\sum_{n=1}^{\infty} \epsilon_n < \infty.$$

Here $\|\cdot\|$ denotes total variation distance. Then if $x \in V$,

$$\mathbb{P}^x\{B[0,T^n] \subset V\} \sim c(x)\, e^{-n\theta}, \qquad n \to \infty, \tag{8.5}$$

where θ is defined by

$$\int_D \mathbb{P}^y\{B[0,T^1] \subset V\}\, d\mu(y) = e^{-\theta}.$$

In fact, if

$$\delta_n = \sum_{m \ge n} \epsilon_m,$$

then for $x \in V_1$, we have

$$\mathbb{P}^x\{B[0,T^n] \subset V\} = c(x)\, e^{-n\theta}\,[1 + O(\delta_n)].$$

Proof By scaling, we may assume $x \in V_1$. Let E_n be the event $\{B[0,T^n] \subset V\}$. Then

$$\mathbb{P}^x(E_{n+1} \mid E_n) = \mathbb{P}^{xe^{-n}}(E_1 \mid E_0) = \int_D \mathbb{P}^y\{B[0,T^1] \subset V\}\, d\mu_x(y) = e^{-\theta}\,[1+O(\epsilon_n)].$$

Therefore,

$$\mathbb{P}^x(E_n) = e^{-n\theta} \prod_{j=0}^{n-1} [1+O(\epsilon_j)].$$

This tells us that the limit

$$c(x) = \lim_{n\to\infty} \mathbb{P}^x(E_n)$$

exists and

$$\mathbb{P}^x(E_n) = c(x)\, e^{n\theta} \prod_{j=n}^{\infty} [1+O(\epsilon_j)] = c(x) e^{n\theta}\,[1+O(\delta_n)].$$

□

Remark. This proposition assumed that n was integer, but it is not very difficult to show that (8.5) holds for noninteger n as well.

It is easier to study cone exponents for Brownian motion than for random walk because Brownian motion satisfies an exact scaling rule. In order to derive estimates for random walks it is often easiest to first derive them for Brownian motion and then use coupling. The next proposition gives one form of such a result. If V is a cone and $\delta > 0$, we let

$$V_-(\delta) = \{z \in V : \mathrm{dist}(z, \partial V) \geq \delta\},$$

$$V_+(\delta) = V \cup \{z \in \mathbb{R}^d \setminus V : \mathrm{dist}(z, \partial V) \leq \delta\}.$$

Note that $V_-(\delta) \subset V \subset V_+(\delta)$.

8.5.1 The Brownian exponent

Throughout this section we assume that V is a nice cone satisfying (8.2) and (8.3). All constants may depend on V. Without loss of generality we will assume that $\mathbf{e}_1 \in V$. We let $\epsilon_1 = \mathrm{dist}(\mathbf{e}_1, \partial V)/2$ and U the closed ball of radius ϵ_1 about $\mathbf{e}_1$.

Comment. The results and proofs in this section can be compared to those for Markov chains conditioned to stay in a set forever; see Section 11.4

Theorem 8.5.2 *Suppose V is a nice cone with $\mathbf{e}_1 \in V$. Then there exist $\theta, \alpha > 0$ such that for all $z \in V_1$, there is a $c(z) > 0$ such that if $z \in V_1$,*

$$\mathbb{P}^z\{B[0,T_n] \subset V\} = c(z)\, n^{-\theta}\,[1+O(n^{-\alpha})].$$

Moreover there is a c such that for all $z \in V_1, n > 0$,

$$\mathbb{P}^z\{B[0,T_n] \subset V\} \leq c\, n^{-\theta},$$

$$\mathbb{P}^{\mathbf{e}_1}\{B[0,T_n] \subset V\} \geq c^{-1}\, n^{-\theta}.$$

We will prove this in a series of lemmas. The first two lemmas show that a Brownian motion conditioned to stay in V is unlikely to stay close to ∂V for the entire time.

Lemma 8.5.3 *There exist c, a such that for n sufficiently large, if $z \in V_{1/2} \cap V_-(2^{-(n+1)})$, then*

$$\mathbb{P}^z\{B[0, \sigma_{2^{-n}n^2}] \cap V_-(2^{-(n-1)}) = \emptyset \mid B[0, T_1] \subset V\} \le e^{-an^2}.$$

Proof Choose n sufficiently large so that $2^{-n}n^2 \le 1/4$. Then by iterating (8.4), we can see that there is an a such that

$$\mathbb{P}^z\{B[0, \sigma_{2^{-n}n^2}] \cap V_-(2^{-(n-1)}) = \emptyset; B[0, \sigma_{2^{-n}n^2}] \subset V\} \le e^{-2an^2}.$$

But, (8.2) gives

$$\mathbb{P}^z\{B[0, T_1] \subset V\} \ge c_1\, 2^{-\alpha_1(n+1)}.$$

Hence,

$$\mathbb{P}^z\{B[0, \sigma_{2^{-n}n^2}] \cap V_-(2^{-(n-1)}) = \emptyset \mid B[0, T_1] \subset V\} \le \frac{e^{-2an^2}}{c_1\, 2^{-\alpha_1(n+1)}} \le e^{-an^2},$$

for n sufficiently large. □

Lemma 8.5.4 *There exists δ_1 such that for all $z \in V_{1/2}$,*

$$\mathbb{P}^z\{B[0, T_{3/4}] \cap V_-(\delta_1) = \emptyset \mid B[0, T_1] \subset V\} \le 1 - \delta_1.$$

Proof Choose N so that $\sum_{n \ge N} n^2\, 2^{-n} \le 1/8$. For $n \ge N$, let

$$r_n = \inf \mathbb{P}^z\{B[0, T_{3/4}] \cap V_-(2^{-N}) = \emptyset \mid B[0, T_1] \subset V\},$$

where the infimum is over all $z \in V_-(2^{-n})$ with

$$|z| \le \frac{3}{4} - \sum_{j=N+1}^{n} j^2\, 2^{-j}.$$

The Harnack inequality for Brownian motion shows that $r_n > 0$ for each n. But the previous lemma and the strong Markov property (applied to Brownian motion conditioned so that $B[0, T_1] \subset V$) imply that

$$r_n \ge r_{n-1}\, [1 - O(e^{-an^2})],$$

from which we can conclude $\inf_n r_n > 0$. □

The next lemma is an example of a *boundary Harnack principle* and is analogous to results sometimes called *separation lemmas*. It shows that Brownian motion conditioned to stay in V has a reasonable chance of being near $\mathbf{e}_1$.

Lemma 8.5.5 *There exists an $s > 0$ such that if $z \in V_{1/2}$,*

$$\mathbb{P}^z\{B(T_1) \in U \mid S[0, T_1] \subset V\} \ge s.$$

In fact, if $g(z,\cdot)$ denotes the conditional density of $B(T_1)$ with respect to surface measure on $\partial\mathcal{B}_1$ given $B[0,T_1] \subset V$, then there exist c_1, c_2 such that for all $z \in V_{1/2}$ and all $w \in U$,

$$c_1 \leq g(z,w) \leq c_2.$$

Proof For $z \in V_-(\delta_1) \cap \partial\mathcal{B}_{3/4}$, this is a straightforward application of the Harnack principle. For other z we can use the previous lemma. □

In particular, since $\inf_{w\in U} \mathbb{P}^w\{B[0,T^1] \subset V\} > 0$, we see that there is a c_4 such that for all $z \in \mathcal{B}_1$ and all $n \geq 1$,

$$\mathbb{P}^z\{B[0,T^{n+1}] \subset V \mid B[0,T^n] \subset V\} \geq c_4.$$

Let

$$u_n = \sup_{|z|\leq 1} \mathbb{P}^z\{B[0,T^n] \subset V\},$$

Lemma 8.5.6 *There exists a δ such that for $n \geq 1$,*

$$u_n = \sup_{|z|=1, z\in V_-(\delta)} \mathbb{P}^z\{B[0,T^n] \subset V\}.$$

Moreover, there exists c_1, c_2 such that for $|z| = 1, z \in V_-(\delta)$,

$$c_1\, u_n \leq \mathbb{P}^z\{B[0,T^n] \subset V\} \leq c_2\, u_n.$$

Proof The strong Markov property and scaling imply that

$$\mathbb{P}^z\{S[0,T^n] \subset V\} \leq \mathbb{P}^z\{S[0,T^1] \subset V\}\, u_{n-1} \leq c_3\, \mathbb{P}^z\{S[0,T^1] \subset V\} \leq u_n.$$

But we can find δ such that for all $z \in V \setminus V_-(\delta)$, $\mathbb{P}^z\{B[0,T^1] \subset V\} < 1/c_3$. This gives the first assertion and the second assertion follows from the Harnack principle. □

Combining the last two lemmas we get

Lemma 8.5.7 *There exists an $s > 0$ such that if $z \in V$ with $|z| \leq 1$ and $n \geq 1$,*

$$\mathbb{P}^z\{S(T_1) \in U \mid S[0,T^n] \subset V\} \geq s.$$

In fact, if $g(z,\cdot)$ denotes the conditional density of $S(T_1)$ with respect to surface measure on $\partial\mathcal{B}_1$ given $S[0,T_1] \subset V$, then there exist c_1, c_2 such that for all $z \in V, |z| \leq 1/2$ and all $w \in U$,

$$c_1 \leq g(z,w) \leq c_2.$$

We now finish the proof of Theorem 8.5.2. Consider Brownian motions started at $z, w \in V_{e^{-n}}$ conditioned so that $B[0,T^0] \subset V$. Let $\mu_{k,z}, \mu_{k,w}$ denote the conditional distribution of $B_{T^{-k}}$ given $B[0,T^0] \subset V$. Then the previous lemma tells us that there is a $\rho > 0$ such that

$$\|\mu_{k,z} - \mu_{k,w}\| \leq (1-\rho)\, \|\mu_{k+1,z} - \mu_{k+1,w}\|.$$

Iterating this, we get

$$\|\mu_{0,z} - \mu_{0,w}\| \leq (1-\rho)^n.$$

Using this, we can deduce that the hypotheses of Proposition 8.5.1 hold.

8.5.2 Random walk exponent

To estimate the probability that a random walk stays in the cone V, we use a coupling with Brownian motion to reduce the question to an estimate about Brownian motions.

Proposition 8.5.8 *For every V and every $\alpha > 0$, there exists a c such that the following is true. Suppose n is a positive integer, and define the events $E_+ = E_+(V,n), E_- = E_-(V,n)$ by*

$$E_\pm = \{B[0, T_{n \mp c \log n}] \subset V_\pm(c \log n)\}.$$

Then for every $x \in V \cap \mathcal{B}_n$,

$$\mathbb{P}^x(E_-) - \frac{c}{n^\alpha} \leq \mathbb{P}^x\{S[0,\xi_n] \subset U\} \leq \mathbb{P}^x(E_+) + \frac{c}{n^\alpha}.$$

Proof It suffices to prove the result for n sufficiently large and for the continuous time random walk $\tilde{S}$ with stopping times $\tilde{\xi}_n$. By the dyadic coupling we can define $\tilde{S}_t, B_t$ on the same probability space so that if $J = J_c$ is the event

$$J = \left\{ \sup_{0 \leq t \leq \tilde{\xi}_{2n} \vee T_{2n}} |\tilde{S}_t - B_t| \geq c \log n \right\},$$

then for some c, $\mathbb{P}(J) \geq 1 - cn^{-\alpha}$. But for n sufficiently large,

$$E_- \cap J \subset \{\tilde{S}[0, \tilde{\xi}_n] \subset V\} \cap J \subset E_+ \cap J.$$

$\square$

Proposition 8.5.9 *There exist c_1, β such that if $z \in V_1$, then*

$$\mathbb{P}^z(E_+ \setminus E_-) \leq c_1 \, n^{-\theta-\beta}.$$

9

Intersection Probabilities for Random Walks

9.1 Long range estimate

In this section we will prove a fundamental inequality concerning the probability of intersection of the paths of two random walks. If S_n is a random walk, we will write

$$S[n_1, n_2] = \{S_n : n_1 \leq n \leq n_2\}.$$

Proposition 9.1.1 *If $p \in \mathcal{P}_d$, there exist c_1, c_2 such that for all n,*

$$\begin{aligned} c_1\, \phi(n) &\leq \mathbb{P}\{S[0,n] \cap S[2n,3n] \neq \emptyset\} \\ &\leq \mathbb{P}\{S[0,n] \cap S[2n,\infty) \neq \emptyset\} \leq c_2\, \phi(n), \end{aligned}$$

where

$$\phi(n) = \begin{cases} 1, & d < 4, \\ (\log n)^{-1}, & d = 4, \\ n^{(4-d)/2}, & d > 4. \end{cases}$$

Proof The upper bound is trivial for $d \leq 3$, and the lower bound for $d \leq 2$ follows from the lower bound for $d = 3$. Hence we can assume that $d \geq 3$. We will assume the walk is aperiodic (only a trivial modification is needed for the bipartite case). The basic strategy is to consider the number of intersections of the paths,

$$J_n = \sum_{j=0}^{n} \sum_{k=2n}^{3n} 1\{S_j = S_k\}, \qquad K_n = \sum_{j=0}^{n} \sum_{k=2n}^{\infty} 1\{S_j = S_k\}.$$

Note that

$$\mathbb{P}\{S[0,n] \cap S[2n,3n] \neq \emptyset\} = \mathbb{P}\{J_n \geq 1\},$$

$$\mathbb{P}\{S[0,n] \cap S[2n,\infty) \neq \emptyset\} = \mathbb{P}\{K_n \geq 1\}.$$

We will derive the following inequalities for $d \geq 3$,

$$c_1\, n^{(4-d)/2} \leq \mathbb{E}(J_n) \leq \mathbb{E}(K_n) \leq c_2\, n^{(4-d)/2}, \tag{9.1}$$

$$\mathbb{E}(J_n^2) \leq \begin{cases} c\,n, & d=3, \\ c\,\log n, & d=4, \\ c\,n^{(4-d)/2}, & d \geq 5. \end{cases} \tag{9.2}$$

Once these are established, the lower bound follows by the second moment lemma (Lemma 11.6.1),

$$\mathbb{P}\{J_n > 0\} \geq \frac{\mathbb{E}(J_n)^2}{\mathbb{E}(J_n^2)}.$$

Let us write $p(n)$ for $\mathbb{P}\{S_n = 0\}$. Then,

$$\mathbb{E}(J_n) = \sum_{j=0}^{n} \sum_{k=2n}^{3n} p(k-j),$$

and similarly for $\mathbb{E}(K_n)$. Since $p(k-j) \asymp (k-j)^{-d/2}$, we get

$$\mathbb{E}(J_n) \asymp \sum_{j=0}^{n} \sum_{k=2n}^{3n} \frac{1}{(k-j)^{d/2}} \asymp \sum_{j=0}^{n} \sum_{k=2n}^{3n} \frac{1}{(k-n)^{d/2}} \asymp \sum_{j=0}^{n} n^{1-(d/2)} \asymp n^{2-(d/2)},$$

and similarly for $\mathbb{E}(K_n)$. This gives (9.1). To bound the second moments, note that

$$\begin{aligned} \mathbb{E}(J_n^2) &= \sum_{0 \leq j,i \leq n} \; \sum_{2n \leq k,m \leq 3n} \mathbb{P}\{S_j = S_k, S_i = S_m\} \\ &\leq 2 \sum_{0 \leq j \leq i \leq n} \; \sum_{2n \leq k \leq m \leq 3n} [\mathbb{P}\{S_j = S_k, S_i = S_m\} + \mathbb{P}\{S_j = S_m, S_i = S_k\}]. \end{aligned}$$

If $0 \leq i,j \leq n$ and $2n \leq k \leq m \leq 3n$, then

$$\begin{aligned} \mathbb{P}\{S_j = S_k, S_i = S_m\} &\leq \left[\max_{l \geq n, x \in \mathbb{Z}^d} \mathbb{P}\{S_l = x\}\right] \left[\max_{x \in \mathbb{Z}^d} \mathbb{P}\{S_{m-k} = x\}\right] \\ &\leq \frac{c}{n^{d/2}\,(m-k+1)^{d/2}}. \end{aligned}$$

Therefore,

$$\mathbb{E}(J_n^2) \leq c\,n^2 \sum_{2n \leq k \leq m \leq 3n} \frac{1}{n^{d/2}\,(m-k+1)^{d/2}} \leq c\,n^{2-(d/2)} \sum_{0 \leq k \leq m \leq n} \frac{1}{(m-k+1)^{d/2}}.$$

This yields (9.2).

The upper bound is trivial for $d = 3$ and for $d \geq 5$ it follows from (9.1) and the inequality $\mathbb{P}\{K_n \geq 1\} \leq \mathbb{E}[K_n]$. Assume $d = 4$. We will consider $\mathbb{E}[K_n \mid K_n \geq 1]$. On the event $\{K_n \geq 1\}$, let k the the smallest integer $\geq 2n$ such that $S_k \in S[0,n]$. Let j be the smallest index such that $S_k = S_j$. Then by the Markov property, given $[S_0, \ldots, S_k]$, the expected value of K_{2n} is

$$\sum_{i=0}^{n} \sum_{l=k}^{\infty} \mathbb{P}\{S_l = S_i \mid S_k = S_j\} = \sum_{i=0}^{n} G(S_i - S_j).$$

Define a random variable, depending on $S_0, \ldots, S_n$,

$$Y_n = \min_{j=0,\ldots,n} \sum_{i=0}^{n} G(S_i - S_j).$$

For any $r > 0$, we can see that

$$\mathbb{E}[K_n \mid K_n \geq 1, Y_n \geq r \log n] \geq r \log n.$$

Note that for each r,

$$\mathbb{P}\{Y_n < r \log n\} \leq (n+1)\, \mathbb{P}\left\{ \sum_{i \leq n/2} G(S_i) < r \log n \right\}.$$

Using Lemma 9.1.2 below, we can find an r such that $\mathbb{P}\{Y_n < r \log n\} = o(1/\log n)$ But,

$$\begin{aligned} c \geq \mathbb{E}[K_n] &\geq \mathbb{P}\{K_n \geq 1; Y_n \geq r \log n\} \mathbb{E}[K_n \mid K_n \geq 1, Y_n \geq r \log n] \\ &\geq \mathbb{P}\{K_n \geq 1; Y_n \geq r \log n\}[r \log n]. \end{aligned}$$

Therefore,

$$\mathbb{P}\{K_n \geq 1\} \leq \mathbb{P}\{Y_n < r \log n\} + \mathbb{P}\{K_n \geq 1; Y_n \geq r \log n\} \leq \frac{c}{\log n}.$$

This finishes the proof except for the one lemma that we will now prove. □

Lemma 9.1.2 *Let $p \in \mathcal{P}_4$.*

(a) For every $\alpha > 0$, there exist c, r such that for all n sufficiently large,

$$\mathbb{P}\left\{ \sum_{j=0}^{\xi_n - 1} G(S_j) \leq r \log n \right\} \leq c\, n^{-\alpha}.$$

(b) For every $\alpha > 0$, there exist c, r such that for all n sufficiently large,

$$\mathbb{P}\left\{ \sum_{j=0}^{n} G(S_j) \leq r \log n \right\} \leq c\, n^{-\alpha}.$$

Proof It suffices to prove (a) when $n = 2^l$ for integer l and we write $\xi^k = \xi_{2^k}$. Since $G(x) \geq c/|x|^2$, we have

$$\sum_{j=0}^{\xi^l - 1} G(S_j) \geq \sum_{k=1}^{l} \sum_{j=\xi^{k-1}}^{\xi^k - 1} G(S_j) \geq c \sum_{k=1}^{l} 2^{-2k}\, [\xi^k - \xi^{k-1}].$$

The reflection principle (Proposition 1.5.2) and the central limit theorem show that for every $\epsilon > 0$, there is a $\delta > 0$ such that if n is sufficiently large, and $x \in \overline{\mathcal{C}}_{n/2}$, then $\mathbb{P}^x\{\xi_n \leq \delta\, n^2\} \leq \epsilon$. Let I_k denote the indicator function of the event $\{\xi^k - \xi^{k-1} \leq \delta\, 2^{2k}\}$. Then we know that

$$\mathbb{P}(I_k = 1 \mid S_0, \ldots, S_{\xi^{k-1}}) \leq \epsilon.$$

Therefore, $J_l := \sum_{k=1}^{l} I_k$ is stochastically bounded by a binomial random variable with parameters l and ϵ. By exponential estimates for binomial random variables (see Lemma 11.2.9), we can find an ϵ such that

$$\mathbb{P}\{J_l \geq l/2\} \leq c\, 2^{-\alpha l}.$$

But on the event $\{J_l < l/2\}$ we know that

$$G(S_j) \geq c(l/2)\,\delta \geq r\,\log n,$$

where the r depends on α.

For part (b) we need only note that $\mathbb{P}\{n < \xi_{n^{1/4}}\}$ decays faster than any power of n and

$$\mathbb{P}\left\{\sum_{j=0}^{n} G(S_j) \leq \frac{r}{4}\,\log n\right\} \leq \mathbb{P}\left\{\sum_{j=0}^{\xi_{n^{1/4}}} G(S_j) \leq r\,\log n^{1/4}\right\} + \mathbb{P}\{n < \xi_{n^{1/4}}\}.$$

$\square$

9.2 For 4d section

Let S denote a random walk with increment distribution $p \in \mathcal{P}_4$. Let $\xi_n = \inf\{j : S_j \notin \mathcal{C}_n\}$ and

$$H_n = \sum_{j=0}^{n} G(S_j).$$

Let C_4, C_4^* be as in Theorem 4.3.1. and note that

$$\overline{p}_n(0) = \frac{C_4^*}{2n^2} = \frac{C_4}{8n^2}.$$

Lemma 9.2.1 *Suppose S, S^1 are independent random walks starting at the origin with increment distribution p. Let*

$$J_n = \sum_{j=0}^{n}\sum_{k=0}^{\infty} 1\{S_j = S_k^1\}.$$

Then

$$\mathbb{E}[H_n] = \mathbb{E}[J_n] = \frac{C_4^*}{2}\,\log n + O(1).$$

Proof For ease we will assume that p is aperiodic. Then

$$\mathbb{E}[J_n] = \sum_{j=0}^{n}\sum_{k=0}^{\infty} \mathbb{P}\{S_j^1 = S_k^2\} = \sum_{j=0}^{n}\sum_{k=0}^{n} p_{j+k}(0) = O(1) + \sum_{j=1}^{n}\sum_{k=0}^{\infty} \frac{C_4^*}{2(j+k)^2} = \frac{C_4^*}{2}\,\log n + O(1).$$

Note that H_n is the conditional expectation of J_n given $S_0, \ldots, S_n$ and hence $\mathbb{E}[H_n] = \mathbb{E}[J_n]$. $\square$

In the next lemma, we show that $H_n/\log n$ is very close to a constant random variable.

Lemma 9.2.2 *If $p \in \mathcal{P}_4$, there exists a c such that*

$$\mathbb{P}\{|H_n - (C_4^*/2)\,\log n| \geq \log^{3/4} n\} \leq c\,\exp\{-\log^{1/8} n\},$$

$$\mathbb{P}\{|H_{\xi_n} - C_4^*\,\log n| \geq \log^{3/4} n\} \leq c\,\exp\{-\log^{1/8} n\},$$

$$\mathbb{P}\{H_{\xi_n} \leq c\,\log n\} \leq cn^{-8}.$$

Proof. Let $\xi^m = \xi_{e^m}$, let $\mathcal{G}_m$ denote the σ-algebra generated by $\{S_k : k \le \xi^m\}$ and let Y_m be the $\mathcal{G}_m$-measurable random varible

$$Y_m = \sum_{\xi^{m-1} \le k < \xi^m} G(S_k).$$

If we set $\xi^{-1} = 0$, then we can write

$$\sum_{0 \le k < \xi^n} G(S_k) = \sum_{m=0}^{n} Y_n.$$

By Theorem 4.3.1,

$$G(x) = \frac{C_4}{\mathcal{J}(x)^2} + O(|x|^{-4}),$$

which implies

$$\sum_{0 \le k < \xi^n} G(S_k) = \sum_{0 \le k < \xi^n} \left[\frac{C_4}{\mathcal{J}(S_k)^2 \vee 1} + O((|S_k|^2 + 1)^{-4}) \right].$$

Using Proposition 4.6.3, we can see that

$$\mathbb{E}[Y_m \mid \mathcal{G}_{m-1}] = 4C_4 + O(e^{-m}) = C_4^* + O(e^{-m}). \tag{9.3}$$

We will show that there exist b, c_1 such that for all m,

$$\mathbb{E}[e^{bY_m} \mid \mathcal{G}_{m-1}] \le c_1. \tag{9.4}$$

Before proving this, we will show how this estimate gives the lemma. We will start with the third estimate. From (9.4) we see that

$$\mathbb{E}[e^{bY_n}] \le c_1^{n+1} = c\, e^{nr}.$$

Hence

$$\mathbb{P}\{Y_n \ge \beta n\} \le c\, e^{n(r - b\beta)} \le c\, e^{-8n},$$

provieded β is chosen large enough. For the second estimate, let $M_{-1} = 0$ and

$$M_n = \sum_{m=0}^{n} (Y_m - \mathbb{E}[Y_m \mid \mathcal{G}_{m-1}]).$$

Then M_n is a martingale satisfying Corollary 11.2.8. Hence there is a t with

$$\mathbb{P}\{|M_n| \ge a\sqrt{n}\} \le ce^{-at}.$$

In particular,

$$\mathbb{P}\{|M_n| \ge n^{2/3}\} \le ce^{-n^{1/8}}.$$

However, (9.3) shows that

$$\sum_{m=0}^{n} \mathbb{E}[Y_m \mid \mathcal{G}_{m-1}] = C_4^* n + O(1),$$

and hence

$$\mathbb{P}\{|M_n - C_4^* n| \geq n^{3/4}\} \leq c e^{-n^{1/8}}.$$

This gives the second inequality and the first can be derived easily from this using (2.34) and (2.35).

It remains to prove (9.3) which is equivalent to showing that there is a b such that

$$\sup_n \sup_{x \in \partial \mathcal{C}_{e^{n-1}}} \mathbb{E}^x[e^{bY_{\xi^n}}] < \infty.$$

We will not give the complete details, but we will sketch the argument. Suppose $x \in \partial C_{e^{n-1}}$. Let $\sigma = \sigma_n$ be the first time that it reaches $\partial \mathcal{C}_{e^n}$ or $\partial \mathcal{C}_{e^{n-2}}$. Note that

$$\sum_{j=0}^{\sigma} G(S_k) \leq c\, \sigma\, e^{-2n}.$$

Using this and standard estimates (such as (2.35)), we can find a b_1 such that

$$\mathbb{E}\left[\exp\left\{b_1 \sum_{j=0}^{\sigma} G(S_k)\right\}\right]$$

is uniformly bounded. If $S_\sigma \in \partial \mathcal{C}_{e^n}$, the random walk stops. Otherwise, we are now on $\partial \mathcal{C}_{e^{n-2}}$ and we run until we reach $\partial \mathcal{C}_{e^{n-3}}$ or $\partial \mathcal{C}_{e^{n-1}}$. We keep doing this until we finally reach $\partial \mathcal{C}_{e^n}$. There is a $\rho > 1/2$ such that for all m sufficiently large and all $x \in \partial \mathcal{C}_m$, the probability of reaching $\partial \mathcal{C}_{m+1}$ before $\partial \mathcal{C}_{m-1}$ is greater than ρ. Hence this process is stochastically bounded by a one-dimensional random walk with positive drift. The estimate boils down to a one-dimensional estimate that we give as Exercise 9.1.

Proposition 9.2.3 *Suppose S, S^1 are simple random walks with increment distribution p. Then if $S_0 \in \partial \mathcal{C}_{e^n}, S_0^1 = 0$,*

$$\mathbb{P}\{S[0, \xi^{n+1}) \cap S^1[0, \infty) = \emptyset\} = \frac{1}{2n} + O(n^{-3/2}).$$

Proof Let

$$K = \sum_{j=0}^{\xi^{n+1}-1} \sum_{k=0}^{\infty} 1\{S_j = S_k^1\}.$$

Note that

$$\mathbb{E}[K] = \mathbb{E}\left[\sum_{j=0}^{\xi^{n+1}-1} G(S_j)\right]$$

denote the expected number of intersections of the paths. By (9.3),

$$\mathbb{E}[\hat{J}_n] = C_4^* + o(n^{-1}).$$

The basic idea of the proof is to show that

$$\mathbb{E}[\hat{J}_n \mid \hat{J}_n > 0] \approx 2\, C_4^*\, n$$

and hence

$$\mathbb{P}\{S[0,\xi^{n+1}) \cap S^1[0,\infty) = \emptyset\} = \mathbb{P}(J_n > 0) = \frac{\mathbb{E}[\hat{J}_n]}{\mathbb{E}[\hat{J}_n \mid \hat{J}_n > 0]} \approx \frac{1}{2n}.$$

For each $x \in \mathcal{C}^{n+1}$, let

$$H(x) = H_n(x) = \sum_{j=0}^{\xi^{n+1}-1} G(x, S_j).$$

Let us partition $\mathcal{C}^{n+1}$ into three sets (depending on n):

$$V_1 = \{x : |H_n(x) - 2\,C_4\,n| \leq n^{3/4}\},$$

$$V_2 = \{x \in \mathcal{C}_{e^{n+1}} \setminus V_1 : H_n(x) \leq \hat{c}n\},$$

$$V_3 = \mathcal{C}_{e^{n+1}} \setminus (V_1 \cup V_2).$$

Here $\hat{c}$ is chosen sufficiently large so that

$$\mathbb{P}\{V_3 \neq \emptyset\} \leq e^{-4n}.$$

□

9.2.1 For later

In this section we will show the following.

Theorem 9.2.4 *If $p \in \mathcal{P}_4$, there exists c, c_1 such that if S^1, S^2, S^3 are independent random walk starting at the origin with corresponding stopping times $\xi^1_m, \xi^2_m, \xi^3_m$, then*

$$\mathbb{P}\{S^1(0,\xi^1_m] \cap S^2[0,\xi^2_m] = \emptyset\} \sim \frac{c}{\sqrt{\log m}},$$

$$\mathbb{P}\{S^1(0,\xi^1_m] \cap (S^2[0,\xi^2_m] \cup S^3[0,\xi^3_m]) = \emptyset\} \sim \frac{c_1}{\log m},$$

.

Let us start by making some simplifications. Clearly it suffices to prove the result for $m = e^n$. Let us write $\rho^j_n = \xi^j_{e^n}$ and let $A_n, \hat{A}_n$ be the events

$$A_n = \{S^1(0,\rho^1_n] \cap S^2[0,\rho^2_n] = \emptyset\}, \quad \hat{A}_n = \{S^1(0,\rho^1_n] \cap S^3[0,\rho^3_n] = \emptyset\}.$$

Then the theorem states that

$$\mathbb{P}(A_n) = \mathbb{P}(\hat{A}_n) \sim \frac{c}{\sqrt{n}}, \qquad \mathbb{P}(A_n \cap \hat{A}_n) \sim \frac{c'}{n}.$$

In particular, this shows that $A_n, \hat{A}_n$ are "asymptotically independent up to a multiplicative constant" in the sense

$$\mathbb{P}(A_n \cap \hat{A}_n) \asymp \mathbb{P}(A_n)\,\mathbb{P}(\hat{A}_n).$$

To prove the theorem it suffices (see (11.29)) to show that

$$\mathbb{P}(A_n^c \mid A_{n-1}) = \frac{1}{2n} + b_n, \tag{9.5}$$

$$\mathbb{P}(A_n^c \cup \hat{A}_n^c \mid A_{n-1} \cap \hat{A}_{n-1}) = 2\,\mathbb{P}(A_n^c \mid A_{n-1} \cap \hat{A}_{n-1}) + \hat{b}_n = \frac{1}{n} + b_n^* + \hat{b}_n, \tag{9.6}$$

where $\sum |b_n| + |b_n^*| + |\hat{b}_n| < \infty$.

Proposition 9.2.5 *There exist $r > 0$, $c < \infty$ such that if*

$$Y_n = \sum_{j=0}^{\xi_n} \frac{1}{1 \wedge |\mathcal{J}^*(S_j)^2|},$$

then

$$\left| \mathbb{E}[Y_n] - \frac{2}{d-2}\,\log n \right| \le c,$$

$$E[e^{rY_n}] \le c\, n^c.$$

Exercises

Exercise 9.1 Suppose $X_1, X_2, \ldots$ are independent random variables with distribution $\mathbb{P}\{X_j = 1\} = 1 - \mathbb{P}\{X_j = 0\} = q$ where $q > 1/2$. Let $S_n = X_1 + \cdots + X_n$ and let $\sigma = \min\{j : S_j = 1\}$.

- Show that there exist $0 < b, c < \infty$ such that for each k, $\mathbb{P}\{\tau \ge k\} \le c\, e^{-bk}$.
- Suppose $Z_0, Z_1, Z_2, \ldots$ are independent nonnegative random variables, independent of X_0, X_1, X_2, ..., and that there exist $0 < t, C < \infty$ such that for all n, $\mathbb{E}[e^{tZ_n}] \le C$. Let

$$W = \sum_{j=0}^{\sigma} Z_j.$$

Show there is a $b_2 > 0$ such that $\mathbb{E}[e^{b_2 W}] < \infty$.
- Finish the details of Lemma 9.2.2 by proving (9.4).

10
Loop-erased random walk

Throughtout this section we fix $p \in \mathcal{P}_d$.

10.1 Excursions, loops, and h-processes

10.1.1 Excursion measure

Definition. If $A \subsetneqq \mathbb{Z}^d$, the *boundary Poisson kernel* is defined as the function $H_{\partial A} : \partial A \times \partial A \to [0,1]$ given by

$$H_{\partial A}(z, z') = \mathbb{P}^z\{S_{\tau_A} = z'\}.$$

Recall that $\tau_A = \min\{j \geq 1 : S_j \notin A\}$. In other words $H_{\partial A}(z, z')$ is the probability that the first point in $\mathbb{Z}^d \setminus A$ visited by a random walk starting at $z \in \partial A$ is z'.

Note that $H_{\partial A}(z, w) = H_{\partial A}(w, z)$ for $z, z' \in \partial A$. If $d \geq 3$, we also define

$$H_{\partial A}(z, \infty) = H_{\partial A}(\infty, z) = \mathbb{P}^z\{\tau_A = \infty\}.$$

For $d = 2$, we define $H_{\partial A}(z, \infty) = H_{\partial A}(\infty, z) = 0$. Note that if $z \in \partial A$,

$$\sum_{w \in \partial A \cup \{\infty\}} H_{\partial A}(z, w) = \mathbb{P}^z\{S_1 \in \overline{A}\},$$

and if ∂A is finite and $d \geq 3$,

$$\sum_{z \in \partial A} H_A(\infty, z) = \mathrm{cap}(\partial A).$$

If $\omega = [\omega_0, \omega_1, \ldots, \omega_n]$ is a finite sequence of points in $\mathbb{Z}^d$ we define $p(\omega)$ by

$$p(\omega) = \mathbb{P}^{\omega_0}\{S_j = \omega_j, j = 0, \ldots, n\} = \prod_{j=1}^{n} p(\omega_{j-1}, \omega_j).$$

Definition. We will call ω a *(finite) path (of length n)* if $p(\omega) > 0$. We sometimes write $|\omega| = n$ for the length of the walk. If $n = 0$, we call ω a *trivial* path and set $p(\omega) = 1$. The points ω_0, ω_n are the *endpoints* of the path. We call the path a *loop* if $\omega_0 = \omega_n$; in particular, the trivial path is a loop. If ω is a loop, then ω_0 will be called the *root* of the loop. A loop is called *basic* if $\omega_j \neq \omega_0, 1 \leq j \leq n-1$. If $n \geq 1$, we call ω an *A-excursion* if $\omega_0, \omega_n \in \partial A$ and $\omega_j \in A$ for $1 \leq j < n$.

The *excursion measure* on A is the measure supported on A-excursions that gives each A-excursion ω measure $p(\omega)$.

The total mass of the excursion measure is

$$\sum_{z,z' \in \partial A} H_{\partial A}(z, z').$$

This is finite if ∂A is finite, but is infinite if ∂A is infinite. If $z, z' \in \partial A$ are distinct points with $H_{\partial A}(z, z') > 0$, we can consider the conditional measure on A-excursions from z to z',

$$p_A^*(\omega) = \frac{p(\omega)}{H_{\partial A}(z, z')}.$$

This is random walk starting at z conditioned so that the next visit to ∂A is at z'. It is a Markov chain stopped when it reaches state w with transitions

$$P^*(x, y) = \frac{p(x, y)\, H_A(y, z')}{\sum_{y'} p(x, y')\, H_A(y', z')}.$$

Here $H_A(\cdot, \cdot)$ denotes the Poisson kernel. (Recall that if $z, z' \in \partial A$, then $H_A(z, z') = \delta(z - z')$.) If $x \in A$, the denominator equals $H_A(x, z')$. If $x = z$, the denominator equals $H_{\partial A}(z, z')$. This process is called an h-process and described more in Section 10.1.3.

10.1.2 Loop measure

The *loop measure in A rooted at x* is the measure that assigns weight $p(\omega)$ to every loop $\omega = [\omega_0, \ldots, \omega_n]$ rooted at x with $\omega_j \in A$ for $j \geq 1$. The trivial loop rooted at x gets weight one even if x is not in A. We let $F_A(x)$ denote the total mass of this measure. Then

$$F_A(x) = 1 + \sum_{\omega \subset A; |\omega| \geq 1; \omega_0 = \omega_n = x} p(\omega) = \begin{cases} G_A(x, x) & x \in A \\ 1 & x \notin A. \end{cases}$$

Let $\mathcal{R}(A; x)$ denote the set of basic loops in A rooted at x of length at least one. A standard Markov chain argument shows that

$$\frac{1}{F_A(x)} = \mathbb{P}^x\{T_x > \overline{\tau}_A\} = 1 - \sum_{\omega \in \mathcal{R}(A;x)} p(\omega), \tag{10.1}$$

where the last sum is over all basic loops ω in A rooted at x of length at least one. If $|s| < 1$, then

$$\sum_{n=0}^{\infty} s^n = \frac{1}{1-s} = e^{-\log(1-s)} = \sum_{k=0}^{\infty} \frac{1}{k!} \left(\sum_{j=0}^{\infty} \frac{s^j}{j} \right)^k. \tag{10.2}$$

The *normalized loop measure* at x in A is the loop measure normalized to be a probability measure, i.e., for each A-loop rooted at x,

$$p^{\#}(\omega; x, A) = \frac{p(\omega)}{F_A(x)}.$$

Lemma 10.1.1 *Suppose $d \geq 3$ or $A \subsetneqq \mathbb{Z}^d$. Suppose $x \in \mathbb{Z}^d$ and S_n is a random walk starting at x. Let*

$$k = \max\{j \leq \overline{\tau}_A : S_j = x\}.$$

Then the distribution of $[S_0, S_1, \ldots, S_k]$ is that of the normalized loop measure at x in A.

Proof This is immediate and left to the reader. If $d \geq 3$, it is possible that $\overline{\tau}_A = \infty$; however k is finite with probability one. □

10.1.3 h-processes

Suppose $A \subset \mathbb{Z}^d$ is connected and $h : \overline{A} \to [0, \infty)$ satisfies

$$h(x) > 0, \quad \mathcal{L}h(x) = 0, \quad x \in A.$$

The *h-process* associated to h is the Markov chain Y_j with state space $\overline{A}$ and transition probabilities

$$q(x, y) = p(x, y)\,\frac{h(y)}{h(x)}, \qquad x \in A.$$

$$q(x, x) = 1 \text{ if } x \in \partial A \text{ with } h(x) > 0,$$

$$q(x, y) = \frac{p(x, y)\,h(y)}{\mathcal{L}h(x)} \text{ if } x \in \partial A \text{ with } h(x) = 0.$$

Note that if $x \in A$,

$$\sum_{y \in \mathbb{Z}^d} q(x, y) = \frac{\sum_{y \in \mathbb{Z}^d} p(x, y)\,h(y)}{h(x)} = 1,$$

since $\mathcal{L}h(x) = 0$. Also if $x \in \partial A$ with $h(x) = 0$, then

$$\sum_{y \in Z^d} q(x, y) = \frac{\sum_{y \in \mathbb{Z}^d} p(x, y)\,h(y)}{\mathcal{L}h(x)} = \frac{\sum_{y \in \mathbb{Z}^d} p(x, y)\,[h(y) - h(x)]}{\mathcal{L}h(x)} = 1.$$

Hence q is a valid transition probability with state space $\overline{A}$.

Suppose $\omega = [\omega_0, \ldots, \omega_n]$ is a path with $\omega_0, \ldots, \omega_{n-1} \in A$ and $\omega_n \in \overline{A}$. Then it is easy to check that the h-process Y satisfies

$$\mathbb{P}\{[Y_0, \ldots, Y_n] = \omega \mid Y_0 = \omega_0\} = p(\omega)\,\frac{h(\omega_n)}{h(\omega_0)}. \tag{10.3}$$

Here are some important examples.

- If $w \in \partial A$, and we let $h(x) = H_A(x, w)$, then the h-process started at $x \in A$ gives the same measure on paths as that of the random walk conditioned on the event $\{S_{\tau_A} = w\}$. If $z \in \partial A \setminus \{w\}$, then $\mathcal{L}h(z) = H_{\partial A}(z, w)$.
- If $d \geq 3$ and $h(x) := \mathbb{P}^x\{\overline{\tau}_A = \infty\}$ is strictly positive on A, the associated h-process started at $x \in A$ gives the same measure on paths as that of the random walk conditioned on the event $\{\tau_A = \infty\}$. Loosely speaking, we can call this process random walk in A conditioned to exit A "at infinity".

- If $d = 2$ and $B := \mathbb{Z}^d \setminus A$ is finite, let

$$h(x) = g_B(x) = \lim_{n\to\infty} C_2 \,(\log n)\mathbb{P}^x\{\xi_n < \overline{\tau}_A\}.$$

This measure can be considered random walk "conditioned to never reach B". Note that we cannot simply condition on the event that the path never reaches B since this is an event of probability zero. Loosely speaking, we can call this process random walk in A conditioned to exit A "at infinity". In the special case $B = \{0\}$, we have $g_B(x) = a(x)$ and we call the process random walk conditioned to avoid the origin. As the next lemma shows, this process is transient.

Lemma 10.1.2 *Suppose $x \in \mathbb{Z}^2$ and $Y_0, Y_1, \ldots$ is random walk conditioned to avoid the origin starting at x. Suppose $0 < r < |x|$. Then,*

$$\mathbb{P}\{Y_n \in \mathcal{C}_r \text{ for some } n \geq 0\} = \frac{\log r + O(1)}{\log |x|}$$

Proof If q denotes the probability on the left, then (10.3) implies that

$$\min_{y\in\partial_i\mathcal{C}_r} a(y) \leq q\, a(x) \leq \max_{y\in\partial_i\mathcal{C}_r} a(y).$$

Therefore,

$$q = \frac{C_2\, \log r + O(1)}{C_2\, \log |x| + O(1)}.$$

□

10.2 Loop erasure

A *self-avoiding walk (SAW)* is a path that visits no point more than once, i.e., that satisfies $\omega_j \neq \omega_k$ for $0 \leq j < k \leq n$. *(Chronological) loop-erasure* is a way to assign to each path $\omega = [\omega_0, \ldots, \omega_n]$ a self-avoiding subpath with the same endpoints. To be precise, for any path ω we define $\mathrm{LE}(\omega)$ as follows.

- Let $j_{-1} = -1$ and $j_0 = \max\{j \leq n : \omega_j = \omega_0\}$
- If $j_0 = n$, we stop. Otherwise, we let $j_1 = \max\{j \leq n : \omega_j = \omega_{j_0+1}\}$.
- Recursively, if $j_i = n$, we stop. Otherwise, we let $j_{i+1} = \max\{j \leq n : \omega_j = \omega_{j_i+1}\}$.
- Let k be the smallest i such that $j_i = n$. Then

$$\mathrm{LE}(\omega) = [\omega_{j_0}, \omega_{j_1}, \ldots, \omega_{j_k}] = [\omega_{j_{-1}+1}, \omega_{j_0+1}, \ldots, \omega_{j_{i-1}+1}].$$

Let $L^i = [\omega_{j_{i-1}+1}, \ldots, \omega_{j_i}]$. We have divided the path ω into a collection of loops, $[L^0, \ldots, L^k]$. This division is uniquely determined by the following conditions.

- The path ω is exactly the path $[L^0, \ldots, L^k]$.
- Each L^i is a loop rooted at ω_{j_i}
- For each $i \geq 1$, the loop L^i does not include any of the points $\{\omega_{j_0}, \omega_{j_1}, \ldots, \omega_{j_{i-1}}\}$.

The *loop-erased excursion measure* on A is the measure $\hat{p}_A$ that assigns to each self-avoiding excursion with distinct endpoints the excursion measure of the set of paths whose loop-erasure is that path. In other words, if $\eta = [\eta_0, \ldots, \eta_k]$ is a self-avoiding A-excursion, then

$$\hat{p}_A(\eta) = \sum p(\omega),$$

where the sum is over all A-excursions ω with $\mathrm{LE}(\omega) = \eta$. Note that if z, z' are distinct points in ∂A, then the loop-erased excursion measure gives measure $H_{\partial A}(z, z')$ to the set of paths with endpoints z, z'. If $z \in \partial A$, and ω is the trivial path at z, we set $\hat{p}_A(\omega) = 1$.

Note that

$$\mathrm{LE}([\omega_{j_0}, \omega_{j_0+1}, \ldots, \omega_n]) = \mathrm{LE}(\omega).$$

In other words, in order to produce a walk with a certain loop-erasure, we can divide the walk into two parts: a beginning loop and a walk that never returns to the starting point. Using this, we can see that if $k \geq 1$, and $\eta = [\eta_0, \eta_1, \ldots, \eta_k]$ is a self-avoiding excursion in A, then

$$\hat{p}_A(\eta) = p(\eta_0, \eta_1)\, F_A(\eta_1)\, \hat{p}_{A \setminus \{\eta_1\}}([\eta_1, \ldots, \eta_k]).$$

By continuing, we get for $1 \leq l \leq k$,

$$\hat{p}_A(\eta) = p([\eta_0, \ldots, \eta_l]) \left[\prod_{j=0}^{l} F_{A_j}(\eta_j)\right] \hat{p}_{A_l}([\eta_l, \ldots, \eta_k]), \tag{10.4}$$

where $A_j = A \setminus \{\eta_1, \ldots, \eta_{j-1}\}$. When $l = k$, $\hat{p}_{A \setminus A_k}([\eta_k]) = 1$.

Lemma 10.2.1 *If $A \subsetneq \mathbb{Z}^d$ and $\{x_1, \ldots, x_n\} \subset \mathbb{Z}^d$, let*

$$F_A(x_1, \ldots, x_n) = \prod_{j=1}^{n} F_{A_j}(x_j),$$

where $A_1 = A$ and $A_j = A \setminus \{x_1, \ldots, x_{j-1}\}$. Then for any permutation σ of $\{1, \ldots, n\}$,

$$F_A(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = F_A(x_1, \ldots, x_n).$$

Proof It suffices to prove the result for transpositions (permutations that switch j and $j+1$ and leave all other points fixed) since all permutations can be generated as a product of transpositions. If either x_j or x_{j+1} is not in A, this is trivial, so it suffices to show that for any A and any distinct $x, y \in A$, that

$$G_{A \setminus \{y\}}(x, x)\, G_A(y, y) = G_A(x, x)\, G_{A \setminus \{x\}}(y, y). \tag{10.5}$$

By splitting the number of returns to x before leaving A into those that occur before visiting y and those after visiting y we get

$$\begin{aligned} G_A(x, x) &= G_{A \setminus \{y\}}(x, x) + \mathbb{P}^x\{T_y < \tau_A\}\, G_A(y, x) \\ &= G_{A \setminus \{y\}}(x, x) + \mathbb{P}^x\{T_y < \tau_A\}\, \mathbb{P}^y\{T_x < \tau_A\}\, G_A(x, x). \end{aligned}$$

Hence,

$$G_{A \setminus \{y\}}(x, x) = [1 - \mathbb{P}^x\{T_y < \tau_A\}\, \mathbb{P}^y\{T_x < \tau_A\}]\, G_A(x, x),$$

from which (10.5) follows readily. □

If $A \subsetneqq \mathbb{Z}^d$ and $B = \{x_1, \ldots, x_n\}$ we define

$$F_A(B) = \prod_{j=1}^{n} F_{A_j}(x_j), \tag{10.6}$$

where $A_1 = A$ and $A_j = A \setminus \{x_1, \ldots, x_{j-1}\}$. The last lemma shows that this quanity is independent of the ordering of the points. Note that $F_A(B) = F_A(B \cap A)$. If $\eta = [\eta_0, \ldots, \eta_k]$ is a self-avoiding path, then we write $F_A(\eta)$ for $F_A(\eta_0, \ldots, \eta_k) = F_A(\{\eta_0, \ldots, \eta_k\})$. We can rewrite (10.4) as

$$\hat{p}_A(\eta) = p(\eta)\, F_A(\eta).$$

In particular if $\eta^R = [\eta_k, \eta_{k-1}, \ldots, \eta_0]$ is the "reversed" path, then $\hat{p}_A(\eta^R) = \hat{p}_A(\eta)$.

Comment. It is easy to give examples of paths ω such that $\mathrm{LE}(\omega^R)$ is not the same as $\mathrm{LE}(\omega)$.

The measures p and $\hat{p}$ are measures on A-excursions. Given p we get $\hat{p}$ by erasing loops. The next proposition discusses how to get p from $\hat{p}$ by adding loops.

Proposition 10.2.2 *Suppose $d \geq 3$ or $A \subsetneqq \mathbb{Z}^d$ and $\eta = [\eta_0, \eta_1, \ldots, \eta_k]$ is a self-avoiding A-excursion. For each j, let*

$$A_j = A \setminus \{\eta_i : i < j\}.$$

For $j = 0, \ldots, n-1$, let $\omega^j = [\omega^j_0, \ldots, \omega^j_{r_j}]$ be chosen independently from the normalized loop measure at η_j in A_j. Let ω be the excursion obtained from adding on these loops,

$$\omega = [\eta_0, \omega^0, \eta_1, \omega^1, \ldots, \eta_{n-1}, \omega^{n-1}, \eta_n].$$

Then the distribution of ω is that of ω conditioned on $LE(\omega) = \eta$.

Proof Straightforward and left to the reader. □

10.3 Unrooted loop soup

A fundamental property of loops is that the quantity $F_A(B)$ does not depend on the order of the points in B. We will describe another way to "add loops" to a set which helps illustrate why this quantity is independent of the order.

An *unrooted loop* is an equivalence class of loops under translation of the root, i.e., it is an equivalence class under the equivalence relation

$$[\omega_0, \omega_1, \ldots, \omega_n] \sim [\omega_1, \omega_2, \ldots, \omega_n, \omega_1] \sim \cdots \sim [\omega_{n-1}, \omega_n, \omega_1, \ldots, \omega_{n-2}, \omega_{n-1}].$$

We will use $\overline{\omega}$ to denote the unrooted loop represented by the rooted loop ω. We define $p(\overline{\omega})$ to be $p(\omega)$; note that $p(\omega)$ is the same for all representatives of $\overline{\omega}$ so this is well defined. We write $\overline{\omega} \subset A$ if $\{\omega_0, \ldots, \omega_n\} \subset A$. Let $\kappa(\overline{\omega})$ be the largest integer m that divides n such that $\omega_k = \omega_{k+(n/m)}$ for all $k \leq n - (n/m)$. Note that $\kappa(\overline{\omega})$ is independent of the choice of representative of $\overline{\omega}$, and the number of distinct rooted representatives of $\overline{\omega}$ is $n/\kappa(\overline{\omega})$.

Example. Suppose x_1, x_2, x_3, x_4 are distinct points. if

$$\overline{\omega} = [x_1, x_2, x_1, x_2, x_1, x_2, x_1],$$

then $\kappa(\overline{\omega}) = 3$ and $n/\kappa(\overline{\omega}) = 2$. If

$$\overline{\omega} = [x_1, x_2, x_3, x_4, x_1, x_2, x_1],$$

then $\kappa(\overline{\omega}) = 1$ and $n/\kappa(\overline{\omega}) = 6$.

Definition. The *unrooted loop measure* is the measure that assigns weight $p(\overline{\omega})/\kappa(\overline{\omega})$ to each unrooted loop $\overline{\omega}$. Equivalently, it is the measure that assigns $\overline{\omega}$ weight

$$\frac{1}{|\overline{\omega}|} \sum_{\omega} p(\omega)$$

where the sum is over rooted loops ω that are representatives of $\overline{\omega}$.

Comment. If $|\overline{\omega}| = n$ and $\kappa(\overline{\omega}) = 1$ (which for n large is true for all but an exponentially small fraction of unroooted loops), then the unrooted loop measure gives measure $p(\overline{\omega})$ to each unrooted loop. We can also think of this as a measure on rooted loops that gives measure $p(\omega)/n$ to each of the n rooted loops that are representatives of $\overline{\omega}$.

Definition. The *(unrooted) loop soup* with intensity t is a realization $\mathcal{U}_t$ from a Poisson point process on unrooted loops with the unrooted loop measure.

To be more precise, let $\{N_t^{\overline{\omega}}\}$ denote a collection of independent Poisson processes indexed by the set of unrooted loops $\overline{\omega}$. The process $N_t^{\overline{\omega}}$ has parameter $p(\overline{\omega})/\kappa(\overline{\omega})$. For each $t > 0$, let $\mathcal{U}_t$ denote the set of loops that have been created by time t. The set $\mathcal{U}_t$ is a multi-set (i.e., a set in which elements can appear more than once) in which the loop $\overline{\omega}$ appears $N_t^{\overline{\omega}}$ times. If $N_t^{\overline{\omega}} > N_{t-}^{\overline{\omega}}$, i.e. a jump of this process occurs at time t, we say that the loop $\overline{\omega}$ is created at time t. Note that with probability one, there is no time at which more than one loop is created.

If $A \subset \mathbb{Z}^d$, we let $\mathcal{U}_t(A) = \{\overline{\omega} \in \mathcal{U}_t : \overline{\omega} \subset A\}$. Suppose an ordering of A is given, $A = \{z_1, z_2, \ldots\}$. Then given $\mathcal{U}_t(A)$ we can associate rooted loops as follows.

- Given $\overline{\omega}$ represented by $\omega = [\omega_0, \ldots, \omega_n]$, let j be the smallest integer such that $z_j \in \{\omega_1, \ldots, \omega_n\}$.
- Let $i_1 < i_2 < \cdots < i_k$ denote all the indices $1 \leq i \leq n$ such that $\omega_i = z_j$.
- Let $Y = Y(\overline{\omega})$ be a random variable, independent of all other random variables, uniformly distributed on $\{1, \ldots, k\}$. Then choose the (rooted) loop

$$[\omega_{i_Y}, \omega_{i_Y+1}, \ldots, \omega_n, \omega_1, \ldots, \omega_{i_Y-1}, \omega_{i_Y}].$$

In other words, given an unrooted loop $\overline{\omega}$, we choose a root for $\overline{\omega}$ by finding the smallest point in the loop, with smallest defined by the given ordering on A. If this gives more than one choice for the root, we choose randomly.

After doing this process we have a realization $\tilde{\mathcal{U}}_\lambda(A)$ of rooted loops. Let $\mathcal{A}_k$ denote the set of all rooted loops ω with $\omega_0 = z_k$, all of whose points lie in $A \setminus \{z_1, \ldots, z_{k-1}\}$. It is easy to see that we now have an independent collection of Poisson processes $\{\tilde{N}_t^{\omega}\}$ indexed by rooted loops $\omega \in \cup_k \mathcal{A}_k$. A moment's thought will show that if $\omega = [\omega_0, \ldots, \omega_n] \in \mathcal{A}_k$, then the parameter of the process $\tilde{N}_t^{\omega}$ is $p(\omega)/J(\omega)$ where $J = J(\omega)$ denotes the number of indices i with $1 \leq i \leq n$ and $\omega_i = z_k$. Note that for each z, there are only a finite number of loops in $\tilde{\mathcal{U}}_\lambda(A)$ rooted at z. Suppose these loops are

$$\omega^1 = [\omega_0^1 = z, \omega_1^1, \ldots, \omega_{n_1}^1 = z], \quad \ldots, \quad \omega^k = [\omega_0^k = z, \omega_1^k, \ldots, \omega_{n_k}^k = z].$$

These are written down in the order that they appeared in the realization. Then we form a single loop at z by

$$L_z = L_z(A) = [z, \omega_1^1, \ldots, \omega_{n_1}^1, \omega_1^2, \ldots, \omega_{n_2}^2, \ldots, \omega_1^k, \ldots, \omega_{n_k}^k].$$

If no loop rooted at z appears in the realization, then L_z is the trivial loop.

Proposition 10.3.1 *If $A = \{z_1, z_2, \ldots\}$ and $t = 1$, then the loops $L_{z_1}, L_{z_2}, \ldots$ are independent, with L_{z_j} chosen from the normalized loop measure in $A \setminus \{z_1, \ldots, z_{j-1}\}$.*

Proof Since the $\mathcal{A}_k$ are disjoint and the loop L_{z_k} is formed from $\mathcal{U}_1(A) \cap \mathcal{A}_k$; the $L_{z_1}, L_{z_2}, \ldots$ are independent. Let

$$\tilde{N}_t^{(k)} = \sum_{\omega \in \mathcal{A}_k} \tilde{N}_t^{\omega}.$$

Then $N_t^{(k)}$ is a Poisson process with parameter

$$\lambda_k := \sum_{\omega \in \mathcal{A}_k} \frac{p(\omega)}{J(\omega)} = \sum_{j=1}^{\infty} \frac{1}{j} \sum_{\omega \in \mathcal{A}_k; J(\omega)=j} p(\omega) = \sum_{j=1}^{\infty} \frac{1}{j} \left[\sum_{\omega \in \mathcal{A}_k; J(\omega)=1} p(\omega) \right]^j = \log F_{A \setminus \{z_1, \ldots, z_{k-1}\}}(z_k).$$

The last equality uses (10.1). The middle equality can be shown by writing each loop as a concatenation of basic loops.

Now suppose that $\omega \in \mathcal{A}_k$. Let $0 = i_0 < i_1 < i_2 < \cdots < i_l = n$, denote all of the indicies with $\omega_{i_j} = z_k$. Then for any collection of positive integers $j_1, \ldots, j_m$ with $j_1 + \cdots + j_m = l$, we can divide ω into m loops of the form

$$\omega^r := [\omega_{i_{j_1+\cdots+j_{r-1}}}, \ldots, \omega_{i_{j_1+\cdots+j_r}}], \qquad r = 1, \ldots, m.$$

Note that $J(\omega_r) = j_r$. Using standard properties about Poisson processes, we can see that the probability that the loop soup restricted to $\mathcal{A}_k$ produces exactly m loops which are in order $\omega^1, \ldots, \omega^m$ as above is

$$e^{-\lambda_k} \frac{p(\omega)}{m!\, j_1 \cdots j_m}.$$

Therefore the probability that the concatenation of the loops is exactly ω is equal to

$$\sum e^{-\lambda_k} \frac{p(\omega)}{m!\, j_1 \cdots j_m}$$

where the sum is over all $j_1, \ldots, j_m$ with $j_1 + \cdots + j_m = l$. We claim that

$$\sum_{j_1, \ldots, j_m > 0, \;\; j_1 + \cdots + j_m = l} \frac{1}{m!} \frac{1}{j_1 \cdots j_m} = 1. \tag{10.7}$$

This follows from

$$\sum_{n=0}^{\infty} s^n = \frac{1}{1-s} = e^{-\log(1-s)} = \sum_{k=0}^{\infty} \frac{1}{k!} \left(\sum_{j=0}^{\infty} \frac{s^j}{j} \right)^k,$$

which is valid for $|s| < 1$. The left-hand side of (10.7) is the coefficient of s^l in the right-hand side of the equation above. Therefore the probability that $L_{z_k} = \omega$ is

$$e^{-\lambda_m} \, p(\omega) = e^{-\log F_{A \setminus \{z_1,\ldots,z_{k-1}\}}(z_k)} \, p(\omega)$$

which is exactly the probability of choosing ω from the normalized loop measure. $\square$

10.4 Loop-erased random walk (LERW)

If $A \subset \mathbb{Z}^d$ and $z, w \in \partial A$ are distinct with $H_{\partial A}(z,w) > 0$, then the *loop-erased random walk (LERW)* from z to w in ∂A is the probability distribution on self-avoiding paths induced by $\hat{p}$ restricted to self-avoiding A-excursions from z to w. In other words, if η is a self-avoiding A-excursion from z to w, then the probability of that η is given by

$$\hat{p}_A^{\#}(\eta; z, w) = \frac{\hat{p}(\eta)}{H_{\partial A}(z,w)} = \frac{p(\eta)\, F_A(\eta)}{H_{\partial A}(z,w)}.$$

The LERW from z to z is defined to be the distribution supported on the trivial walk.

If $n \leq m$ and $\omega = [\omega_0, \ldots, \omega_n], \tilde{\omega} = [\tilde{\omega}_0, \ldots, \tilde{\omega}_m]$ are two paths we write $\omega \prec \tilde{\omega}$ if $\omega_j = \tilde{\omega}_j$ for $0 \leq j \leq n$. The following proposition follows from (10.4).

Proposition 10.4.1 *Suppose z, w are distinct and $\overline{\eta} = [\eta_0, \ldots, \eta_k]$ is a self-avoiding path with $\eta_0 = z$; $\eta_1, \ldots, \eta_{k-1} \in A$; and $\eta_k \in \overline{A}$. Then,*

$$\sum_{\overline{\eta} \prec \eta} \hat{p}_A(\eta; z, w) = p(\overline{\eta})\, F_A(\overline{\eta})\, H_{\partial(A \setminus \overline{\eta})}(\eta_k, w).$$

In particular, if $\eta_k \in A$ and $[\hat{S}_0, \ldots, \hat{S}_m]$ denotes the LERW from z to w in A, then

$$\mathbb{P}\{\hat{S}_{k+1} = x \mid [\hat{S}_0, \ldots, \hat{S}_k] = \overline{\eta}\} = \frac{p(\eta_k, x)\, H_{A \setminus \{\overline{\eta}}(x,w)}{\sum_y p(\eta_k, y)\, H_{A \setminus \overline{\eta}}(y,w)}.$$

The quantity on the right-hand side of the last equation is zero for x not in $(A \setminus \overline{\eta}) \cup \{w\}$. Another way to think of the LERW is as a Markov chain taking values in ordered pairs (V, x) where $V \subset \mathbb{Z}^d$ and $x \in \mathbb{Z}^d \setminus V$. If $p(x,y) > 0$, the probability of a transition from (V, x) to $(V \setminus \{y\}, y)$ is the probability that the LERW in V from x to w takes its first step to y which we have stated is proportional to $p(x,y)\, H_V(y,w)$. All other transitions have probability zero. The initial state is (A, z), and the chain is stopped when the second component of the process equals w.

There are a number of similar LERWs.

- If $x \in A, w \in \partial A$ with $H_A(x,w) > 0$, we define the LERW from x to w in A to be be the LERW from x to w in $A \setminus \{x\}$. This is equivalent to taking a random walk starting at x conditioned to leave A at w and performing loop-erasure. The distribution on loop-erasure is the same if we condition the random walk starting at x never to return to x (see Proposition 4.6.5). If $\eta = [\eta_0, \ldots, \eta_k]$ is a self-avoiding path with $\eta_0 = x; \eta_1, \ldots, \eta_{k-1} \in A$; $\eta_k \in \overline{A}$, then

$$\mathbb{P}\{[\hat{S}_0, \ldots, \hat{S}_k] = \eta\} = \frac{p(\eta)\, F_A(\eta)\, H_A(\eta_k, w)}{H_A(x, w)}.$$

- If $x \in A$ with $\mathbb{P}^x\{\tau_A < \infty\} = 1$, we define LERW from x to ∂A to be the measure on paths obtained from $\mathrm{LE}([S_0, \ldots, S_{\tau_A}])$. If $\eta = [\eta_0, \ldots, \eta_k]$ is a self-avoiding path with $\eta_0 = x; \eta_1, \ldots, \eta_{k-1} \in A$; $\eta_k \in \overline{A}$,

$$\mathbb{P}\{[\hat{S}_0, \ldots, \hat{S}_k] = \eta\} = p(\eta)\, F_A(\eta) \mathbb{P}^{\eta_k}\{S_j \notin \eta : j = 1, 2, \ldots, \overline{\tau}_A\}.$$

  The last probability on the right-hand side equals one if $\eta_k \in \partial A$.
- If $d \geq 3$, the LERW starting at the origin in $\mathbb{Z}^d$ can be defined by doing the loop-erasure on the infinite path. This can be considered as a consistent family of measures on paths of length n. If $\eta = [0 = \eta_0, \ldots, \eta_k]$ is a self-avoiding path starting at the origin, then

$$\mathbb{P}\{[\hat{S}_0, \ldots, \hat{S}_k] = \eta\} = p(\eta)\, F_{\mathbb{Z}^d}(\eta)\, \mathrm{Es}_\eta(\eta_n) = p(\eta)\, F_{\mathbb{Z}^d}(\eta)\, \mathcal{L} g_\eta(\eta_n),$$

  where $g_\eta(x) = \mathbb{P}^x\{\overline{T}_\eta = \infty\}$ is the unique function on $\mathbb{Z}^d$ that is harmonic on $\mathbb{Z}^d \setminus \eta$; equals zero on η; and tends to 1 at infinity.
- We can define the LERW starting at the origin for $d = 2$ by erasing loops from the h-process corresponding to random walk conditioned to never return to the origin. If $\eta = [0 = \eta_0, \ldots, \eta_k]$ is a self-avoiding path starting at the origin, then

$$\mathbb{P}\{[\hat{S}_0, \ldots, \hat{S}_k] = \eta\} = p(\eta)\, F_{\mathbb{Z}^d}(\eta)\, \mathrm{Es}_\eta(\eta_n) = p(\eta)\, F_{\mathbb{Z}^d}(\eta)\, \mathcal{L} g_\eta(\eta_n),$$

  where

$$g_\eta(x) = \lim_{n \to \infty} C_2\, (\log n)\, \mathbb{P}^x\{\xi_n < \tau_\eta\}$$

  as in Chapter 6.

We can summarize the probability distributions above as follows. Suppose $[\hat{S}_0, \hat{S}_1, \ldots]$ has the distribution of LERW from z to w in A. (LERW in $\mathbb{Z}^d$ starting at 0 can be called LERW from 0 to $w = \infty$ in $\mathbb{Z}^d \setminus \{0\}$.) Then the conditional distribution of $\hat{S}_{k+1}$ given $[\hat{S}_0, \ldots, \hat{S}_k]$ is the same as the first step of the h-process in $A \setminus \{\hat{S}_0, \ldots, \hat{S}_k\}$ corresponding to random walk starting at $\hat{S}_k$ conditoned to leave $A \setminus \{\hat{S}_0, \ldots, \hat{S}_k\}$ at w. Because of the relationship between the loop-erased walk and harmonic functions, the loop-erased random walk is sometimes called the *Laplacian random walk.*

10.5 Spanning trees and Wilson's algorithm

Loop-erasure of walks gives an efficient way to find spanning trees of graphs. We give an example in this section. Suppose $A \subset \mathbb{Z}^d$ is finite and connected. Define the weighted undirected graph A_W (the W stands for "wired") by:

- The vertices of A_W are $A \cup \{\partial\}$ where ∂ is a single boundary point.
- If $x, y \in A$ are distinct with $p(x, y) > 0$, then $\{x, y\}$ is an edge with weight $p(x, y)$.

- If $x \in A$ and $z \in \partial A$, there is an an edge between x and ∂ with weight $p(x,z)$. Note that there can be more than one edge between x and ∂.

A *spanning tree* $\mathcal{T}$ of A_W is a connected subgraph containing all the vertices with no cycles. In other words, it is a subset of the edges of A_W with the property that if x, y are distinct vertices, then there exists a unique path $x = x_0, x_1, \ldots, x_k = y$ with $x_0, \ldots, x_k$ distinct vertices in $A \cup \{\partial\}$ such that for $j = 1, \ldots, k$, $\{x_{j-1}, x_j\}$ is an edge of $\mathcal{T}$. Note that the number of edges in any spanning tree of $A \cup \{\partial\}$ is $\#(A)$. We define the weight of the tree by

$$p(\mathcal{T}) = \prod_{\{x,y\} \text{ edge of } \mathcal{T}} p(x,y).$$

Wilson's algorithm is a way to choose a random spanning tree of A_W using loop-erased walks. We state one form of it here.

- Let $A_0 = A$.
- Order the points in A (arbitrarily): $x_1, \ldots, x_n$.
- Start a random walk at x_1 and stop it when it leaves A. Take the loop-erasure of the path producing $\eta = [\eta_0, \ldots, \eta_k]$ with $k > 0$. The point η_k is in ∂A and rename it ∂. Include in the spanning tree the edges $\{\eta_0, \eta_1\}, \{\eta_1, \eta_2\}, \ldots, \{\eta_{k-1}, \eta_k\}$.
- Let $\tilde{A} = A \setminus \{\eta_1, \ldots, \eta_{k-1}\}$. If $\tilde{A} = \emptyset$, stop. Otherwise continue.
- Let j be the smallest index with $x_j \in \tilde{A}$. Start a random walk at x_j and stop it when it leaves $\tilde{A}$. Take the loop-erasure of the path producing $\eta = [\eta_0, \ldots, \eta_k]$ with $k > 0$. If the point η_k is in ∂A, we identify it with ∂. Add to the tree the edges $\{\eta_0, \eta_1\}, \{\eta_1, \eta_2\}, \ldots, \{\eta_{k-1}, \eta_k\}$.
- Let $A' = \tilde{A} \setminus \{\eta_1, \ldots, \eta_{k-1}\}$. If $A' = \emptyset$ stop. Otherwise, set $\tilde{A} = A'$ and return to the previous step.

It is easy to see that this algorithm produces a spanning tree and hence generates a distribution on spanning trees. The next proposition describes the distribution and shows that it is independent of the ordering of the points $\{x_1, \ldots, x_n\}$.

Proposition 10.5.1 *For any spanning tree $\mathcal{T}$ of A_W, the probability that $\mathcal{T}$ is output by Wilson's algorithm is $p(\mathcal{T})\, F_A(A)$, where $F_A(A)$ is defined as in (10.6).*

Proof Suppose $\mathcal{T}$ is a spanning tree. Then $\mathcal{T}$ produces an ordering of the vertices of A, $\{y_1, \ldots, y_n\}$ as follows.

- The unique path from x_1 to ∂ is given by $\eta^1 := [y_1, \ldots, y_{j_1}, \partial]$.
- If $j_i < n$, let $m = m_i$ be the smallest index m such that $x_m \notin \{y_1, \ldots, y_{j_i}\}$. Then $\eta^{i+1} := [y_{j_i+1}, \ldots, y_{j_{i+1}}, z]$ is the unique path from x_m to $\{\partial, y_1, \ldots, y_{j_i}\}$ in $\mathcal{T}$.

We have partitioned $\mathcal{T}$ into l self-avoiding paths , $\eta^1, \ldots, \eta^l$, with disjoint edges. In order for $\mathcal{T}$ to be produced from the algorithm, it must be the case that at the first step η^1 is created, and then η^2 at the second step, etc. Therefore, the probability that $\mathcal{T}$ is created is

$$\hat{p}(\eta^1)\, \hat{p}_{A^1}(\eta^2) \,\cdots\, \hat{p}_{A^{l-1}}(\eta^l),$$

where $A^i = A \setminus \{y_0, \ldots, y_{j_i}\}$. Using (10.4) and Lemma 10.2.1, we get the proposition. $\square$

Corollary 10.5.2 (Uniform Spanning Tree) *Suppose p is simple random walk in $\mathbb{Z}^d$ and $A \subset \mathbb{Z}^d$ is finite and connected. Then Wilson's algorithm chooses a spanning tree for A_W from the uniform measure on spanning trees. The number of spanning trees is $(2d)^{\#(A)}\, F_A(A)^{-1}$.*

Proof If p is simple random walk, then the probability that any particular spanning tree $\mathcal{T}$ is produced equals $\hat{p}(\mathcal{T}) = (2d)^{-\#(A)}\, F_A(A)$. □

10.6 Fomin's identity

In this section, we prove a result that relates the probability that the paths of a loop-erased walk and a random walk do not intersect to a determinant involving only the random walk.

Theorem 10.6.1 (Fomin's identity) *Suppose $A \subset \mathbb{Z}^d$ and $z_1, z_2, w_1, w_2 \in \partial A$. For $j = 1, 2$, let*

$$q_j = q_j(A; z_1, z_2, w_1, w_2) = \sum p(\eta^1)\, p(\eta^2), \tag{10.8}$$

where the sum is over all A-excursions such that η^1 has endpoints z_1, w_j; η^2 has endpoints z_1, w_{3-j}; and $\mathrm{LE}(\eta_1) \cap \eta_2 \cap A = \emptyset$. Then,

$$q_1 - q_2 = \det \begin{bmatrix} H_{\partial A}(z_1, w_1) & H_{\partial A}(z_1, w_2) \\ H_{\partial A}(z_2, w_1) & H_{\partial A}(z_2, w_2) \end{bmatrix}.$$

Proof Since

$$H_{\partial A}(z_i, w_j) = \sum p(\omega^{ij})$$

where the sum is over A-excursions ω^{ij} from z_i to w_j, we have

$$\det \begin{bmatrix} H_{\partial A}(z_1, w_1) & H_{\partial A}(z_1, w_2) \\ H_{\partial A}(z_2, w_1) & H_{\partial A}(z_2, w_2) \end{bmatrix} = \sum_{\omega^{11}} \sum_{\omega^{22}} p(\omega^{11})\, p(\omega^{22}) - \sum_{\omega^{12}} \sum_{\omega^{21}} p(\omega^{12})\, p(\omega^{21}).$$

Let $\mathcal{A}^1$ denote the set of ordered pairs $(\omega^{11}, \omega^{22})$ such that $\mathrm{LE}(\omega^{11}) \cap \omega^{22} \cap A \neq \emptyset$ and $\mathcal{A}^2$ the set of ordered pairs $(\omega^{12}, \omega^{21})$ such that $\mathrm{LE}(\omega^{12}) \cap \omega^{21} \cap A \neq \emptyset$. We will find a one-to-one correspondence between $\mathcal{A}^1$ and $\mathcal{A}^2$, $(\omega^{11}, \omega^{22}) \leftrightarrow (\Lambda\omega^{11}, \Lambda\omega^{22})$ such that

$$p(\omega^{11})\, p(\omega^{22}) = p(\Lambda\omega^{11})\, p(\Lambda\omega^{22}).$$

Once we have done this, we can see that all these terms will cancel in the sum and we get (10.8).

Suppose $\omega^{11} = [z_1, x_1, \ldots, x_m, w_1]$ and $\omega^{22} = [z_2, y_1, \ldots, y_n, w_2]$ are given and suppose that $\mathrm{LE}(\omega^{11}) = [z_1, \hat{x}_1, \ldots, \hat{x}_k, w_1]$. If $\mathrm{LE}(\omega^{11}) \cap \omega^{22} \cap A \neq \emptyset$, we can let j_1 be the smallest positive integer j such that $\hat{x}_j \in \{y_1, \ldots, y_n\}$ and l_1 the largest integer l such that $x_l = \hat{x}_j$. Let l_2 be the largest integer l such that $y_j = \hat{x}_j$. We now set

$$\omega^{12} = [z_1, x_1, \ldots, x_{l_1}, y_{l_2+1}, \ldots, y_n, w_2],$$

$$\omega^{21} = [z_2, y_1, \ldots, y_{l_2}, x_{l_1+1}, \ldots, x_m, w_1].$$

One can check that this correspondence works. □

10.7 Long time behavior

Let S_n denote a random walk with increment distribution $p \in \mathcal{P}_d, d \geq 3$. Let I_n denote the indicator function that S_n is "erased" in loop-erasing. To be more precise, I_n is the indicator function of the event

$$\mathrm{LE}[S_0, \ldots, S_n] \cap S[n+1, \infty) = \emptyset.$$

11
Appendix

11.1 Some expansions

11.1.1 Riemann sums

In this book we often approximate sums by integrals. Here we give bounds on the size of the error in such approximation.

Lemma 11.1.1 *If $f : (0,\infty) \to \mathbb{R}$ is a C^2 function, and b_n is defined by*

$$b_n = f(n) - \int_{n-(1/2)}^{n+(1/2)} f(s)\, ds,$$

then

$$|b_n| \leq \frac{1}{24} \sup\left\{ |f''(r)| : |n-r| \leq \frac{1}{2} \right\}. \tag{11.1}$$

If $\sum |b_n| < \infty$, let

$$C = \sum_{n=1}^{\infty} b_n, \qquad B_n = \sum_{j=n+1}^{\infty} |b_n|.$$

Then

$$\sum_{j=1}^{n} f(j) = \int_{1/2}^{n+(1/2)} f(s)\, ds + C + O(|B_n|).$$

Also, for all $m < n$

$$\left| \sum_{j=m}^{n} f(j) - \int_{m-(1/2)}^{n+(1/2)} f(s)\, ds \right| \leq B_m.$$

Proof Taylor's theorem shows that for $|s-n| \leq 1/2$,

$$f(s) = f(n) + (s-n)\, f'(n) + \frac{1}{2}\, f''(r_s)\, (r_s - n)^2,$$

for some $|n - r_s| < 1/2$. Hence, for such s,

$$|f(s) + f(-s) - 2f(n)| \leq \frac{s^2}{2} \sup\{|f''(r)| : |n-r| \leq s\}.$$

Integrating gives (11.1). The rest is straightforward. □

Example. Suppose $\alpha < 1, \beta \in \mathbb{R}$ and

$$f(n) = n^\alpha \log^\beta n.$$

Note that for $t \geq 2$,

$$|f''(t)| \leq c\, t^{\alpha-2} \log^\beta t.$$

Therefore, there is a $C(\alpha,\beta)$ such that

$$\begin{aligned}
\sum_{j=2}^{n} n^\alpha \log^\beta n &= \int_2^{n+(1/2)} t^\alpha \log^\beta t\, dt + C(\alpha,\beta) + O(n^{\alpha-1} \log^\beta n) \\
&= \int_2^{n} t^\alpha \log^\beta t\, dt + \frac{1}{2}\, n^\alpha \log^\beta n + C(\alpha,\beta) + O(n^{\alpha-1} \log^\beta n) \qquad (11.2)
\end{aligned}$$

11.1.2 Logarithm

Here we collect some facts about the (complex) logarithm. The starting point is the series

$$\log(1+z) = \left[\sum_{j=1}^{k} \frac{(-1)^{j+1} z^j}{j}\right] + O_\epsilon(|z|^{k+1}), \quad |z| \leq 1-\epsilon.$$

Therefore, If $r \in (0,1)$ and $|\xi| \leq rt$,

$$\log\left(1+\frac{\xi}{t}\right)^t = \xi - \frac{\xi^2}{2t} + \frac{\xi^3}{3t^2} + \cdots + (-1)^{k+1} \frac{\xi^k}{kt^{k-1}} + O_r\left(\frac{|\xi|^{k+1}}{t^k}\right),$$

$$\left(1+\frac{\xi}{t}\right)^t = e^\xi \, \exp\left\{-\frac{\xi^2}{2t} + \frac{\xi^3}{3t^2} + \cdots + (-1)^{k+1} \frac{\xi^k}{kt^{k-1}} + O_r\left(\frac{|\xi|^{k+1}}{t^k}\right)\right\}. \qquad (11.3)$$

If $|\xi|^2/t$ is not too big, we can expand the exponential in a Taylor series. Recall that for fixed $R < \infty$, we can write

$$e^z = 1 + z + \frac{z^2}{2!} + \cdots + \frac{z^k}{k!} + O_R(|z|^{k+1}), \quad |z| \leq R.$$

Therefore, if $r \in (0,1), R < \infty$, $|\xi| \leq rt, |\xi|^2 \leq Rt$, we can write

$$\left(1+\frac{\xi}{t}\right)^t = e^\xi \left[1 - \frac{\xi^2}{2t} + \frac{8\xi^3 + 3\xi^4}{24t^2} + \cdots + \frac{f_k(\xi)}{t^{k-1}} + O\left(\frac{|\xi|^{2k+2}}{t^k}\right)\right], \qquad (11.4)$$

where f_k is a polynomial of degree $2(k-1)$ and the implicit constant in the $O(\cdot)$ term depends only on r, R and k. In particular,

$$\left(1+\frac{1}{n}\right)^n = e \left[1 - \frac{1}{2n} + \frac{11}{24\, n^2} + \cdots + \frac{b_k}{n^k} + O\left(\frac{1}{n^{k+1}}\right)\right]. \qquad (11.5)$$

Lemma 11.1.2 *There exists a constant γ (called Euler's constant) and $b_2, b_3, \ldots$ such that for every integer $k \geq 2$,*

$$\sum_{j=1}^{n} \frac{1}{j} = \log n + \gamma + \frac{1}{2n} + \sum_{l=2}^{k} \frac{b_l}{n^l} + O\left(\frac{1}{n^{k+1}}\right).$$

In fact,

$$\gamma = \lim_{n\to\infty} \left[\sum_{j=1}^{n} \frac{1}{j}\right] - \log n = \int_0^1 (1-e^{-t})\,\frac{1}{t}\,dt - \int_1^\infty e^{-t}\,\frac{1}{t}\,dt. \tag{11.6}$$

Proof We start by writing a telescoping sum

$$\sum_{j=1}^{n} \frac{1}{j} = \log\left(n + \frac{1}{2}\right) + \log 2 + \sum_{j=1}^{n} \beta_j,$$

where

$$\beta_j = \frac{1}{j} - \log\left(j + \frac{1}{2}\right) + \log\left(j - \frac{1}{2}\right) = \sum_{k=1}^{\infty} \frac{2}{(2k+1)\,(2j)^{2k+1}}.$$

In particular, $\beta_j = O(j^{-3})$, and hence $\sum \beta_j < \infty$. We can write

$$\sum_{j=1}^{n} \frac{1}{j} = \log\left(n + \frac{1}{2}\right) + \gamma - \sum_{j=n+1}^{\infty} \beta_j = \log n + \gamma + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n\,2^n} - \sum_{j=n+1}^{\infty} \beta_j,$$

where γ is the constant

$$\gamma = \log 2 + \sum_{j=1}^{\infty} \beta_j.$$

We can take as many terms as we wish of this expansion.

We will sketch the proof of (11.6) leaving the details to the reader. By Taylor's series, we know that

$$\log n = -\log\left[1 - \left(1 - \frac{1}{n}\right)\right] = \sum_{j=1}^{\infty} \left(1 - \frac{1}{n}\right)^j \frac{1}{j}.$$

Therefore,

$$\begin{aligned} \gamma &= \lim_{n\to\infty} \left[\sum_{j=1}^{n} \frac{1}{j}\right] - \log n \\ &= \lim_{n\to\infty} \sum_{j=1}^{n} \left[1 - \left(1 - \frac{1}{n}\right)^j\right] \frac{1}{j} \;-\; \lim_{n\to\infty} \sum_{n+1}^{\infty} \left(1 - \frac{1}{n}\right)^j \frac{1}{j}. \end{aligned}$$

We now use the approximation $(1 - n^{-1})^n \sim e^{-1}$ to write

$$\lim_{n\to\infty} \sum_{j=1}^{n} \left[1 - \left(1 - \frac{1}{n}\right)^j\right] \frac{1}{j} = \lim_{n\to\infty} \sum_{j=1}^{n} \frac{1}{n}\,(1 - e^{-j/n})\,\frac{1}{j/n} = \int_0^1 (1-e^{-t})\,\frac{1}{t}\,dt,$$

$$\lim_{n\to\infty}\sum_{n+1}^{\infty}\left(1-\frac{1}{n}\right)^j\frac{1}{j}=\lim_{n\to\infty}\sum_{n+1}^{\infty}\frac{1}{n}\,e^{-j/n}\,\frac{1}{j/n}=\int_1^{\infty}e^{-t}\,\frac{1}{t}\,dt.$$

□

Lemma 11.1.3 *Suppose $\alpha\in\mathbb{R}$ and m is a positive integer. There exist constants $r_0,r_1,\ldots,$ such that if k is a positive integer and $n\geq m$,*

$$\prod_{j=m}^{n}\left(1-\frac{\alpha}{m}\right)=r_0\,n^{-\alpha}\left[1+\frac{r_1}{n}+\cdots+\frac{r_k}{n^k}+O\left(\frac{1}{n^{k+1}}\right)\right].$$

Proof Without loss of generality we assume that $|\alpha|\leq 2m$; if this does not hold we can factor out the first few terms of the product and then analyze the remaining terms. Note that

$$\log\prod_{j=m}^{n}\left(1-\frac{\alpha}{m}\right)=\sum_{j=m}^{n}\log\left(1-\frac{\alpha}{m}\right)\;=-\sum_{j=m}^{n}\sum_{l=1}^{\infty}\frac{\alpha^l}{l\,m^l}=-\sum_{l=1}^{\infty}\sum_{j=m}^{n}\frac{\alpha^l}{l\,m^l}.$$

For the $l=1$ term we have

$$\sum_{j=m}^{n}\frac{\alpha}{m}=-\sum_{j=1}^{m-1}\frac{\alpha}{m}+\alpha\left[\log n+\gamma+\frac{1}{2n}+\sum_{l=2}^{k}\frac{b_l}{n^l}+O\left(\frac{1}{n^{k+1}}\right)\right].$$

All of the other terms can be written in powers of $(1/m)$. Therefore, we can write

$$\log\prod_{j=m}^{n}\left(1-\frac{\alpha}{m}\right)=-\alpha\,\log n+\sum_{l=0}^{k}\frac{C_l}{n^l}+O\left(\frac{1}{n^{k+1}}\right).$$

The lemma is then obtained by exponentiating both sides. □

11.2 Martingales

Martingales are one of the most important concepts in probability. A filtration $\mathcal{F}_0\subset\mathcal{F}_1\subset\cdots$ is an increasing sequence of σ-algebras.

Definition. A sequence of integrable random variables $M_0,M_1,\ldots$ is called a *martingale* with respect to the filtration $\{\mathcal{F}_n\}$ if each M_n is $\mathcal{F}_n$-measurable and for each $m\leq n$,

$$\mathbb{E}[M_n\mid\mathcal{F}_m]=M_m. \tag{11.7}$$

If (11.7) is replaced with $\mathbb{E}[M_n\mid\mathcal{F}_m]\geq M_m$ the sequence is called a *submartingale*. If (11.7) is replaced with $\mathbb{E}[M_n\mid\mathcal{F}_m]\leq M_m$ the sequence is called a *supermartingale*

Using properties of conditional expectation, it is easy to see that to verify (11.7) it suffices to show for each n that $\mathbb{E}[M_{n+1}\mid\mathcal{F}_n]=M_n$. This equality only needs to hold up to an event of probability zero; in fact, the conditional expectation is only defined up to events of probability zero. If the filtration is not specified, then the assumption is that $\mathcal{F}_n$ is the σ-algebra generated by $M_0,\ldots,M_n$. If $M_0,X_1,X_2,\ldots$ are independent random variables with $\mathbb{E}[|M_0|]<\infty$ and $\mathbb{E}[X_j]=0$ for $j\geq 1$, and

$$M_n=M_0+X_1+\cdots+X_n,$$

then $M_0, M_1, \ldots$ is a martingale. We omit the proof of the next lemma which is the conditional expectation version of Jensen's inequality.

Lemma 11.2.1 (Jensen's inequality) *If X is an integrable random variable; $f : \mathbb{R} \to \mathbb{R}$ is convex with $\mathbb{E}[|f(X)|] < \infty$; and $\mathcal{F}$ is a σ-algebra, then $\mathbb{E}[f(X) \mid \mathcal{F}] \geq f(\mathbb{E}[X \mid \mathcal{F}])$. In particular, if $M_0, M_1, \ldots$ is a martingale; $f : \mathbb{R} \to \mathbb{R}$ is convex with $\mathbb{E}[|f(M_n)|] < \infty$ for all n; and $Y_n = f(M_n)$; then $Y_0, Y_1, \ldots$ is a submartingale.*

In particular, if $M_0, M_1, \ldots$ is a martingale then

- if $\alpha \geq 1$, $Y_n := |M_n|^\alpha$ is a submartingale;
- if $b \in \mathbb{R}$, then $Y_n := e^{bM_n}$ is a submartingale.

In both cases, this is assuming that $\mathbb{E}[Y_n] < \infty$.

11.2.1 Optional Sampling Theorem

A *stopping time* with respect to a filtration $\{\mathcal{F}_n\}$ is a $\{0, 1, \ldots\} \cup \{\infty\}$ valued random variable T such that for each n, $\{T \leq n\}$ is $\mathcal{F}_n$-measurable. If T is a stopping time, and n is a positive integer, then $T_n := T \wedge n$ is a stopping time satisfying $T_n \leq n$.

Proposition 11.2.2 *Suppose $M_0, M_1, \ldots$ is a martingale and T is a stopping time each with respect to the filtration $\mathcal{F}_n$. Then $Y_n := M_{T_n}$ is a martingale with respect to $\mathcal{F}_n$. In particular,*

$$\mathbb{E}[M_0] = \mathbb{E}[M_{T_n}].$$

Proof Note that

$$Y_{n+1} = M_{T_n} \, 1\{T \leq n\} + M_{n+1} \, 1\{T \geq n+1\}.$$

The event $\{T \geq n+1\}$ is the complement of the event $\{T \leq n\}$ and hence is $\mathcal{F}_n$-measuable. Therefore, by properties of conditional expecation,

$$\mathbb{E}[M_{n+1} \, 1\{T \geq n+1\} \mid \mathcal{F}_n] = 1\{T \geq n+1\} \, \mathbb{E}[M_{n+1} \mid \mathcal{F}_n] = 1\{T \geq n+1\} \, M_n.$$

Therefore,

$$\mathbb{E}[Y_{n+1} \mid \mathcal{F}_n] = M_{T_n} \, 1\{T \leq n\} + M_n \, 1\{T \geq n+1\} = Y_n.$$

□

The optional sampling theorem states that under certain condtions, if $\mathbb{P}\{T < \infty\} = 1$, then $\mathbb{E}[M_0] = \mathbb{E}[M_T]$. However, this does not hold without some further assumptions. For example, if M_n is one-dimensional simple random walk starting at the origin and T is the first n such that $M_n = 1$, then $\mathbb{P}\{T < \infty\} = 1$, $M_T = 1$, and hence $\mathbb{E}[M_0] \neq \mathbb{E}[M_T]$. In the next theorem we include a number of sufficient conditions under which we can conclude that $\mathbb{E}[M_0] = \mathbb{E}[M_T]$.

Theorem 11.2.3 (Optional Sampling Theorem) *Suppose $M_0, M_1, \ldots$ is a martingale and T is a stopping time with respect to the filtration $\{\mathcal{F}_n\}$. Suppose that $\mathbb{P}\{T < \infty\} = 1$ and $\mathbb{E}[|M_T|] < \infty$. Suppose also that at least one of the following conditions holds:*

- *There is a* $K < \infty$ *such that* $\mathbb{P}\{T \le K\} = 1$.
- *There exists an integrable random variable* Y *such that for all* n, $|M_{T_n}| \le Y$.
- $\mathbb{E}\,[M_T] < \infty$ *and* $\lim_{n\to\infty} \mathbb{E}[|M_n|; T > n] = 0$.
- *The random variables* $M_0, M_1, \ldots$ *are uniformly integrable, i.e., for every* $\epsilon > 0$ *there is a* $K_\epsilon < \infty$ *such that for all* n,

$$\mathbb{E}[|M_n|; |M_n| > K_\epsilon] < \epsilon.$$

- *There exists an* $\alpha > 1$ *and a* $K < \infty$ *such that for all* n, $\mathbb{E}[|M_n|^\alpha] \le K$.

Then $\mathbb{E}[M_0] = \mathbb{E}[M_T]$.

Proof We will consider the conditions in order. The sufficiency of the first follows immediately from Proposition 11.2.2. We know that $M_{T_n} \to M_T$ with probability one. Proposition 11.2.2 gives $\mathbb{E}[M_{T_n}] = \mathbb{E}[M_0]$. Hence we need to show that

$$\lim_{n\to\infty} \mathbb{E}[M_{T_n}] = \mathbb{E}[M_T]. \tag{11.8}$$

If the second condition holds, then this limit is justified by the dominated convergence theorem. Now assume the third condtion. Note that

$$M_T = M_{T_n} + M_T\, 1\{T > n\} - M_n\, 1\{T > n\}.$$

Since $\mathbb{P}\{T > n\} \to 0$, and $\mathbb{E}[|M_T|] < \infty$, it follows from standard arguments that

$$\lim_{n\to\infty} \mathbb{E}[M_T\, 1\{T > n\}] = 0.$$

Hence if $\mathbb{E}[M_n\, 1\{T > n\}] \to 0$, we have (11.8). Standard exercises show that the fourth implies the third and the fifth condition implies the fourth, so either the fourth or fifth condition is sufficient. □

11.2.2 Maximal inequality

Here we prove an important maximal inequality that can be considered a generalization of Chebyshev's inequality.

Theorem 11.2.4 (Maximal inequality) *Suppose* $M_0, M_1, \ldots$ *is a nonnegative submartingale with respect to* $\{\mathcal{F}_n\}$ *and* $\lambda > 0$. *Then*

$$\mathbb{P}\left\{\max_{0\le j\le n} M_j \ge \lambda\right\} \le \frac{\mathbb{E}[M_n]}{\lambda}.$$

Proof Let $T = \min\{j \ge 0 : M_j \ge \lambda\}$. Then,

$$\mathbb{P}\left\{\max_{0\le j\le n} M_j \ge \lambda\right\} = \sum_{j=0}^{n} \mathbb{P}\{T = j\},$$

$$\mathbb{E}[M_n] \ge \mathbb{E}[M_n; T \le n] = \sum_{j=0}^{n} \mathbb{E}[M_n; T = j].$$

Since M_n is a submartingale and $\{T = j\}$ is $\mathcal{F}_j$-measurable,

$$\mathbb{E}[M_n; T = j] = \mathbb{E}[\mathbb{E}[M_n \mid \mathcal{F}_j]; T = j] \geq \mathbb{E}[M_j; T = j] \geq \lambda\, \mathbb{P}\{T = j\}.$$

Combining these estimates gives the theorem. □

By combining Theorem 11.2.4 with Lemma 11.2.1 gives the following theorem as a corollary.

Theorem 11.2.5 (Martingale maximal inequalities) *Suppose $M_0, M_1, \ldots$ is a martingale with respect to $\{\mathcal{F}_n\}$ and $\lambda > 0$. Then if $\alpha \geq 1, b \geq 0$,*

$$\mathbb{P}\left\{\max_{0 \leq j \leq n} |M_j| \geq \lambda\right\} \leq \frac{\mathbb{E}[|M_n|^\alpha]}{\lambda^\alpha}, \tag{11.9}$$

$$\mathbb{P}\left\{\max_{0 \leq j \leq n} M_j \geq \lambda\right\} \leq \frac{\mathbb{E}[e^{bM_n}]}{e^{b\lambda}}.$$

Corollary 11.2.6 *Let $X_1, X_2, \ldots$ be independent, identically distributed random variables in $\mathbb{R}$ with mean zero, and let k be a positive integer for which $\mathbb{E}[|X_1|^{2k}] < \infty$. There exists $c < \infty$ such that for all $\lambda > 0$,*

$$\mathbb{P}\left\{\max_{0 \leq j \leq n} |S_j| \geq \lambda \sqrt{n}\right\} \leq c\, \lambda^{-2k}. \tag{11.10}$$

Proof Fix k and allow constants to depend on k. Note that

$$\mathbb{E}[S_n^{2k}] = \sum \mathbb{E}[X_{j_1} \cdots X_{j_{2k}}],$$

where the sum is over all $(j_1, \ldots, j_{2k}) \in \{1, \ldots, n\}^{2k}$. If there exists an l such that $j_i \neq j_l$ for $i \neq l$, then we can use independence and $\mathbb{E}[X_{j_l}] = 0$ to see that $\mathbb{E}[X_{j_1} \cdots X_{j_{2k}}] = 0$. Hence

$$\mathbb{E}[S_n^{2k}] = \sum \mathbb{E}[X_{j_1} \cdots X_{j_{2k}}],$$

where the sum is over all $(2k)$-tuples such that if $l \in \{j_1, \ldots, j_{2k}\}$, then l appears at least twice. The number of such $(2k)$-tuples is $O(n^k)$ and hence we can see that

$$\mathbb{E}\left[\left(\frac{S_n}{\sqrt{n}}\right)^{2k}\right] \leq c.$$

Hence we can apply (11.9) to the martingale $M_j = S_j/\sqrt{n}$.

□

Comment. The last corollary holds for all real $k \geq 1$ but a different proof is needed. See the proof of Proposition 2.4.6.

Corollary 11.2.7 *Let $X_1, X_2, \ldots$ be independent, identically distributed random variables in $\mathbb{R}$ with mean zero, variance σ^2, and such that for some $\delta > 0$, the moment generating function $\psi(t) = \mathbb{E}[e^{tX_j}]$ exists for $|t| < \delta$. Let $S_n = X_1 + \cdots + X_n$. There is an $\epsilon_0 > 0$ such that for all n and all $0 \leq r \leq \epsilon_0 n$,*

$$\mathbb{P}\left\{\max_{0 \leq j \leq n} S_j \geq r\, \sigma \sqrt{n}\right\} \leq e^{-r^2/2} \exp\left\{O\left(\frac{r^3}{\sqrt{n}}\right)\right\}. \tag{11.11}$$

if $\mathbb{P}\{X_1 \geq R\} = 0$ *for some* R, *we may choose* $\epsilon_0 = \infty$.

Proof Without loss of generality, we may assume $\sigma^2 = 1$. Let $\epsilon_0 = \delta/2$. The moment generating function of $S_n = X_1 + \cdots + X_n$ is $\psi(t)^n$. Hence, if $0 \leq a \leq \epsilon_0 n$,

$$\mathbb{P}\left\{\max_{0\leq j\leq n} S_j \geq r\sqrt{n}\right\} \leq e^{-a^2}\,\psi(r/\sqrt{n})^n.$$

Using the expansion for $\psi(t)$ at zero,

$$\psi(t) = 1 + \frac{t^2}{2} + O(t^3),$$

we see that for $0 \leq r \leq \epsilon_0 n$,

$$\psi(r/\sqrt{n})^n = \left[1 + \frac{r^2}{2n} + O\left(\frac{r^3}{n^{3/2}}\right)\right]^n \leq e^{r^2/2}\exp\left\{O\left(\frac{r^3}{\sqrt{n}}\right)\right\}.$$

This gives (11.11). If $\mathbb{P}\{X_1 > R\} = 0$, then (11.11) holds for $r \geq R\sqrt{n}$ trivially. For $\epsilon_0\sqrt{n} \leq r \leq R\sqrt{n}$, we need only change the constant in the $O(\cdot)$ term. □

Remark. From the last corollary, we also get the following modulus of continuity result for random walk. Let $X_1, X_2, \ldots$ and S_n be as in the previous lemma. There exist c, b, ϵ such that for every $m \leq n$ and every $0 \leq r \leq \epsilon\sqrt{m}$

$$\mathbb{P}\{\max_{0\leq j\leq n}\max_{1\leq k\leq m} |S_{k+j} - S_j| \geq r\sqrt{m}\} \leq c\,n\,e^{-br^2}. \tag{11.12}$$

The following generalization is proved similarly.

Corollary 11.2.8 *Suppose* $M_0, M_1, M_2, \ldots$ *is a martingale with respect to* $\mathcal{F}_n$ *with* $M_0 = 0$, *and suppose there exist* $C < \infty$ *and* $t > 0$ *such that for all* $n \geq 1$

$$\mathbb{E}[e^{t|M_n - M_{n-1}|} \mid \mathcal{F}_{n-1}] \leq C. \tag{11.13}$$

Then there exists c, N *such that if* $n \geq N$ *and* $a > 0$,

$$\mathbb{P}\{|M_n| \geq a\sqrt{n}\} \leq c\,e^{-at}.$$

Proof Using (11.13), we can see that

$$\sup_n \mathbb{E}[e^{tM_n/\sqrt{n}} + e^{-tM_n/\sqrt{n}}] < \infty.$$

□

Although the next simple lemma is not about martingales, it will be useful to include it here.

Lemma 11.2.9 *If* $0 < \alpha < \infty$, $0 < r < 1$ *and* X_n *is a binomial random variable with parameters* n *and* $\alpha\,e^{-2\alpha/r}$, *then*

$$\mathbb{P}\{X_n \geq rn\} \leq e^{-\alpha n}.$$

Proof

$$\mathbb{P}\{X_n \geq rn\} \leq e^{-2\alpha n}\,\mathbb{E}[e^{(2\alpha/r)X_n}] \leq e^{-2\alpha n}\,[1+\alpha]^n \leq e^{-\alpha n}.$$

□

11.2.3 Continuous martingales

A process M_t adapted to a filtration $\mathcal{F}_t$ is called a *continuous martingale* if for each $s < t$, $\mathbb{E}[M_t \mid M_s] = M_s$ and with probability one the function $t \mapsto M_t$ is continuous. If M_t is a continuous martingale, and $\delta > 0$, then

$$M_n^{(\delta)} := M_{\delta n}$$

is a discrete time martingale. Using this, we can extend results about discrete time martingales to continuous martingales. We state one such result here.

Theorem 11.2.10 (Optional Sampling Theorem) *Suppose M_t is a uniformly integrable continuous martingale and τ is a stopping time with $\mathbb{P}\{\tau < \infty\} = 1$ and $\mathbb{E}[|M_\tau|] < \infty$. Suppose that*

$$\lim_{t\to\infty} \mathbb{E}[|M_t| : \tau > t] = 0.$$

Then

$$\mathbb{E}[M_T] = \mathbb{E}[M_0].$$

11.3 Joint normal distributions

A random vector $Z = (Z_1, \ldots, Z_d) \in \mathbb{R}^d$ is said to have a *(mean zero) joint normal distribution* if there exist independent (one-dimensional) mean zero, variance one normal random variables $N_1, \ldots, N_n$ and scalars a_{jk} such that

$$Z_j = a_{j1}\,N_1 + \cdots + a_{jn}\,N_n, \qquad j = 1, \ldots, d,$$

or in matrix form

$$Z = AN.$$

Here $A = (a_{jk})$ is a $d \times n$ matrix and Z, N are column vectors. Note that

$$\mathbb{E}(Z_j Z_k) = \sum_{m=1}^{n} a_{jm}\,a_{km}.$$

In other words, the covariance matrix $\Gamma = [\,\mathbb{E}(Z_j Z_k)\,]$ is the $d \times d$ symmetric matrix

$$\Gamma = AA^T.$$

We say Z has a *nondegenerate* distribution if Γ is invertible.

The characteristic function of Z can be computed easily using the known formula for the characteristic function of N_k,

$$\mathbb{E}[e^{itN_k}] = e^{-t^2/2},$$

$$\begin{aligned}
\mathbb{E}[\exp\{i\theta\cdot Z\}] &= \mathbb{E}\left[\exp\left\{i\sum_{j=1}^d \theta_j \sum_{k=1}^n a_{jk}\,N_k\right\}\right] \\
&= \mathbb{E}\left[\exp\left\{i\sum_{k=1}^n N_k \sum_{j=1}^d \theta_j a_{jk}\right\}\right] \\
&= \prod_{k=1}^n \mathbb{E}\left[\exp\left\{iN_k \sum_{j=1}^d \theta_j a_{jk}\right\}\right] \\
&= \prod_{k=1}^n \exp\left\{-\frac{1}{2}\left(\sum_{j=1}^d \theta_j a_{jk}\right)^2\right\} \\
&= \exp\left\{-\frac{1}{2}\sum_{k=1}^n\sum_{j=1}^d\sum_{l=1}^d \theta_j\,\theta_l\,a_{jk}\,a_{lk}\right\} \\
&= \exp\left\{-\frac{1}{2}\,\theta AA^T\theta^T\right\} = \exp\left\{-\frac{1}{2}\,\theta\Gamma\theta^T\right\}.
\end{aligned}$$

Since the characteristic function determines the distribution, we see that the distribution of Z depends only on Γ.

The matrix Γ is symmetric and nonnegative definite. Hence we can find an orthogonal basis $u_1,\ldots,u_d$ of unit vectors in $\mathbb{R}^d$ that are eigenvectors of Γ with nonnegative eigenvalues $\alpha_1,\ldots,\alpha_d$. The random variable

$$Z = \sqrt{\alpha_1}\,N_1\,u_1 + \cdots + \sqrt{\alpha_d}\,N_d\,u_d$$

has a joint normal distribution with covariance matrix Γ. In matrix languauge, we have written $\Gamma = \Lambda\,\Lambda^T = \Lambda^2$ for a $d\times d$ nonnegative definite symmetric matrix Λ. The distribution is nondegenerate if and only if all of the α_j are strictly positive.

Comment. Although we allow the matrix A to have n columns, what we have shown is that there is a symmetric, positive definite $d\times d$ matrix Λ which gives the same distribution. Hence joint normal distribution in $\mathbb{R}^d$ can be described as linear combinations of d independent one-dimensional normals. Moreover, if we choose the correct orthogonal basis for $\mathbb{R}^d$, the components of Z with respect to that basis are independent normals.

If Γ is invertible, then Z has a density $f(z^1,\ldots,z^d)$ with respect to Lebesque measure that can be computed using the inversion formula

$$\begin{aligned}
f(z^1,\ldots,z^d) &= \frac{1}{(2\pi)^d}\int e^{-i\theta\cdot z}\,\mathbb{E}[\exp\{i\theta\cdot Z\}]\,d\theta \\
&= \frac{1}{(2\pi)^d}\int \exp\left\{-i\theta\cdot z - \frac{1}{2}\,\theta\Gamma\theta^T\right\}\,d\theta.
\end{aligned}$$

(Here and for the remainder of this paragraph the integrals are over $\mathbb{R}^d$ and $d\theta$ represents $d^d\theta$.) To evaluate the integral, we start with the substitution $\theta_1 = \Lambda\theta$ which gives

$$\int \exp\left\{-i\theta\cdot z - \frac{1}{2}\,\theta\Gamma\theta^T\right\}\,d\theta = \frac{1}{\det\Lambda}\int e^{-|\theta_1|^2/2}\,e^{-i(\theta_1\cdot\Lambda^{-1}z)}\,d\theta_1.$$

By completing the square we see that the right-hand side equals

$$\frac{e^{-|\Lambda^{-1}z|^2/2}}{\det \Lambda} \int \exp\left\{\frac{1}{2}\,(i\theta_1 - \Lambda^{-1}z)\cdot(i\theta_1 - \Lambda^{-1}z)\right\}\, d\theta_1.$$

The substitution $\theta_2 = \theta_1 - i\Lambda^{-1}z$ gives

$$\int \exp\left\{\frac{1}{2}\,(i\theta_1 - \Lambda^{-1}z)\cdot(i\theta_1 - \Lambda^{-1}z)\right\}\, d\theta_1 = \int e^{-|\theta_2|^2/2}\, d\theta_2 = (2\pi)^{d/2}.$$

Hence, the density of Z is

$$f(z) = \frac{1}{(2\pi)^{d/2}\sqrt{\det \Gamma}}\, e^{-|\Lambda^{-1}z|^2/2} = \frac{1}{(2\pi)^{d/2}\sqrt{\det \Gamma}}\, e^{-(z\cdot\Gamma^{-1}z)/2}. \tag{11.14}$$

Corollary 11.3.1 *Suppose $Z = (Z_1, \ldots, Z_d)$ has a mean zero, joint normal distribution such that $\mathbb{E}[Z_j Z_k] = 0$ for all $j \neq k$. Then $Z_1, \ldots, Z_d$ are independent.*

Proof Suppose $\mathbb{E}[Z_j Z_k] = 0$ for all $j \neq k$. Then Z has the same distribution as

$$(b_1 N_1, \ldots, b_d N_d),$$

where $b_j = \sqrt{\mathbb{E}[Z_j^2]}$. In this representation, the components are obviously independent. □

Comment. If $Z_1, \ldots, Z_d$ are mean zero random varaibles satisfying $\mathbb{E}[Z_j Z_k] = 0$ for all $j \neq k$, they are called orthogonal. Independence implies orthogonality but the converse is not always true. However, the corollary tells us that the converse is true in the case of joint normal random varaibles. Orthogonality is often easier to verify than independence.

11.4 Markov chains

A *(time-homogeneous) Markov chain* on a countable state space D is a process X_n taking values in D whose transitions satisfy

$$\mathbb{P}\{X_{n+1} = x_{n+1} \mid X_0 = x_0, \ldots, X_n = x_n\} = p(x_n, x_{n+1})$$

where $p : D \times D \to [0,1]$ is the *transition function* satisfying $\sum_{y\in D} p(x,y) = 1$ for each x. If A is finite, we call the transition function the transition matrix $P = [p(x,y)]_{x,y\in A}$. The n-step transitions are given by the matrix P^n. In other words, if $p_n(x,y)$ is defined to be $\mathbb{P}\{X_n = y \mid X_0 = x\}$, then

$$p_n(x,y) = \sum_{z\in D} p(x,z)\, p_{n-1}(z,y) = \sum_{z\in D} p_{n-1}(x,z)\, p(z,y).$$

A Markov chain is called *irreducible* if for each $x, y \in A$, there exists an $n = n(x,y) \geq 0$ with $p_n(x,y) > 0$. The chain is *aperiodic* if for each x there is an N_x such that for $n \geq N_x$, $p_n(x,x) > 0$. If D is finite, then the chain is irreducible and aperiodic if and only if there exists an n such that P^n has strictly positive entries.

Theorem 11.4.1 *[Perron-Froebenius Theorem] If P is an $m \times m$ matrix such that for some positive integer n, P^n has all entries strictly positive, then there exists $\alpha > 0$ and vectors $\mathbf{v}, \mathbf{w}$, with strictly positive entries such that*

$$\mathbf{v}\, P = \alpha\, \mathbf{v}, \qquad P\, \mathbf{w} = \alpha\, \mathbf{w}.$$

This eigenvalue is simple and all other eigenvalues of P have absolute value strictly less than α. In particular, if P is the transition matrix for an irreducible aperiodic Markov chain there is a unique invariant probability π satisfying

$$\sum_{x \in D} \pi(x) = 1, \qquad \pi(x) = \sum_{y \in D} \pi(y)\, P(y, x).$$

Proof We first assume that P has all strictly positive entries. It suffices to find a right eigenvector, since the left eigenvector can be handled by considering the transpose of P. We write $\mathbf{w}_1 \geq \mathbf{w}_2$ if every component of $\mathbf{w}_1$ is greater than or equal to the corresponding component of $\mathbf{w}_2$. Similarly, we write $\mathbf{w}_1 > \mathbf{w}_2$ if all the components of $\mathbf{w}_1$ are greater than the corresponding components of $\mathbf{w}_2$. We let $\mathbf{0}$ denote the zero vector and $\mathbf{e}_j$ the vector whose jth component is 1 and whose other components are 0. If $\mathbf{w} \geq \mathbf{0}$, let

$$\lambda_{\mathbf{w}} = \sup\{\lambda : P\mathbf{w} \geq \lambda \mathbf{w}\}.$$

Clearly $\lambda_{\mathbf{w}} < \infty$, and since P has strictly positive entries, $\lambda_{\mathbf{w}} > 0$ for all $\mathbf{w} > 0$. Let

$$\alpha = \sup\{\lambda_{\mathbf{w}} : \mathbf{w} \geq \mathbf{0}, \sum_{j=1}^{m} [\mathbf{w}]_j = 1\}.$$

By compactness and continuity arguments we can see that there exists a $\mathbf{w}$ with $\mathbf{w} \geq \mathbf{0}$, $\sum_j [\mathbf{w}]_j = 1$ and $\lambda_{\mathbf{w}} = \alpha$. We claim that $P\mathbf{w} = \alpha \mathbf{w}$. Indeed, if $[P\mathbf{w}]_j > \alpha [\mathbf{w}]_j$ for some j, one can check that there exist positive ϵ such that $P[\mathbf{w} + \epsilon \mathbf{e}_j] \geq (\alpha + \epsilon)\, [\mathbf{w} + \epsilon \mathbf{e}_j]$, which contradicts the maximality of α. If $\mathbf{v}$ is a vector with both positive and negative component, then for each j,

$$|[P\mathbf{v}]_j| < [P|\mathbf{v}|]_j \leq \alpha\, [|\mathbf{v}|]_j.$$

Here we write $|\mathbf{v}|$ for the vector whose components are the absolute values of the components of $\mathbf{v}$. Hence any eigenvector with both positive and negative values has an eigenvalue with absolute value strictly less than α. Also, if $\mathbf{w}_1, \mathbf{w}_2$ are positive eigenvectors with eigenvalue α, then $\mathbf{w}_1 - t\mathbf{w}_2$ is an eigenvector for each t. If $\mathbf{w}_1$ is not a multiple of $\mathbf{w}_2$ then there is some value of t such that $\mathbf{w}_1 - t\mathbf{w}_2$ has both positive and negative values. Since this is impossible, we conclude that the eigenvector $\mathbf{w}$ is unique. If $\mathbf{w} \geq 0$ is an eigenvector, then the eigenvalue must be positive. Therefore, α is a simple eigenvalue and all other eigenvalues have absolute value strictly less than α. Note that if $\mathbf{v} > 0$, then $P\mathbf{v}$ has all entries strictly positive; hence the eigenvector $\mathbf{W}$ must have all entries strictly positive.

If P is a matrix such that P^n has all entries stricly positive, and $\mathbf{w}$ is an eigenvector of P with eigenvalue α, then $\mathbf{w}$ is an eigenvector for P^n with eigenvalue α^n. Using this, we can conclude the result for P. The final assertion follows by noting that the vector of all 1s is a right eigenvector for a stochastic matrix. $\square$

If P is the transition matrix for a irreducible, aperiodic Markov chain, then $p_n(x,y) \to \pi(y)$ as $n \to \infty$. In fact, this holds for countable state space provided the chain is positive recurrent, i.e., if there exists an invariant probability measure. The next proposition gives a simple, quantitative version of this fact provided the chain satisfies a certain condition which always holds for the finite irreducible, aperiodic case.

Proposition 11.4.2 *Suppose $p : D \times D \to [0,1]$ is the transition probability for a positive recurrent, irreducible, aperiodic Markov chain on a countable state space D. Let π denote the invariant probability measure. Suppose there exists $\epsilon > 0$ and a positive integer k such that for all $x, y \in D$,*

$$\frac{1}{2} \sum_{z \in D} |p_k(x,z) - p_k(y,z)| \leq 1 - \epsilon. \tag{11.15}$$

Then for all positive integers j and all $x \in A$,

$$\frac{1}{2} \sum_{z \in D} |p_j(x,z) - \pi(z)| \leq c\, e^{-\beta j},$$

where $c = (1-\epsilon)^{-1}$ and $e^{-\beta} = (1-\epsilon)^{1/k}$.

Proof If ν is any probability distribution, let us write

$$\nu_j(x) = \sum_{y \in D} \nu(y) p_j(y,x).$$

Then (11.15) implies that for every ν,

$$\frac{1}{2} \sum_{z \in D} |\nu_k(z) - \pi(z)| \leq 1 - \epsilon.$$

In other words we can write $\nu_k = \epsilon\, \pi + (1-\epsilon)\, \nu^{(1)}$ for some probabiility measure $\nu^{(1)}$. By iterating (11.15), we can see that for every integer $i \geq 1$ we can write $\nu_{ik} = (1-\epsilon)^i\, \nu^{(i)} + [1 - (1-\epsilon)^i]\pi$ for some probabiility measure $\nu^{(i)}$. This establishes the result for $j = ki$ (with $c = 1$ for these values of j) and for other j we find i with $ik \leq j < (i+1)k$. □

11.4.1 Chains restricted to subsets

We will now consider Markov chains restricted to a subset of the original state space. If X_n is an irreducible, aperiodic Markov chain with state space D and A is a finite proper subset of D, we write $P_A = [p(x,y)]_{x,y \in A}$. Note that $(P_A)^n = [p_n^A(x,y)]_{x,y \in A}$ where

$$p_n^A(x,y) = \mathbb{P}\{X_n = y; X_0, \ldots, X_n \in A \mid X_0 = x\} = \mathbb{P}^x\{X_n = y, \tau_A > n\}, \tag{11.16}$$

where $\tau_A = \inf\{n : X_n \notin A\}$. Note that

$$\mathbb{P}^x\{\tau_A < \infty\} = \sum_{y \in A} p_n^A(x,y).$$

We call A *connected and aperiodic (with respect to P)* if for each $x, y \in A$, there is an N such that for $n \geq N$, $p_n^A(x,y) > 0$. If A is finite, then A is connected and aperiodic if and only if there exists an n such that $(P_A)^n$ has all entries strictly positive. In this case all of the row sums of P_A are less than or equal to one and (since A is a proper subset) there is at least one row whose sum is strictly less than one.

Suppose X_n is an irreducible, aperiodic Markov chain with state space D and A is a finite, connected, aperiodic proper subset of D. Let α be as in the Perron-Froebenius Theorem. It is easy to see that $0 < \alpha < 1$. Let v, w be the corresponding eigenvectors which we write functions,

$$\sum_{x \in A} v(x)\, p(x,y) = \alpha\, v(y), \qquad \sum_{y \in A} w(y)\, p(x,y) = \alpha\, w(x).$$

We normalize the functions so that

$$\sum_{x \in A} v(x) = 1, \qquad \sum_{x \in A} v(x)\, w(x) = 1,$$

and we let $\pi(x) = v(x)\, w(x)$. Let

$$q^A(x,y) = \alpha^{-1}\, p(x,y)\, \frac{w(y)}{w(x)}.$$

Note that

$$\sum_{y \in A} q^A(x,y) = \frac{\sum_{y \in A} p(x,y)\, w(y)}{\alpha\, w(x)} = 1.$$

In other words, $Q^A := [q^A(x,y)]_{x,y \in A}$ is the transition matrix for a Markov chain which we will denote by Y_n. Note that $(Q^A)^n = [q_n^A(x,y)]_{x,y \in A}$ where

$$q_n^A(x,y) = \alpha^{-n}\, p_n^A(x,y)\, \frac{w(y)}{w(x)}$$

and $p_n^A(x,y)$ is as in (11.16). From this we see that the chain is irreducible and aperiodic. Since

$$\sum_{x \in A} \pi(x)\, q^A(x,y) = \sum_{x \in A} v(x)\, w(x)\, \alpha^{-1}\, p(x,y)\, \frac{w(y)}{w(x)} = \pi(y),$$

we see that π is the invariant probability for this chain.

Proposition 11.4.3 *Under the assumptions above, there exist c, β such that for all n,*

$$|\alpha^{-n}\, p_n^A(x,y) - w(x)\, v(y)| \le c\, e^{-\beta n}.$$

In particular,

$$\mathbb{P}\{X_0, \ldots, X_n \in A \mid X_0 = x\} = w(x)\, \alpha^n\, [1 + O(e^{-\beta n})].$$

Proof Consider the Markov chain with transition matrix Q^A. Choose positive integer k and $\epsilon > 0$ such that $q_k^A(x,y) \ge \epsilon\, \pi(y)$ for all $x, y \in A$. Proposition 11.4.2 gives

$$|q_n^A(x,y) - \pi(y)| \le c\, e^{-\beta n},$$

for some c, β. Since $\pi(y) = v(y)\,w(y)$ and $q_n^A(x,y) = \alpha^{-n}\,p_n^A(x,y)\,w(y)/w(x)$, we get the first assertion. The second assertion follows from the first using $\sum_y v(y) = 1$ and

$$\mathbb{P}\{X_0, \ldots, X_n \in A \mid X_0 = x\} = \sum_{y \in A} p_n^A(x,y).$$

□

Comment. The chain Y_n can be considered the chain derived from X_n by conditioning the chain to "stay in A forever". The probability measures v, π are both "invariant probability measures", but with different interpretations. Roughly speaking, the three measures v, w, π can be described as follows.

- Suppose the chain X_n is observed at a large time n and it known that the chain has stayed in A for all times up to n. Then the conditional distribution on X_n given this information approaches v.
- For $x \in A$, the probability that the chain stays in A up to time n is asymptotic to $w(x)\,\alpha^n$.
- Suppose the chain X_n is observed at a large time n and it is known that the chain has stayed in A and will stay in A for all times up to N where $N \gg n$. Then the conditional distribution on X_n given this information approaches π. We can think of the first term of the product $v(x)\,w(x)$ as the conditional probability of being at x given that the walk has stayed in A up to time n and the second part of the product is the conditional probability given this that the walk stays in A for times between n and N.

The next proposition gives a criterion for determining the rate of convergence to the invariant distribution v. Let us write

$$\hat{p}_n^A(x,y) = \frac{p_n^A(x,y)}{\sum_{z \in A} p_n^A(x,z)} = \mathbb{P}^x\{X_n = y \mid \tau_A > n\}.$$

Proposition 11.4.4 *Suppose X_n is an irreducible, aperiodic Markov chain on the countable state space D. Suppose A is a finite, proper subset of D and $A' \subset A$. Suppose there exist $\epsilon > 0$ and integer $k > 1$ such that the following is true.*

- *If $x \in A$,*
$$\sum_{y \in A'} \hat{p}_k(x,y) \geq \epsilon. \tag{11.17}$$
- *If $x, x' \in A$,*
$$\sum_{y \in A'} |\hat{p}_k^A(x,y) - \hat{p}_k^A(x',y)| \geq \epsilon. \tag{11.18}$$
- *If $x \in A, y \in A'$ and n is a positive integer*
$$\mathbb{P}^y\{\tau_A > n\} \geq \epsilon\,\mathbb{P}^x\{\tau_A > n\}. \tag{11.19}$$

Then there exists $\delta > 0$, depending only on ϵ, such that for all $x, z \in A$,

$$\frac{1}{2}\sum_{y \in A} \left|\hat{p}_{km}^A(x,y) - \hat{p}_{km}^A(z,y)\right| \leq (1-\delta)^m.$$

Proof We fix ϵ and allow all constants in this proof to depend on ϵ. Let $q_n = \max_{y \in A} \mathbb{P}^y\{\tau_A > n\}$. Then (11.19) implies that for all $y \in A'$ and all n, $\mathbb{P}^y\{\tau_A > n\} \geq \epsilon\, q_n$. Combining this with (11.17) gives for all positive integers n,

$$c\, q_n\, \mathbb{P}^x\{\tau_A > k\} \leq \mathbb{P}^x\{\tau_A > k+n\} \leq q_n\, \mathbb{P}^x\{\tau_A > k\}. \tag{11.20}$$

Let m be a positive integer and let

$$Y_0, Y_1, Y_2, \ldots Y_m$$

be the process corresponding to $X_0, X_k, X_{2k}, \ldots, X_{mk}$ conditioned so that $\tau_A > mk$. This is a time *inhomogeneous* Markov chain with transition probabilities

$$\mathbb{P}\{Y_j = y \mid Y_{j-1} = x\} = \frac{p_k^A(x,y)\,\mathbb{P}^y\{\tau_A > (m-j)k\}}{\mathbb{P}^x\{\tau_A > (m-j-1)k\}}, \qquad j = 1, 2, \ldots, m.$$

Note that (11.20 implies that for all $y \in A$,

$$\mathbb{P}\{Y_j = y \mid Y_{j-1} = x\} \le c_2\,\hat{p}_k^A(x,y),$$

and if $y \in A'$

$$\mathbb{P}\{Y_j = y \mid Y_{j-1} = x\} \ge c_1\,\hat{p}_k^A(x,y).$$

Using this and (11.18), we can see that there is a $\delta > 0$ such that if $x, z \in A$ and $j \le m$,

$$\frac{1}{2}\sum_{y \in A} |\mathbb{P}\{Y_j = y \mid Y_{j-1} = x\} - \mathbb{P}\{Y_j = y \mid Y_{j-1} = z\}| \le 1 - \delta,$$

and using an argument as in the proof of Proposition 11.4.2 we can see that

$$\frac{1}{2}\sum_{y \in A} |\mathbb{P}\{Y_m = y \mid Y_0 = x\} - \mathbb{P}\{Y_m = y \mid Y_0 = z\}| \le (1-\delta)^m.$$

□

11.4.1.1 Symmetric case

When the Markov chain is symmetric ($p(y,x) = p(x,y)$), then $w(x) = c\,v(x), \pi(x) = c\,v(x)^2$. The function $g(x) = \sqrt{c}\,v(x)$ can be characterized by the fact that g is strictly positive and satisfies

$$P^A g(x) = \alpha\, g(x), \qquad \sum_{x \in A} g(x)^2 = 1.$$

If $f_1, f_2 : A \to \mathbb{R}$, we define the quadratic form $Q = Q_p$ by

$$Q(f_1, f_2) = \frac{1}{2}\sum_{x,y \in D} p(x,y)\,[f_1(y) - f_1(x)]\,[f_2(y) - f_2(x)],$$

where f_1, f_2 are extended to D by setting $f_j \equiv 0$ on $D \setminus A$. The terms in the sum are zero unless one of x, y is in A. We write just $Q(f)$ for $Q(f,f)$. Note that

$$\begin{aligned} Q(f) &= \frac{1}{2}\sum_{x,y \in D} p(x,y)[f(x) - f(y)]\,[f(x) - f(y)] \\ &= \sum_{x \in A, y \in D} p(x,y)\,f(x)\,[f(x) - f(y)] \\ &= \sum_{x \in A} f(x) \sum_{y \in D} p(x,y)\,[f(x) - f(y)] \\ &= \sum_{x \in A} f(x)\,(I - P^A) f(x). \end{aligned} \tag{11.21}$$

Proposition 11.4.5 *Under the assumptions above,*

$$1 - \alpha = Q(g) = \inf Q(f),$$

where the infimum is over all $f : A \to \mathbb{R}$ with $\sum f(x)^2 = 1$.

Proof To get the first equality, we use (11.21),

$$Q(g_A) = \sum_{x \in A} g(x)\,(I - P^A)g(x) = (1-\alpha) \sum_{x \in A} g(x)^2 = 1 - \alpha.$$

Write $A = \{x_1, \ldots, x_m\}, D \setminus A = \{x_{m+1}, \ldots\}$. Then the goal is to minimize the function

$$F(t_1, \ldots, t_m) = \sum_{x_j, x_k \in A} p(x_j, x_k)\,(t_k - t_j)^2.$$

subject to $\sum t_k^2 = 1$ and $t_j = 0$ for $j > m$. It is easy to see that the minimum is obtained at a local minimum rather than on the boundary. Using Lagrange multipliers, we set

$$L(t_1, \ldots, t_m, \lambda) = F(t_1, \ldots, t_m) - \lambda \left[(\sum_{j=1}^{m} t_j^2 - 1) \right].$$

At the local minimum we must have

$$\partial_{t_j} L(t_1, \ldots, t_m, \lambda) = 0, \quad \partial_\lambda L(t_1, \ldots, t_m, \lambda) = 0,$$

which translates to

$$-2 \sum_{x_k \in A} p(x_j, x_k)(t_k - t_j) = 2\lambda\, t_j.$$

From this we see that any f that minimizes must be an eigenvector. However, it is also easy to see that any minimizing f must have a constant sign, for otherwise $Q(|f|) < Q(f)$. Since every eigenvector of constant sign is a multiple of g (see the proof of Theorem 11.4.1), this must be the minimizer. □

11.4.2 Maximal coupling of Markov chains

Here we will describe the maximal coupling of a Markov chain. Suppose that $p : D \times D \to [0,1]$ is the transition probability function for an irreducible, aperiodic Markov chain with countable state space D. Assume that g_0^1, g_0^2 are two initial probability distributions on D. Let g_n^j denote the corresponding distribution at time n, given recursively by

$$g_n^j(x) = \sum_{x \in D} g_{n-1}^j(z)\,p(z, x).$$

Let $\|\cdot\|$ denote the total variation distance,

$$\|g_n^1 - g_n^2\| = \frac{1}{2} \sum_{x \in D} |g_n^1(x) - g_n^2(x)| = 1 - \sum_{x \in D} [g_n^1(x) \wedge g_n^2(x)].$$

Suppose

$$X_0^1, X_1^1, X_2^1, \ldots, \qquad X_0^2, X_1^2, X_2^2, \ldots$$

are defined on the same probability space such that for each j, $\{X_n^j : n = 0, 1, \ldots\}$ has the distribution of the Markov chain with initial distribution g_0^j. Then it is clear that

$$\mathbb{P}\{X_n^1 = X_n^2\} \leq 1 - \|g_n^1 - g_n^2\| = \sum_{x \in D} g_n^1(x) \wedge g_n^2(x). \tag{11.22}$$

The following theorem shows that there is a way to define the chains on the same probability space so that equality is obtained in (11.22). This theorem gives one example of the powerful probabilistic technique called *coupling*. Coupling refers to the defining of two or more processes on the same probability space in a way so that each individual process has a certain distribution but the joint distribution has some particularly nice properties. Often, as in this case, the two processes are equal except for an event of small probability. This example is also typical in that the definition of the coupled processes is not very natural.

Theorem 11.4.6 *Suppose p, g_n^1, g_n^2 are as defined in the previous paragraph. We can define $(X_n^1, X_n^2), n = 0, 1, 2, \ldots$ on the same probability space such that:*

- *for each j, $X_0^j, X_1^j, \ldots$ has the distribution of the Markov chain with initial distribution g_0^j;*
- *for each integer $n \geq 0$,*

$$\mathbb{P}\{X_m^1 = X_m^2 \text{ for all } m \geq n\} = 1 - \|g_n^1 - g_n^2\|.$$

Comment. Before doing this proof, we might wish to consider an easier problem of defining (X^1, X^2) on the same probability space so that X^j has distribution g_0^j and

$$\mathbb{P}\{X^1 = X^2\} = 1 - \|g_0^1 - g_0^2\|.$$

Let $f^j(x) = g_0^j(x) - [g_0^1(x) \wedge g_0^2(x)]$; $\rho^j(x) = [g_0^1(x) \wedge g_0^2(x)]/g_0^j(x)$ if $g_0^j(x) > 0$, and $\rho^j(x) = 0$ otherwise. Start with independent Z^1, Z^2 from the distributions g_o^1, g_0^2, respectively. For each x, let $Y^j(x)$ be a 0-1 random variable with $\mathbb{P}\{Y^j(x) = 1\} = \rho^j(x)$. Let $J^j = Y^j(Z^j)$. Then for each x, $\mathbb{P}\{J^1 = 1, Z^1 = x\} = \mathbb{P}\{J^2 = 1, Z^2 = x\} = g_0^1(x) \wedge g_0^2(x)$. We can then do the following construction.

- Choose a realization of the 0-1 random variable J with $\mathbb{P}\{J = 0\} = \|g_0^1 - g_0^2\|$.
- Let (X, W^1, W^2) be independent random varaibles independent of J. X has the distribution

$$\mathbb{P}\{X = x\} = \mathbb{P}\{Z^j = x \mid J^j = 1\} = \frac{g_1(x) \wedge g_2(x)}{1 - \|g_0^1 - g_0^2\|},$$

  and W^j has the distribution

$$\mathbb{P}\{W^j = x\} = \mathbb{P}\{Z^j = x \mid J^j = 0\} = \frac{f^j(x)}{\|g_0^1 - g_0^2\|}.$$

- Let $X^j = 1\{J = 1\}\, X + 1\{J = 0\} W^j$.

It is easy to check that this construction works.

Proof For ease, we will assume that $\|g_n^1 - g_n^2\| \to 0$ as $n \to \infty$; the adjustment needed if this does not hold is left to the reader. Let (Z_n^1, Z_n^2) be independent Markov chains with the appropriate distributions. Let $f_n^j(x) = g_n^j(x) - [g_n^1(x) \wedge g_n^2(x)]$ and define h_n^j by $h_0^j(x) = g_0^j(x)$ and for $n > 1$,

$$h_n^j(x) = \sum_{z \in S} f_{n-1}^j(z)\, p(z, x).$$

Note that $f_{n+1}^j(x) = h_{n+1}^j(x) - [h_n^1(x) \wedge h_n^2(x)]$. Let

$$\rho_n^j(x) = \frac{h_n^1(x) \wedge h_n^2(x)}{h_n^j(x)} \quad \text{if } h_n^j(x) \neq 0.$$

We set $\rho_n^j(x) = 0$ if $h_n^j(x) = 0$. We let $\{Y^j(n,x) : j = 1,2; n = 0,1,2,\ldots; x \in D\}$ be independent 0-1 random variables, independent of (Z_n^1, Z_n^2), with $\mathbb{P}\{Y^j(n,x) = 1\} = \rho_n^j(x)$.

We now define 0-1 random variables J_n^j as follows:

- $J_0^j = Y^j(0, Z_0^j)$.
- If $J_n^j = 1$, then $J_m^j = 1$ for all $m \geq n$.
- If $J_n^j = 0$, then $J_{n+1}^j = Y^j(n+1, Z_{n+1}^j)$.

We claim that

$$\mathbb{P}\{J_n^j = 0; Z_n^j = x\} = f_n^j(x).$$

For $n = 0$, this follows immediately from the definition. Also,

$$\mathbb{P}\{J_{n+1}^j = 0; Z_{n+1}^j = x\} =$$

$$\sum_{z \in D} \mathbb{P}\{J_n^j = 0; Z_n^j = z\} \mathbb{P}\{Z_{n+1}^j = x, Y^j(n+1,x) = 0 \mid J_n^j = 0; Z_n^j = z\}.$$

The random variable $Y^j(n+1,x)$ is independent of the Markov chain, and the event $\{J_n^j = 0; Z_n^j = z\}$ depends only on the chain up to time n and the values of $\{Y^j(k,y) : k \leq n\}$. Therefore,

$$\mathbb{P}\{Z_{n+1}^j = x, Y^j(n+1,x) = 0 \mid J_n^j = 0; Z_n^j = z\} = p(z,x)\,[1 - \rho_{n+1}^j(x)].$$

Therefore, we have the inductive argument

$$\begin{aligned}
\mathbb{P}\{J_{n+1}^j = 0; Z_{n+1}^j = x\} &= \sum_{z \in D} f_n^j(z)\, p(z,x)\,[1 - \rho_{n+1}^j(x)] \\
&= h_{n+1}^j(x)\,[1 - \rho_{n+1}^j(x)] \\
&= h_{n+1}^j(x) - [h_{n+1}^1(x) \wedge h_{n+1}^2(x)] = f_{n+1}^j(x),
\end{aligned}$$

which establishes the claim.

Let K^j denote the smallest n such that $J_n^j = 1$. The condition $\|g_0^1 - g_0^2\| \to 0$ implies that $K^j < \infty$ with probability one. A key fact is that for each n and each x,

$$\mathbb{P}\{K^1 = n+1; Z_n^1 = x\} = \mathbb{P}\{K^2 = n+1; Z_n^2 = x\} = h_{n+1}^1(x) \wedge h_{n+1}^2(x).$$

This is immediate for $n = 0$ and for $n > 0$,

$$\begin{aligned}
&\mathbb{P}\{K^j = n+1; Z_{n+1}^j = x\} \\
&\quad = \sum_{z \in D} \mathbb{P}\{J_n = 0; Z_n^j = z\}\, \mathbb{P}\{Y^j(n+1,x) = 1; Z_{n+1}^j = x \mid J_n = 0; Z_n^j = z\}
\end{aligned}$$

$$= \sum_{z\in D} f_n^j(z)\, p(z,x)\, \rho_{n+1}^j(x)$$
$$= h_{n+1}^j(x)\, \rho_{n+1}^j(x) = h_{n+1}^1(x) \wedge h_{n+1}^2(x).$$

The last important observation is that the distribution of $W_m := X_{m-n}^j$ given the event $\{K^j = n; X_n^j = x\}$ is that of a Markov chain with transition probability p starting at x.

The reader may note that for each j, the process (Z_n^j, J_n^j) is a time-inhomogeneous Markov chain with transition probabilities

$$\mathbb{P}\{(Z_{n+1}^j, J_{n+1}^j) = (y,1) \mid (Z_n^j, J_n^j) = (x,1)\} = p(x,y),$$

$$\mathbb{P}\{(Z_{n+1}^j, J_{n+1}) = (y,0) \mid (Z_n^j, J_n^j) = (x,0)\} = p(x,y)\,[1-\rho_{n+1}^j(y)],$$

$$\mathbb{P}\{(Z_{n+1}^j, J_{n+1}) = (y,1) \mid (Z_n^j, J_n^j) = (x,0)\} = p(x,y)\,\rho_{n+1}^j(y).$$

The chains (Z_n^1, J_n^1) and (Z_n^2, J_n^2) are independent. However, the transition probabilities for these chains depend on both initial distributions and p.

We are now ready to make our construction of (X_n^1, X_n^2).

- Define for each (n,x) a process $\{W_m^{n,x} : m = 0,1,2,\ldots\}$ that has the distribution of the Markov chain with initial point x. Assume that all these processes are independent.
- Choose (n,x) according to the probability distribution
$$h_{n+1}^1(x) \wedge h_{n+1}^2(x) = \mathbb{P}\{K^j = n; Z_n^j = x\}.$$
Set $J_m^j = 1$ for $m \geq n$ and $J_m^j = 0$ for $m < n$, and $K^1 = K^2 = n$. Note that K^j is the smallest n such that $J_n^j = 1$.
- Given (n,x), choose $X_0^1, \ldots, X_n^1$ from the conditional distribution of the Markov chain with initial distribution g_0^1 conditioned on the event $\{K^1 = n; Z_n^1 = x\}$.
- Given (n,x), choose $X_0^2, \ldots, X_n^2$ (conditionally) independent of $X_0^1, \ldots, X_n^1$ from the conditional distribution of the Markov chain with initial distribution g_0^2 conditioned on the event $\{K^2 = n; Z_n^2 = x\}$.
- Let
$$X_m^j = W_{m-n}^{n,x}, \quad m = n, n+1, \ldots.$$

The two conditional distributions above are not easy to express explicity; fortunately, we do not need to do so.

To finish the proof, we need only check that the above construction satisfies the conditions. For fixed j, the fact that $X_0^j, X_1^j, \ldots$ has the distribution of the chain with initial distribution g_0^j is immediate from construction and the eariler observation that the distribution of $\{X_n^j, X_{n+1}^j, \ldots\}$ given $\{K^j = n; Z_n^j = x\}$ is that of the Markov chain starting at x. Also, the construction immediately gives $X_m^1 = X_m^2$ if $m \geq K^1 = K^2$. Also,

$$\mathbb{P}\{J_n^j = 0\} = \sum_{x\in D} f_n^j(z) = \|g_1^n - g_2^n\|.$$

$\square$

Remark. A review of the proof of Theorem 11.4.6 shows that we do not need to assume that the Markov chain is time-homogeneous. However, time-homgeneity makes the notation a little simpler and we only use the result for time-homogenous chains.

11.5 Some Tauberian theory

Lemma 11.5.1 *Suppose $\alpha > 0$. Then as $\xi \to 1-$,*

$$\sum_{n=2}^{\infty} \xi^n \, n^{\alpha-1} = \frac{\mathbf{\Gamma}(\alpha)}{(1-\xi)^\alpha} \, [1 + O(1-\xi)].$$

Proof Let $\epsilon = 1 - \xi$. First note that

$$\sum_{n \geq \epsilon^{-2}} \xi^n \, n^{\alpha-1} = \sum_{n \geq \epsilon^{-2}} [(1-\epsilon)^{1/\epsilon}]^{n\epsilon} \, n^{\alpha-1} \leq \sum_{n \geq \epsilon^{-2}} e^{-n\epsilon} \, n^{\alpha-1},$$

and the right-hand side decays faster than every power of ϵ. For $n \leq \epsilon^{-2}$ we can do the asymptotics

$$\xi^n = \exp\{n \log(1-\epsilon)\} = \exp\{n(-\epsilon - O(\epsilon^2))\} = e^{-n\epsilon}[1 + O(n\epsilon^2)].$$

Hence,

$$\sum_{n \leq \epsilon^{-2}} \xi^n \, n^{\alpha-1} = \epsilon^{-\alpha} \sum_{n \leq \epsilon^{-2}} \epsilon \, e^{-n\epsilon} (n\epsilon)^{\alpha-1} \, [1 + (n\epsilon) \, O(\epsilon)].$$

Using Riemann sum approximations we see that

$$\sum_{n=1}^{\infty} \epsilon \, e^{-n\epsilon} (n\epsilon)^{\alpha-1} = \int_0^\infty e^{-t} \, t^{\alpha-1} \, dt + O(\epsilon) = \mathbf{\Gamma}(\alpha) + O(\epsilon).$$

□

Proposition 11.5.2 *Suppose u_n is a sequence of nonnegative real numbers. If $\alpha > 0$, the following two statements are equivalent:*

$$\sum_{n=0}^{\infty} \xi^n \, u_n \sim \frac{\mathbf{\Gamma}(\alpha)}{(1-\xi)^\alpha}, \qquad \xi \to 1-, \tag{11.23}$$

$$\sum_{n=1}^{N} u_n \sim \alpha^{-1} \, N^\alpha, \qquad N \to \infty. \tag{11.24}$$

Moreover, if the sequence is monotone, either of these statements implies

$$u_n \sim n^{\alpha-1}, \qquad n \to \infty.$$

Proof Let $U_n = \sum_{j \leq n} u_j$ where $U_{-1} = 0$. Note that

$$\sum_{n=0}^{\infty} \xi^n \, u_n = \sum_{n=0}^{\infty} \xi^n \, [U_n - U_{n-1}] = (1-\xi) \sum_{n=0}^{\infty} \xi^n \, U_n. \tag{11.25}$$

If (11.24) holds, then by the previous lemma

$$\sum_{n=0}^{\infty} \xi^n u_n \sim (1-\xi) \sum_{n=0}^{\infty} \xi^n \alpha^{-1} n^{\alpha} \sim \frac{\mathbf{\Gamma}(\alpha+1)}{\alpha (1-\xi)^{\alpha}} = \frac{\mathbf{\Gamma}(\alpha)}{(1-\xi)^{\alpha}}.$$

Now suppose (11.23) holds. We first give an upper bound on U_n. Using $1-\xi = 1/n$, we can see as $n \to \infty$,

$$\begin{aligned} U_n &\leq n^{-1} \left(1 - \frac{1}{n}\right)^{-2n} \sum_{j=n}^{2n-1} \left(1 - \frac{1}{n}\right)^{j} U_j \\ &\leq n^{-1} \left(1 - \frac{1}{n}\right)^{-2n} \sum_{j=0}^{\infty} \left(1 - \frac{1}{n}\right)^{j} U_j \sim e^2 \mathbf{\Gamma}(\alpha) n^{\alpha}. \end{aligned}$$

The last relation uses (11.25). Let $\nu^{(j)}$ denote the measure on $[0,\infty)$ that gives measure $j^{-\alpha} u_n$ to the point n/j. Then the last estimate shows that the total mass of $\nu^{(j)}$ is uniformly bounded on each compact interval and hence there is a subsequence that converges weakly to a measure ν that is finite on each compact interval. Using (11.23) we can see that that for each $\lambda > 0$,

$$\int_0^{\infty} e^{-\lambda x} \nu(dx) = \int_0^{\infty} e^{-\lambda x} x^{\alpha-1} dx.$$

This implies that ν is $x^{\alpha-1} dx$. Since the limit is independent of the subseqence, we can conclude that $\nu^{(j)} \to \nu$ and this implies (11.24).

The fact that (11.24) implies the last assertion if u_n is monotone is straightforward using

$$U_{n(1+\epsilon)} - U_n \sim \alpha^{-1} \left[(n(1+\epsilon))^{\alpha} - n^{\alpha} \right], \quad n \to \infty,$$

□

The following is proved similarly.

Proposition 11.5.3 *Suppose u_n is a sequence of nonnegative real numbers. If $\alpha \in \mathbb{R}$, the following two statements are equivalent:*

$$\sum_{n=0}^{\infty} \xi^n u_n = \left(\frac{1}{1-\xi}\right) \log^{\alpha} \left(\frac{1}{1-\xi}\right), \quad \xi \to 1-, \tag{11.26}$$

$$\sum_{n=1}^{N} u_n \sim N \log^{\alpha} N \quad N \to \infty. \tag{11.27}$$

Moreover, if the sequence is monotone, either of these statements implies

$$u_n \sim \log^{\alpha} n, \quad n \to \infty.$$

11.6 Second moment method

Lemma 11.6.1 *Suppose X is a nonnegative random variable with $\mathbb{E}[X^2] < \infty$ and $0 < r < 1$. Then*

$$\mathbb{P}\{X \geq r\mathbb{E}(X)\} \geq \frac{(1-r)^2\,\mathbb{E}(X)^2}{\mathbb{E}(X^2)}.$$

Proof Without loss of generality, we may assume that $\mathbb{E}(X) = 1$. Since $\mathbb{E}[X; X < r] \leq r$, we know that $\mathbb{E}[X; X \geq r] \geq (1-r)$. Then,

$$\begin{aligned} \mathbb{E}(X^2) \geq \mathbb{E}[X^2; X \geq r] &= \mathbb{P}\{X \geq r\}\,\mathbb{E}[X^2 \mid X \geq r] \\ &\geq \mathbb{P}\{X \geq r\}\,(\mathbb{E}[X \mid X \geq r])^2 \\ &\geq \frac{\mathbb{E}[X; X \geq r]^2}{\mathbb{P}\{X \geq r\}} \\ &\geq \frac{(1-r)^2}{\mathbb{P}\{X \geq r\}}. \end{aligned}$$

□

Corollary 11.6.2 *Suppose $E_1, E_2, \ldots$ is a collection of events with $\sum \mathbb{P}(E_n) = \infty$. Suppose there is a $K < \infty$ such that for all $j \neq k$, $\mathbb{P}(E_j \cap E_k) \leq K\,\mathbb{P}(E_j)\,\mathbb{P}(E_k)$. Then*

$$\mathbb{P}\{E_k \text{ i.o.}\} \geq \frac{1}{K}.$$

Proof Let $V_n = \sum_{k=1}^n 1_{E_k}$. Then the assumptions imply that

$$\lim_{n\to\infty} \mathbb{E}(V_n) = \infty,$$

and

$$\mathbb{E}(V_n^2) \leq \sum_{j=1}^{k} \mathbb{P}(E_j) + \sum_{j\neq k} K\,\mathbb{P}(E_j)\,\mathbb{P}(E_k) \leq \mathbb{E}(V_n) + K\,\mathbb{E}(V_n)^2 = \left[\frac{1}{\mathbb{E}(V_n)} + K\right] \mathbb{E}(V_n)^2.$$

By Lemma 11.6.1, for every $r > 0$,

$$\mathbb{P}\{V_n \geq r\mathbb{E}(V_n)\} \geq \frac{(1-r)^2}{K + \mathbb{E}(V_n)^{-1}}.$$

Since $\mathbb{E}(V_n) \to \infty$, this implies

$$\mathbb{P}\{V_\infty = \infty\} \geq \frac{(1-r)^2}{K}.$$

Since this holds for every $r > 0$, we get the result. □

11.7 Subadditivity

Lemma 11.7.1 (Subadditivity lemma) *Suppose* $f : \{1, 2, \ldots\} \to \mathbb{R}$ *is subadditive, i.e., for all* n, m, $f(n+m) \le f(n) + f(m)$. *Then,*

$$\lim_{n\to\infty} \frac{f(n)}{n} = \inf_{n>0} \frac{f(n)}{n}.$$

Proof Fix integer $N > 0$. We can write any integer n as $jN + k$ where j is a nonnegative integer and $k \in \{1, \ldots, N\}$. Let $b_N = \max\{f(1), \ldots, f(N)\}$. Then subadditivity implies

$$\frac{f(n)}{n} \le \frac{jf(N) + f(k)}{jN} \le \frac{f(N)}{N} + \frac{b_N}{jN}.$$

Therefore,

$$\limsup_{n\to\infty} \frac{f(n)}{n} \le \frac{f(N)}{N}.$$

Since this is true for every N, we get the lemma. □

Corollary 11.7.2 *Suppose* r_n *is a sequence of positive numbers and* $b_1, b_2 > 0$ *such that for every* n, m,

$$b_1\, r_n\, r_m \le r_{n+m} \le b_2\, r_n\, r_m. \tag{11.28}$$

Then there exists $\alpha > 0$ *such that for all* n,

$$b_2^{-1}\, \alpha^n \le r_n \le b_1^{-1}\, \alpha^n.$$

Proof Let $f(n) = \log r_n + \log b_2$. Then f is subadditive and hence

$$\lim_{n\to\infty} \frac{f(n)}{n} = \frac{f(n)}{n} := \alpha.$$

This shows that $r_n \ge \alpha^n / b_2$. Similarly, by considering the subadditive function $g(n) = -\log r_n - \log b_1$, we get $r_n \le b_1^{-1}\, \alpha^n$. □

Remark. Note that if r_n satisfies (11.28), then so does $\beta^n r_n$ for each $\beta > 0$. Therefore, we cannot determine the value of α from (11.28).

Exercises

Exercise 11.1 Find $f_3(\xi), f_4(\xi)$ in (11.4).

Exercise 11.2 Suppose $E_1 \supset E_2 \supset \cdots$ is a decreasing sequnce of events with $\mathbb{P}(E_n) > 0$ for each n. Suppose there exist $\alpha > 0$ such that

$$\sum_{n=1}^{\infty} |\mathbb{P}(E_n \mid E_{n-1}) - (1 - \alpha n^{-1})| < \infty.$$

Show there exists c such that

$$\mathbb{P}(E_n) \sim c\, n^{-\alpha}. \tag{11.29}$$

(Hint: use Lemma 11.1.3.)