1 Simple Random Walk

We consider one of the basic models for random walk, *simple random walk on the integer lattice* $\mathbb{Z}^d$. At each time step, a random walker makes a random move of length one in one of the lattice directions.

1.1 One dimension

We start by studying simple random walk on the integers. At each time unit, a walker flips a fair coin and moves one step to the right or one step to the left depending on whether the coin comes up heads or tails. We let S_n denote the position of the walker at time n. If we assume that the walker starts at x, we can write

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

where $X_1, X_2, \ldots$ are independent random variables with $\mathbb{P}\{X_j = 1\} = \mathbb{P}\{X_j\} = -1 = 1/2$.

In this chapter we assume that the walker starts at the origin ($x = 0$) and investigate where we expect the walker to be. The main questions are:

- On the average, how far is the walker from the starting point?
- What is the probability distribution for the position of the walker?
- What is the probability that at a particular time the walker is at the origin?
- Does the random walker keep returning to the origin or does the walker eventually leave forever?

Probabilists use the notation $\mathbb{E}$ for *expectation* (also called *expected value, mean, average value*) defined for discrete random variables by

$$\mathbb{E}[X] = \sum_z z\,\mathbb{P}\{X = z\}.$$

The random walk satisifes $\mathbb{E}[S_n] = 0$ since steps of $+1$ and -1 are equally likely. To compute the average distance, one might try to compute $\mathbb{E}\left[|S_n|\right]$. It turns out to be much easier to compute $\mathbb{E}[S_n^2]$,

$$\begin{aligned}
\mathbb{E}[S_n^2] &= \mathbb{E}\left[\left(\sum_{j=1}^n X_j\right)^2\right] \\
&= \mathbb{E}\left[\sum_{j=1}^n \sum_{k=1}^n X_j X_k\right] \\
&= \sum_{j=1}^n \sum_{k=1}^n \mathbb{E}[X_j X_k] = n + \sum_{j \neq k} \mathbb{E}[X_j X_k].
\end{aligned}$$

This calculation uses an easy (but very important) property of average values: $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$. If $j \neq k$, then $X_j X_k$ is equally likely to be $+1$ and -1 and so the expectation is zero. We therefore get

$$\mathrm{Var}[S_n] = \mathbb{E}[S_n^2] = n.$$

The variance of a random variable is defined by

$$\mathrm{Var}[X] = \mathbb{E}\left[(X - \mathbb{E}X)^2\right] = \mathbb{E}[X^2] - (\mathbb{E}X)^2$$

(a simple calculation establishes the second equality). Our calculation illustrates an important fact about variances of sums: if $X_1, \ldots, X_n$ are independent, then

$$\mathrm{Var}[X_1 + \cdots + X_n] = \mathrm{Var}[X_1] + \cdots + \mathrm{Var}[X_n].$$

Sum rules like this make it much easier to compute averages of the *square* of a random variable than other powers. In many ways, this is just an analogy of the Pythagorean theorem from geometry: the independence of the random variables is the analogue of the idea of perpendicular or orthogonal vectors.

The next question is: what is the probability that the walker is at the origin after n steps? Before computing this very accurately, let us do a heuristic argument. Since $\mathbb{E}[S_n^2] = n$, the typical distance away from the origin is of order $\sqrt{n}$. There are about $\sqrt{n}$ even integers that are distance $\sqrt{n}$ from the starting point, so one might guess that the probability for being at a particular one should decay like $n^{-1/2}$. This is indeed the case as we demonstrate.

In fact, it is easy to give an exact formula for the distribution of the walker after n steps. It is easy to see that after an odd number of steps the walker is at at odd integer and after an even number of steps the walker is at an even integer. Therefore, $\mathbb{P}\{S_n = x\} = 0$ if $n + x$ is odd. Let us suppose the walker has taken an even number of steps, $2n$. In order for the walker to be back at the origin at time $2n$ there need to be exactly n "$+1$" steps and n "-1" steps. The number of ways to choose which n steps are $+1$ is $\binom{2n}{n}$ and each particular choice of $2n$ $+1$s and -1s has probability 2^{-2n} of occuring. Therfore,

$$\mathbb{P}\{S_{2n} = 0\} = \binom{2n}{n} 2^{-2n} = \frac{(2n)!}{n!\, n!} 2^{-2n}.$$

More generally,

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

While these formulas are exact, it is not obvious how to use them because they contain ratios of very large numbers. Trying to understand this quantity leads to studying the behavior of $n!$ as n gets large. This we discuss in the next section.

1.2 Stirling's formula

Stirling's formula states that as $n \to \infty$,

$$n! \sim \sqrt{2\pi}\, n^{n+\frac{1}{2}}\, e^{-n},$$

where $\sim$ means that the ratio of the two sides tends to 1. We will prove this in the next two subsections. In this subsection we will prove that there is a positive number C_0 such that

$$\lim_{n\to\infty} b_n = C_0, \qquad \text{where} \qquad b_n = \frac{n!}{n^{n+\frac{1}{2}} e^{-n}}, \tag{1}$$

and in Section 1.3 we show that $C_0 = \sqrt{2\pi}$. To start, it is easy to check that $b_1 = e$ and if $n \geq 2$,

$$\frac{b_n}{b_{n-1}} = e \left(\frac{n-1}{n}\right)^{n-\frac{1}{2}} = e \left(1 - \frac{1}{n}\right)^{n} \left(1 - \frac{1}{n}\right)^{-1/2}.$$

Let $\delta_n = (b_n/b_{n-1}) - 1$ so that for $n \geq 2$,

$$b_n = b_1 \prod_{m=2}^{n} \frac{b_m}{b_{m-1}} = e \prod_{m=2}^{n} [1 + \delta_m].$$

Then (1) can be restated as saying that the infinite product

$$e \prod_{n=2}^{\infty} [1 + \delta_n] = \lim_{N\to\infty} e \prod_{n=2}^{N} [1 + \delta_n]$$

converges to a positive constant that we call C_0. By taking the logarithm of both sides, we see that convergence to a positive constant is equivalent to

$$\sum_{n=2}^{\infty} \log[1 + \delta_n] < \infty. \tag{2}$$

One of the most important tools for determining limits is Taylor's theorem with remainder, a version of which we now recall. Suppose f is a C^{k+1} function. Let $P_k(x)$ denote the kth order Taylor series polynomial about the origin. Then, for $x > 0$

$$|f(x) - P_k(x)| \leq a_k\, x^{k+1},$$

where

$$a_k = \frac{1}{(k+1)!} \max_{0 \leq t \leq x} |f^{(k+1)}(t)|.$$

A similar estimate is derived for negative x by considering $\tilde{f}(x) = f(-x)$. The Taylor series for the logarithm gives

$$\log(1 + u) = u - \frac{u^2}{2} + \frac{u^3}{3} - \cdots,$$

which is valid for $|u| < 1$. In fact, the Taylor series with remainder tells us that for every positive integer k

$$\log(1+u) = P_k(u) + O(|u|^{k+1}), \tag{3}$$

where $P_k(u) = u - (u^2/2) + \cdots + (-1)^{k+1}(u^k/k)$. The $O(|u|^{k+1})$ denotes a term that is bounded by a constant time $|u|^{k+1}$ for small u. The Taylor series with remainder implies that there is a constant c_k such that for all $|u| \leq 1/2$,

$$|\log(1+u) - P_k(u)| \leq c_k\, |u|^{k+1}. \tag{4}$$

We will use the $O(\cdot)$ notation as in (3) when doing asymptotics — in all cases this will be shorthand for a more precise statement as in (4).

By taking the Taylor series, we can see that to prove (2) assuming $1 + \delta_n > 0$ for all n, it suffices to show that $\sum |\delta_n| < \infty$. We will establish this by showing that $\delta_n = O(n^{-2})$, i.e., there is a c such that

$$|\delta_n| \leq \frac{c}{n^2}.$$

To see this consider $(1 - \frac{1}{n})^n$ which we know approaches e^{-1} as n gets large. We use the Taylor series to estimate how fast it converges. We write

$$\log\left(1 - \frac{1}{n}\right)^n = n \log\left(1 - \frac{1}{n}\right) = n\left[-\frac{1}{n} - \frac{1}{2n^2} + O(n^{-3})\right] = -1 - \frac{1}{2n} + O(n^{-2}).$$

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

By adding the terms we finish the proof of (1). In fact (see Exercise 6.19) we can show that

$$n! = C_0\, n^{n+\frac{1}{2}}\, e^{-n}\left[1 + O(n^{-1})\right]. \tag{5}$$

1.3 Central limit theorem

Let S_n be the position of a simple random walker on the integers assuming $S_0 = 0$. For every integer j,

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

Let us assume that $|j| \leq n/2$. Then plugging into Stirling's formula and simplifying gives

$$\mathbb{P}\{S_{2n} = 2j\} \sim \frac{\sqrt{2}}{C_0}\left(1 - \frac{j^2}{n^2}\right)^{-n}\left(1 + \frac{j}{n}\right)^{-j}\left(1 - \frac{j}{n}\right)^{j}\left(\frac{n}{n^2 - j^2}\right)^{1/2}. \tag{6}$$

In fact (if one uses (5)), there is a c such that the ratio of the two sides is within distance c/n of 1 (assuming $|j| \leq n/2$).

What does this look like as n tends to infinity? Let us first consider the case $j = 0$. Then we get that

$$\mathbb{P}\{S_{2n} = 0\} \sim \frac{\sqrt{2}}{C_0\, n^{1/2}}.$$

We now consider j of order $\sqrt{n}$. If we write $j = r\sqrt{n}$, the right hand side of (6) becomes

$$\frac{\sqrt{2}}{C_0\,\sqrt{n}} \left(1 - \frac{r^2}{n}\right)^{-n} \left[\left(1 + \frac{r}{\sqrt{n}}\right)^{-\sqrt{n}}\right]^{r} \left[\left(1 - \frac{r}{\sqrt{n}}\right)^{-\sqrt{n}}\right]^{-r} \left(\frac{1}{1-(r^2/n)}\right)^{1/2}.$$

As $n \to \infty$, this is asymptotic to

$$\frac{\sqrt{2}}{C_0\,\sqrt{n}}\, e^{r^2}\, e^{-r^2}\, e^{-r^2} = \frac{\sqrt{2}}{C_0\,\sqrt{n}}\, e^{-j^2/n}.$$

For every $a < b$,

$$\lim_{n\to\infty} \mathbb{P}\{a\sqrt{2n} \le S_{2n} \le b\sqrt{2n}\} = \lim_{n\to\infty} \sum \frac{\sqrt{2}}{C_0\,\sqrt{n}}\, e^{-j^2/n},$$

where the sum is over all j with $a\sqrt{2n} \le 2j \le b\sqrt{2n}$. The right hand side is the Riemann sum approximation of an integral where the intervals in the sum have length $\sqrt{2/n}$. Hence the limit is

$$\int_a^b \frac{1}{C_0}\, e^{-x^2/2}\, dx.$$

This limiting distribution must be a probability distribution, so we can see that

$$\int_{-\infty}^{\infty} \frac{1}{C_0}\, e^{-x^2/2}\, dx = 1.$$

This gives the value $C_0 = \sqrt{2\pi}$ (see Exercise 6.21), and hence Stirling's formula can be written as

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

where this means that there exists a c such that for all n,

$$\left|\frac{n!}{\sqrt{2\pi}\, n^{n+\frac{1}{2}}\, e^{-n}} - 1\right| \le \frac{c}{n}.$$

1.4 Returns to the origin

We now consider the number of times that the random walker returns to the origin. Let $J_n = 1\{S_n = 0\}$. Here we use the *indicator function* notation: if E is an event, then 1_E or

$1(E)$ is the random variable that takes the value 1 if the event occurs and 0 if it does not occur. The total number of visits to the origin by the random walker is

$$V = \sum_{n=0}^{\infty} J_{2n}.$$

Note that

$$\mathbb{E}[V] = \sum_{n=0}^{\infty} \mathbb{E}[J_{2n}] = \sum_{n=0}^{\infty} \mathbb{P}\{S_{2n} = 0\}.$$

We know that $\mathbb{P}\{S_{2n} = 0\} \sim c/\sqrt{n}$ as $n \to \infty$. Also, $\sum n^{-a}$ converges if and only if $a > 1$. Therefore,

$$\mathbb{E}[V] = \infty.$$

Let q be the probability that the random walker ever returns to the origin after time 0. We will show that $q = 1$ by first assuming $q < 1$ and deriving a contradiction. Suppose that $q < 1$. Then we can give the distribution for V. For example, $\mathbb{P}\{V = 1\} = (1 - q)$ since $V = 0$ if and only if the walker never returns after time zero. More generally,

$$\mathbb{P}\{V = k\} = q^{k-1}\,(1 - q), \quad k = 1, 2, \ldots$$

This tells us that

$$\mathbb{E}[V] = \sum_{k=1}^{\infty} k\,\mathbb{P}\{V = k\} = \sum_{k=1}^{\infty} k\,q^k\,(1 - q) = \frac{1}{1 - q} < \infty.$$

But, we know that $\mathbb{E}[V] = \infty$. Hence it must be the case that $q = 1$. We have established the following.

Theorem 1.1. *The probability that a (one-dimensional) simple random walker returns to the origin infinitely often is one.*

Note that this also implies that if the random walker starts at $x \neq 0$, then the probability that it will get to the origin is one.

1.5 Several dimensions

We now consider a random walker on the d-dimensional integer grid

$$\mathbb{Z}^d = \{(x_1, \ldots, x_d) : x_j \text{ integers}\}.$$

At each time step, the random walker chooses one of its $2d$ nearest neighbors, each with probability $1/2d$, and moves to that site. We again let

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

denote the position of the particle. Here $x, X_1, \ldots, X_n, S_n$ represent points in $\mathbb{Z}^d$. If $S_0 = 0$, then $\mathbb{E}[S_n] = 0$. Note that $\mathbb{E}[X_j \cdot X_j] = 1$ and $\mathbb{E}[X_j \cdot X_k] = 0$ if $j \neq k$. A calculation as in the one-dimensional case gives

$$\mathbb{E}[|S_n|^2] = \mathbb{E}[S_n \cdot S_n] = \mathbb{E}\left[\left(\sum_{j=1}^n X_j\right) \cdot \left(\sum_{j=1}^n X_j\right)\right] = n.$$

What is the probability that we are at the origin after n steps assuming $S_0 = 0$? This is zero if n is odd. If n is even, let us give a heuristic argument. The typical distance from the origin of S_n is of order $\sqrt{n}$. In d dimensions that number of lattice points with distance $\sqrt{n}$ grows like $(\sqrt{n})^d$. Hence the probability that we choose a particular point should decay like a constant times $n^{-d/2}$.

The combinatorics for justifying this is a little more complicated than in the one dimensional case so we will just wave our hands to get the right behavior. In $2n$ steps, we expect that approximately $2n/d$ of them will be taken in each of the d possible directions (e.g., if $d = 2$ we expect about n horizontal and n vertical steps). In order to be at the origin, we need to take an even number of steps in each of the d-directions. The probability of this (Exercise 6.17) is $2^{-(d-1)}$. Given that each of these numbers is even, the probability that each individual component is at the origin is the probability that a one dimensional walk is at the origin at time $2n/d$ (or, more precisely, an even integer very close to $2n/d$). Using this idea we get the asymptotics

$$\mathbb{P}\{S_{2n} = 0\} \sim \frac{c_d}{n^{d/2}}, \qquad c_d = \frac{d^{d/2}}{\pi^{d/2}\, 2^{d-1}}.$$

Consider the expected number of returns to the origin. If V is the number of visits to the origin,

$$\mathbb{E}[V] = \sum_{n=0}^{\infty} \mathbb{P}\{S_{2n} = 0\}.$$

Since $\mathbb{P}\{S_{2n} = 0\} \sim c/n^{d/2}$,

$$\mathbb{E}[V] \begin{cases} < \infty, & d \geq 3 \\ = \infty, & d = 2 \end{cases}.$$

Theorem 1.2. *Suppose S_n is simple random walk in $\mathbb{Z}^d$. If $d = 1, 2$, the random walk is* recurrent*, i.e., with probability one it returns to the origin infinitely often. If $d \geq 3$, the random walk is* transient*, i.e., with probability one it returns to the origin only finitely often. Also,*

$$\mathbb{P}\{S_n \neq 0 \text{ for all } n > 0 \mid S_0 = 0\} > 0 \text{ if } d \geq 3.$$

2 Dirichlet problem

2.1 One-dimension: gambler's ruin

Let us fix a positive integer N and suppose that a random walker on the integers starts at $x \in \{0, 1, \ldots, N\}$. Let S_n denote the position of the walker at time n. We will stop the walker when the walker reaches 0 or N. To be more precise, we let

$$T = \min \{n : S_n = 0 \text{ or } N\},$$

and consider $\hat{S}_n = S_{n \wedge T}$. Here $n \wedge T$ means the minimum of n and T. We know that with probability one $T < \infty$. Define the function $F : \{0, \ldots, N\} \to [0, 1]$ by

$$F(x) = \mathbb{P}\{S_T = N\}.$$

We can give a gambling interpretation to this by viewing S_n as the current winnings of a gambler who at each time step plays a fair game winning or losing one unit. The gambler starts with a bankroll of x and plays until he or she has N or has gone bankrupt. The chance that the gambler does not go bankrupt before attaining N is $F(x)$. Clearly, $F(0) = 0$ and $F(N) = 1$. Suppose $0 < x < N$. After the first game, the bankroll of the gambler is either $x - 1$ or $x + 1$, and each of these outcomes is equally likely. Therefore,

$$F(x) = \frac{1}{2} F(x+1) + \frac{1}{2} F(x-1), \quad x = 1, \ldots, N-1. \tag{7}$$

One function F that satisfies (7) with the boundary conditions $F(0) = 0, F(N) = 1$ is $F(x) = x/N$. In fact, this is the only solution as we show now see.

Theorem 2.1. *Suppose a, b are real numbers and N is a positive integer. Then the only function $F : \{0, \ldots, N\} \to \mathbb{R}$ satisfying (7) with $F(0) = a$ and $F(N) = b$ is the linear function*

$$F_0(x) = a + \frac{x(b-a)}{N}.$$

This is a fairly easy theorem to prove. In fact, we will give several proofs. It is often useful to give different proofs to the same theorem because it gives us a number of different approaches to trying to prove generalizations. It is immediate that F_0 satisfies the conditions; the real question is one of uniqueness.

Proof 1. Consider the set $\mathcal{V}$ of all functions $F : \{0, \ldots, N\} \to \mathbb{R}$ that satisfy (7). It is easy to check that $\mathcal{V}$ is a vector space, i.e., if $f, g \in \mathcal{V}$ and c_1, c_2 are real numbers, then $c_1 f + c_2 g \in \mathcal{V}$. In fact, we claim that this vector space has dimension two. To see this, we will give a basis. Let f_1 be the function defined by $f_1(0) = 0, f_1(1) = 1$ and then extended in the unique way to satisfy (7). In other words, we define $f_1(x)$ for $x > 1$ by

$$f_1(x) = 2f_1(x-1) - f_1(x-2).$$

It is easy to see that f_1 is the only solution to (7) satisfying $f_1(0) = 0, f_1(1) = 1$. We define f_2 similarly with initial conditions $f_2(0) = 1, f_2(1) = 0$. Then $c_1 f_1 + c_2 f_2$ is the unique solution to (7) satisfying $f_1(0) = c_2, f_1(1) = c_1$. The set of functions of the form F_0 as a, b vary form a two dimensional subspace of $\mathcal{V}$ and hence must be all of $\mathcal{V}$.

Proof 2. Suppose F is a solution to (7). Then for each $0 < x < N$,

$$F(x) \leq \max\{F(x-1), F(x+1)\}.$$

Using this we can see that the maximum of F is obtained either at 0 or at N. Similarly, the minimum of F is obtained on $\{0, N\}$. Suppose $F(0) = 0, F(N) = 0$. Then the minimum and the maximum of the function are both 0 which means that $F \equiv 0$. Suppose $F(0) = a, F(N) = b$ and let F_0 be the linear function with these same boundary values. Then $F - F_0$ satisfies (7) with boundary value 0, and hence is identically zero. This implies that $F = F_0$.

Proof 3. Consider the equations (7) as $N-1$ linear equations in $N-1$ unknowns, $F(1), \ldots, F(N-1)$. We can write this as

$$\mathbf{A}\mathbf{v} = \mathbf{w},$$

where

$$\mathbf{A} = \begin{bmatrix} -1 & \frac{1}{2} & 0 & 0 & \cdots & 0 & 0 \\ \frac{1}{2} & -1 & \frac{1}{2} & 0 & \cdots & 0 & 0 \\ 0 & \frac{1}{2} & -1 & \frac{1}{2} & \cdots & 0 & 0 \\ & & & & \vdots & & \\ 0 & 0 & 0 & 0 & \cdots & -1 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \cdots & \frac{1}{2} & -1 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} -F(0) \\ 0 \\ 0 \\ \vdots \\ 0 \\ -F(N) \end{bmatrix}.$$

If we prove that $\mathbf{A}$ is invertible, then the unique solution is $\mathbf{v} = \mathbf{A}^{-1}\mathbf{w}$. To prove invertibility it suffices to show that $\mathbf{A}\mathbf{v} = 0$ has a unique solution and this can be done by an argument as in the previous proof.

Proof 4. Suppose F is a solution to (7). Let S_n be the random walk starting at x. We claim that for all n, $\mathbb{E}[F(S_{n\wedge T})] = F(x)$. We will show this by induction. For $n = 0$, $F(S_0) = F(x)$ and hence $\mathbb{E}[F(S_0)] = x$. To do the inductive step, we use a rule for expectation in terms of conditional expectations:

$$\mathbb{E}[F(S_{(n+1)\wedge T})] = \sum_{y=0}^{N} \mathbb{P}\{S_{n\wedge T} = y\}\, \mathbb{E}[F(S_{(n+1)\wedge T}) \mid S_{n\wedge T} = y].$$

If $y = 0$ or $y = N$ and $S_{n\wedge T} = y$, then $S_{(n+1)\wedge T} = y$ and hence $\mathbb{E}[F(S_{(n+1)\wedge T}) \mid S_{n\wedge T} = y] = F(y)$. If $0 < y < x$ and $S_{n\wedge T} = y$, then

$$\mathbb{E}[F(S_{(n+1)\wedge T}) \mid S_{n\wedge T} = y] = \frac{1}{2}\, F(y+1) + \frac{1}{2}\, F(y-1) = F(y).$$

Therefore,

$$\mathbb{E}[F(S_{(n+1)\wedge T})] = \sum_{y=0}^{N} \mathbb{P}\{S_{n\wedge T} = y\}\, F(y) = \mathbb{E}[F(S_{n\wedge T})] = F(x),$$

with the last inequality holding by the inductive hypothesis. Therefore,

$$\begin{aligned}
F(x) &= \lim_{n\to\infty} \mathbb{E}[F(S_{(n+1)\wedge T})] \\
&= \lim_{n\to\infty} \sum_{y=0}^{N} \mathbb{P}\{S_{n\wedge T} = y\}\, F(y) \\
&= \mathbb{P}\{S_T = 0\}\, F(0) + \mathbb{P}\{S_T = N\}\, F(N) \\
&= [1 - \mathbb{P}\{S_T = N\}]\, F(0) + \mathbb{P}\{S_T = N\}\, F(N).
\end{aligned}$$

Considering the case $F(0) = 0, F(N) = 1$ gives $\mathbb{P}\{S_T = N \mid S_0 = x\} = x/N$ and for more general boundary conditions,

$$F(x) = F(0) + \frac{x}{N}\,[F(N) - F(0)].$$

One nice thing about the last proof is that we did not need to have already guessed the linear functions as solutions. The proof produces these solutions.

2.2 Higher dimensions

We will generalize this result to higher dimensions. We replace the interval $\{1, \ldots, N\}$ with an arbitrary finite subset A of $\mathbb{Z}^d$. We let ∂A be the *(outer) boundary* of A defined by

$$\partial A = \{z \in \mathbb{Z}^d \setminus A : \mathrm{dist}(z, A) = 1\},$$

and we let $\overline{A} = A \cup \partial A$ be the *"closure"*[1] of A. We define the linear operators $\mathbf{Q}, \mathcal{L}$ on functions by

$$\mathbf{Q}F(x) = \frac{1}{2d} \sum_{y\in\mathbb{Z}^d, |x-y|=1} F(y),$$

$$\mathcal{L}F(x) = (\mathbf{Q} - \mathbf{I})F(x) = \frac{1}{2d} \sum_{y\in\mathbb{Z}^d, |x-y|=1} [F(y) - F(x)]$$

(The term linear operator is often used for a linear function whose domain is a space of functions. In our case, the domain is the space of functions on the finite set A which is isomorphic to $\mathbb{R}^K$ where $K = \#(A)$. Hence a linear operator is the same as a linear transformation from linear algebra. We can think of $\mathbf{Q}$ and $\mathcal{L}$ as matrices.) The operator $\mathcal{L}$

[1] This terminology may seem strange, but in the continuous analogue, A will be an open set, ∂A its topological boundary and $\overline{A} = A \cup \partial A$ its topological closure.

is often called the *(discrete) Laplacian*. We let S_n be a simple random walk in $\mathbb{Z}^d$. Then we can write

$$\mathcal{L}F(x) = \mathbb{E}[F(S_1) - F(S_0) \mid S_0 = x].$$

We say that F is *(discrete) harmonic* at x if $\mathcal{L}F(x) = 0$; this is an example of a *mean-value property.*

Dirichlet problem for harmonic functions. Given a set $A \subset \mathbb{Z}^d$ and a function $F : \partial A \to \mathbb{R}$ find an extension of F to $\overline{A}$ such that is harmonic in A, i.e.,

$$\mathcal{L}F(x) = 0 \text{ for all } x \in A. \tag{8}$$

We let $T_A = \min\{n : S_n \notin A\}$.

Theorem 2.2. *If $A \subset \mathbb{Z}^d$ is finite, then for every $F : \partial A \to \mathbb{R}$, there is a unique extension of F to $\overline{A}$ that satisfies* (8). *It is given by*

$$F_0(x) = \mathbb{E}[F(S_{T_A}) \mid S_0 = x] = \sum_{y \in \partial A} \mathbb{P}\{S_{T_A} = y \mid S_0 = x\}\, F(y).$$

It is not difficult to verify that F_0 as defined above is a solution to the Dirichlet problem. The problem is to show that it is unique. Let us consider the last method of proof in the previous section. Suppose F is harmonic on A; $S_0 = x \in \overline{A}$; and let

$$M_n = F(S_{n \wedge T_A}).$$

Then (8) can be rewritten as

$$\mathbb{E}[M_{n+1} \mid S_0, \ldots, S_n] = F(S_{n \wedge T_A}) = M_n. \tag{9}$$

A process that satisfies $\mathbb{E}[M_{n+1} \mid S_0, \ldots, S_n] = M_n$ is called a *martingale (with respect to the random walk).* It is easy to see that $F(S_{n \wedge T_A})$ being a martingale is essentially equivalent to F being harmonic on A. It is easy to check that martingales satisfy $\mathbb{E}[M_n] = \mathbb{E}[M_0]$, and hence if $S_0 = x$,

$$\mathbb{E}[M_n] = \sum_{y \in \overline{A}} \mathbb{P}\{S_{n \wedge T_A} = y\}\, F(y) = M_0 = F(x).$$

An easy argument shows that with probability one $T_A < \infty$. We can take limits and get

$$F(x) = \lim_{n \to \infty} \sum_{y \in \overline{A}} \mathbb{P}\{S_{n \wedge T_A} = y\}\, F(y) = \sum_{y \in \partial A} \mathbb{P}\{S_{T_A} = y\}\, F(y).$$

(There is no problem interchanging the limit and the sum because it is a finite sum.)

Let us consider this from the perspective of linear algebra. Suppose that A has N elements and ∂A has K elements. The solution of the Dirichlet problem can be considered as a linear function from the set of functions on ∂A to the set of functions on A (the reader should check that this is a linear transformation). Any such function can be written in terms of an $N \times K$ matrix which we write as

$$\mathbf{H}_A = [H_A(x,y)]_{x \in A, y \in \partial A}$$

What we have shown is that

$$H_A(x,y) = \mathbb{P}\left\{S_{T_A} = y \mid S_0 = x\right\}.$$

This matrix is often called the *Poisson kernel*. For a given set A, we can solve the Dirichlet problem for any boundary function in terms of the Poisson kernel.

What happens if we allow A to be an infinite set? In this case it is not always true that the solution is unique. Let us consider the one-dimensional example with $A = \{1, 2, 3, \ldots\}$ and $\partial A = \{0\}$. Then for every $c \in \mathbb{R}$, the function $F(x) = cx$ is harmonic in A with boundary value 0 at the origin. Where does our proof break down? This depends on which proof we consider (they all break down!), but let us consider the martingale version. Suppose F is harmonic on A with $F(0) = 0$ and suppose S_n is a simple random walk starting at positive integer x. As before, we let $T = \min\{n : S_n = 0\}$ and $M_n = F(S_{n \wedge T})$. The same argument shows that M_n is a martingale and

$$F(x) = \mathbb{E}[M_n] = \sum_{y=0}^{\infty} F(y)\, \mathbb{P}\{S_{n \wedge T} = y\}.$$

We have shown in a previous chapter that with probability one $T < \infty$. This implies

$$\lim_{n \to \infty} \sum_{y > 0} \mathbb{P}\{S_{n \wedge T} = y\} = 0.$$

However, if F is unbounded, we cannot conclude from this that

$$\lim_{n \to \infty} \sum_{y > 0} F(y)\, \mathbb{P}\{S_{n \wedge T} = y\} = 0.$$

However, we do see from this that there is only one *bounded* function that is harmonic on A with a given boundary value at 0.

3 Heat equation

We will now introduce the discrete heat equation which is a model for heat flow. Let A be a finite subset of $\mathbb{Z}^d$ with boundary ∂A. Suppose we set the temperature at the boundary to be zero. Consider the temperature in A to be controlled by a very large number of "heat particles". These particles perform random walks on A until they leave A at which time they are killed. The temperature at x at time n, $p_n(x)$ is given by the density of particles at x. The random walk rule gives the relation

$$p_{n+1}(x) = \frac{1}{2d} \sum_{|y-x|=1} p_n(y).$$

If we introduce the notation $\partial_n p_n(x) = p_{n+1}(x) - p_n(x)$, we get the *heat equation*

$$\partial_n p_n(x) = \mathcal{L} p_n(x), \quad x \in A. \tag{10}$$

The initial temperature is given as an initial condition

$$p_0(x) = f(x), \quad x \in A. \tag{11}$$

The boundary condition is

$$p_n(x) = 0, \quad x \in \partial A. \tag{12}$$

If $x \in A$ and the initial condition is $f(x) = 1$ and $f(z) = 0$ for $z \neq x$, then

$$p_n(y) = \mathbb{P}\{S_{n \wedge T_A} = y \mid S_0 = x\}.$$

Given any initial condition f, it is easy to see that there is a unique function p_n satisfing (10)–(12). Indeed, we just set $p_0(x) = f(x), x \in A$ and for $n > 0$, we define $p_n(x), x \in A$ recursively by (10). This tells us that set of functions satisfying (10) and (12) is a vector space of dimension $\#(A)$. In fact, $\{p_n(x) : x \in A\}$ is the vector $\mathbf{Q}^n f$.

Once we have existence and uniqueness, the problem remains to find the function. For a bounded set A, this is a problem in linear algebra and essentially becomes the question of diagonalizing the matrix $\mathbf{Q}$. We will state an important fact from linear algebra.

Theorem 3.1. *Suppose $\mathbf{A}$ is a $k \times k$ symmetric matrix. Then we can find k (not necessarily distinct) real eigenvalues*

$$\lambda_k \leq \lambda_{k-1} \leq \cdots \leq \lambda_1,$$

and k orthogonal vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ that are eigenvectors,

$$\mathbf{A}\mathbf{v}_j = \lambda_j \, \mathbf{v}_j.$$

Example Let us compute the function p_n in the case $d = 1$, $A = \{1, \ldots, N-1\}$. We start by looking for functions satisfying (10) of the form

$$p_n(x) = \lambda^n \, \phi(x). \tag{13}$$

If p_n is of this form, then

$$\partial_n p_n(x) = (\lambda - 1) \, p_n(x).$$

This leads us to look for eigenvalues and eigenfunctions of $\mathbf{Q}$, i.e., to find λ, ϕ such that

$$\mathbf{Q}\phi(x) = \lambda \, \phi(x), \tag{14}$$

with $\phi \equiv 0$ on ∂A. The sum rule for sine

$$\sin((x \pm 1)\theta) = \sin(x\theta)\cos(\theta) \pm \cos(x\theta)\sin(\theta),$$

tells us that

$$\mathbf{Q}[\sin(\theta x)] = \lambda_\theta \, \sin(\theta x), \qquad \lambda_\theta = \cos\theta.$$

If we choose $\phi_j(x) = \sin(\pi j x/N)$, then we satisfy the boundary condition $\phi_j(0) = \phi_j(x) = 0$. Since these are eigenvectors with different eigenvalues, we know that they are orthogonal, and hence linearly independent. Hence every function f on A can be written in a unique way as a finite *Fourier series*

$$f(x) = \sum_{j=1}^{N-1} c_j \, \sin\left(\frac{\pi j x}{N}\right), \tag{15}$$

and the solution to the heat equation with initial condition f is

$$p_n(y) = \sum_{j=1}^{N-1} c_j \, \left[\cos\left(\frac{j\pi}{N}\right)\right]^n \, \phi_j(y).$$

It is not difficult to check that

$$\sum_{x=1}^{N-1} \sin\left(\frac{\pi j x}{N}\right) \, \sin\left(\frac{\pi k x}{N}\right) = \begin{cases} \frac{N-1}{2} & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases}.$$

(For $j \neq k$ this is orthogonality of different eigenvectors. For $j = k$ one can show that the sum is the same if we replace sin wth cos and we can add the sums.) In particular, if we choose the solution with initial condition $f(x) = 1; f(z) = 0, z \neq x$ is

$$\mathbb{P}\{S_{n\wedge T_A} = y \mid S_0 = x\} \frac{2}{N-1} \sum_{j=1}^{N-1} \phi_j(x) \, \left[\cos\left(\frac{j\pi}{N}\right)\right]^n \, \phi_j(y).$$

As $n \to \infty$, this sum is dominated by the $j = 1$ and $j = N-1$ terms for which the eigenvalue has maximal absolute value. These two terms give

$$\frac{2}{N-1} \cos^n\left(\frac{2\pi}{N}\right) \, \left[\sin\left(\frac{\pi x}{N}\right) \, \sin\left(\frac{\pi y}{N}\right) + (-1)^n \, \sin\left(\frac{x\pi(N-1)}{N}\right) \, \sin\left(\frac{y\pi(N-1)}{N}\right)\right].$$

One can check that

$$\sin\left(\frac{x\pi(N-1)}{N}\right) = (-1)^x \, \sin\left(\frac{\pi x}{N}\right),$$

and hence if $x, y \in \{1, \ldots, N-1\}$, as $n \to \infty$,

$$\mathbb{P}\{S_{n\wedge T_A} = y \mid S_0 = x\} \sim \frac{4}{N-1} \cos^n\left(\frac{2\pi}{N}\right) \, (-1)^{n+x+y} \, \sin\left(\frac{\pi x}{N}\right) \, \sin\left(\frac{\pi y}{N}\right).$$

The technique of finding solutions of the form (13) goes under the name of *separation of variables*. In the case of finite A this is essentially the same as computing powers of a matrix by diagonalization. We summarize here.

Theorem 3.2. *If A is a finite subset of $\mathbb{Z}^d$ with N elements, then we can find N linearly independent functions $\phi_1, \ldots, \phi_N$ that satisfy* (14) *with real eigenvalues $\lambda_1, \ldots, \lambda_N$. The solution to* (10)–(12) *is given by*

$$p_n(x) = \sum_{j=1}^{N} c_j \, \lambda_j^n \, \phi_j(x),$$

where c_j are chosen so that

$$f(x) = \sum_{j=1}^{N} c_j \, \phi_j(x).$$

In fact, the ϕ_j can be chosen to be orthonormal,

$$\langle \phi_j, \phi_k \rangle := \sum_{x \in A} \phi_j(x) \, \phi_k(x) = \delta(k - j).$$

Here we have introduced the *delta function* notation, $\delta(z) = 1$ if $z = 0$ and $\delta(z) = 0$ if $z \neq 0$. Since $p_n(x) \to 0$ as $n \to \infty$, we know that the eigenvalues have absolute value strictly less than one. We can order the egienvalues

$$1 > \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N > -1.$$

We will write $p(x, y; A)$ to be the solution of the heat equation with initial condition equal to one at x and 0 otherwise. In other words,

$$p_n(x, y; A) = \mathbb{P}\{S_n = y, T_A > n \mid S_0 = x\}, \qquad x, y \in A.$$

Then if $\#(A) = N$,

$$p_n(x, y; A) = \sum_{j=1}^{N} c_j(x) \, \lambda_j^n \, \phi_j(y)$$

where $c_j(x)$ have been chosen so that

$$\sum_{j=1}^{N} c_j(x) \phi_j(y) = \delta(y - x).$$

In fact, this tells us that $c_j(x) = \phi_j(x)$. Hence

$$p_n(x, y; A) = \sum_{j=1}^{N} \lambda_j^n \, \phi_j(x) \, \phi_j(y).$$

Note that the quantity on the right is symmetric in x, y. One can check that the symmetry also follows from the defintion of $p_n(x, y; A)$.

The largest eigenvalue λ_1 is often denoted λ_A. We can give a "variational" definition of λ_A as follows.

Theorem 3.3. *If A is a finite subset of $\mathbb{Z}^d$, then λ_A is given by*

$$\lambda_A = \sup_f \frac{\langle \mathbf{Q} f, f\rangle}{\langle f, f\rangle},$$

where the supremum is over all functions f on A, and $\langle \cdot, \cdot \rangle$ denotes inner product

$$\langle f, g\rangle = \sum_{x \in A} f(x)\, g(x).$$

The proof of this is easy. First, if ϕ is an eigenvector with eigenvalue λ_1, then $\mathbf{Q}\phi = \lambda_1 \phi$ and plugging in ϕ shows that the supremum is at least as large as λ_1. Conversely, there is an orthogonal basis of eigenfunctions $\phi_1, \ldots, \phi_N$ and we can write any f as

$$f = \sum_{j=1}^{N} c_j\, \phi_j.$$

Then

$$\begin{aligned} \langle \mathbf{Q} f, f\rangle = \left\langle \mathbf{Q} \sum_{j=1}^{N} c_j\, \phi_j, \sum_{j=1}^{N} c_j\, \phi_j \right\rangle &= \left\langle \sum_{j=1}^{N} c_j \mathbf{Q}\phi_j, \sum_{j=1}^{N} c_j\, \phi_j \right\rangle \\ &= \sum_{j=1} c_j\, \lambda_j\, \langle \phi_j, \phi_j\rangle \\ &\leq \lambda_1 \sum_{j=1} c_j\, \langle \phi_j, \phi_j\rangle = \lambda_1\, \langle f, f\rangle. \end{aligned}$$

The reader should check that the computation above uses the orthogonality of the eigenfunctions and also the fact that $\langle \phi_j, \phi_j\rangle \geq 0$.

Using this variational formulation, we can see that the eigenfunction for λ_1 can be chosen so that $\phi_1(x) \geq 0$ for each x (since if ϕ_1 took on both positive and negative values, we would have $\langle \mathbf{Q}|\phi_1|, |\phi_1|\rangle > \langle \phi_1, \phi_1\rangle$). The eigenfunction is unique, i.e., $\lambda_2 < \lambda_1$, provided we put an additional condition on A. We say that a subset A on $\mathbb{Z}^d$ is *connected* if any two points in A are connected by a nearest neighbor path that stays entirely in A. Equivalently, A is connected if for each $x, y \in A$ there exists an n such that $p_n(x, y; A) > 0$. We leave it as Exercise 6.23 to show that this implies that $\lambda_1 > \lambda_2$.

Before stating the final theorem, we need to discuss some parity (even/odd) issues. If $x = (x_1, \ldots, x_d) \in \mathbb{Z}^d$ we let $e(x) = (-1)^{x_1 + \cdots + x_d}$. We call x *even* if $e(x) = 1$ and otherwise x is odd. If n is a nonnegative integer, then

$$p_n(x, y; A) = 0 \;\text{ if }\; (-1)^n e(x + y) = -1.$$

If $\mathbf{Q}\phi = \lambda\phi$, then $\mathbf{Q}[e\phi] = -\lambda e\phi$.

Theorem 3.4. *Suppose A is a finite connected subset of $\mathbb{Z}^d$ with at least two points. Then $\lambda_1 > \lambda_2$, $\lambda_N = -\lambda_1 < \lambda_{N-1}$. The eigenfunction ϕ_1 can be chosen so that $\phi_1(x) > 0$ for all $x \in A$.*

$$\lim_{n\to\infty} \lambda_1^{-n}\, p_n(x,y;A) = [1 + (-1)^n\, e(x+y)]\, \phi_1(x)\, \phi_1(y).$$

Example One set in $\mathbb{Z}^d$ for which we can compute the eigenfunctions and eigenvalues exactly is a d-dimensional rectangle

$$A = \{(x_1, \ldots, x_d) \in \mathbb{Z}^d : 1 \le x_j \le N_j - 1\}.$$

The eigenfunctions are indexed by $\bar{k} = (k_1, \ldots, k_d) \in A$,

$$\phi_{\bar{k}}(x_1, \ldots, x_d) = \sin\left(\frac{k_1 \pi x_1}{N_1}\right) \sin\left(\frac{k_2 \pi x_2}{N_2}\right) \cdots \sin\left(\frac{k_d \pi x_d}{N_d}\right),$$

with eigenvalue

$$\lambda_{\bar{k}} = \frac{1}{d}\left[\cos\left(\frac{k_1 \pi}{N_1}\right) + \cdots + \cos\left(\frac{k_d \pi}{N_d}\right)\right].$$

4 Expected time to escape

4.1 One dimension

Let S_n denote the position of a one-dimensional random walk starting at $x \in \{0, \ldots, N\}$ and let T be the first time that the walker reaches $\{0, N\}$. Here we study the expected time to reach 0 or N,

$$e(x) = \mathbb{E}[T \mid S_0 = x].$$

Clearly $e(0) = e(N) = 0$. Now suppose $x \in \{1, \ldots, N-1\}$. Then the walker takes one step which goes to either $x-1$ to $x+1$. Using this we get the relation

$$e(x) = 1 + \frac{1}{2}\left[e(x+1) + e(x-1)\right].$$

Hence e satisfies

$$e(0) = e(N) = 0, \quad \mathcal{L}e(x) = -1, \quad x = 1, \ldots, N-1. \tag{16}$$

A simple calculation shows that if $f(x) = x^2$, then $\mathcal{L}f(x) = 1$. Using this we can see that one solution to (16) is

$$e(x) = x\,(N-x).$$

In fact it is the unique solution. To see this, assume that e_1 were another solution. Then $e - e_1$ is harmonic on $\{1, \ldots, N-1\}$ and vanishes at 0 and N which implies that $e - e_1 = 0$. Suppose $N = 2m$ is even. Then we get

$$e(m) = N^2/4 = m^2.$$

In other words, the expected time for a random walker to go distance m is *exactly* m^2. Note that the expected time starting at $x = 1$ to leave the interval is N. While this is an expected value, it is not necessarily a "typical" value. Most of the time the random walker will leave quickly. However, there is a probability of $1/m$ that the random walker will reach m before leaving the interval. If that happens then the walker will still have on the order of N^2 steps before leaving.

One other interesting fact concerns the time until a walker starting at 1 reaches the origin. Let T_0 be the first n such that $S_n = 0$. If $S_0 = 1$, we know that $T_0 < \infty$ with probability one. However, the amount of time to reach 0 is at least as large as the amount of time to reach 0 or N. Therefore $\mathbb{E}[T_0] \geq N$. Since this is true for every N, we must have $\mathbb{E}[T_0] = \infty$. In other words, while it is guaranteed that a random walker will return to the origin *the expected amount of time until it happens is infinite!*

4.2 Several dimensions

Let A be a finite subset of $\mathbb{Z}^d$; S_n a simple random walker starting at $x \in \overline{A}$; and T_A the first time that the walker is not in A. Let

$$e_A(x) = \mathbb{E}[T_A \mid S_0 = x].$$

Then just as in the one-dimensional case we can see that $f(x) = e_A(x)$ satisfies

$$f(x) = 0, \quad x \in \partial A \tag{17}$$

$$\mathcal{L}f(x) = -1, \quad x \in A. \tag{18}$$

We can argue in the same as in the one-dimensional case that there is at most one function satisfying these equations.

Let $f(x) = d|x|^2 = d(x_1^2 + \cdots + x_d^2)$. Then a simple calculation shows that $\mathcal{L}f(x) = 1$. Let us consider the process

$$M_n = dS_{n\wedge T_A}^2 - (n \wedge T_A).$$

Then, we can see that

$$\mathbb{E}[M_{n+1} \mid S_0, \ldots, S_n] = M_n,$$

and hence M_n is a martingale. This implies

$$\mathbb{E}[M_n] = \mathbb{E}[M_0] = d|S_0|^2, \quad \mathbb{E}[n \wedge T_A] = d\,\left[\mathbb{E}[|S_{n\wedge T_A}|^2] - |S_0|^2\right].$$

In fact, we claim we can take the limit to assert

$$\mathbb{E}[T_A] = d\,\left[\mathbb{E}[|S_{T_A}|^2] - |S_0|^2\right].$$

To prove this we use the *monotone convergence theorem*, see Exercise 6.6. This justifies the step

$$\lim_{n\to\infty} \left[\mathbb{E}[|S_{T_A}|^2\, 1\{T_A \leq n\}\right] = \mathbb{E}[|S_{T_A}|^2].$$

Also,

$$\left[\mathbb{E}[|S_{T_A}|^2\, 1\{T_A > n\}\right] \leq \mathbb{P}\{T_A > n\} \left[\max_{x\in A} |x|^2\right] \to 0.$$

This is a generalization of the formula we derived in the one-dimensional case. If $d = 1$ and $A = \{1, \ldots, N-1\}$, and

$$\mathbb{E}[|S_{T_A}|^2] = N^2\, \mathbb{P}\{S_{T_A} = N \mid S_0 = x\} = N\, x,$$

$$\mathbb{E}[T_A] = \mathbb{E}[|S_{T_A}|^2] - x^2 = x\,(N-x).$$

Example Suppose that A is the "discrete ball" of radius r about the origin,

$$A = \{x \in \mathbb{Z}^d : |x| < r\}.$$

Then for every $y \in \partial A$ satisfies $r \leq |y| < r+1$. Suppose we start the random walk at the origin.

$$d\, r^2 \leq \mathbb{E}[T_A] < d\,(r+1)^2.$$

For any $y \in A$, let V_y denote the number of visits to y before leaving A,

$$V_y = \sum_{n=0}^{T_A - 1} 1\{S_n = y\} = \sum_{n=0}^{\infty} 1\{S_n = y, T_A > n\}.$$

Here we again use the indicator function notation. Note that

$$\mathbb{E}[V_y \mid S_0 = x] = \sum_{n=0}^{\infty} \mathbb{P}\{S_n = y, T_A > n \mid S_0 = x\} = \sum_{n=0}^{\infty} p_n(x, y; A).$$

This quantity is of sufficient interest that it is given a name. The *Green's function* $G_A(x,y)$ is the function on $A \times A$ given by

$$G_A(x,y) = \mathbb{E}[V_y \mid S_0 = x] = \sum_{n=0}^{\infty} p_n(x,y;A).$$

We define $G_A(x,y) = 0$ if $x \notin A$ or $y \notin A$. The Green's function satisfies $G_A(x,y) = G_A(y,x)$. This is not immediately obvious from the first equality but follows from the symmetry of $p_n(x,y;A)$. If we fix $y \in A$, then the function $f(x) = G_A(x,y)$ satisfies the following:

$$\mathcal{L}f(y) = -1,$$

$$\mathcal{L}f(x) = 0, \quad x \in A \setminus \{y\},$$

$$f(x) = 0, \quad x \in \partial A.$$

Note that

$$T_A = \sum_{y \in A} V_y,$$

and hence

$$\mathbb{E}[T_A \mid S_0 = x] = \sum_{y\in A} G_A(x,y).$$

Theorem 4.1. *Suppose A is a bounded subset of $\mathbb{Z}^d$; and $g : A \to \mathbb{R}$ is a given function. Then the unique function $F : \overline{A} \to \mathbb{R}$ satisfying*

$$F(x) = 0, \qquad x \in \partial A,$$

$$\mathcal{L}F(x) = -g(x), \quad x \in A,$$

is

$$F(x) = \mathbb{E}\left[\sum_{j=0}^{T_A-1} g(S_j) \mid S_0 = x\right] = \sum_{y\in A} g(y)\, G_A(x,y). \tag{19}$$

We have essentially already proved this. Uniqueness follows from the fact that if F, F_1 are both solutions, then $F - F_1$ is harmonic in A with boundary value 0 and hence equals 0 everywhere. Linearity of $\mathcal{L}$ shows that

$$\mathcal{L}\left[\sum_{y\in A} g(y)\, G_A(x,y)\right] = \sum_{y\in A} g(y)\, \mathcal{L}G_A(x,y) = -g(x). \tag{20}$$

The second equality in (19) follows by writing

$$\sum_{j=0}^{T_A-1} g(S_j) = \sum_{j=0}^{T_A-1}\sum_{y\in A} g(y)\, 1\{S_j = y\} = \sum_{y\in A} g(y) \sum_{j=0}^{T_A-1} 1\{S_j = y\} = \sum_{y\in A} g(y)\, V_y.$$

We can consider the Green's function as a matrix or operator,

$$G_A g(x) = \sum_{x\in A} G_A(x,y)\, g(y).$$

Then (20) can be written as

$$-\mathcal{L}G_A g(x) = g(x),$$

or $G_A = [-\mathcal{L}]^{-1}$ For this reason the Green's function is often referred to as the *inverse of (the negative of) the Laplacian.*

If $d \geq 3$, then the expected number of visits to a point is finite and we can define the *(whole space) Green's function*

$$G(x,y) = \lim_{A\uparrow\mathbb{Z}^d} G_A(x,y) = \mathbb{E}\left[\sum_{n=0}^{\infty} 1\{S_n = y\} \mid S_0 = x\right] = \sum_{n=0}^{\infty} \mathbb{P}\{S_n = y \mid S_0 = x\}.$$

It is a bounded function. In fact, if τ_y denotes the smallest $n \geq 0$ such that $S_n = y$, then

$$G(x,y) = \mathbb{P}\{\tau_y < \infty \mid S_0 = x\}\, G(y,y) = \mathbb{P}\{\tau_y < \infty \mid S_0 = x\}\, G(0,0) \leq G(0,0) < \infty.$$

The function G is symmetric and satisfies a translation invariance property: G(x,y) = $G(0, y-x)$. for fixed y, $f(x) = G(x,y)$ satisfies

$$\mathcal{L}f(y) = -1, \quad \mathcal{L}f(x) = 0,\, x \neq y, \quad f(x) \to 0 \text{ as } x \to \infty.$$

5 Space of harmonic functions

If A is a finite subset of $\mathbb{Z}^d$, then the space of functions on $\overline{A}$ that are harmonic on A has dimension $\#(\partial A)$. In fact, as we have seen, there is a linear isomorphism between this space and the set of all functions on $\#(\partial A)$. One standard basis for the space is the Poisson kernel, i.e., for each $z \in \partial A$, we consdier the function

$$H_z(x) = H_A(x, y) = \mathbb{P}\{S_{T_A} = z \mid S_0 = x\}.$$

Sometimes other bases are useful.

Example Let A be the rectangle in $\mathbb{Z}^2$,

$$A = \{(x_1, x_2) : x_j = 1, \ldots, N - 1\}.$$

We will find a basis for the space of harmonic functions by another form of "separation of variables". We can write $\partial A = \partial_{1,0} \cup \partial_{1,N} \cup \partial_{2,0} \cup \partial_{2,N}$ where $\partial_{1,0} = \{(0, x_2) : x_2 = 1, \ldots, N - 1\}$, etc. Consider the function

$$h_j(x) = h_{j,1,N}(x) = \sinh\left(\frac{\beta_j \, x_1}{N}\right) \sin\left(\frac{j\pi x_2}{N}\right).$$

Using the sum formulas for sinh and sin we get

$$\mathcal{L}h_j(x) = \left[\frac{1}{2}\cosh\left(\frac{\beta_j}{N}\right) + \frac{1}{2}\cos\left(\frac{j\pi}{N}\right) - 1\right] h_j(x).$$

Since $\cosh(0) = 1$ and $\cosh(x)$ increases to infinity for $0 \leq x < \infty$, we can see that there is a unique number which we call β_j such that

$$\cosh\left(\frac{\beta_j}{N}\right) + \cos\left(\frac{j\pi}{N}\right) = 2,$$

When we choose this β_j, h_j is a harmonic function. Note that h_j vanishes on three of the four parts of the boundary and

$$h_j(N, y) = \sinh(\beta_j) \sin\left(\frac{j\pi y}{N}\right).$$

If we choose $y \in \{1, \ldots, N - 1\}$ and find constants $c_1, \ldots, c_{N-1}$ such that

$$\sum_{j=1}^{N-1} c_j \sinh(\beta_j) \sin\left(\frac{j\pi k}{N}\right) = \delta(y - k),$$

Then,

$$H_{(N,y)}(x) = \sum_{j=1}^{N-1} c_j \, h_j(x).$$

But we have already seen that the correct choice is

$$c_j = \frac{2}{(N-1)\,\sinh(\beta_j)}\,\sin\left(\frac{\pi j y}{N}\right).$$

Therefore,

$$H_{(N,y)}(x_1,x_2) = \frac{2}{N-1}\sum_{j=1}^{N-1}\frac{1}{\sinh(\beta_j)}\,\sin\left(\frac{\pi j y}{N}\right)\,\sinh\left(\frac{\beta_j\,x_1}{N}\right)\,\sin\left(\frac{j\pi x_2}{N}\right). \tag{21}$$

The formula (21) is somewhat complicated, but there are some nice things that can be proved using this formula. Let A_N denote the rectangle and let

$$\hat{A}_N = \left\{(x_1,x_2)\in A : \frac{N}{4}\leq x_j \leq \frac{3N}{4}\right\}.$$

Note that $\hat{A}_N$ is a cube of half the side length of A_N in the middle of A_N. Let $y\in\{1,\ldots,N-1\}$ and consider $H_{(N,y)}$. In Exercise 6.13 you are asked to show the following: there exist $c, c_1 < \infty$ such that the following is true for every N and every y and every $x,\tilde{x}\in\hat{A}_N$:

- $$c^{-1}\,N^{-1}\,\sin(\pi y/N) \leq H_{N,y}(x) \leq c\,N^{-1}\,\sin(\pi y/N). \tag{22}$$

  In particular,
  $$H_{N,y}(x) \leq 2c\,H_{N,y}(\tilde{x}). \tag{23}$$

- $$|H_{N,y}(x) - H_{N,y}(\tilde{x})| \leq c_1\,\frac{|x-\tilde{x}|}{N}\,H_{n,y}(x) \leq c_1\,c\,\frac{|x-\tilde{x}|}{N}\,N^{-1}\,\sin(\pi y/N). \tag{24}$$

The argument uses the explicit formula that we derive for the rectangle. Although we cannot get such a nice formula in general, we can derive two important facts. Suppose A is a finite subset of $\mathbb{Z}^2$ containing A_N. Then for $x\in A_N$, $z\in\partial A$,

$$H_A(x,z) = \sum_{y\in\partial A_N} H_{A_N}(x,y)\,H_A(y,z).$$

Using this and (23) we get for $x,\tilde{x}\in\hat{A}_N$,

$$H_A(x,z) \leq 2c\,H_A(\tilde{x},z),$$

$$|H_A(x,z) - H_A(\tilde{x},z)| \leq c_1\,\frac{|x-\tilde{x}|}{N}\,H_A(x,z).$$

We can extend this to harmonic functions.

Theorem 5.1 (Difference estimates)**.** *There is a $c < \infty$ such that if A is a finite subset of $\mathbb{Z}^d$ and $F : \overline{A} \to [-M, M]$ is harmonic on A, then if $x, z \in A$ with $|z - x| = 1$,*

$$|F(z) - F(x)| \leq \frac{c\,M}{\mathrm{dist}(x, \partial A)}. \tag{25}$$

Theorem 5.2 (Harnack principle)**.** . *Suppose K is a compact subset of $\mathbb{R}^d$ and U is an open set containing K. There is a $c = c(K, U) < \infty$ such that the following holds. Suppose N is a positive integer; A is a finite subset of $\mathbb{Z}^d$ contained in $NU = \{z \in \mathbb{R}^d : z/N \in U\}$; and $\hat{A}$ is a subset of A contained in NK. Suppose $F : \overline{A} \to [0, \infty)$ is a harmonic function. Then for all $x, z \in \hat{A}$,*

$$F(x) \leq c\,F(z).$$

As an application of this, let us show that that: *the only bounded functions on $\mathbb{Z}^d$ that are harmonic everywhere are constants.* For $d = 1$, this is immediate from the fact that the only harmonic functions are the linear functions. For $d \geq 2$, we suppose that F is a harmonic function on $\mathbb{Z}^d$ with $|F(z)| \leq M$ for all z. If $x \in \mathbb{Z}^d$ and A_R is a bounded subset of $\mathbb{Z}^d$ containing all the points within distance R of the origin, then (25) shows that

$$|F(x) - F(0)| \leq c\,M\,\frac{|x|}{R - |x|}.$$

(Although (25) gives this only for $|x| = 1$, we can apply the estimate $O(|x|)$ times to get this estimate.) By letting $R \to \infty$ we see that $F(x) = F(0)$. Since this is true for every F, F must be constant.

5.1 Exterior Dirichlet problem

We consider the following question. Suppose A is a cofinite subset of $\mathbb{Z}^d$, i.e., a subset such that $\mathbb{Z}^d \setminus A$ is finite. Suppose $F : \mathbb{Z}^d \setminus A \to \mathbb{R}$ is given. Find all bounded functions on $\mathbb{Z}^d$ that are harmonic on A and take on the boundary value F on $\mathbb{Z}^d \setminus A$. If $A = \mathbb{Z}^d$, then this was answered at the end of the last section; the only possible functions are constants. For the remainder of this section we assume that A is nontrivial, i.e., $\mathbb{Z}^d \setminus A$ is nonempty.

For $d = 1, 2$, there is, in fact, only a single solution. Suppose F is such a function with $L = \sup |F(x)| < \infty$. Let S_n be a simple random walk starting at $x \in \mathbb{Z}^d$, and let $T = T_A$ be the first time n with $S_n \notin A$. If $d \leq 2$, we know that the random walk is recurrent and hence $T < \infty$ with probability one. As done before, we can see that $M_n = F(S_{n \wedge T})$ is a martingale and hence

$$F(x) = M_0 = \mathbb{E}[M_n] = \mathbb{E}[F(S_T)\,1\{T \leq n\}] + \mathbb{E}\left[F(S_n)\,1\{T > n\}\right].$$

The monotone convergence theorem tells us that

$$\lim_{n \to \infty} \mathbb{E}[F(S_T)\,1\{T \leq n\}] = \mathbb{E}[F(S_T)].$$

Also

$$\lim_{n\to\infty} |\mathbb{E}\,[F(S_n)\; 1\{T > n\}]| \le \lim_{n\to\infty} L\,\mathbb{P}\{T > n\} = 0.$$

Therefore,

$$F(x) = \mathbb{E}[F(S_T) \mid S_0 = x],$$

which is exactly the same solution as we had for bounded A.

If $d \ge 3$, there is more than one solution. In fact,

$$f(x) = \mathbb{P}\{T_A = \infty \mid S_0 = x\},$$

is a bounded function that is harmonic in A and equals zero on $\mathbb{Z}^d \setminus A$. The next theorem shows that this is essentially the only new function that we get. We can interpret the theorem as saying that the boundary value determines the function *if* we include ∞ as a boundary point.

Theorem 5.3. *If A is a nontrivial cofinite subset of $\mathbb{Z}^d$, then the only bounded functions on $\mathbb{Z}^d$ that vanish on $\mathbb{Z}^d \setminus A$ and are harmonic on A are of the form*

$$F(x) = r\,\mathbb{P}\{T_A = \infty \mid S_0 = x\}, \quad r \in \mathbb{R}. \tag{26}$$

We will first consider the case $A = \mathbb{Z}^d \setminus \{0\}$ and assume that $F : \mathbb{Z}^d \to [-M, M]$ is a function satisfying $\mathcal{L}F(x) = 0$ for $x \neq 0$. Let $\alpha = \mathcal{L}F(0)$ and let

$$f(x) = F(x) - \alpha\, G(x, 0).$$

Then f is a bounded harmonic function and hence must be equal to a constant, say C. Since $G(x, 0) \to 0$ as $x \to \infty$, the constant must be r and hence

$$F(x) = r - \alpha\, G(x, 0) = r\,\mathbb{P}\{\tau_0 = \infty \mid S_0 = x\} + \mathbb{P}\{\tau_0 < \infty \mid S_0 = x\}[r - \alpha G(0, 0)].$$

Since $F(0) = 0$ and $\mathbb{P}\{\tau_0 = \infty \mid S_0 = 0\} = 0$, we know that $r - \alpha G(0, 0) = 0$ and hence F is of the form (26).

For other cofinite A, assume F is such a function with $|F| \le 1$. Then F satisfies

$$\mathcal{L}F(x) = -g(x), \quad x \in A$$

for some function g that vanishes on A. In particular,

$$f(x) = F(x) + \sum_{y \in \mathbb{Z}^d \setminus A} G(x, y)\, g(x),$$

is a bounded harmonic function (why is it bounded?) and hence constant. This tells us that there is an r such that

$$F(x) = r - \sum_{y \in \mathbb{Z}^d \setminus A} G(x, y)\, g(x),$$

which implies, in particular, that $F(x) \to r$ as $r \to \infty$. Also, if $x \in \mathbb{Z}^d \setminus A$, $F(x) = 0$ which implies

$$\sum_{y \in \mathbb{Z}^d \setminus A} G(x,y)\, g(x) = r.$$

If we show that $G(x,y)$ is invertible on $\mathbb{Z}^d \setminus A$, then we know there is a unique solution to this equation, which would determine g, and hence F.

To do this, assume $\#(\mathbb{Z}^d \setminus A) = K$; let $\tilde{T}_A = \min\{n \geq 1 : S_n \notin A\}$; and for $x, y \in \mathbb{Z}^d \setminus A$, we define

$$J(x,y) = \mathbb{P}\{\tilde{T}_A < \infty, S_{\tilde{T}_A} = y \mid S_0 = x\}.$$

then J is a $K \times K$ matrix. In fact (Exercise 6.25),

$$(J - I)\, G = -I. \tag{27}$$

In particular, G is invertible.

6 Exercises

Exercise 6.1. *Suppose that $X_1, X_2, \ldots$ are independent, identically distributed random variables such that*

$$\mathbb{E}[X_j] = 0, \quad \mathrm{Var}[X_j] = \mathbb{E}[X_j^2] = 0 \quad \mathbb{P}\{|X_j| > K\} = 0,$$

for some $K < \infty$.

- *Let $M(t) = \mathbb{E}[e^{tX_j}]$ denote the moment generating function of X_j. Show that for every $t > 0, \epsilon > 0$,*
$$\mathbb{P}\{X_1 + \cdots + X_n \geq \epsilon n\} \leq [M(t)\, e^{-t\epsilon}]^n.$$
- *Show that for each $\epsilon > 0$, there is a $t > 0$ such that $M(t)\, e^{-\epsilon t} < 1$. Conclude the following: for every $\epsilon > 0$, there is a $\rho = \rho(\epsilon) < 1$ such that for all n*
$$\mathbb{P}\{|X_1 + \cdots + X_n| \geq \epsilon n\} \leq 2\, \rho^n.$$
- *Show that we can prove the last result with the boundedness assumption replaced by the following: there exists a $\delta > 0$ such that for all $|t| < \delta$, $\mathbb{E}[e^{tX_j}] < \infty$.*

Exercise 6.2. *Prove the following: there is a constant γ (called Euler's constant) and a $c < \infty$ such that for all positive integers n,*

$$\left| \left(\sum_{j=1}^{n} \frac{1}{j} \right) - \gamma - \log n \right| \leq \frac{c}{n}.$$

Hint: write

$$\log\left(n+\frac{1}{2}\right)-\log\left(\frac{1}{2}\right)=\int_{\frac{1}{2}}^{n+\frac{1}{2}}\frac{1}{x}\,dx,$$

and estimate

$$\left|\frac{1}{j}-\int_{j-\frac{1}{2}}^{j+\frac{1}{2}}\frac{dx}{x}\right|.$$

Exercise 6.3. *Show that there is a $c>0$ such that the following is true. For every real number r and every integer n,*

$$e^{-cr^2/n}\le e^r\left(1-\frac{r}{n}\right)^n\le e^{cr^2/n}. \tag{28}$$

Exercise 6.4. *Find constants a_1, a_2 such that the following is true as $n\to\infty$,*

$$\left(1-\frac{1}{n}\right)^n=e^{-1}\left[1+\frac{a_1}{n}+\frac{a_2}{n^2}+O\left(n^{-3}\right)\right].$$

Exercise 6.5. *Let S_n be a one-dimensional simple random walk and let*

$$p_n=\mathbb{P}\{S_{2n}=0\mid S_0=0\}.$$

- *Show that*

  $$p_{n+1}=p_n\,\frac{2n+1}{2n+2}, \tag{29}$$

  and hence

  $$p_n=\frac{1\cdot 3\cdot 5\cdots(2n-1)}{2\cdot 4\cdot 6\cdots(2n)}.$$

- *Use the relation (29) to give another proof that there is a c such that as $n\to\infty$*

  $$p_n\sim\frac{c}{\sqrt{n}}.$$

(Our work in this chapter shows in fact that $c=1/\sqrt{\pi}$, but you do not need to prove this here.)

Exercise 6.6.

- *Show that if X is a nonnegative random variable, then*

  $$\lim_{n\to\infty}\mathbb{E}[X\,1\{X\le n\}]=\lim_{n\to\infty}\mathbb{E}[X\wedge n]=\mathbb{E}[X].$$

- *(Monotone Convergence Theorem] Show that if $0\le X_1\le X_2\le\cdots$, then*

  $$\mathbb{E}\left[\lim_{n\to\infty}X_n\right]=\lim_{n\to\infty}\mathbb{E}[X_n].$$

In both parts, the limits and the expectations are allowed to take on the value infinity.

Exercise 6.7. *Suppose $X_1, X_2, \ldots$ are independent random variables each of whose distribution is symmetric about 0. Show that for every $a > 0$,*

$$\mathbb{P}\left\{\max_{1\leq j\leq n} X_1 + \cdots + X_j \geq a\right\} \leq 2\,\mathbb{P}\{X_n \geq a\}.$$

(Hint: Let K be the smallest j with $X_1 + \cdots + X_j \geq a$ and consider

$$\mathbb{P}\{X_1 + \cdots + X_n \geq a \mid K = j\}.\ )$$

Exercise 6.8. *Suppose X is a random variable taking values in $\mathbb{Z}$. Let*

$$\phi(t) = \mathbb{E}[e^{itX}] = \mathbb{E}[\cos(tX)] + i\,\mathbb{E}[\sin(tX)] = \sum_{x\in\mathbb{Z}} e^{itx}\,\mathbb{P}\{X = x\},$$

be its characteristic function. *Prove the following facts.*

- $\phi(0) = 1$, $|\phi(t)| \leq 1$ *for all t and* $\phi(t + 2\pi) = \phi(t)$.
- *If the distribution of X is symmetric about the origin, then $\phi(t) \in \mathbb{R}$ for all t.*
- *For all integers x,*
$$\mathbb{P}\{X = x\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi(t)\, e^{-ixt}\, dt.$$
- *Let k be the greatest common divisor of the set of all integers n with $\mathbb{P}\{|X| = n\} > 0$. Show that $\phi(t + (2\pi/k)) = \phi(t)$ and $|\phi(t)| < 1$ for $0 < t < (2\pi/k)$.*
- *Show that ϕ is a continuous (in fact, uniformly continuous) function of t. Hint: use*
$$|e^{i\theta} - e^{i\theta'}| \leq \min\{2, |\theta - \theta'|\}.$$

Exercise 6.9. *Suppose $X_1, X_2, \ldots$ are independent, identically distributed random variables taking values in the integers with characteristic function ϕ. Suppose that the distribution of X_j is symmetric about the origin,* $\mathrm{Var}[X_j] = \mathbb{E}[X_j^2] = \pi, \mathbb{E}[|X_j|^3] < \infty$. *Also assume,*

$$\mathbb{P}\{X_j = 0\} > 0, \quad \mathbb{P}\{X_j = 1\} > 0.$$

The goal of this exercise is to prove

$$\lim_{n\to\infty} \sqrt{2\pi\sigma^2 n}\,\mathbb{P}\{S_n = 0\} = 1.$$

Prove the following facts.

- *The characteristic function of $X_1 + \cdots + X_n$ is ϕ^n.*

- *For every* $0 < \epsilon \leq \pi$ *there is a* $\rho < 1$ *such that* $|\phi(t)| < \rho$ *for* $\epsilon \leq |t| \leq \pi$.
- $$\mathbb{P}\{S_n = 0\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi(t)^n \, dt = \frac{1}{2\pi\sqrt{n}} \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} \phi(t/\sqrt{n})^n \, dt.$$
- *There is a* c *such that for* $|t| \leq \pi$,
$$\left| \phi(t) - 1 - \frac{\sigma^2 \, t^2}{2} \right| \leq c \, t^3.$$
- $$\lim_{n \to \infty} \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} \phi(t/\sqrt{n})^n \, dt = \int_{-\infty}^{\infty} e^{-\sigma^2 t^2/2} \, dt = \frac{\sqrt{2\pi}}{\sigma}.$$

  Hint: you will probably want to use (28).

Exercise 6.10. *Suppose* A *is a bounded subset of* $\mathbb{Z}^d$ *and*

$$F : \partial A \to \mathbb{R}, \qquad g : A \to \mathbb{R}$$

are given functions. Show that there is a unique extension of F *to* $\overline{A}$ *such that*

$$\mathcal{L}F(x) = -g(x), \quad x \in A.$$

Give a formula for F.

Exercise 6.11. *Suppose* A *is a bounded subset of* $\mathbb{Z}^d$ *and*

$$F : \partial A \to \mathbb{R}, \qquad f : A \to \mathbb{R}$$

are given functions. Show that there is a unique function $p_n(x), n = 0, 1, 2 \ldots, x \in \overline{A}$ *satisfying the following:*

$$p_n(x) = F(x), \quad x \in \partial A,$$
$$\partial p_n(x) = \mathcal{L}F(x), \quad x \in A.$$

Show that $p(x) = \lim_{n\to\infty} p_n(x)$ *exists and describe the limit function* p.

Exercise 6.12. *Consider* $\beta(j, N), j = 1, \ldots, N-1$ *where* $\beta(j, N)$ *is the unique positive number satisfying*

$$\cosh\left(\frac{\beta(j,N)}{N}\right) + \cos\left(\frac{j\pi}{N}\right) = 2.$$

Prove the following estimates. The constants c_1, c_2, c_3 *are positive constants independent of* N *and the estimates should hold for all* N *and all* $j = 1, \ldots, N-1$.

- $$\beta(1, N) < \beta(2, N) < \cdots < \beta(N-1, N) \leq N \, \cosh^{-1}(2).$$

- *There is a* c_1 *such that*

$$\left|\cosh\left(\frac{j\pi}{N}\right)+\cos\left(\frac{j\pi}{N}\right)-2\right|\leq\frac{c_1\,j^4}{N^4}.$$

- *There is a* c_2 *such that*

$$|\beta(j,N)-\pi j|\leq\frac{c_2\,j^4}{N^3}.$$

- *There is a* c_3 *such that*

$$\beta(j,N)\geq c_3\,j.$$

Exercise 6.13. *Prove* (22) *and* (24).

Exercise 6.14. *Find the analogue of the formula* (21) *for the* d*-dimensional cube*

$$A=\{(x_1,\ldots,x_d)\in\mathbb{Z}^d : x_j=1,\ldots,N-1\}$$

Exercise 6.15. *Suppose* F *is a harmonic function on* $\mathbb{Z}^d$ *such that*

$$\lim_{|x|\to\infty}\frac{|F(x)|}{|x|}=0.$$

Show that F *is constant.*

Exercise 6.16. *The* relaxation method *for solving the Dirichlet problem is the following. Suppose* A *is a bounded subset of* $\mathbb{Z}^d$ *and* $F:\partial A\to\mathbb{R}$ *is a given function. Define the functions* $F_n(x), x\in\overline{A}$ *as follows.*

$$F_n(x)=F(x) \text{ for all } n \text{ if } x\in\partial A.$$

$$F_0(x), x\in A, \text{ is defined arbitrarily },$$

and for $n\geq 0$,

$$F_{n+1}(x)=\sum_{|x-y|=1}F_n(y),\qquad x\in A.$$

Show that for any choice of initial function F_0 *on* A,

$$\lim_{n\to\infty}F_n(x)=F(x),\qquad x\in A,$$

where F *is the solution to the Dirichlet problem with the given boundary value. (Hint: compare this to Exercise 6.11.)*

Exercise 6.17. *Let* S_n *denote a* d*-dimensional simple random walk and let* $R_n^1,\ldots,R_n^d$ *denote the number of steps taken in each of the* d*-directions. Show that for all* $n>0$*, the probability that* $R_{2n}^1,\ldots,R_{2n}^d$ *are all even is* $2^{-(d-1)}$.

Exercise 6.18. *Suppose that S_n is a biased one-dimensional random walker. To be more specific, let $p > 1/2$ and*

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

where $X_1, \ldots, X_n$ are independent with

$$\mathbb{P}\{X_j = 1\} = 1 - \mathbb{P}\{X_j = -1\} = p.$$

Show that there is a $\rho < 1$ such that as $n \to \infty$,

$$\mathbb{P}\{S_{2n} = 0\} \sim \rho^n \, \frac{1}{\sqrt{\pi n}}.$$

Find ρ explicitly. Use this to show that with probability one the random walk does not return to the origin infinitely often.

Exercise 6.19. *Suppose δ_n is a sequence of real numbers with $|\delta_n| < 1$ and such that*

$$\sum_{j=1}^{\infty} |\delta_n| < \infty.$$

Let

$$s_n = \prod_{j=1}^{n} (1 + \delta_j).$$

Show that the limit $s_\infty = \lim_{n\to\infty} s_n$ exists and is strictly positive. Moreover, there exists an N such that for all $n \geq N$,

$$\left|1 - \frac{s_n}{s_\infty}\right| \leq 2 \sum_{j=n+1}^{\infty} |\delta_j|.$$

Exercise 6.20. *Find the number t such that*

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

Exercise 6.21. *Prove that*

$$\int_{-\infty}^{\infty} e^{-x^2/2}\, dx = \sqrt{2\pi}.$$

(Hint: there are many ways to do this but direct antidifferentiation is not one of them. One approach is to consider the square of the left hand side, write it as a double (iterated) integral, and then use polar coordinates.)

Exercise 6.22. *Suppose S_n is a simple random walk in $\mathbb{Z}^d$ and $A \subset \mathbb{Z}^d$ is finite with N points. Let T_A be the smallest n such that $S_n \notin A$. Show that*

$$\mathbb{P}\{T_A > kN\} \leq \left(1 - \frac{1}{2d}\right)^k.$$

Exercise 6.23. *Finish the proof of Theorem 3.4 by doing the following.*

- *Use connectedness of A to show that any nonzero eigenfunction ϕ with every component nonnegative must actually have every component strictly positive.*
- *Give an example of a nonconnected A such that λ_1 has multiplicity greater than one.*
- *Given an example of a disconnected A such that λ_1 has multiplicity one. Does Theorem 3.4 hold in this case?*

Exercise 6.24. *Suppose A is a bounded subset of $\mathbb{Z}^d$. We call $\{x, y\}$ an* edge *of $\overline{A}$ if $x, y \in \overline{A}$, $|x - y| = 1$ and at least one of x, y is in A. If $F : \overline{A} \to \mathbb{R}$ is a function, we define its energy by*

$$\mathcal{E}(f) = \sum [f(x) - f(y)]^2,$$

where the sum is over the edges of $\overline{A}$. For any $F : \partial A \to \mathbb{R}$, define $\mathcal{E}(F)$ to be the infimum of $\mathcal{E}(f)$ where the infimum is over all f on $\overline{A}$ that agree with F on ∂A. Show that if f agrees with F on ∂A, then $\mathcal{E}(f) = \mathcal{E}(F)$ if and only if f is harmonic in A.

Exercise 6.25. *Verify* (27).

Exercise 6.26. *We will construct a "tree" each of whose vertices has three neighbors. We start by constructing $\mathcal{T}_1$ as follows: the vertices of $\mathcal{T}_1$ are the "empty word", denoted by o, and all finite sequences of the letters a, b, i.e., "words" $x_1 \ldots, x_n$ where $x_1, x_2, \ldots, x_n \in \{a, b\}$. Both words of one letter are adjacent to o. We say that a word of length $n-1$ and of length n are adjacent if they have the exact same letters, in order, in the first $n-1$ positions. Note that each word of positive length is adjacent to three words and the root is adjacent to only two words. We construct another tree $\mathcal{T}_2$ similarly, calling the root $\tilde{o}$ and using the letters $\tilde{a}, \tilde{b}$. Finally we make a tree $\mathcal{T}$ by taking the union of $\mathcal{T}_1$ and $\mathcal{T}_2$ and adding one more connection: we say that o and $\tilde{o}$ are adjacent.*

- *Convince yourself that $\mathcal{T}$ is a connected tree, i.e., between any two points of $\mathcal{T}$ there is a unique path in the tree that does not go through any point more than once.*
- *Let S_n denote simple random walk on the tree, i.e., the process that at each step chooses one of the three nearest neighbors at random, each with probability $1/3$, with the choice being independent of all the previous moves. Show that S_n is transient, i.e., with probability one S_n visits the origin only finitely often. (Hint: Exercise 6.18 could be helpful.)*
- *Show that with probability one the random walk does one of the two following things: either the random walk visits $\mathcal{T}_1$ only finitely oftern or it visits $\mathcal{T}_2$ only finitely often. Let $f(x)$ be the probability that the walk visits $\mathcal{T}_1$ finitely often. Show that f is a nonconstant bounded harmonic function. (A function f on $\mathcal{T}$ is harmonic if for every $x \in \mathcal{T}$ $f(x)$ equals the average value of f on the nearest neighbors of x.)*

- *Consider the space of bounded harmonic functions on $\mathcal{T}$. Show that this is an infinite dimensional vector space.*

Exercise 6.27. *Show that if $A \subset A_1$ are two subsets of $\mathbb{Z}^d$, then $\lambda_A \leq \lambda_{A_1}$. Show that if A_1 is connected and $A \neq A_1$, then $\lambda_A < \lambda_{A_1}$. Give an example with A_1 disconnected and A a strict subset of A_1 for which $\lambda_A = \lambda_{A_1}$.*