RANDOM WALKS IN $\mathbb{Z}^d$ AND THE DIRICHLET PROBLEM

ERIC GUAN

Abstract. Random walks can be used to solve the Dirichlet problem – the boundary value problem for harmonic functions. We begin by constructing the random walk in $\mathbb{Z}^d$ and showing some of its properties. Later, we introduce and examine harmonic functions in $\mathbb{Z}^d$ in order to set up the discrete Dirichlet problem. Finally, we solve the Dirichlet problem using random walks. Throughout the paper, we discuss connections with the continuous analog.

Contents

Introduction

Random walks are a fascinating probabilistic phenomenon. Despite their simple construction, random walks display a wide array of complex properties and have applications in fields ranging from chemistry to economics. One notable mathematical application is in solving the Dirichlet problem – the boundary value problem for harmonic functions. This paper focuses on random walks in $\mathbb{Z}^d$, but virtually every statement made is also true for the continuous analog in $\mathbb{R}^d$.

This paper follows content from the first chapter of Lawler [1], albeit at a slower pace and occasionally with different approaches. Readers should have a handle on basic probability, which will be briefly reviewed in the Background section.

1. Background

Before we can begin discussing random walks, we must first establish a footing in probability theory. First, a few basic definitions. (Further expositions and proofs of all following definitions and theorems can be found in most introductory probability books such as Ross [2], as well as from online resources.)

Definition 1.1. A *Probability Space* is a triple $\{\Omega, \mathcal{F}, \mathbb{P}\}$ consisting of a non-empty sample space Ω, a σ-algebra $\mathcal{F}$ on Ω, and a probability measure $\mathbb{P} : \mathcal{F} \to [0, 1]$.

Definition 1.2. A *random variable* X is a measurable function from a probability space $\{\Omega, \mathcal{F}, \mathbb{P}\}$ to a topological space (S, τ).

Date: September 4, 2011.

Random variables are generally not written explicitly as functions on the state space. For example, when using the probability measure, we write

$$\mathbb{P}(X^{-1}(B)) = \mathbb{P}\{X = B\}.$$

When we say "random variable X in $\mathbb{Z}$," we mean that the range of X is $\mathbb{Z}$.

Also important to the discussion of probability is the notion of independence – whether or not two events or random variables are linked in some way.

Definition 1.3. Events A and B are *independent* if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

Theorem 1.4. *Random variables X and Y in $\mathbb{R}$ are independent if for all x, y in $\mathbb{R}$, we have*

$$\mathbb{P}(\{X \geq x\} \cap \{Y \geq y\}) = \mathbb{P}\{X \geq x\}\mathbb{P}\{Y \geq y\}.$$

Finally, conditional probability and an important theorem that will be useful later.

Notation 1.5. The probability of event A, given event B, is written as $\mathbb{P}(A|B)$.

Theorem 1.6 (Conditional Probability)**.** *For events A and B, the conditional probability of A occurring given that B has occurred can be computed as*

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Theorem 1.7 (Law of Total Probability)**.** *For event A and a set of disjoint events $\{B_i\}_{i \in I}$ such that $\mathbb{P}(\bigcup_{i \in I} B_i) = 1$, the probability of A can be computed as*

$$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A|B_i)\mathbb{P}(B_i).$$

In the theorem above, "disjoint" means that no two events B_i and B_j, may simultaneously occur unless $i = j$. The criterion that the probability of the union of all of B_i equals 1 means that at least one event in the set occurs with probability 1.

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

The idea of a random walk is fairly intuitive; given some space, take a unit step in a random direction every unit time. Figures 1, 2, and 3 illustrate examples of random walks on the integer lattice beginning at the origin. Different colors represent different random walks, and the colors lighten over time. A more formal definition of the random walk is as follows.

Definition 2.1. A *random walk* in $\mathbb{Z}^d$ is a sequence of random variables S_n starting at $S_0 = x \in \mathbb{Z}^d$ with probability 1, and for each $n \in \mathbb{N}$ we define S_n to be

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

Here the X_j are independent and identically distributed (i.i.d.) random variables such that, for each time index j and dimension index k,

$$\mathbb{P}\{X_j = \mathbf{e}_k\} = \mathbb{P}\{X_j = -\mathbf{e}_k\} = \frac{1}{2d}$$

with $\mathbf{e}_k$ denoting the kth unit basis vector.

Unless otherwise stated, we will usually consider random walks starting at the origin – that is, $S_0 = 0$.

In this construction of the random walk, each individual step is represented by a random variable X_j. Because each step is i.i.d., each individual step "looks the same" independent of how many steps were taken before. This property is known as *time homogeneity*, and is defined as follows.

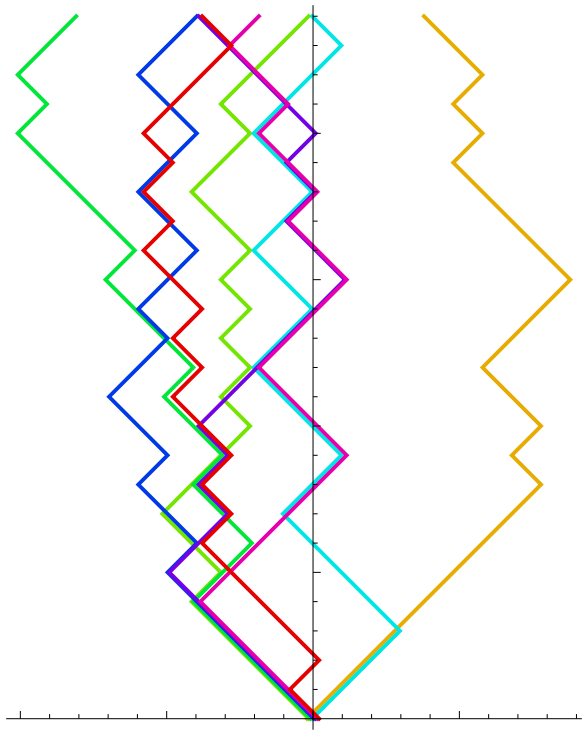

FIGURE 1. 8 Random Walks in $\mathbb{Z}$ of length 25, plotted over time (vertical axis).

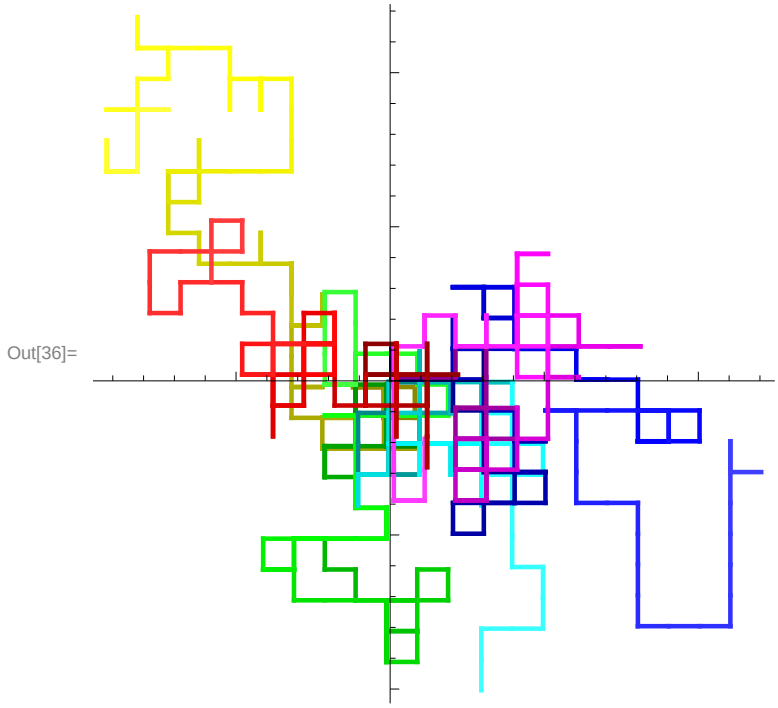

FIGURE 2. 6 Random Walks in $\mathbb{Z}^2$ of length 100. The walks lighten over time.

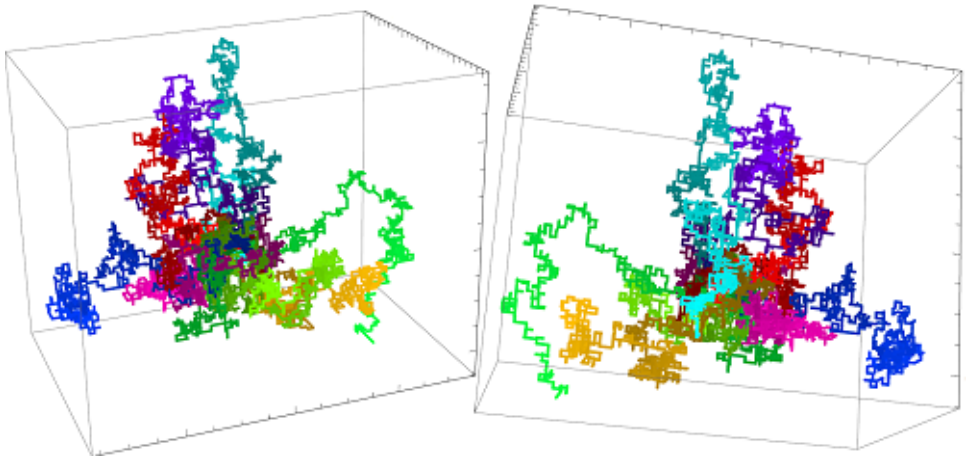

FIGURE 3. 8 Random Walks in $\mathbb{Z}^3$ of length 1000. The walks lighten over time. These two images are of the same walks from different perspectives.

Theorem 2.2 (Time Homogeneity)**.** *Let S_n be a random walk in $\mathbb{Z}^d$. For all $n, m, k \in \mathbb{N}$,*

$$\mathbb{P}\{S_{n+k} = y \mid S_n = x\} = \mathbb{P}\{S_{m+k} = y \mid S_m = x\}.$$

Proof. We can expand the expression to read

$$\mathbb{P}\{S_{n+k} = y | S_n = x\} = \mathbb{P}\left\{\sum_{j=1}^{n+k} X_j = y \mid \sum_{j=1}^{n} X_j = x\right\}$$

Expanding the RHS with Theorem 1.6, we obtain

$$\mathbb{P}\left\{\sum_{j=1}^{n+k} X_j = y \mid \sum_{j=1}^{n} X_j = x\right\} = \frac{\mathbb{P}\left\{(\sum_{j=1}^{n+k} X_j = y) \bigcap (\sum_{j=1}^{n} X_j = x)\right\}}{\mathbb{P}\left\{\sum_{j=1}^{n} X_j = x\right\}}$$

Note that, by manipulating the summation, we get

$$\left(\sum_{j=1}^{n+k} X_j = y\right) \bigcap \left(\sum_{j=1}^{n} X_j = x\right) = \left(\sum_{j=n+1}^{n+k} X_j = y - x\right) \bigcap \left(\sum_{j=1}^{n} X_j = x\right).$$

Since steps of the random walk X_i, X_j are independent if $i \neq j$, then we can simplify the expression using Theorem 1.4 to read

$$\frac{\mathbb{P}\left\{(\sum_{j=1}^{n+k} X_j = y) \bigcap (\sum_{j=1}^{n} X_j = x)\right\}}{\mathbb{P}\left\{\sum_{j=1}^{n} X_j = x\right\}} = \frac{\mathbb{P}\left\{\sum_{j=1}^{n+k} X_j = y - x)\right\} \mathbb{P}\left\{\sum_{j=1}^{n} X_j = x)\right\}}{\mathbb{P}\left\{\sum_{j=1}^{n} X_j = x\right\}}$$

$$\implies \mathbb{P}\left\{\sum_{j=1}^{n+k} X_j = y \mid \sum_{j=1}^{n} X_j = x\right\} = \mathbb{P}\left\{\sum_{j=1}^{n+k} X_j = y - x\right\} \tag{2.3}$$

Because each step X_j of the random walk is identically distributed regardless of j, we can rewrite the expression as

$$\mathbb{P}\left\{\sum_{j=n+1}^{n+k} X_j = y - x\right\} = \mathbb{P}\left\{\sum_{j=1}^{k} X_j = y - x\right\} = \mathbb{P}\{S_k = y - x \mid S_0 = 0\}.$$

Thus, the value of $\mathbb{P}\{S_{n+k} = y \mid S_n = x\}$ is independent of n. By symmetry, the same is true for m. Hence the two expressions are equal. $\square$

Notation 2.4. For the sake of brevity, we write $P_k(x, y) \equiv \mathbb{P}\{S_{n+k} = y \mid S_n = x\}$.

In addition to time independence, each step X_j is also independent of direction; the random walk moves forward and backward with equal probability. This symmetric property can be formally expressed as follows.

Theorem 2.5 (Symmetry)**.** *Let S_n be a random walk in $\mathbb{Z}^d$. For all $n, k \in \mathbb{N}$,*

$$\mathbb{P}\{S_{n+k} = y|S_n = x\} = \mathbb{P}\{S_{n+k} = x|S_n = y\}.$$

Proof. Using the notation from 2.4, we rewrite the statement to obtain

$$P_k(x, y) = P_k(y, x).$$

We recall that in the construction of the random walk, each component step is symmetric. That is to say, if $x = e_i$ for some i, then we get

$$\mathbb{P}\{X_j = x\} = \frac{1}{2d} = \mathbb{P}\{X_j = -x\}.$$

Also, if $x \neq e_i$ for some i then we get

$$\mathbb{P}\{X_j = x\} = 0 = \mathbb{P}\{X_j = -x\}.$$

More generally, for every x in $\mathbb{Z}^d$, we have

$$\mathbb{P}\{X_j = x\} = \mathbb{P}\{X_j = -x\} \tag{2.6}$$

Let us apply induction on the time shift k. First, examine the base case $k = 1$. Using (2.3), we obtain

$$P_1(x, y) = \mathbb{P}\{S_1 = y|S_0 = x\} = \mathbb{P}\{X_1 = y - x\}$$

Applying (2.6) gives us

$$P_1(x, y) = \mathbb{P}\{X_1 = y - x\} = \mathbb{P}\{X_1 = x - y\} = P_1(y, x).$$

So it is true when $k = 1$.

Now for the inductive step. Given the statement is true for k, let us show it for $k + 1$.

By the law of total probability (Theorem 1.7) over the entire state space $\mathbb{Z}^d$, we can expand the expression to read

$$P_{k+1}(x, y) = \mathbb{P}\{S_{n+k+1} = y \mid S_n = x\} = \sum_{z \in \mathbb{Z}^d} \mathbb{P}\{S_{n+k} = z \mid S_n = x\}\mathbb{P}\{S_{n+k+1} = y \mid S_{n+k} = z\}.$$

$$P_{k+1}(x, y) = \sum_{z \in \mathbb{Z}^d} P_k(x, z)P_1(z, y).$$

Applying the inductive hypothesis $P_k(x, z) = P_k(z, x)$ and the base case $P_1(z, y) = P_1(y, z)$ yields

$$P_{k+1}(x, y) = \sum_{z \in \mathbb{Z}^d} P_k(x, z)P_1(z, y) = \sum_{z \in \mathbb{Z}^d} P_k(z, x)P_1(y, z)$$

$$P_{k+1}(x, y) = \sum_{z \in \mathbb{Z}^d} \mathbb{P}\{S_{n+k+1} = x \mid S_{n+1} = z\}\mathbb{P}\{S_{n+1} = z \mid S_n = y\}.$$

Finally, using the law of total probability to collapse the sum, we get

$$P_{k+1}(x, y) = \mathbb{P}\{S_{n+k+1} = x \mid S_n = y\}.$$

By induction, the random walk is symmetric. $P_k(x, y) = P_k(y, x)$. $\square$

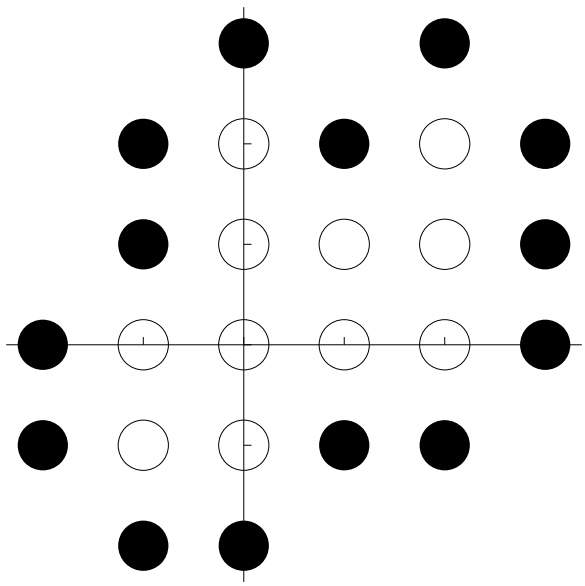

FIGURE 4. The boundary of a set in $\mathbb{Z}^2$.

We now see several nice properties of the random walk. Not only is it homogenous in time with self-similar branches, but it is also symmetric in space, with any particular movement and its inverse occurring with equal probability. These symmetries in the random walk line up with conventional notions of randomness as impartially distributed.

3. The Mean Value Property and Harmonic Functions

We shall now depart briefly from random walks to discuss functions on $\mathbb{Z}^d$ and the mean value property. To have a function over a region in $\mathbb{Z}^d$ is to assign a value to each point in the region. The significance of the value is arbitrary, but in the context of random walks, it is useful to think of the value as temperature or density. First, we must define regions and boundaries in $\mathbb{Z}^d$.

Definition 3.1. Let A be a subset of $\mathbb{Z}^d$. The *boundary* of A is the set

$$\{z \in \mathbb{Z}^d | z \notin A, \exists x \in A \ \ s.t. \ |z - x| = 1\}.$$

We denote it by ∂A. The *closure* $\bar{A}$ is $\bar{A} = A \cup \partial A$. The set A is also called the *interior* of $\bar{A}$.

These are just the discrete analogues of the standard metric definitions of boundary and closure. In $\mathbb{Z}^d$ we cannot appeal to arbitrarily small open balls, and thus the nearest-neighbor condition $|z - x| = 1$ is used instead. See Figure 4 for a basic example.

The mean value property can be understood as having the value at each point being the average of surrounding points in the region. More formally, for $A \subset \mathbb{R}^d$, the function $f : A \to \mathbb{R}$ satisfies the (continuous) mean value property if for suitable $x \in A$ and $r \in \mathbb{R}$,

$$f(x) = \int_{\partial B_r(x)} f(y) ds(y).$$

The continuous definition is complex, but constructing a discrete analog is fairly simple; take the average over all adjacent points rather than a sphere.

Definition 3.2. Let A be a subset of $\mathbb{Z}^d$. A function $F : \bar{A} \to \mathbb{R}$ has the *mean value property* if for all x in A,

$$F(x) = \sum_{\substack{y \in \mathbb{Z}^d \\ |x-y|=1}} \frac{1}{2d} F(y). \tag{3.3}$$

This discrete definition specifically averages nearest neighbors; it lacks the flexible radius found in the continuous definition. Expanding the radius of the discrete mean value property in $\mathbb{Z}$ is fairly simple, the proof of which is left to the reader. Similar expansions in higher dimensions also work, but we will not attempt to prove them because they are computationally tedious.

Example 3.4. Let $F : \mathbb{Z} \to \mathbb{R}$ be a function satisfying the mean value property. For all $x \in \mathbb{Z}$, $k \in \mathbb{N}$, $F(x) = \frac{1}{2}F(x-k) + \frac{1}{2}F(x+k)$

It turns out that the mean value property is closely related to harmonic functions. For continuous functions, a function f is harmonic if

$$\sum_{k=1}^{d} \frac{\partial^2 f}{\partial x_k^2} = \Delta f = 0,$$

where Δ is called the Laplacian.

We can construct an analog in $\mathbb{Z}^d$. Because $\mathbb{Z}$ lacks derivatives, we approximate

$$\frac{\partial F}{\partial x}(c) \approx F(c+1) - F(c).$$

Similarly, we can approximate the second derivative

$$\frac{\partial^2 F}{\partial x^2}(c) \approx \frac{\partial F}{\partial x}(c) - \frac{\partial F}{\partial x}(c-1) \approx F(c+1) - 2F(c) + F(c-1).$$

This motivates the following definition.

Definition 3.5. On $\mathbb{Z}^d$ the *(discrete) Laplacian*, $\mathcal{L}$, of a function $F : \mathbb{Z}^d \to \mathbb{R}$ is defined as follows.

$$\mathcal{L}F(x) = \sum_{\substack{y \in \mathbb{Z}^d \\ |x-y|=1}} \frac{1}{2d}(F(y) - F(x)).$$

Definition 3.6. Let A be a subset of $\mathbb{Z}^d$. We say a function $F : \bar{A} \to \mathbb{R}$ is *harmonic* if for all x in A,

$$\mathcal{L}F(x) = 0.$$

For $d = 1$, we obtain $\mathcal{L}F(c) = \frac{1}{2}(F(c+1) - 2F(c) + F(c-1))$, which is one-half of the formula we constructed before. Note that with harmonics, we are only concerned with $\mathcal{L}F = 0$, and scaling by a constant factor does not affect this equation.

The formula in 3.3 is remarkably similar to that of the Laplacian. This is no coincidence; in fact, a function in $\mathbb{R}^d$ is harmonic if and only if it satisfies the mean value property. From the continuous definitions of harmonic and mean value property, this is somewhat surprising. However, the proof linking harmonic functions and the mean value property in $\mathbb{Z}^d$ is straightforward.

Theorem 3.7. *A function satisfies the mean value property if and only if it is harmonic over the entire domain.*

Proof. First, we expand the Laplacian

$$\mathcal{L}F(x) = \sum_{\substack{y\in\mathbb{Z}^d \\ |x-y|=1}} \frac{1}{2d}(F(y) - F(x)) = \sum_{\substack{y\in\mathbb{Z}^d \\ |x-y|=1}} \frac{1}{2d}F(y) - \sum_{\substack{y\in\mathbb{Z}^d \\ |x-y|=1}} \frac{1}{2d}F(x).$$

Since in $A \subseteq \mathbb{Z}^d$ each point is distance one away from exactly $2d$ other points in $\bar{A}$, then for all x in A,

$$\#\{y \in \mathbb{Z}^d \mid |x-y| = 1\} = 2d.$$

Thus, we can simplify the sum

$$\sum_{\substack{y\in\mathbb{Z}^d \\ |x-y|=1}} \frac{1}{2d}F(y) - \sum_{\substack{y\in\mathbb{Z}^d \\ |x-y|=1}} \frac{1}{2d}F(x) = \sum_{\substack{y\in\mathbb{Z}^d \\ |x-y|=1}} \frac{1}{2d}F(y) - 2d\left(\frac{1}{2d}F(x)\right) = -F(x) + \sum_{\substack{y\in\mathbb{Z}^d \\ |x-y|=1}} \frac{1}{2d}F(y).$$

Now, if F is harmonic, then by definition, $\mathcal{L}F(x) = 0$. Thus,

$$-F(x) + \sum_{\substack{y\in\mathbb{Z}^d \\ |x-y|=1}} \frac{1}{2d}F(y) = 0$$

$$\implies \sum_{\substack{y\in\mathbb{Z}^d \\ |x-y|=1}} \frac{1}{2d}F(y) = F(x).$$

The above expression is exactly the mean value property. Since each step is reversible, then the converse follows. Therefore, satisfying the mean value property is equivalent to being harmonic. □

Functions with zero Laplacian are common in mathematics, and harmonic functions are often found in physical phenomena as well. The mean value property is a useful mathematical tool, but directly proving a function satisfies the property is difficult. The revelation that these two properties are the same allows for more versatile applications. From now on, we may use "harmonic" and "mean value property" interchangeably.

Harmonic functions in $\mathbb{R}^d$ have a variety of properties. One notable property is that harmonics cannot have extrema on the interior of their domains. Let us show this applies in $\mathbb{Z}^d$ as well.

Definition 3.8. The set $A \subset \mathbb{Z}^d$ is *connected* if for any two points x, y in A, there exists a finite sequence of points $x_1, \cdots, x_n$ such that $|x - x_1| = 1 = |x_n - y|$ and for all suitable i,

$$|x_i - x_{i+1}| = 1.$$

Theorem 3.9. *Let A be a connected subset of $\mathbb{Z}^d$. If a harmonic function $F : \bar{A} \to \mathbb{R}$ has a maximum or minimum, then F is constant or has its maximum and minimum on the boundary ∂A.*

Proof. There must be some $x_0 \in \bar{A}$ such that $F(x_0) = \max\{F(x) \mid x \in \bar{A}\}$. For the sake of contradiction, assume x_0 is in the interior A. Note that

$$F(x_0) = \sum_{\substack{y\in\bar{A} \\ |x_0-y|=1}} \frac{1}{2d}F(y).$$

Note that $F(y) \leq F(x_0)$ for all y, and that there are $2d$ such y in the summation. This implies equality.

$$F(y) = F(x_0) \quad \text{for every } y \in \bar{A},\ |x_0 - y| = 1.$$

Applying this reasoning repeatedly, we can extend this equality to the entire connected domain,

$$F(x_0) = F(y) \quad \text{for every } y \in \bar{A}.$$

Thus, if F has a maximum on the interior of the domain, then F is constant. Similarly, the same proof applies for minima. Disregarding the constant case, F cannot have any extrema on the interior of its domain. □

When we say 'connected,' we mean that between any two points in A, there is a nearest-neighbor path of points in A connecting them.

Corollary 3.10. *Let A be a connected, bounded subset of $\mathbb{Z}^d$. If a harmonic function $F : \bar{A} \to \mathbb{R}$ is constant on the boundary ∂A, then F is constant on the entire domain.*

Proof. Note that because $\mathbb{Z}^d$ is not dense, A must be finite.

Let $F(z) = c$ for all $z \in \partial A$. For the sake of contradiction, assume there exists an interior point $x \in A$ such that $F(x) \neq c$.

This implies $F(x) > c$ or $F(x) < c$. Either way, there exists an extremum not on the boundary, contradicting Theorem 3.9. Therefore F is constant. □

Later, we will find ourselves subtracting harmonic functions; the difference is also a harmonic function. The proof is fairly simple and is left as an excercise.

Example 3.11. Any linear combination of harmonic functions on the same domain is also harmonic.

Finally, an important property of harmonic functions.

Theorem 3.12. *Let A be a finite subset of $\mathbb{Z}^d$, and let $F : \bar{A} \to \mathbb{R}$ be a harmonic function. Then the interior values $\{F(x) \mid x \in A\}$ are uniquely determined by the boundary values $\{F(z) \mid z \in \partial A\}$.*

Proof. Suppose there exists another harmonic function $G : \bar{A} \to \mathbb{R}$ with the same boundary values as F. By Remark 3.11, the function $(F - G)$ is also harmonic.

Because F and G are equal on the boundary, $(F - G)$ must be 0 for all boundary points. Thus, by Corollary 3.10, $(F - G)$ is 0 for all interior points as well, so F and G are identical.

Therefore, F is unique. □

This result is somewhat surprising. Regardless of the size of the interior, the harmonic function on a region is uniquely determined by the boundary; solving for a harmonic function on the interior of its domain will always produce the unique solution. This allows us to solve harmonic functions by constructing an example and showing that it works, since uniqueness is no longer an issue.

4. The Dirichlet Problem

The Dirichlet problem is as follows: Given a bounded region and a harmonic function defined on the boundary, find an extension to the interior. Before we can solve this problem, we must first construct some tools.

Up to now, we have only discussed unbounded random walks. A random walk in a boundary continues until the walk hits the boundary. The time at which this happens is the stopping time, formally defined below. Figure 5 illustrates multiple random walks in a bounded domain with various stopping times.

Definition 4.1. Let S_n be a random walk in $A \in \mathbb{Z}^d$. The *stopping time* T is

$$T = \min\{k \in \mathbb{N} \mid S_k \in \partial A\}.$$

Because T depends on S_0 and A, we may write $T_A^{S_0}$ to denote the stopping time of the random walk starting at S_0 and in the domain A to avoid ambiguity.

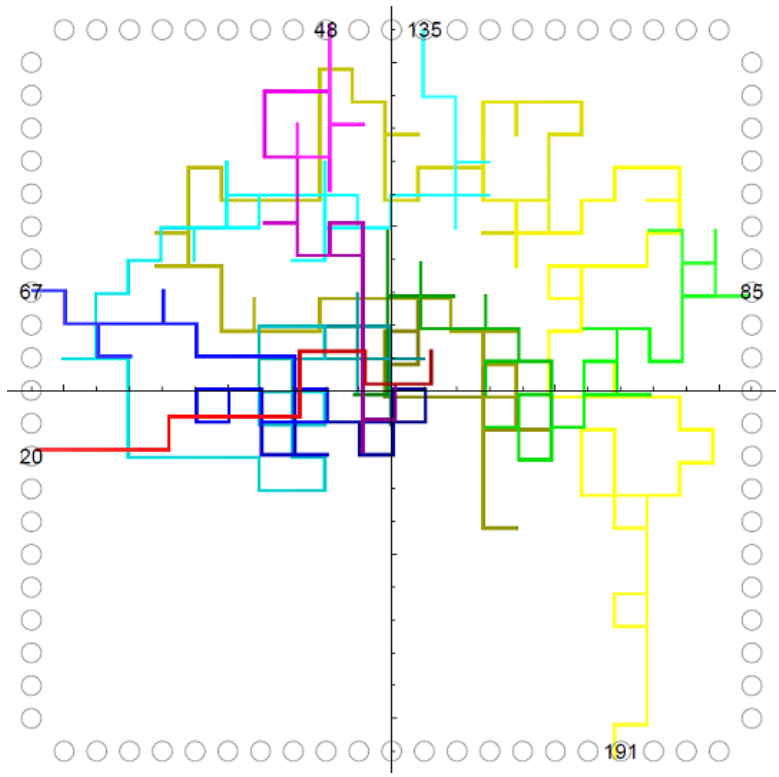

FIGURE 5. 6 Random Walks in $\mathbb{Z}^2$ in a 20×20 square boundary, with stopping time displayed.

Since stopping time T_A is when the random walk reaches the boundary, then S_{T_A} is located on the boundary of A.

Remark 4.2. For any $x \in A$ and $n \in \mathbb{N}$, time homogeneity (Theorem 2.2) tells us,

$$\mathbb{P}\{T_A^x = n\} = \mathbb{P}\{T_A^z = n+1 | S_1 = x\},$$

where $z \in A$ such that $|z - x| = 1$.

We can get a similar statement for S_T. Indeed for any $x \in A$ and $y \in \partial A$, time homogeneity also gives us

$$\mathbb{P}\{S_{T_A^x} = y\} = \mathbb{P}\{S_{T_A^z} = y | S_1 = x\},$$

where $z \in A$ and $|x - z| = 1$.

Lemma 4.3. *Let A be a bounded subset of $\mathbb{Z}^d$. For any random walk in A, stopping time T_A is finite with probability 1.*

Proof. For some $x \in A$, let the stopping time T_A^x be finite for some probability $q(x) \leq 1$. For boundary points $z \in \partial A$, stopping time is zero, so $q(z) = 1$. Note that, by the law of total probability (Theorem 1.7), for all x in A,

$$\mathbb{P}\{T_A^x = n\} = \sum_{\substack{y \in \bar{A} \\ |x-y|=1}} \mathbb{P}\{S_1 = y \mid S_0 = x\} \mathbb{P}\{T_A^x = n \mid S_1 = y\} = \sum_{\substack{y \in \bar{A} \\ |x-y|=1}} \frac{1}{2d} \mathbb{P}\{T_A^x = n \mid S_1 = y\}.$$

Using Remark 4.2, we have $\mathbb{P}\{T_A^y = n-1\} = \mathbb{P}\{T_A^x = n \mid S_1 = y\}$. Thus,

$$\mathbb{P}\{T_A^x = n\} = \sum_{\substack{y \in \bar{A} \\ |x-y|=1}} \frac{1}{2d} \mathbb{P}\{T_A^y = n-1\}.$$

Since this is true for all finite n, then we can extend this to our function q.

$$q(x) = \sum_{\substack{y \in \bar{A} \\ |x-y|=1}} \frac{1}{2d} q(y).$$

Thus q satisfies the mean value property. Because $q(z) = 1$ uniformly on the boundary, Theorem 3.10 implies that $q(x) = 1$ everywhere. □

The finitude of stopping time is needed in order to apply functions to stopping times. If stopping time were infinite, then S_T would be undefined, and functions applied to S_T would also be undefined; applying a function to the end point of a random walk will be key to solving the Dirichlet problem.

Now we are ready to tackle the Dirichlet problem. In $\mathbb{R}^d$, the Dirichlet problem is a boundary value problem in PDE theory. Since we have uniqueness from Theorem 3.12, it is easier to define a function and show that it satisfies the necessary properties rather than explicitly solving the equations. Classical PDE arguments will not work in the discrete case; we will instead approach the problem using random walks. Since random walks have a uniform nearest-neighbor distribution similar to the mean value property, we can use random walks to construct a harmonic function.

Theorem 4.4 (Dirichlet Problem in $\mathbb{Z}^d$)**.** *Let A be a bounded subset of $\mathbb{Z}^d$. Given a function $F : \partial A \to \mathbb{R}$, there exists a unique extension $\tilde{F} : \bar{A} \to \mathbb{R}$ that satisfies the mean value property, and it is*

$$\tilde{F}(x) = \mathbb{E}[F(S_{T_A^x})] = \sum_{z \in \partial A} F(z) \mathbb{P}\{S_{T_A^x} = z\} \tag{4.5}$$

Proof. Note that $\tilde{F}$ is well defined due to finitude of stopping time (Lemma 4.3). Let us show that $\tilde{F}$ defined as such satisfies the mean value property. Uniqueness will follow from Thm 3.12. First, some observations.

We see that remark 4.2 and the fact that $S_0 \notin \partial A$ implies

$$\mathbb{P}\{S_{T_A^x} = z\} = \mathbb{P}\{S_{T_A^{S_0}} = z | S_1 = x\}. \tag{4.6}$$

Also recall the definition of the random walk and remember that, if $|y - z| = 1$, $y \in A$, then,

$$\mathbb{P}\{S_{n+1} = y | S_n = x\} = \frac{1}{2d}. \tag{4.7}$$

Now we analyze the behavior of $\tilde{F}$. As defined,

$$\tilde{F}(x) = \sum_{z \in \partial A} F(z) \mathbb{P}\{S_{T_A^x} = z\}.$$

Applying the law of total probability (Thm 1.7),

$$\mathbb{P}\{S_{T_A^x} = z\} = \sum_{\substack{y \in \mathbb{Z}^d \\ |x-y|=1}} \mathbb{P}\{S_1 = y | S_0 = x\} \mathbb{P}\{S_{T_A^x} = z | S_1 = y\}.$$

Plugging in and rearranging,

$$\tilde{F}(x) = \sum_{\substack{y \in \mathbb{Z}^d \\ |x-y|=1}} \mathbb{P}\{S_1 = y | S_0 = x\} \sum_{z \in \partial A} F(z) \mathbb{P}\{S_{T_A^x} = z | S_1 = y\}.$$

Using (4.6), we get

$$\tilde{F}(x) = \sum_{\substack{y \in \mathbb{Z}^d \\ |x-y|=1}} \mathbb{P}\{S_1 = y | S_0 = x\} \sum_{z \in \partial A} F(z) \mathbb{P}\{S_{T_A^y} = z\},$$

and then (4.7) yields,

$$\tilde{F}(x) = \sum_{\substack{y \in \mathbb{Z}^d \\ |x-y|=1}} \frac{1}{2d} \sum_{z \in \partial A} F(z) \mathbb{P}\{S_{T_A^y} = z\}.$$

Finally, substituting in $\tilde{F}(y)$ as defined earlier in (4.5),

$$\tilde{F}(x) = \sum_{\substack{y \in \mathbb{Z}^d \\ |x-y|=1}} \frac{1}{2d} \tilde{F}(y).$$

Therefore, $\tilde{F}$ satisfies the mean value property.

Finally, by Theorem 3.12, $\tilde{F}$ is unique, so we are done. □

Like many other ideas presented earlier, this theorem also holds true in the continuous case, with partial derivative Laplacians and continuous random walks (Brownian motion). This result is incredible; such a pervasive mathematical phenomenon and a major object of study in PDEs can be modelled and solved using random walks. A probabilistic approach is not at all obvious from the statement of the Dirichlet problem, yet it offers a very neat solution.

Acknowledgments. I cannot properly conclude this paper without thanking my mentors, Max Engelstein and Daphne Kao, whose experienced guidance, devoted enthusiasm, and relentless nagging have whipped my paper into respectable shape. Additionally, I would like to take Greg Lawler for a fascinating series of talks for the REU, as well as lending me his book which served as much of the basis for this paper. Finally, none of this would have been possible without Peter May, whose enthusiasm and commitment have made the REU the fantastic program it is.

References

[1] G. F. Lawler. Random Walk and the Heat Equation. University of Chicago Press. 2010.
[2] S. M. Ross A First Course in Probability Prentice Hall. 1998.