Chapter 4

Random Walk in $\mathbb{Z}^d$

4.1. Introduction

In this chapter we will focus on the integer lattice

$$\mathbb{Z}^d = \{(z_1, \ldots, z_d) : z_j \in \mathbb{Z}\}$$

viewed as an undirected graph where two vertices z, w are adjacent if they are nearest neighbors, that is, $|z - w| = 1$. Here and throughout this chapter we use $|\cdot|$ to denote the usual Euclidean distance. If $A \subset \mathbb{Z}^d$, we write

$$\partial A = \{z \in \mathbb{Z}^d \setminus A : |z - w| = 1 \text{ for some } w \in A\},$$

$$\overline{A} = A \cup \partial A.$$

We let $\mathcal{B}_n$ denote the discrete ball of radius n about the origin

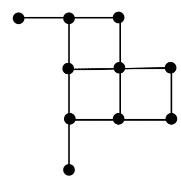
$$\mathcal{B}_n = \{z \in \mathbb{Z}^d : |z| < n\},$$

and note that for all $w \in \partial \mathcal{B}_n$,

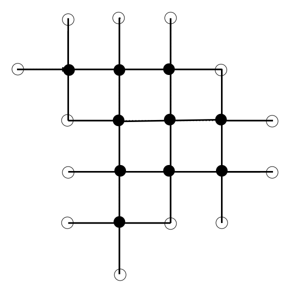
$$n \leq |w| < n + 1.$$

There are three natural "subgraphs" of $\mathbb{Z}^d$ associated to a subset A:

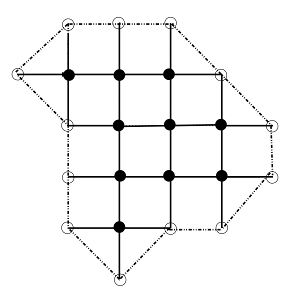
- **Free boundary:** The vertices are A, and the edges are the edges of $\mathbb{Z}^d$ with both endpoints in A.



- **Closure:** The vertices are $\overline{A}$, and the edges are the edges of $\mathbb{Z}^d$ with at least one endpoint in A.



- **Wired boundary:** The vertices are $A \cup \{\partial A\}$ where all the points on the boundary are identified ("wired") and considered as a single point. The edges are the same edges as in the closure but now there can be multiple edges from a vertex $z \in A$ to the boundary ∂A.



Simple random walk on $\mathbb{Z}^d$ starting at the origin can be written as

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

where $X_1, \ldots, X_n$ are independent random variables with distribution $\mathbb{P}\{X_j = w\} = 1/(2d)$ for all $|w| = 1$. We will write $p_n(z, w)$ for the corresponding n-step transition probabilities

$$p_n(z, w) = \mathbb{P}^z\{S_n = w\},$$

and $p_n(w) = p_n(0, w) = p_n(z, z + w)$. The transition probabilities are symmetric, $p_n(z, w) = p_n(w, z)$. We write $\mathcal{L}$ for the Laplacian

$$\mathcal{L}f(z) = (I - P)\, f(z) = f(z) - \frac{1}{2d} \sum_{|w-z|=1} f(w).$$

The transition probabilities $p_n(z)$ satisfy the "discrete heat equation"

$$p_{n+1}(z) = \frac{1}{2d} \sum_{|z-w|=1} p_n(w), \tag{4.1}$$

which can also be written as

$$p_{n+1}(z) - p_n(z) = -\mathcal{L}p_n(z),$$

where the Laplacian is with respect to the variable z. Simple random walk is a Markov chain with period 2. We can divide $\mathbb{Z}^d$ into the "even" points and the "odd" points where the even points are the $(z_1, \ldots, z_d)$ with $z_1 + \cdots + z_d$ even. If one starts at an even point, then after an odd number of steps one is at an odd point, and after an even number of steps one is at an even point.

There are other variations of simple random walk that get rid of this periodicity. Two standard ones are:

- *Lazy walker:* Let $0 < p < 1$. At each time step the walker chooses with probability p to not move. If the walker moves, then it chooses its new site as in the simple random walk.
- *Continuous time:* Let S_t be a continuous time walk that waits for an exponential amount of time and then takes a step. In this model the components of the walk are independent.

These models are the same if we view it only at the times the walker chooses a new site. There are advantages and disadvantages to each of these.

Our discrete heat equation is a discretization of the usual (continuous) heat equation

$$\partial_t \, p(t,x) = \frac{1}{2} \, \Delta_x \, p(t,x).$$

The latter describes the evolution of the probability density function of the continuous analogue of random walk which is called Brownian motion.

4.2. Local central limit theorem

Let S_n denote the position of a simple random walk starting at the origin in $\mathbb{Z}^d$. The central limit theorem states that the distribution of $n^{-1/2} \, S_n$ converges to a normal distribution; in this case, if U is an open ball in $\mathbb{R}^d$,

$$\lim_{n\to\infty} \mathbb{P}\left\{ \frac{S_n}{\sqrt{n}} \in U \right\} = \int_U \overline{p}(x) \, dx.$$

where

$$\begin{aligned} \overline{p}(x) &= \prod_{j=1}^{d} \left[\frac{1}{\sqrt{2\pi(1/d)}} \exp\left\{ -\frac{x_j^2}{2(1/d))} \right\} \right] \\ &= \left(\frac{d}{2\pi}\right)^{d/2} \exp\left\{ -\frac{d|x|^2}{2} \right\}. \end{aligned}$$

This should be familiar at least for $d = 1$. For general d, $\overline{p}(x)$ is the density of independent normal random variables with mean 0 and variance $1/d$. The variance is $1/d$ because that is the variance of one step for each component; for example, each step in the first component equals 1 with probability $1/2d$; -1 with probability $1/2d$; and 0 otherwise.

We define

$$\begin{aligned} \overline{p}_n(x) &= n^{-d/2} \, \overline{p}(x/\sqrt{n}) \\ &= \frac{1}{n^{d/2}} \left(\frac{d}{2\pi}\right)^{d/2} \exp\left\{ -\frac{d|x|^2}{2n} \right\}. \end{aligned} \tag{4.2}$$

Using the central limit theorem as a guide we might conjecture that if n and $x = (x_1, \ldots, x_d)$ have the same "parity", that is, if $n + x_1 + x_2 + \cdots + x_d$ is even, then $p_n(x) \sim 2\,\overline{p}_n(x)$. Statements of this kind are called *local central limit theorems (LCLT)*. theorems are stronger than the usual central limit theorem which is not sufficiently precise to estimate probabilities at points.

We will state one strong (although not the strongest) version of the LCLT for simple random walk. The basic idea and proof work for all d, but for ease we will discuss the full details of the proof only for $d = 1$. Let

$$\mathcal{T}_d = \underbrace{[-\pi,\pi] \times [-\pi,\pi] \times \cdots \times [-\pi,\pi]}_{d}.$$

If $\theta = (\theta_1, \ldots, \theta_d) \in \mathcal{T}_d$ we write $d\theta$ for $d\theta_1 \cdots d\theta_d$. If X is a random variable taking values in $\mathbb{Z}^d$, its *characteristic function* is is the function $\phi : \mathbb{R}^d \to \mathbb{C}$ defined by

$$\phi(\theta) = \phi_X(\theta) = \mathbb{E}[e^{i\theta\cdot X}] = \sum_{x \in \mathbb{Z}^d} e^{i\theta\cdot x}\, \mathbb{P}\{X = x\}.$$

Note that since X takes values in $\mathbb{Z}^d$ and $e^{2\pi i} = 1$, the function ϕ is periodic of period 2π in each variable. In other words, if $y \in \mathbb{Z}^d$, then $e^{i2\pi y\cdot X} = 1$ and hence

$$\phi(\theta) = \phi(\theta + 2\pi y).$$

The next proposition shows that we can give the distribution of S_n in terms of the characteristic function: the idea is a version of "Fourier inversion".

Proposition 4.1. *Suppose X is a random variable taking values in $\mathbb{Z}^d$ with characteristic function ϕ. Then,*

$$\mathbb{P}\{X = z\} = \frac{1}{(2\pi)^d} \int_{\mathcal{T}_d} e^{-iz\cdot\theta}\, \phi(\theta)\, d\theta.$$

Proof.

$$\begin{aligned}\int_{\mathcal{T}_d} e^{-iz\cdot\theta}\,\phi(\theta)\,d\theta &= \int_{\mathcal{T}_d} e^{-iz\cdot\theta}\left[\sum_{w\in\mathbb{Z}^d} e^{iw\cdot\theta}\,\mathbb{P}\{X=w\}\right]d\theta \\ &= \sum_{w\in\mathbb{Z}^d}\mathbb{P}\{X=w\}\int_{\mathcal{T}_d} e^{-iz\cdot\theta}e^{iw\cdot\theta}\,d\theta \\ &= (2\pi)^d\,\mathbb{P}\{X=w\}.\end{aligned}$$

The third equality uses the identity

$$\int_{\mathcal{T}_d} e^{iw\cdot\theta}\,d\theta = \begin{cases}(2\pi)^d, & w=0,\\ 0, & w\in\mathbb{Z}^d\setminus\{0\}\end{cases}.$$

The interchange of sum and integral in the second equality is valid since

$$\sum_{w\in\mathbb{Z}^d}\int_{\mathcal{T}_d}\left|e^{-iz\cdot\theta}e^{iw\cdot\theta}\,\mathbb{P}\{X=w\}\right|\,d\theta = \sum_{w\in\mathbb{Z}^d}\mathbb{P}\{X=w\}(2\pi)^d<\infty.$$

□

If $S_n = X_1+\cdots+X_n$ is simple random walk in $\mathbb{Z}^d$, then

$$\mathbb{E}[e^{i\theta\cdot X_1}] = \frac{1}{2d}\sum_{j=1}^d[e^{i\theta_j}+e^{-i\theta_j}] = \frac{1}{d}\sum_{j=1}^d\cos\theta_j.$$

$$\mathbb{E}[e^{i\theta\cdot S_n}] = \mathbb{E}\left[\prod_{k=1}^n e^{i\theta\cdot X_k}\right] = \prod_{k=1}^n\mathbb{E}\left[e^{i\theta\cdot X_k}\right] = \left[\frac{1}{d}\sum_{j=1}^d\cos\theta_j\right]^n.$$

The second equality uses the independence of $X_1,\ldots,X_n$. Combining this with the last proposition, we get an exact expression for the distribution of S_n.

Proposition 4.2. *The n-step transition probabilities are given by*

$$p_n(z) = \frac{1}{(2\pi)^d}\int_{\mathcal{T}_d} e^{-iz\cdot\theta}\,\phi(\theta)^n\,d\theta \tag{4.3}$$

where

$$\phi(\theta) = \frac{1}{d}\sum_{j=1}^d\cos\theta_j.$$

While (4.3) is an exact expression, the integrand is highly oscillatory for large $|z|$ which means that there is a lot of cancellation. Hence it takes work to estimate the integral.

Theorem 4.3 (Local Central Limit Theorem(LCLT))**.** *For every integer $d \geq 1$, there exists $c < \infty$ such that for every positive integer n and $x = (x_1, \ldots, x_d) \in \mathbb{Z}^d$ with $n + x_1 + \cdots + x_d$ even,*

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

Here $\overline{p}_n(x)$ is as in (4.2)*.*

Remark 4.4. For a "typical" x with $|x| \leq \sqrt{n}$, $\overline{p}_n(x)$ is of order $n^{-d/2}$ and hence we can write

$$p_n(x) = 2\,\overline{p}_n(x)\,[1 + O(n^{-1})].$$

However, if $|x| \gg \sqrt{n}$, then $\overline{p}_n(x)$ is of smaller order and the error $n^{-(\frac{d}{2}+1)}$ can be larger than the dominant term. In this case, while the theorem is valid, it is not very useful. There are other versions of the LCLT that give better estimates for these atypical values of x, but we will not discuss them.

The proof of Theorem 4.3 is similar in all dimensions and involves estimating the integral in (4.3). We will only discuss it in the case $d = 1$ and n, x are even for which

$$\overline{p}_n(x) = \frac{1}{\sqrt{2\pi n}}\, e^{-x^2/2n},$$

and hence (4.3) gives

$$\mathbb{P}\{S_n = x\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ix\theta}\, \cos^n\theta\, d\theta.$$

Since n and x are both even integers, the function $v(\theta) = e^{-ix\theta}\, \cos^n\theta$ has period π and hence the integral from $-\pi$ to π is the same as twice the integral from $-\pi/2$ to $\pi/2$,

$$p_n(x) = \frac{2}{2\pi} \int_{-\pi/2}^{\pi/2} e^{-ix\theta}\, \cos^n\theta\, d\theta.$$

Let us consider this integral. We know that we expect the left-hand side to be of order $n^{-1/2}$ at least if x is not too far away from the origin. We also know that $\cos\theta$ goes from 1 to 0 as $|\theta|$ goes from 0 to

$\pi/2$. Unless $\cos\theta$ is very near one, $\cos^n\theta$ will be very small for large n. To make this observation precise, we will use the Taylor polynomial approximation of $\cos y$. By Taylor's theorem with remainder we know that there exist $C < \infty$ such that

$$\left|\cos y - 1 + \frac{y^2}{2}\right| \leq C\, y^4, \quad |y| \leq \pi/2. \tag{4.4}$$

Indeed, we could give an explicit C but we will not need it. We are letting n go to infinity, so we only need consider n sufficiently large that $C \leq \sqrt{n}/4$. We claim that

$$p_n(x) + o(n^{-3/2}) = \frac{1}{\pi}\int_{-n^{-1/4}}^{n^{-1/4}} e^{-ix\theta}\, \cos^n\theta\, d\theta. \tag{4.5}$$

To see this, we use (4.4) to see that

$$\cos n^{-1/4} \leq 1 - \frac{(n^{-1/4})^2}{2} + C\,(n^{-1/4})^2\,(n^{-1/4})^2 \leq 1 - \frac{1}{4\sqrt{n}},$$

and hence

$$\begin{aligned}\left|\int_{n^{1/4}\leq|y|\leq\pi/2} e^{-ixy}\cos^n y\, dy\right| &\leq 2\int_{n^{1/4}}^{\pi/2}\cos^n y\, dy \\ &\leq \pi\left[1 - \frac{1}{4\sqrt{n}}\right]^n \\ &\leq \pi\, e^{-n^{1/2}/4}\end{aligned}$$

The first inequality is immediate since $|e^{-ixy}| = 1$. Note that $e^{-n^{1/2}/4} = o(n^{-3/2})$.

If we do the change of variables $\theta = s/\sqrt{n}$, the right-hand side of (4.5) becomes

$$\frac{2}{\sqrt{2\pi n}}\, I \quad \text{where} \quad I = \frac{1}{\sqrt{2\pi}}\int_{-n^{1/4}}^{n^{1/4}} e^{-ixs/\sqrt{n}}\, \cos^n(s/\sqrt{n})\, ds.$$

Note that $I = I_1 - I_2 + I_3$ where

$$I_1 = \int_{-\infty}^{\infty} e^{-ixs/\sqrt{n}}\, \frac{1}{\sqrt{2\pi}} e^{-s^2/2}\, ds.$$

$$I_2 = \int_{|s|\geq n^{1/4}} e^{-ixs/\sqrt{n}}\, \frac{1}{\sqrt{2\pi}} e^{-s^2/2}\, ds.$$

$$I_3 = \frac{1}{\sqrt{2\pi}} \int_{-n^{1/4}}^{n^{1/4}} e^{-ixs/\sqrt{n}} \, [\cos^n(s/\sqrt{n}) - e^{-s^2/2}] \, ds.$$

The integral I_1 is the characteristic function of a standard normal random variable evaluated at $-x/\sqrt{n}$; one can compute this or look it up to see that $I_1 = e^{-x^2/2n}$. Using $|e^{-iy}| = 1$, we see that

$$|I_2| \leq \int_{|s| \geq n^{1/4}} \frac{1}{\sqrt{2\pi}} e^{-s^2/2} \, ds \leq O(e^{-\sqrt{n}/2}) = o(n^{-1}).$$

Similarly,

$$\sqrt{2\pi} \, I_3 \leq \int_{-n^{1/4}}^{n^{1/4}} \left| \cos^n(s/\sqrt{n}) - e^{-s^2/2} \right| \, ds.$$

Using the expansion for the cosine (details omitted) we see that

$$\left| \cos^n(s/\sqrt{n}) - e^{-s^2/2} \right| \leq c \, e^{-s^2/2} \, \frac{s^4}{n}.$$

Hence,

$$I_3 \leq \frac{c}{n} \int_{-\infty}^{\infty} s^4 \, e^{-s^2/2} \, ds = O(1/n).$$

The error term I_3 is the largest of the error terms and indeed can be as large as c/n.

There is another approach to the LCLT in one dimension that uses another exact expression

$$\mathbb{P}\{S_{2n} = x\} = \binom{2n}{n+x} 2^{-2n}$$

and then uses Stirling formula (with error terms) to evaluate the right-hand side. This is easier than our proof, if one knows Stirling's formula, but the proof we give is easier to adapt to higher dimensions and also can be used for random walks other than the simple walk.

The LCLT implies that

$$p_{2n}(0) = \frac{C_d}{n^{d/2}} + O(n^{-\frac{d}{2}-1}).$$

where $C_d = 2^{1-d}\,(d/\pi)^{d/2}$. In particular,

$$\sum_{n=0}^{\infty} p_n(x) \left\{ \begin{array}{ll} < \infty & \text{if } d \geq 3 \\ = \infty & \text{if } d \leq 2 \end{array} \right. .$$

the expected number of returns to the origin is infinite if $d \leq 2$ and finite for $d \geq 3$.

Theorem 4.5 (Pólya)**.** *With probability one, simple random walk walk in $\mathbb{Z}^1$ and $\mathbb{Z}^2$ is recurrent. If $d \geq 3$, the random walk is transient.*

Exercise 4.6. Use Proposition 1.2 to prove this theorem.

4.3. Green's function

If $d \geq 3$, simple random walk is transient, and the *(whole space) Green's function*

$$G(z,w) = \sum_{n=0}^{\infty} \mathbb{P}^z\{S_n = w\} = \sum_{n=0}^{\infty} p_n(w-z),$$

is well defined. Note that $G(z,w) = G(w,z) = G(w-z)$ where we write $G(z)$ for $G(0,z)$. Analysts think of the Green's function as the "fundamental solution of the Laplacian". The discrete analogue of this viewpoint is the statement in the next proposition. We write $\delta(z)$ for the delta-function in $\mathbb{Z}^d$ defined by

$$\delta(z) = \left\{ \begin{array}{ll} 1 & z = 0 \\ 0 & z \neq 0. \end{array} \right.$$

Proposition 4.7. *The Green's function G satisfies*

$$\mathcal{L}G(x) = \delta(z).$$

Proof. Using (4.1), we can see that

$$\begin{aligned} G(z) = \sum_{n=0}^{\infty} p_n(z) &= \delta(z) + \sum_{n=1}^{\infty} p_n(z) \\ &= \delta(z) + \sum_{n=1}^{\infty} \frac{1}{2d} \sum_{|w-z|=1} p_{n-1}(w) \\ &= \delta(z) + \frac{1}{2d} \sum_{|w-z|=1} \sum_{n=1}^{\infty} p_{n-1}(w) \\ &= \delta(z) + \frac{1}{2d} \sum_{|w-z|=1} G(w) \\ &= \delta(z) + G(z) - \mathcal{L}G(z). \end{aligned}$$

□

We will give the asymptotics of the Green's function as $|x| \to \infty$. This can be deduced from local central limit theorems although we would need a stronger version than we have proved here. For this reason, we will not give a complete proof of the asymptotics, but we will show how the leading order term arises from a computation using the LCLT. For ease let us assume that $x \in \mathbb{Z}^d \setminus \{0\}$ and that the sum of the components of x is even. We start with

$$\begin{aligned} G(x) = \sum_{n=1}^{\infty} p_{2n}(x) \sim \sum_{n=1}^{\infty} 2\,\overline{p}_{2n}(x) \\ \sim \sum_{n=1}^{\infty} \overline{p}_n(x) = \frac{(d/2\pi)^{d/2}}{n^{d/2}} \sum_{n=1}^{\infty} e^{-y/n} \end{aligned}$$

where $y = d\,|x|^2/2$. We write the right-hand side as

$$\frac{d}{2}\,\frac{|x|^{2-d}}{\pi^{d/2}} \left[\frac{1}{y} \sum_{n=1}^{\infty} (n/y)^{-d/2}\, e^{-y/n}\right]. \tag{4.6}$$

We write it this way because the quantity in the brackets is a Riemann sum approximation using intervals of length y^{-1} of the integral

$$\int_0^{\infty} t^{-d/2}\, e^{-1/t}\, dt.$$

To compute the integral we use the substitution

$$t = 1/s, \quad dt = -s^{-2}\, ds$$

to make it

$$\int_0^\infty s^{\frac{d}{2}}\, e^{-s}\, s^{-2}\, ds = \Gamma\left(\frac{d}{2} - 1\right) = \Gamma(d/2)\, \frac{2}{d-2},$$

where

$$\Gamma(r) = \int_0^\infty z^{r-1}\, e^{-z}\, dz$$

is the Gamma function which satisfies $r\Gamma(r) = \Gamma(r+1)$. Combining with (4.6) we see that as $|x| \to \infty$,

$$G(x) \sim \frac{d\,\Gamma(d/2)}{(d-2)\,\pi^{d/2}}\, |x|^{2-d}.$$

By more careful analysis which we omit one can give a sharp bound on the error in the above asymptotics.

Proposition 4.8. *If $d \geq 3$, then as $|x| \to \infty$*

$$G(x) \sim \beta_d\, |x|^{2-d}, \qquad \textit{where } \beta_d = \frac{d\,\Gamma(d/2)}{(d-2)\,\pi^{d/2}}.$$

In fact,

$$G(x) = \beta_d\, |x|^{2-d} + O(|x|^{-d}). \tag{4.7}$$

The statement of this proposition uses a convenient shorthand. The conclusion can be written more precisely as: there exists $c < \infty$ such that for all x,

$$\left|\, G(x) - \beta_d\, |x|^{2-d} \,\right| \leq \frac{c}{|x|^d}.$$

Writing things like this is a bit bulky so we will use the $O(\cdot)$ and $o(\cdot)$ notation. It is important to remember that there is an implicit constant in this notation and that this constant is uniform over all $x \in \mathbb{Z}^d$.

It is useful to know what is worth memorizing and what is not so critical. In the case of the last proposition, the exponent $2 - d$ is worth committing to memory but not the value of the constant β_d. The function $f(x) = |x|^{2-d}$ is harmonic in $\mathbb{R}^d \setminus \{0\}$ and is the fundamental solution of the continuous Laplacian.

One way to remember the exponent is to use the following heuristic derivation. If $S_n = x$ then we would expect n to be of order $|x|^2$. So there are about $|x|^2$ possible times that contribute to the sum. For each of these values, the probability of being at x is of order $|x|^{-d}$. Therefore the Green's function is of order $|x|^2\,|x|^{-d}$.

The Green's function as defined above does not exist if $d = 2$ because the simple random walk is recurrent. However, there is another quantity that has many of the same properties, the potential kernel. Some authors refer to this as the Green's function.

Definition 4.9. If $d = 2$ the *potential kernel* is defined by

$$a(x) = \lim_{n\to\infty} \left[\sum_{j=0}^{n} p_j(0) - \sum_{j=0}^{n} p_j(x)\right].$$

One has to be careful with this definition. We cannot naively write the limit as

$$\sum_{j=0}^{\infty} p_n(0) - \sum_{j=0}^{\infty} p_n(x), \tag{4.8}$$

since both of these sums are infinite.

Let us show why the limit exists. We will do the case where $x = (x_1, x_2)$ with $x_1 + x_2$ even. We write

$$a(x) = \lim_{n\to\infty} \sum_{j=0}^{n} [p_{2j}(0) - p_{2j}(x)]. \tag{4.9}$$

Using the local central limit theorem. we can write

$$p_{2n}(0) - p_{2n}(x) = \frac{C_2}{n}\left[1 - e^{-|x|^2/n}\right] + O(n^{-2}).$$

If we fix x and let $n \to \infty$, we see that

$$1 - e^{-|x|^2/n} = \frac{|x|^2}{n} + O(|x|^4/n^2).$$

Hence there exists a constant c_x such that for all n,

$$|p_n(0) - p_n(x)| \le \frac{c_x}{n^2}.$$

This shows that the sum in (4.9) is absolutely convergent and we can write

$$a(x) = \sum_{j=0}^{\infty} [p_j(0) - p_j(x)].$$

If $x_1 + x_2$ is odd, we can similarly write

$$a(x) = \sum_{j=0}^{\infty} [p_j(0) - p_{j+1}(x)].$$

The next proposition shows that this is the fundamental solution of the Laplacian with $d = 2$ although we get a change in sign.

Proposition 4.10. *If $d = 2$, $a(0) = 0$, and $a(y) = 1$ if $|y| = 1$. Moreover, for all $x \in \mathbb{Z}^2$, $\mathcal{L}a(x) = -\delta(x)$.*

Exercise 4.11. Prove Proposition 4.10.

We could have defined a for $d \geq 3$ using the same definition, but in that case the naive expression (4.8) is valid and

$$a(x) = G(0) - G(x).$$

It is more convenient to use $G(x)$ rather than $a(x)$.

We now consider the asymptotics of the potential kernel in $\mathbb{Z}^2$ as $|x| \to \infty$. We will consider the case where $x_1 + x_2$ is even and let $y = |x|^2$ so that

$$p_{2n}(x) = \frac{1}{\pi n} e^{-y/n} + O(n^{-2}).$$

We will ignore the error term for the moment and consider

$$\sum_{n=0}^{\infty} \frac{1}{n} \left[1 - e^{-y/n}\right].$$

Note that

$$\sum_{n \geq y} \frac{1}{n} \left[1 - e^{-y/n}\right] \leq c \sum_{n \geq y} \frac{y}{n^2} \leq c_0,$$

$$\sum_{n \leq y} \frac{1}{n} e^{-y/n} = \sum_{n \leq y} \frac{1}{y} \frac{1}{n/y} e^{-y/n} \sim \int_0^1 \frac{e^{-1/t}}{t} \, dt \leq c_0.$$

Therefore,

$$a(x) = O(1) + \frac{1}{\pi}\sum_{n\leq y}\frac{1}{n} = \frac{1}{\pi}\log y + O(1) = \frac{2}{\pi}\log|x| + O(1).$$

The next proposition gives a more precise version. As in the case for the Green's function for $d \geq 3$, this can be proved from a sufficiently strong LCLT, but we will not prove it here.

Proposition 4.12. *If $d = 2$, as $|x| \to \infty$,*

$$a(x) = \frac{2}{\pi}\log|x| + k_0 + O(|x|^{-2}), \tag{4.10}$$

where

$$k_0 = \frac{1}{\pi}\,\log 8 + \frac{2}{\pi}\,\gamma$$

and γ is Euler's constant.

Euler's constant is defined by

$$\gamma = \lim_{n\to\infty}\left[-\log n + \sum_{j=1}^{n}\frac{1}{j}\right].$$

The actual value $\frac{1}{\pi}\,\log 8 + \frac{2}{\pi}\,\gamma$ is not so important but just the fact that there exists k_0 such that

$$a(x) = \frac{2}{\pi}\log|x| + k_0 + O(|x|^{-2}).$$

4.4. Harmonic functions

The study of simple random walk is very closely related to the study of harmonic functions on the lattice $\mathbb{Z}^d$. A good starting point for understanding harmonic function is the sharp estimates of the Green's function and potential kernel, (4.7) and (4.10). We will assume these even though we have not given complete proofs.

Suppose $A \subset \mathbb{Z}^d$ with $\partial A \neq \emptyset$. Let $T_A = \min\{n : S_n \notin A\}$. Recall that the Poisson kernel $H_A(z,w)$ for $z \in A, w \in \partial A$, is defined by

$$H_A(z,w) = \mathbb{P}^z\{S_{T_A} = w\}.$$

For fixed z, this gives a probability measure on ∂A provided that $\mathbb{P}^z\{T_A < \infty\} = 1$. This will always be true if $d \leq 2$ or if A is finite. There are examples with $d \geq 3$ such that $\mathbb{P}^z\{T_A < \infty\} < 1$, for example, if $\mathbb{Z}^d \setminus A$ is finite. The next proposition is a particular case of Proposition 1.10 so we do not need to prove it again.

Proposition 4.13. *Suppose $A \subset \mathbb{Z}^d$ such that for each $x \in A$, $\mathbb{P}^x\{T_A < \infty\} = 1$. Suppose $F : \partial A \to \mathbb{R}$ is a bounded function. Then there exists a unique bounded function $f : \overline{A} \to \mathbb{R}$ satisfying*

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

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

It is given by

$$f(x) = \mathbb{E}^x[F(S_{T_A})] = \sum_{y \in \partial A} F(y)\, H_A(x,y). \tag{4.11}$$

Exercise 4.14. Suppose $d \geq 3$ and $\mathbb{Z}^d \setminus A$ is finite. Show that (4.11) gives the unique function that is harmonic in A, equals zero on ∂A, and satisfies the extra condition

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

If $\mathbb{P}^x\{T_A < \infty\} < 1$ for some x we can get a similar result by adding the point "∞" to ∂A and setting

$$H_A(z, \infty) = \mathbb{P}^z\{T_A = \infty\}.$$

In this case we must also give the boundary value $F(\infty)$. See Exercise 4.31 for a proof of this.

Recall that $\mathcal{B}_n$ is the discrete ball of radius n about the origin and for $z \in \partial \mathcal{B}_n$, $n \leq |z| < n+1$. Propositions 4.7 and 4.10 imply that for all n and all $z \in \partial \mathcal{B}_n$,

$$G(z) = \beta_d\, n^{2-d} + O(n^{1-d}), \quad d \geq 3,$$

$$a(z) = \frac{2}{\pi} \log n + k_0 + O(n^{-1}), \quad d = 2.$$

Here we do not use the full force of the asymptotics of the Green's function. Although we know $G(z)$ up to an error of $|z|^{-d}$, there is an error of order n^{1-d} when we replace $|z|$ with n since

$$|z|^{2-d} = n^{2-d} + O(n^{1-d}), \quad z \in \partial\mathcal{B}_n,$$
$$\log|z| = \log n + O(n^{-1}), \quad z \in \partial\mathcal{B}_n.$$

The next proposition expresses the Green's function G_A on a finite set in terms of the whole space Green's function or the potential kernel.

Proposition 4.15. *Suppose $A \subset \mathbb{Z}^d$ is finite. Then for all $z, w \in A$,*

- *If $d \geq 3$,*

$$G_A(z,w) = G(z,w) - \sum_{y \in \partial A} H_A(z,y)\, G(y,w)$$
$$= G(w-z) - \sum_{y \in \partial A} H_A(z,y)\, G(w-y).$$

- *If $d = 2$,*

$$G_A(z,w) = -a(w-z) + \sum_{y \in \partial A} H_A(z,y)\, a(w-y).$$

Proof. Without loss of generality we may assume that $z = 0 \in A$ and let $T = T_A$. For $d \geq 3$, we write the total number of visits to w as

$$\sum_{j=0}^{\infty} 1\{S_j = w\} = \sum_{j=0}^{T-1} 1\{S_j = w\} + \sum_{j=T}^{\infty} 1\{S_j = w\}.$$

Taking expectations, we get

$$G(w) = G_A(0,w) + \sum_{y \in \partial A} H_A(0,y)\, G(y,w).$$

A similar proof can be given for $d = 2$, but it takes more work because of the recurrence of the random walk. We give a different proof. Without loss of generality assume that $w = 0$ and let

$$g(z) = -a(-z) + \sum_{y \in \partial A} H_A(z,y)\, a(-y).$$

Note that if $z \in \partial A$, then $g(z) = 0$. Also if $z \in A$,

$$\mathcal{L}g(z) = \delta(z).$$

The unique function satisfying this is $g(w) = G_A(z,0)$. $\square$

As a corollary, we estimate the expected number of returns to the origin before leaving the ball $\mathcal{B}_n$ by a random walker starting at the origin,

Proposition 4.16.

- *If* $d \geq 3$,
$$G_{\mathcal{B}_n}(0,0) = G(0) - O(n^{2-d}).$$
- *If* $d = 2$,
$$G_{\mathcal{B}_n}(0,0) = \frac{2}{\pi}\log n + k_0 + O(n^{-1})$$
where k_0 *is as in* (4.10).
- *If* $d = 2$ *and* $x \in \mathcal{B}_n$,

$$G_{\mathcal{B}_n}(x,0) = \frac{2}{\pi}\log\left(\frac{n}{|x|}\right) + O(n^{-1}) + O(|x|^{-2}). \tag{4.12}$$

Exercise 4.17. Use Proposition 4.15 to prove the last proposition.

Proposition 4.18. *If* $d \geq 3$ *and* $|x| \geq n$, *then the probability that a random walk starting at* x *ever enters* $\mathcal{B}_n$ *equals*

$$\left(\frac{n}{|x|}\right)^{d-2}\left[1 + O(n^{-1})\right].$$

Proof. Let $A = \mathbb{Z}^d \setminus \mathcal{B}_n$ and let $q = q(x,n)$ be this probability. Note that

$$q = \sum_{z \in \partial A} H_A(x,z).$$

The (whole space) Green's function $G(\cdot)$ is a bounded function that is harmonic in A and goes to zero as $x \to \infty$. Therefore (see Exercise 4.14),

$$G(x) = \sum_{z \in \partial A} H_A(x,z)\, G(z).$$

We know that $G(x) = \beta_d\,|x|^{2-d} + O(|x|^{-d})$, and since $|x| \geq n$,

$$G(x) = \beta_d\,|x|^{2-d}\left[1 + O(n^{-2})\right].$$

For $z \in \partial A$,

$$G(z) = \beta_d\, n^{2-d} + O(n^{1-d}) = \beta_d\, n^{2-d}\,[1 + O(n^{-1})],$$

and hence,

$$\sum_{z \in \partial A} H_A(x,z)\, G(z) = q\, \beta_d\, n^{2-d}\, [1 + O(n^{-1})].$$

Therefore,

$$q = \frac{|x|^{2-d}}{n^{2-d}}\, [1 + O(n^{-1})].$$

□

Proposition 4.19. *Suppose $d = 2$, and let $q(n,x)$ be the probability that a simple random walk starting at $x \in \mathbb{Z}^2$ leaves $\mathcal{B}_n$ before reaching the origin. Then for $|x| < n$,*

$$q(n,x) = \frac{a(x)}{\frac{2}{\pi}\log n + k_0 + O(n^{-1})}.$$

In particular,

$$\lim_{n \to \infty} \frac{2}{\pi} (\log n)\, q(n,x) = a(x). \tag{4.13}$$

Proof. Let $A = \mathcal{B}_n \setminus \{0\}$. The potential kernel is a harmonic function on A with $a(0) = 0$, and hence

$$a(x) = \sum_{z \in \partial A} a(z)\, H_A(x,z) = \sum_{z \in \partial \mathcal{B}_n} a(z)\, H_A(x,z).$$

For $z \in \mathcal{B}_n$,

$$a(z) = \frac{2}{\pi} \log n + k_0 + O(n^{-1}).$$

The probability that we want is

$$\sum_{z \in \partial \mathcal{B}_n} H_A(x,z) = \frac{a(x)}{\frac{2}{\pi}\log n + k_0 + O(n^{-1})}.$$

□

Exercise 4.20. Show that if $d = 2$ and $m < |x| < n$, then the probability that a random walk starting at x enters $\mathcal{B}_m$ before leaving $\mathcal{B}_n$ equals

$$\frac{\log n - \log |x| + O(|x|^{-1})}{\log n - \log m + O(m^{-1})}.$$

Hint: The potential kernel $a(\cdot)$ is a harmonic function in $\mathcal{B}_n \setminus \mathcal{B}_m$.

The next proposition gives a difference estimate for harmonic functions. Difference estimates are discrete analogues of estimates of derivatives. We will use the following estimates which follow immediately from Propositions 4.7 and 4.10: if $x, y \in \mathbb{Z}^d$, $|x-y| = 1$, then

$$|G(x) - G(y)| \le c\,|x|^{1-d}, \quad d \ge 3, \tag{4.14}$$

$$|a(x) - a(y)| \le c\,|x|^{-1}, \quad d = 2.$$

If f is a function on a countable set V we write

$$\|f\|_\infty = \sup\{|f(x)| : x \in V\}.$$

If V is finite, the supremum is the same as the maximum of $|f|$. We also write

$$\operatorname{dist}(x, \partial A) = \min_{y \in \partial A} |x - y|.$$

Proposition 4.21. *There exists $c = c(d) < \infty$ such that if $f : \overline{A} \to \mathbb{R}$ is harmonic in A and $x, y \in A$ with $|x-y| = 1$ and $\operatorname{dist}(x, \partial A) \ge n$, then*

$$|f(x) - f(y)| \le \frac{c}{n}\,\|f\|_\infty.$$

It is important to note the order of quantifiers in the proposition. There is a single constant c that works for all subsets $A \subset \mathbb{Z}^d$ and all harmonic function on A.

Proof. Without loss of generality we will assume that $x = 0$, and since $\mathcal{B}_n \subset A$, we can assume $A = \mathcal{B}_n$.

We will show that for every $|y| = 1$ and $z \in \partial A$,

$$H_A(y, z) = H_A(0, z)\,[1 + O(n^{-1})]. \tag{4.15}$$

We recall that this is shorthand for the statement that there exists c such that for every $n > 0$, every $z \in \partial A$, and every $|y| = 1$,

$$|H_A(0, z) - H_A(y, z)| \le \frac{c}{n}\, H_A(0, z).$$

To see that (4.15) suffices, we can use (4.11) to write

$$\begin{aligned} |f(0) - f(y)| &\le \sum_{z \in \partial A} |H_A(0,z) - H_A(y,z)|\,|f(z)| \\ &\le \frac{c}{n}\,\|f\|_\infty \sum_{z \in \partial A} H_A(0,z) = \frac{c}{n}\,\|f\|_\infty. \end{aligned}$$

Let us fix $z \in \partial A$ and write $h(x) = H_A(x, z)$. We will use a technique knows as a *last-exit decomposition*. Let $\tau = \tau_n = \min\{j : S_j \notin \mathcal{B}_n\}$, and let $V = \partial \mathcal{B}_{n/2}$. For $w \in V$, let $q(w)$ be the probability that a random walker starting at w does not return to V before leaving $\mathcal{B}_n$ and that it leaves $\mathcal{B}_n$ at z,

$$q(w) = q_{n,z}(w) = \mathbb{P}^w\{S_\tau = z; S_j \notin V \text{ for } j = 1, 2, \ldots, \tau\}.$$

Then we claim that for all $x \in \mathcal{B}_{n/2}$,

$$h(x) = \sum_{w \in V} G_A(x, w)\, q(w). \tag{4.16}$$

To see this we focus on the *last* visit to V by the random walker before leaving $\mathcal{B}_n$. Note that if we start in $\mathcal{B}_{n/2}$, we must visit V before leaving $\mathcal{B}_n$. Let ρ be the largest k with $S_k \in V$ and $k < \tau$. Then using the total law of probability,

$$h(x) = \sum_{k=1}^{\infty} \sum_{w \in V} \mathbb{P}^x\{S_\tau = z, \rho = k, S_k = w\}.$$

Note that

$$\mathbb{P}^x\{S_\tau = z, \rho = k, S_k = w\} =$$
$$\mathbb{P}^x\{S_k = w, k < \tau\}\, \mathbb{P}^x\{S_\tau = z, \rho = k \mid S_k = w, k < \tau\}.$$

Using the Markov property we can see that

$$\mathbb{P}^x\{S_\tau = z, \rho = k \mid S_k = w, k < \tau\} = q(w).$$

Therefore,

$$\begin{aligned} \mathbb{P}^x\{S_\tau = z\} &= \sum_{w \in V} q(w) \sum_{k=0}^{\infty} \mathbb{P}^x\{S_k = w, k < \tau\} \\ &= \sum_{w \in V} q(w)\, G_A(x, w). \end{aligned}$$

Our next step is to claim that for all $w \in V$, we have

$$G_A(0, w) = G_A(y, w)\,[1 + O(n^{-1})]. \tag{4.17}$$

We will show this in the case $d \geq 3$; the $d = 2$ case is done similarly. Proposition 4.15 gives

$$G_A(x, w) = G_A(w, x) = G(w, x) - \sum_{\zeta \in \partial \mathcal{B}_n} H_A(z, \zeta)\, G_A(\zeta, x).$$

Using this and (4.7), we see for $w \in V$ and $x \in \{0, y\}$

$$G_A(x, w) = [2^{d-2} - 1]\, \beta_d\, n^{2-d} + O(n^{1-d}).$$

Also (4.14) gives

$$|G(\zeta, 0) - G(\zeta, y)| \le c\, n^{1-d}, \quad |\zeta| \ge n/2.$$

Combining these two estimates gives (4.17).

Finally we can write

$$\begin{aligned} h(0) &= \sum_{w \in V} q(w)\, G_A(0, w) \\ &= \sum_{w \in V} q(w)\, G_A(y, w)\, [1 + O(n^{-1})] \\ &= h(y)\, [1 + O(n^{-1})]. \end{aligned}$$

□

Exercise 4.22. Suppose $f : \mathbb{Z}^d \to \mathbb{R}$ is harmonic.

(1) Show that if f is bounded then f is constant.

(2) More generally, show that if

$$\lim_{|x| \to \infty} \frac{|f(x)|}{|x|} = 0,$$

then f is constant.

For nonnegative functions we get another important result that says that values of positive functions in the interior are comparable. The key point in this lemma is that the constant C_r can be chosen so that the inequality holds for all n and all positive harmonic functions in $\mathcal{B}_n$.

Proposition 4.23 (Harnack inequality)**.** *For every $0 < r < 1$, there exists $C_r = C_r(d) < \infty$ such that if $f : \overline{\mathcal{B}}_n \to [0, \infty)$ is harmonic in $\mathcal{B}_n$, then for all $|x|, |y| < rn$, $f(x) \le C_r\, f(y)$.*

Proof. Let $c_r = c/(1 - r)$ where c is from the previous proposition. Then if $|x|, |y| < rn$ with $|x - y| = 1$, (4.15) gives

$$H_{\mathcal{B}_n}(x, w) \le H_{\mathcal{B}_n}(y, w)\, \left[1 + \frac{c_r}{n}\right].$$

Since f is nonnegative,

$$\begin{aligned} f(x) &= \sum_{z \in \partial A} H_{\mathcal{B}_n}(x,z)\, f(z) \\ &\leq \sum_{z \in \partial A} H_{\mathcal{B}_n}(y,z) \left[1 + \frac{c_r}{n}\right] f(z) \\ &= f(y) \left[1 + \frac{c_r}{n}\right]. \end{aligned}$$

If $|x|, |y| \leq rn$, then we can connect x to y by a path staying in $\mathcal{B}_{rn}$ of at most $2r\sqrt{d}n$ steps. Therefore, by repeated application of the above inequality we get

$$f(x) \leq \left[1 + \frac{c_r}{n}\right]^{2r\sqrt{d}n} f(y) \leq C_r\, f(y).$$

where $C_r = \exp\{2\sqrt{d}c_r r\}$. □

Proposition 4.24. *There exists $c < \infty$ such that if $f : \overline{A} \to [0, \infty)$ is harmonic in A then the following holds. Suppose ω is a nearest neighbor path from z to w in A of length k with $\mathrm{dist}(\omega, \partial A) \geq N$. Then,*

$$f(z) \leq f(w) \exp\{c\, k/N\}.$$

Exercise 4.25.

(1) Check that the proof of Proposition 4.22 extends to prove the last proposition.

(2) Use the last proposition to show the following. There exists $c < \infty$ such that if $A = \mathbb{Z}^d \setminus \mathcal{B}_n$ and $f : \overline{A} \to [0, \infty)$ is harmonic in A, then for all $z, w \in \partial \mathcal{B}_{2n}$,

$$f(z) \leq c\, f(w).$$

Exercise 4.26. Show that there exists $c < \infty$ such that the following is true for every $f : \overline{\mathcal{B}_n} \to \mathbb{R}$ that is harmonic in $\mathcal{B}_n$.

- For every $y \in \mathcal{B}_n$,

$$|f(y) - f(0)| \leq c\, \frac{|y|}{n}\, \|f\|_\infty.$$

- If $f \geq 0$ on $\mathcal{B}_n$, then for every $y \in \mathcal{B}_{n/2}$,

$$|f(y) - f(0)| \leq c\, \frac{|y|}{n}\, f(0). \tag{4.18}$$

Hint: The first is a consequence of the difference estimate and the second uses the Harnack inequality as well.

Exercise 4.27. Show that there exists $\alpha > 0$ and $c < \infty$ such that the following is true. Let $A = \mathbb{Z}^d \setminus \mathcal{B}_n$ and $z \in \partial A$. Then if $r \geq 2$ and $x, y \in \mathbb{Z}^d \setminus \mathcal{B}_{rn}$, then

$$|H_A(x,z) - H_A(y,z)| \leq \frac{c}{r^\alpha} H_A(x,z). \tag{4.19}$$

Hint:

(1) Let $V_k = \partial \mathcal{B}_{2^k n}$ for positive integers k. Explain why it suffices to prove (4.19) for $x, y \in V_k$ for all k.

(2) Let

$$\lambda_k = \max \left\{ \frac{|H_A(x,z) - H_A(y,z)|}{H_A(x,z)} : x, y \in V_k \right\}.$$

Show that there exists $\rho < 1$ (independent of z, n, k) such that if $k \geq 1$,

$$\lambda_{k+1} \leq \rho \, \lambda_k.$$

Hint: Use Exercise 4.25.

Exercise 4.28. Suppose $n \geq 2m$ and $A \subset \mathcal{B}_m$. Let τ_A and τ_n denote the first time that a random walk hits A and $\partial \mathcal{B}_n$, respectively. Let $z \in \partial \mathcal{B}_n$. If $x \in \partial \mathcal{B}_{2m}$, define $\epsilon_A(x,z)$ by

$$\mathbb{P}^x\{S(\tau_n) = z \mid \tau_n < \tau_A\} = \mathbb{P}\{S(\tau_n) = z\} \, [1 + \epsilon_A(x,z)] \, .$$

Show that there there exists $c = c(d) < \infty$ such that for all $n \geq 2m$,

$$|\epsilon_A(x,z)| \leq c \, \frac{m}{n}, \quad d \geq 3,$$

$$|\epsilon_A(x,z)| \leq c \, \frac{m}{n} \, \log \, \frac{n}{m}, \quad d = 2.$$

Hint: Use (4.18) to show that

$$\mathbb{P}^x\{S(\tau_n) = z \mid \tau_n > \tau_A\} = \mathbb{P}\{S(\tau_n) = z\} \left[1 + O\left(\frac{m}{n}\right) \right].$$

We will use our work so far to show the existence of harmonic measure from infinity. We start by giving the definition and then we will prove a proposition that shows that the definition is valid.

Definition 4.29. Suppose $A \subset \mathbb{Z}^d, d \geq 2$ is finite and let

$$T = T_A = \min\{j \geq 1 : S_j \in A\}.$$

Then the *harmonic measure (from infinity)* of A is defined by

$$\mathrm{hm}_A(x) = \lim_{w \to \infty} \mathbb{P}^w\{S_T = x \mid T < \infty\}$$

If $d = 2$, then $\mathbb{P}^w\{S_T = x\} = 1$ and we can write simply

$$\mathrm{hm}_A(x) = \lim_{w \to \infty} \mathbb{P}^w\{S_T = x\}.$$

We will now establish the existence of the limit. Before doing so, we note that the limit does not exist for $d = 1$. If we consider the set $A = \{0, 1\}$, then the probability that a random walk "from infinity" first visits A at 0 depends on whether the walker is coming from the right-hand side or the left-hand side. The proposition below shows that in more than one dimension, the hitting probability is the same (in the limit) regardless of the direction one is coming from.

Proposition 4.30. *If $A \subset \mathbb{Z}^d, d \geq 2$, is finite, then for each $x \in A$, the limit*

$$\mathrm{hm}_A(x) = \lim_{w \to \infty} \mathbb{P}^w\{S_T = x \mid T < \infty\}$$

exists and is also given by

$$\lim_{n \to \infty} \frac{\mathbb{P}^x\{\tau_n < T\}}{\sum_{y \in A} \mathbb{P}^y\{\tau_n < T\}},$$

where

$$\tau_n = \min\{j \geq 1 : S_j \in \partial \mathcal{B}_n\}.$$

Moreover, if $A \subset \mathcal{B}_m$ and $|w| \geq 2m$, then

$$\mathbb{P}^w\{S_T = x \mid T < \infty\} = \mathrm{hm}_A(x)\,[1 + O(\epsilon)],$$

where $\epsilon = \epsilon(m, w) = m/|w|$ for $d \geq 3$ and $\epsilon = (m/|w|)\log(|w|/m)$ for $d = 2$.

Proof. For $x \in A, z \in \partial \mathcal{B}_n$, let

$$r_n(x, z) = \mathbb{P}^x\{\tau_n < T, S(\tau_n) = z\}.$$

By reversing paths (check this!) we also see that

$$r_n(x, z) = \mathbb{P}^z\{\tau_n > T, S(T) = x\}.$$

By Exercise 4.28,

$$r_n(x,z) = \mathbb{P}^x\{\tau_n < T\}\,\mathbb{P}^0\{S(\tau_n) = z\}\,[1 + O(\epsilon_n)],$$

where $\epsilon_n = (m/n)$ if $d \geq 3$ and $\epsilon_n = (m/n)\,\log(n/m)$ if $d = 2$. We now use a last-exit decomposition for $|w| > n$ focusing on the last visit to ∂B_n before reaching the set A. More precisely, arguing as in (4.16), we get for $|w| > n$,

$$\begin{aligned} &\mathbb{P}^w\{S(T_A) = x\} \\ &\quad= \sum_{z \in \partial\mathcal{B}_n} G_{\mathbb{Z}^d \setminus A}(w,z)\, r_n(x,z) \\ &\quad= \sum_{z \in \partial\mathcal{B}_n} G_{\mathbb{Z}^d \setminus A}(w,z)\, \mathbb{P}^x\{\tau_n < T\}\,\mathbb{P}^0\{S(\tau_n) = z\}\,[1 + O(\epsilon_n)] \\ &\quad= J_n(w,A)\,\mathbb{P}^x\{\tau_n < T\}\,[1 + O(\epsilon_n)], \end{aligned}$$

where

$$J_n(w,A) = \sum_{z \in \partial\mathcal{B}_n} G_{\mathbb{Z}^d \setminus A}(w,z)\, H_{\mathcal{B}_n}(0,z).$$

The term $J_n(w,A)$ is independent of x, and hence if we set

$$h_n(x) = \frac{\mathbb{P}^x\{\tau_n < T\}}{\sum_{y \in A} \mathbb{P}^y\{\tau_n < T\}}, \tag{4.20}$$

then, we can write the above as

$$\mathbb{P}^w\{S(T_A) = x \mid T_A < \infty\} = h_n(x)\ [1 + O(\epsilon_n)]\,.$$

□

Exercise 4.31. Suppose $A \subset \mathbb{Z}^d$ $(d \geq 2)$ with $\mathbb{Z}^d \setminus A$ finite. Show that if $f : \overline{A} \to \mathbb{R}$ is bounded and harmonic on A, then the limit exists.

$$L = \lim_{z \to \infty} f(z)$$

Hint:

(1) It suffices to prove the result when $\|f\|_\infty = 1$.

(2) Let

$$\hat{f}(z) = \sum_{x \in \partial A} H_A(z,x)\, f(x).$$

Let $\hat{f}(\infty) = 0$ if $d \geq 3$ and for $d = 2$

$$\hat{f}(\infty) = \sum_{x \in \partial A} \mathrm{hm}_{\partial A}(x)\, f(x).$$

Show that

$$\lim_{z \to \infty} \hat{f}(z) = \hat{f}(\infty).$$

(3) Let $g = f - \hat{f}$ and note that this satisfies the hypotheses with $g \equiv 0$ on A.

(4) Use Exercise 4.28 to show that

$$g(z) = C\, H_A(z, \infty).$$

for some C.

4.5. Capacity for $d \geq 3$

If A is a finite subset of $\mathbb{Z}^d$ with $d \geq 3$, there are various ways to describe the size of A. One obvious way is the number of points, but this does not distinguish n points that are close together from n points spread apart. We will consider another notion called *capacity* which is related to the probability that a simple random walker ever visits the set. We will start with the definition and then we will give this interpretation. Let

$$T_A = \min\{j \geq 1 : S_j \in A\},$$

where we set $T_A = \infty$ if $S_j \notin A$ for all $j \geq 1$. Note that $T_A \geq 1$ even if we start in A since we are considering $j \geq 1$. We let

$$\mathrm{Es}_A(x) = \mathbb{P}^x\{T_A = \infty\}$$

and call $\mathrm{Es}_A(x)$ the *escape probability.*

Definition 4.32. If $d \geq 3$ and $A \subset \mathbb{Z}^d$ is finite, the *capacity* of A, is defined by

$$\mathrm{cap}(A) = \sum_{z \in A} \mathrm{Es}_A(z).$$

In this language we can write (4.20) for $d \geq 3$,

$$\mathrm{hm}_A(x) = \frac{\mathrm{Es}_A(x)}{\mathrm{cap}(A)}.$$

We recall that the Green's function satisfies

$$G(x) = \beta_d \, |x|^{2-d} + O(|x|^{-d}).$$

Exercise 4.33. Let $S_j, j \geq 0$ denote simple random walk in $\mathbb{Z}^d$, $d \geq 3$, starting at the origin, and let z be a nearest neighbor of the origin. Let p denote the probability that the random walk returns to the origin. We know that

$$G(0,0) = \frac{1}{1-p}.$$

(1) Show that the probability of ever visiting z is p.

(2) Let $T = \min\{j \geq 1 : S_j \in \{0, z\}\}$ with $T = \infty$ if no such j exists. Show that

$$\mathbb{P}\{S_T = 0\} = \mathbb{P}\{S_T = z\} = \frac{p}{1+p}.$$

(3) Show that

$$G_{\mathbb{Z}^d \setminus \{0\}}(z, z) = 1 + p,$$

and hence

$$F_{\{0,z\}}(\mathbb{Z}^d) = G(0,0) \, G_{Z^d \setminus \{0\}}(z, z) = \frac{1+p}{1-p}.$$

(4) Show that

$$\mathrm{cap}(\{0, z\}) = 2 \, \frac{1-p}{1+p}.$$

Proposition 4.34. *If $A \subset \mathbb{Z}^d, d \geq 3$ is a finite set, then*

$$\mathrm{cap}(A) = \lim_{x \to \infty} \beta_d^{-1} \, |x|^{d-2} \, \mathbb{P}^x\{T_A < \infty\}. \tag{4.21}$$

More precisely, if $A \subset \mathcal{B}_n$, and $|x| \geq 2n$,

$$\mathbb{P}^x\{T_A < \infty\} = \beta_d \, |x|^{2-d} \, \mathrm{cap}(A) \left[1 + O\left(\frac{n}{|x|}\right)\right].$$

Proof. We use a last-exit decomposition. Suppose we start a simple random walk at $x \notin A$ and let

$$\sigma = \max\{k < \infty : S_k \in A\}$$

with $\sigma = \infty$ if there is no such k. Then

$$\begin{aligned}
\mathbb{P}^x\{T_A < \infty\} &= \mathbb{P}^x\{\sigma < \infty\} \\
&= \sum_{k=1}^{\infty} \sum_{z \in A} \mathbb{P}^x\{\sigma = k, S_k = z\} \\
&= \sum_{k=1}^{\infty} \sum_{z \in A} \mathbb{P}^x\{S_k = z\}\, \mathbb{P}^x\{\sigma = k \mid S_k = z\} \\
&= \sum_{z \in A} \sum_{k=1}^{\infty} \mathbb{P}^x\{S_k = z\} \ \mathrm{Es}_A(z) \\
&= \sum_{z \in A} \mathrm{Es}_A(z)\, G(x,z) \\
&= \sum_{z \in A} \mathrm{Es}_A(z)\, \beta_d\, |x|^{d-2} \left[1 + O\left(\frac{n}{|x|}\right)\right] \\
&= \beta_d\, |x|^{d-2} \ \mathrm{cap}(A) \left[1 + O\left(\frac{n}{|x|}\right)\right].
\end{aligned}$$

□

Proposition 4.35. *If* $A, B \subset \mathbb{Z}^d, d \geq 3$, *are finite, then*

$$\mathrm{cap}(A \cup B) \leq \mathrm{cap}(A) + \mathrm{cap}(B) - \mathrm{cap}(A \cap B). \tag{4.22}$$

The inequality (4.22) is what characterizes capacities as opposed to other "measures" of size. Recall that probabilities, and more generally finite measures, satisfy

$$\mathbb{P}(E_1 \cup E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2) - \mathbb{P}(E_1 \cup E_2).$$

so the capacity condition is weaker than the probability (measure) condition.

Proof. Let $x \in \mathbb{Z}^d$, start a random walk at x, and let E_V denote the event that the random walk visits V. Note that $E_{A\cup B} = E_A \cup E_B$ and $E_{A\cap B} \subset E_A \cap E_B$. Since it is possible for the walker to visit both A and B without visiting $A \cap B$, it is not always true that $E_{A\cap B} = E_A \cap E_B$. Then,

$$\begin{aligned}
\mathbb{P}^x(E_{A\cup B}) &= \mathbb{P}^x(E_A) + \mathbb{P}^x(E_B) - \mathbb{P}^x(E_A \cap E_B) \\
&\leq \mathbb{P}^x(E_A) + \mathbb{P}^x(E_B) - \mathbb{P}^x(E_{A\cap B}).
\end{aligned}$$

If we multiply both sides by $\beta_d^{-1}\,|x|^{2-d}$ and take the limit using (4.21), we get the result. □

Proposition 4.36. *If* $A = \mathcal{B}_n$,

$$\operatorname{cap}(A) = \beta_d^{-1}\, n^{d-2} + O(n^{d-1}).$$

Proof. By Proposition 4.18,

$$\mathbb{P}^x\{T_A < \infty\} = \left(\frac{n}{|x|}\right)^{d-2} \left[1 + O\left(\frac{1}{n}\right)\right].$$

Therefore,

$$\begin{aligned}\operatorname{cap}(A) &= \lim_{x\to\infty} \beta_d^{-1}\,|x|^{d-2}\,\mathbb{P}^x\{T_A < \infty\} \\ &= \beta_d^{-1}\, n^{d-2}\left[1 + O\left(\frac{1}{n}\right)\right].\end{aligned}$$

□

Since a transient random walk visits each point only a finite number of times, it also visits every finite set only finitely often. What about infinite sets?

Exercise 4.37. Let $A \subset \mathbb{Z}^d$, and let

$$g(x) = g_A(x) = \mathbb{P}^x\{\text{random walk visits } A \text{ infinitely often}\}.$$

Show that $g \equiv 0$ or $g \equiv 1$.

Hint: Show that g is harmonic and use Exercise 4.22 to conclude that g is constant. Now assume that $g \equiv q \in (0,1)$ and derive a contradiction.

Definition 4.38. A subset $A \subset \mathbb{Z}^d$ is called *transient* if and only if with probability one the simple random walk visits A only finitely many times. Otherwise A is called *recurrent*.

We can construct infinite transient sets by spacing points far apart. Let $\{x_1, x_2, \ldots\}$ be a sequence of points with

$$\sum_{k=1}^{\infty} |x_k|^{2-d} < \infty.$$

Let S be a simple random walk starting at the origin and let

$$V = \sum_{n=0}^{\infty} 1\{S_n \in V\} = \sum_{n=0}^{\infty}\sum_{j=1}^{\infty} 1\{S_n = x_j\}$$

be the number of times that the random walk visits A. Then,

$$\mathbb{E}[V] = \sum_{j=1}^{\infty}\sum_{n=0}^{\infty} \mathbb{P}\{S_n = x_j\} = \sum_{j=1}^{\infty} G(x_j) < \infty.$$

Hence, $\mathbb{P}\{V < \infty\} = 1$.

One can ask the converse: if $\mathbb{E}[V] = \infty$ is it true that $\mathbb{P}\{V = \infty\} = 1$? The answer turns out to be no. Let us do a construction.

Set $b = 1 - \frac{1}{d}$ and let $A = \bigcup_{n=1}^{\infty} A_n$ where A_n is the discrete ball of radius $r_n = 2^{nb}$ centered at the point $z_n = (2^n, 0, 0, \ldots, 0)$. The number of elements of A_n is comparable to $r_n^d = 2^{n(d-1)}$ and

$$\sum_{x \in A} G(x) = \sum_{n=1}^{\infty} \sum_{x \in A_n} G(x) \asymp \sum_{n=1}^{\infty} 2^{n(2-d)}\, r_n^d \asymp \sum_{n=1}^{\infty} 2^n = \infty.$$

Also, Proposition 4.34 shows that $\mathrm{cap}(A_n) \asymp r_n^{d-2}$ and hence the probability of visiting A_n is comparable to

$$2^{-n(d-2)}\, r_n^{d-2} \asymp 2^{n(d-2)(b-1)} = 2^{-n(d-2)/d}.$$

Therefore,

$$\sum_{n=1}^{\infty} \mathbb{P}^x\{\text{random walk visits } A_n\} < \infty.$$

We used the word "comparable" and the notation $\asymp$ in the last paragraph. If a_n, b_n are two sequences of positive numbers, we say that a_n and b_n are *comparable*, written $a_n \asymp b_n$, if there exists $C < \infty$ such that for all n, $C^{-1}\, a_n \leq b_n \leq C\, a_n$.

We will give a criterion to determine whether or not a set is transient. We start with an exercise that we will use. We let $A \subset \mathbb{Z}^d, d \geq 3$ and $A_n = A \cap \{z : 2^n \leq |z| < 2^{n+1}\}$.

Exercise 4.39. Show that there exist $0 < c_1 < c_2 < \infty$ such that for every A and every n, if $|x| \leq 2^{n-1}$ or $2^{n+2} \leq |x| \leq 2^{n+3}$, then

$$c_1 \,\mathrm{cap}(A_n) \leq 2^{n(d-2)} \,\mathbb{P}^x\{T_{A_n} < \infty\} \leq c_2 \,\mathrm{cap}(A_n).$$

Proposition 4.40 (Wiener's Test)**.** *Let $A \subset \mathbb{Z}^d, d \geq 3$ and let*

$$A_n = A \cap \{z : 2^n \leq |z| < 2^{n+1}\}.$$

Then the set A is transient for random walk if and only if

$$\sum_{n=1}^{\infty} 2^{n(2-d)} \,\mathrm{cap}(A_n) < \infty. \tag{4.23}$$

Proof. Let $q_n = 2^{n(2-d)} \,\mathrm{cap}(A_n) \asymp \mathbb{P}\{T_{A_n} < \infty\}$, and let

$$Y = \sum_{n=1}^{\infty} 1\{T_{A_n} < \infty\}$$

denote the number of sets $A_1, A_2, \ldots$ that the random walk visits. The condition (4.23) is equivalent to the condition $\mathbb{E}[Y] < \infty$. If $\mathbb{E}[Y] < \infty$, then with probability one Y is finite and hence the walk is transient. This gives one direction.

To finish we need to show that if $\mathbb{E}[Y] = \infty$, then $\mathbb{P}\{Y = \infty\} = 1$. Using Exercise 4.37, it suffices to show that $\mathbb{P}\{Y = \infty\} > 0$. Assume that the sum in (4.23) is infinite. Then at least one of

$$\sum_{n=1}^{\infty} q_{2n}, \qquad \sum_{n=1}^{\infty} q_{2n-1}$$

is infinite. We will assume the first is infinite; essentially the same argument holds if the second sum is infinite. Let E_n be the event that the random walk visits A_{2n}. Then using the exercises immediately above, we get the relation

$$\mathbb{P}(E_n \cap E_m) \leq c \,\mathbb{P}(E_n)\,\mathbb{P}(E_m)$$

for some c. To see that this suffices, we use the second moment method, see Proposition A.4. □

Proposition 4.41. *Let S_n be a simple random walk in $\mathbb{Z}^d$ and let $A = \{S_j : j = 0, 1, 2, \ldots\}$ be the points visited by the path and let $\hat{A}$ be the set of points visited by the loop erasure of the path. Then with probability one,*

- *If $d \geq 5$, A and $\hat{A}$ are transient sets.*
- *If $d \leq 4$, A and $\hat{A}$ are recurrent sets.*

Exercise 4.42. Show that

$$\sum_{x \in \mathbb{Z}^d \setminus \{0\}} |x|^{-r} < \infty$$

if and only if $r > d$.

Proof. We will only prove the result for A; the results for $\tilde{A}$ is similar but requires some more work. Let

$$Y = \sum_{x \in A} G(x) = \sum_{x \in \mathbb{Z}^d} 1\{x \in A\} \, G(x).$$

which is now a random variable since A is a random set. Note that

$$\mathbb{E}[Y] = \sum_{x \in \mathbb{Z}^d} \mathbb{P}\{x \in A\} \, G(x) = \sum_{x \in \mathbb{Z}^d} \frac{G(x)^2}{G(0)}.$$

Since $G(x) \asymp |x|^{2-d}$, we have $G(x)^2 \asymp |x|^{4-2d}$. Therefore the sum converges if and only if $2d - 4 > d$, that is, $d > 4$. Therefore, if $d > 4$, $\mathbb{E}[Y] < \infty$ and hence with probability one $Y < \infty$. As we have seen, this implies that A is a transient set and since $\hat{A} \subset A$, $\hat{A}$ is also a transient set.

The case $d \leq 4$ takes more work; we will do only the case of A with $d = 4$. Let $S, \tilde{S}$ be two independent simple random walks starting at the origin and let

$$A = \{S_j : j = 0, 1, \ldots\}, \quad \tilde{A} = \{\tilde{S}_j : j = 0, 1, 2, \ldots\}.$$

We will show that with probability one, $A \cap \tilde{A}$ is an infinite set. Let

$$\sigma_n = \min\{j : S_j \geq 2^n\},$$

$$A_n = \mathcal{B}_{2^{n+1}} \cap \{S_j : \sigma_{n-2} \leq j \leq \sigma_{n+2}\}.$$

and let $\tilde{\sigma}_n, \tilde{A}_n$ be the analogous quantities using the walk $\tilde{S}$. We will show that following

- With probability one, there exist infinitely many n with $A_{4n} \cap \tilde{A}_{4n} \neq \emptyset$.

We will not give all the details but leave it as an exercise in the ideas of this chapter to put in all the details, However, we will give the sketch of facts to verify. Let E_n denote the event that $A_{4n} \cap \tilde{A}_{4n} \neq \emptyset$ and let $U_n = \mathcal{B}_{2^{4n+1}} \setminus \mathcal{B}_{2^{4n}}$.

- Show that there exists $c_1 > 0$ such that for all $x \in U_n$, $\mathbb{P}\{x \in A_n\} \geq c_1\, 2^{-8n}$.
- Show that there exists $c_2 < \infty$ such that for all $x, y \in U_n$ distinct,
$$\mathbb{P}\{x, y \in A_n\} \leq c_2\, 2^{-16n}\, |x-y|^{-4}.$$
- Show that if
$$Y_n = \sum_{x \in A_{4n} \cap \tilde{A}_{4n}} 1\{x \in A \cap \tilde{A}\}.$$
then there exist $c_3, c_4 > 0$ such that
$$\mathbb{E}[Y_n] \geq c_3, \quad \mathbb{E}[Y_n^2] \leq c_4\, n.$$
- Use the second moment method (see Section A.2) to conclude that
$$\mathbb{P}(E_n) = \mathbb{P}\{Y_n > 0\} \geq \frac{c_4}{c_3^2\, n}.$$
- Show that there exists c_6 such that for all $m < n$,
$$\mathbb{P}(E_m \cap E_n) \leq c_6\, \mathbb{P}(E_m)\, \mathbb{P}(E_n).$$
- Use the second moment method again to conclude that with probability one
$$\sum_{n=1}^{\infty} 1\{E_m\} = \infty.$$

Exercise 4.43. Put it all the details of the last proof!

Exercise 4.44. Show that if $d \geq 5$, there exists $c < \infty$ such that the following holds. Suppose S^1, S^2 are simple random walks starting at 0 and x respectively. Then,
$$\mathbb{P}\{S^1[0,\infty) \cap S^2[0,\infty) \neq \emptyset\} \leq c\, |x|^{4-d}.$$
Hint: Let I_y be the indicator function of the event that there exist j, k with $S_j^1 = y$ and $S_k^2 = y$. Let $I = \sum_{y \in \mathbb{Z}^d} I(y)$. Show that $\mathbb{E}[I] \leq c\, |x|^{4-d}$.

□

Many of the results about intersection of random walk paths are reflections of the fact that a random walk path in $\mathbb{Z}^d, d \geq 2$ is a "two-dimensional set". This is seen by noting that for large R, the number of points of the path that lie in the ball of radius R is of order R^2. Two two-dimensional sets (think, for example, planes) in $\mathbb{R}^d$ typically do not intersect if $d > 4$ and intersect if $d < 4$ with $d = 4$ being the critical dimension where it is a close call.

4.6. Capacity in two dimensions

There is a also a notion of capacity in two dimensions that relates to the probability of hitting a set, but we cannot use the same definition since every nonempty set is hit with probability one. Instead we will take a limit. If A is a finite set we write τ_A for the first time that we visit A and we write τ_n for $\tau_{\partial \mathcal{B}_n}$. We start with (4.13) which can be rewritten as

$$a(x) = \lim_{n\to\infty} \frac{2}{\pi}\,(\log n)\,\mathbb{P}^x\{\tau_n < \tau_0\}.$$

Proposition 4.45. *If $A \subset \mathbb{Z}^2$ is finite, then the limit*

$$a_A(x) = \lim_{n\to\infty} \frac{2}{\pi}\,(\log n)\,\mathbb{P}^x\{\tau_n < \tau_A\}.$$

exists for every $x \in \mathbb{Z}^2$. Moreover, $a_A \equiv 0$ on A and if $z \in A$, $x \in \mathbb{Z}^2 \setminus A$, then

$$a_A(x) = a(x-z) - \sum_{w\in A} \mathbb{P}^x\{S(T_A) = w\}\, a(w-z).$$

Proof. We will do the case $z = 0$ and leave the general case to the reader. Note that

$$\mathbb{P}^x\{\tau_n < \tau_0\} = \mathbb{P}^x\{\tau_n < \tau_A\} + \mathbb{P}^x\{\tau_A < \tau_n < \tau_0\}, \tag{4.24}$$

and

$$\mathbb{P}^x\{\tau_A < \tau_n < \tau_0\} = \sum_{w \in A} \mathbb{P}^x\{S(\tau_A \wedge \tau_n) = w\}\, \mathbb{P}^w\{\tau_n < \tau_0\}.$$

Using (4.13), we get

$$\lim_{n\to\infty} \frac{2}{\pi}\,(\log n)\,\mathbb{P}^x\{\tau_A < \tau_n < \tau_0\} = \sum_{w \in A} \mathbb{P}^x\{S(\tau_A) = w\}\, a(w).$$

Using this again in (4.24), we get

$$\lim_{n\to\infty} \frac{2}{\pi}\,(\log n)\,\mathbb{P}^x\{\tau_n < \tau_A\} = a(x) - \sum_{w \in A} \mathbb{P}^x\{S(\tau_A) = w\}\, a(w).$$

□

If $0 \in A$, we can write

$$a(x) - a_A(x) = \lim_{n\to\infty} \frac{2}{\pi}\,(\log n)\,\left[\mathbb{P}^x\{\tau_n < \tau_0\} - \mathbb{P}^x\{\tau_n < \tau_A\}\right].$$

Definition 4.46. If $x \in A \subset \mathbb{Z}^2$, then the *capacity* of A is defined by

$$\begin{aligned} \operatorname{cap}(A) &= \lim_{|z|\to\infty} [a(z) - a_A(z)] \\ &= \sum_{y \in A} \operatorname{hm}_A(y)\, a(y - x). \end{aligned}$$

The existence of this limit follows from Exercise 4.31. In some sense this is defined up to an additive constant and we have chosen the constant so that $\operatorname{cap}(\{0\}) = 0$. Another reasonable choice would be to choose $\operatorname{cap}(\{0\}) = -k_0$. We have the expansion

$$a_A(z) = \frac{2}{\pi}\,\log|z| + k_0 - \operatorname{cap}(A) + o(1), \quad z \to \infty.$$

Further reading

The classical book by Frank Spitzer [**18**] includes an extensive bibliography on the early work on random walk. This chapter can be considered as a sampler from [**10**] which is a serious graduate/research level treatment of simple random walk.