Random Walks on $\mathbb{Z}^2$

Kartikesh Mishra, Darren Yao, Bryce Hancock, Wei Zhang, Thomas Pickering

Fall 2023

Abstract

Random walks on integer lattices, a cornerstone of mathematical modeling, find extensive applications across various disciplines, including physics, computer science, finance, and beyond. This paper delves into the intricacies of two-dimensional random walks on the integer lattice. Our research introduces a novel layered lattice structure and delves into the realms of escape probabilities and exit times within constrained spaces. We present compelling evidence that escape probabilities universally converge to 1 within finite regions and seamlessly extend this finding to encompass certain symmetric infinite regions. Additionally, we furnish comprehensive closed-form formulas for expected exit times and boundary exit probabilities across a spectrum of infinite region scenarios, further enhancing our understanding of these fundamental stochastic processes

1 Introduction

Random walks basically refer to random movements of an object in a defined space with a probability specified for each direction. Though an easily explainable stochastic process, it turns out to be a useful mathematical object for simulating real-world complexities in various fields like physics, computer science, finance, and beyond. For example, in physics, random walks provide a model for the intricate movements of particles in systems from crystal lattices to diffusing chemicals. By mapping stochastic trajectories, random walks shed light on the microscopic dynamics of matter in motion. They open a window into the hidden complexity behind deceptively simple diffusion processes.

These real-world applications demonstrate the relevance of random walk models on integer lattices across many scientific disciplines. In this paper, we explore a specific case - the random walk of a particle on a two-dimensional integer lattice where a particle has four equally probable moves at each step, with the ability to traverse left, right, up, or down. For example, a particle starting at the coordinate $(a, b) \in \mathbb{Z}^2$ is equally likely to be in one of $(a-1, b), (a+1, b), (a, b+1)$, or $(a, b-1)$ positions in next step. This fundamental property provides the basis for our analysis of random motion on the lattice.

To further specify this random walk model, we introduce constraints in the form of regions and their boundaries. A region can be defined as a set of points where the particle randomly walks without stopping. The boundaries are the points where the particle should halt its motion. More precisely, for a point p on the integer lattice $\mathbb{Z}^2$, we define the set of its neighboring points (left, right, top, bottom) as N(p). A region D is then defined as a collection of strongly connected points on the lattice, meaning each point p in D has at least one neighbor that also belongs to D: $\forall p \in D, D \cap N(p) \neq \phi$. The boundary B consists of all points surrounding region D - formally, $B = \{q \in N(p) : p \in D, q \notin D\}$. Note that by this definition, a region can have holes in it which should also be bounded by the boundary points.

Together, region D and boundary B form a graph $G = D \cup B$. Graph G provides constraints on the random walk, restricting motion to region D and imposing boundaries at B where the walk terminates. We have shown two such graphs with some walking steps randomly generated using Python script in the figure (1). This introduction of regions and boundaries makes the model more applicable to real-world scenarios.

Now that we have defined the region, boundary, and graph, we can formulate our central questions:

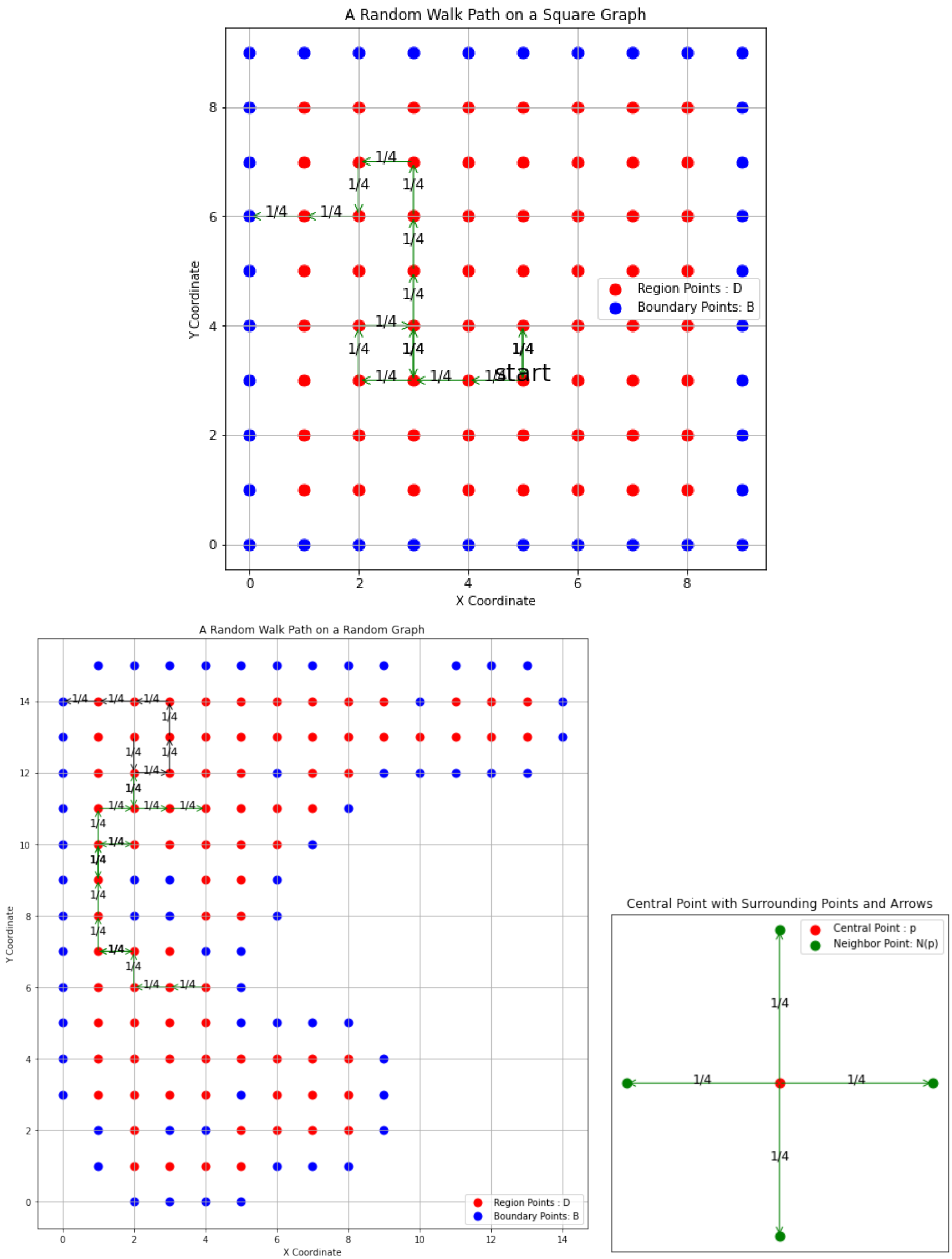


Figure 1: A square region (top) and a random region (bottom) with their boundaries showing random walk with the fundamental property shown on (bottom right).

- Will a particle randomly walking in region D ever escape to boundary B?
- If it does escape, how many steps does it need at minimum, maximum, and on average?

What we have found out as a main result of this paper is that the randomly walking particle actually escapes with certainty i.e. with probability 1 for any finite regions. Furthermore, this result can be extended to some special infinite regions like infinite rectangles or half-planes where some of the movements are symmetric. And in such infinite cases, we also have worked out what the expected number of steps will be required to escape. To present our result more precisely and in a concise manner, we begin by defining what we mean by the terms escape probability and exit time.
An escape probability from any point p, denoted by $Q(p)$ is the probability that a particle starting at point p will eventually escape to the boundary. Then the first result can be shown as a theorem.

Theorem 1. *The escape probability Q(p) is 1 for all points p in any finite region D.*

This theorem can be extended to certain infinite regions with some properties specified in more detail in definition (5).
Similarly, we define the exit time from a point p, denoted by $T(p)$ as the expected number of steps, a particle starting at point p needs to take to escape to the boundary. Next, we define our special regions (which satisfy the definition (5)), an infinite rectangle as a set of all points with one of its two coordinates enclosed within 0 and n for some fixed n i.e $D := 0 < y < n$ or $D := 0 < x < n$, and a half plane as a set of all points with one of its co-ordinates greater than 0 i.e. $D := 0 < y$ or $D := 0 < x$. These regions can be viewed in figure (3). Our results regarding these two special regions are presented as follows.

Theorem 2. *In an infinite rectangle, say $D := 0 < y < n$, the escape probability is 1 for every point, moreover a particle at point p = (x,k) escapes the boundary $y = 0$ with probability $\frac{n-k}{n}$ and the boundary $y = n$ with probability $\frac{k}{n}$ while having the overall exit time as $T(p) = 2k(n - k)$.*

Theorem 3. *In a half-plane defined above, the escape probability from any point is still 1 but the exit time from any point in the half-plane is infinite.*

With these key theorems firmly established, our focus now shifts to their rigorous proofs, commencing with the demonstration of the certain escape from finite regions. These proofs hinge on the meticulous application of our defined escape probability and exit time concepts, often intertwined with the principles of mathematical induction.

The structure of our paper unfolds as follows: we begin with a background overview in Section (1), which sets the stage for the ensuing sections. One such section introduces a lattice structure, to prove Theorem (1). We then investigate extending this theorem to infinite regions; while solving the problems of escape probability and exit time is difficult for general infinite regions, we consider the specific cases of the infinite rectangle and the half-plane; their highly symmetric nature simplifies solutions while maintaining the necessary intuition. Our exploration proceeds to leverage Markov Chains to deepen our understanding of escape probability. Subsequently, we embark on the sections dedicated to the rigorous proofs of Theorems (2) and (3), which offer comprehensive insights into exit times within the context of infinite rectangles and half-plane cases.

2 Layered Lattice Structure

In this section, we attempt to prove theorem (1) and extend it to some special cases of infinity regions. We do so by introducing a novel approach to dealing with random walks over integer lattices. Our method divides a given region into disjoint layers, each characterized by its minimum distance from the boundary. The intuition behind this approach is that a particle needs to cover a certain distance to reach the boundary and if we can show that at each step it gets closer to the boundary then we shall be able to show it will eventually reach the boundary which can play a crucial role in the proof of Theorem 1. We initiate this proof by formally defining the distance between points to the boundaries as its layer value.

Definition 1 (Layer Value)**.** *For any point p in the graph G, we define its layer value, denoted as $l(p)$, as the minimum distance from this point to the boundary. Formally, $l : G \to \mathbb{N} \cup \{0\}$ is defined as $l(p) = \min\{d(p, b) : b \in B\}$, where the distance function $d(x, y) = \max(|x_1 - y_1|, |x_2 - y_2|)$ for any x, $y \in \mathbb{Z}^2$.*

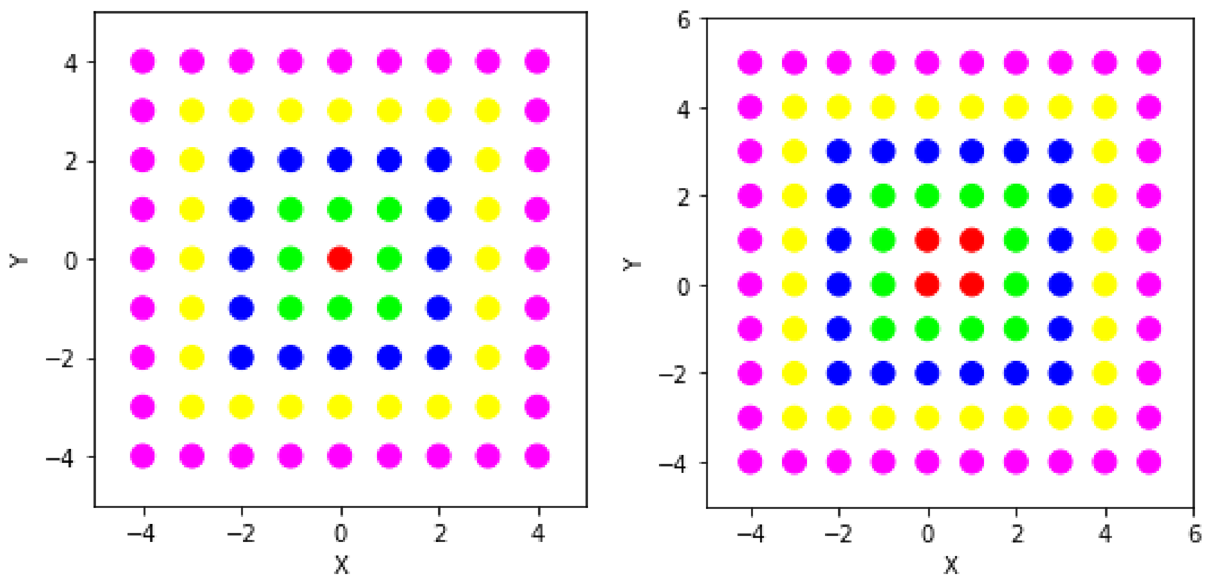


Figure 2: Partitioning of a square region into layers for odd and even side length cases.

Remark: The layer value of a point is also the minimum number of steps required for it to escape. Next, we can collect points with equal layer values as a layer.

Definition 2 (Layers)**.** *A layer of a graph G, denoted as L_k, is a collection of all points in G with a layer value of k.*

In particular, L_0 corresponds to the boundary B and each layer, L_k, is characterized by the points that are precisely at a distance of k from the boundary. These layers can be visualized in figure 2, which illustrates the layering of a square region for cases with odd and even side lengths.
After dividing the regions into these layers, as discussed above, our next step is to establish the relation between the consecutive layers so that we can study the dependency of escape probabilities across the layers. But to do so, we need some established mathematical properties of the escape probabilities. Given that so far, we only know the fundamental property of step transition probability, we begin by formally defining and establishing its properties first which we can use to build the mathematical tools associated with the escape probability.

Definition 3 (Step Transition Probability)**.** *The k-step transition probability, denoted by $P_k(p,q)$ is the probability of a particle starting at point $p \in G$ and ending at $q \in G$ within k steps.*

Using this step transition probability we can define escape probability more precisely as follows.

Definition 4 (Escape Probability)**.** *A k-step escape probability, $Q_k(p)$ is the probability that a particle starting at point p will escape to any boundary point within k steps. i.e. $Q_k : G \to [0,1]$ is defined by $Q_k(p) = \sum_{b \in B} P_k(p,b)$. And when k tends to infinity, there are no step constraints, we call it the escape probability, denoted by $Q_\infty(x) = Q(x)$.*

Now we shift our focus to establish the mathematical properties associated with these probability terms.

Proposition 1. *The k-step transition probability satisfies the following properties.*

- $P_k(p,q) = P_k(q,p) \forall p, q \in D$
- $P_k(p,q) = 0 \forall p \in B, q \in G, p \neq q$
- $P_k(p,q) = \frac{1}{4} \sum_{p' \in N(p)} P_{k-1}(p',q)$

Proof. The symmetric property is in the definition itself and it can be shown as follows

$P_k(p, q) = \mathbb{P}(\text{particle reaches q from p within k steps}) = \mathbb{P}(\text{particle reaches p from q within k steps}) = P_k(q, p)$

The first and last equality follows from the definition of P_k. The second equality holds because the graph is undirected, so k-step paths from p to q are identical to k-step paths from q to p.

To prove the second property, note that by definition, the walk terminates upon reaching any boundary point. Thus, for $p \in B$ and $q \in G$ with $p \neq q$:

$$P_k(p, q) = \mathbb{P}(\text{particle reaches q from p within k steps}) = 0$$

Since no transitions can occur after reaching the boundary.

The third property is derived by applying the law of total probability:

$$P_k(p, q) = \mathbb{P}(\text{particle reaches q from p within k steps}) = \sum_{p' \in N(p)} \mathbb{P}(\text{particle is at p' in 1 step}) \cdot P_{k-1}(p', q)$$

And thus,

$$P_k(p, q) = \sum_{p' \in N(p)} P_{k-1}(p', q) \cdot \frac{1}{4}$$

We use the fact that each neighbor point p' is reached with probability $\frac{1}{4}$ in one step. □

Now, finally, we can extend this property for the escape probability which we will be using continuously throughout this section to prove theorem 1 and its extensions.

Proposition 2. *The k-step escape probability has the following properties.*

- $Q_k(b) = 1 \forall p \in G$
- $Q_k(p) = \frac{1}{4} \sum_{p' \in N(p)} Q_{k-1}(p') \forall p \in D$

Proof. By definition, the escape probability for any boundary point $b \in B$ is:

$$Q_k(b) = \mathbb{P}(\text{particle escapes region within k steps} \mid \text{starts at b}) = 1$$

Since the particle begins on the boundary, escape is guaranteed.

For the second property, use the fact that $Q_k(p) = \sum_{b \in B} P_k(p, b)$ along with the recursive formula for P_k:

$$Q_k(p) = \sum_{b \in B} P_k(p, b) \quad = \sum_{b \in B} \left[\frac{1}{4} \sum_{p' \in N(p)} P_{k-1}(p', b) \right] = \frac{1}{4} \sum_{p' \in N(p)} \left[\sum_{b \in B} P_{k-1}(p', b) \right] \quad = \frac{1}{4} \sum_{p' \in N(p)} Q_{k-1}(p')$$

□

With these tools established, we approach the analysis of escape probability in any finite region. We can now demonstrate that the escape probability from any point is invariably 1 using the following lemmas which bound the neighbors of any point in the region and help us in using the properties related to neighbors.

Lemma 1. *For any point, p in a layer L_k for some $k \geq 1$, all 4 of its neighbors are contained either in the same layer or its neighboring layers i.e. $N(p) \subseteq L_{k-1} \cup L_k \cup L_{k+1}$.*

Proof. Take a point, $p = (p_x, p_y)$ in L_k for some $k \geq 1$ then there exists $b = (b_x, b_y) \in B$ such that $d(p, b) = k$. Here, the neighbors of p are $p_1 = (p_x, p_y - 1), p_2 = (p_x, p_y + 1), p_3 = (p_x - 1, p_y), p_4 = (p_x + 1, p_y)$. Calculating the distance from each of these points to b using the fact $|a| - 1 \leq |a \pm 1| \leq |a| + 1$ from triangle inequality,

$$d(p_1, b) = max(|p_x - b_x|, |p_y - 1 - b_y|) \implies d(p, b) - 1 \leq d(p_1, b) \leq d(p, b) + 1$$

as $|p_x - b_x|$ does not change so the change in max only depends upon $|p_y - 1 - b_y|$ which by triangle inequality is $|p_y - b_y| - 1 \leq |p_y - 1 - b_y| \leq |p_y - b_y| + 1$. The arguments for p_2, p_3 and p_4 are similar to p_1 i.e.

$$d(p_2, b) = max(|p_x - b_x|, |p_y + 1 - b_y|) \implies d(p, b) - 1 \leq d(p_2, b) \leq d(p, b) + 1$$

$$d(p_3, b) = max(|p_x - 1 - b_x|, |p_y - b_y|) \implies d(p, b) - 1 \leq d(p_3, b) \leq d(p, b) + 1$$

$$d(p_4, b) = max(|p_x + 1 - b_x|, |p_y - b_y|) \implies d(p, b) - 1 \leq d(p_4, b) \leq d(p, b) + 1$$

This establishes that for all point $p' \in N(p)$, there exists at least one boundary point b from which, $k - 1 \leq d(p', b) \leq k + 1$, and thus all the neighbors must be contained in $L_{k-1} \cup L_k \cup L_{k+1}$ □

This lemma provides us with some constraints on the neighbors of a point in any region and establishes some relation between the neighboring layers, we can further solidify this relation, at least between the layers L_k and L_{k-1}, which will further help us in mathematically navigating from any point to the boundary.

Lemma 2. *For any point in a region D and in L_k for some $k \geq 1$ then at least one of its 4 neighbors belongs to layer L_{k-1}. Formally, $\forall p \in D \implies p \in L_k$ for some k, $L_{k-1} \cap N(p) \neq \phi$*

Proof. Suppose there exists a point $p = (p_x, p_y) \in D$ and for some k, $p \in L_k$ such that none of its neighboring points is in layer L_{k-1}. Since $p \in L_k$, there must exist, $b = (b_x, b_y) \in B$ s.t. $d(p, b) = k$. This implies either $d(p, b) = |p_x - b_x| = k$ then one of left $(p_x - 1, p_y)$ or right $(p_x + 1, p_y)$ is at $\leq k - 1$ distance from b by the linearity of the points, or it could be $d(p, b) = |p_y - b_y| = k$ then similarly, one of top $(p_x, p_y + 1)$ or bottom $(p_x, p_y - 1)$ is at distance $\leq k - 1$ from b. In all cases, there exists at least one neighbor of p at a distance $\leq k - 1$ from a boundary point proving and from Lemma (1), we know all neighbors have to be in either L_{k-1} or L_k or L_{k+1}. Thus, at least one neighbor of p must be in L_{k-1} proving the lemma. □

Now, with this relation between consecutive layers, theorem (1) can be easily proven with the arguments that the particle traveling from one layer to another must reach the boundary eventually and all we need to do is show that some specific points can transfer its escape probability through layers. In a more rigorous way, the argument is presented as follows.

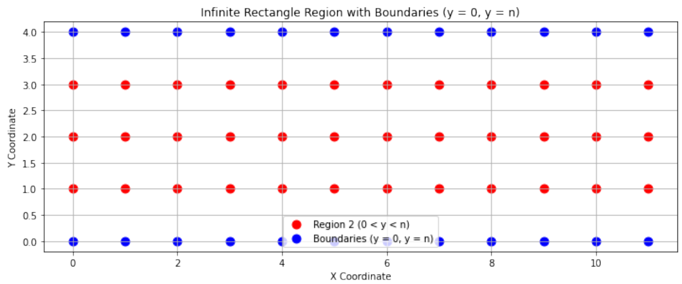
Proof. ***Theorem***(1). Suppose there exist some points in the region D whose escape probability is not 1 then it must be strictly less than 1. Let $p \in D$ be the point with the minimum escape probability $Q(p)$. The existence of such p is guaranteed by D being finite i.e. only finitely many points. Then by proposition (2), $Q(p) = \frac{1}{4}\sum_{p' \in N(p)} Q(p')$. This implies $Q(p) = Q(p') \forall p' \in N(p)$, otherwise there will be at least one neighbor with a lower escape probability and it will violate our choice of minimum. But by lemma (2), now we have at least one point in layer L_{k-1} with the minimum escape probability. By the continuous repetition of the above argument, we conclude that there exists a point in layer L_0 with minimum probability Q(p) but L_0 is the boundary and the escape probability for any boundary points is 1. Thus, $Q(p) = 1$ proving theorem (1). □

2.1 Special Infinite Regions

After establishing that all points in any finite region have escape probability 1, one can naturally ask what about infinite regions. In this subsection, we present that certain infinite regions, though bounded, still guarantee eventual escape. From lemma (2), we know for certain, that a point in layer L_k has a neighbor point in L_{k-1}. With more constraints, we can introduce a class of regions defined as follows.

Definition 5 (Layer Symmetric Regions)**.** *A layer L_k is said to be symmetric if for any point $p \in L_k$, one of its neighbors lies in layer L_{k-1}, one in L_{k+1} and the rest two in L_k itself. A region D is then called layer-symmetric if all of its layers are symmetric.*

The escape probability can be deduced as 1 for all points in such regions. We show this by first establishing that escape probability in such regions only depends upon the layer value and then we prove that the escape probability turns out to be an arithmetic progression that solves to 1. The two canonical bounded domains presented in the figure (3): the infinite rectangle and the half-plane (for which we have computed the exit



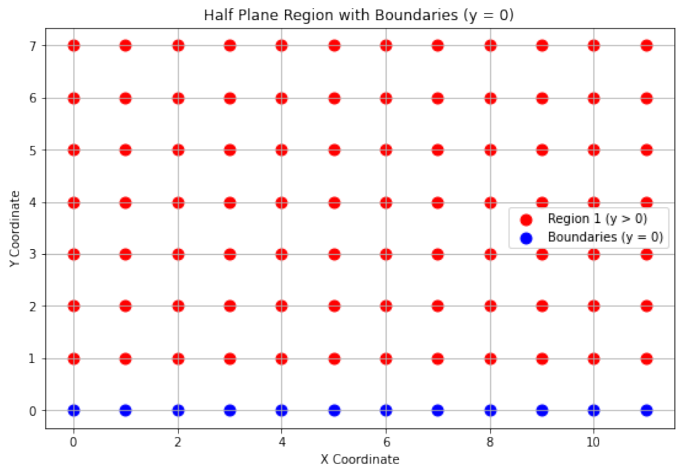


Figure 3: An infinite rectangle region with boundary y=0 and y=n (top) and a half-plane region with boundary y=0 (bottom).

times in section (4)) satisfy the property of layer-symmetric region due to their symmetry along an axis. In the infinite rectangle case, it is a bit tricky to evaluate it as a layered symmetric region, which is further explained below while proving the first half of theorem (2).

Lemma 3. *In a layer-symmetric region, the escape probability from a point only depends upon the layer value of the point.*

Proof. We are going to show this using strong induction. Suppose for all $m \leq k$, we say the escape probability from all points in layer L_m only depends upon the layer value m. Let's denote the escape probability by Q_m then. The base case can be observed from the proposition (2) which implies all points in layer $L_0 = B$, the escape probability is 1. Now, for the induction step, take any two points p and q in layer L_{k+1} then by lemma (2) or simply by definition (5) of the layer-symmetric region, both of these points have one of their neighbors in layer L_k, say p' and q' respectively then the escape probabilities $Q(p') = \frac{1}{4}\sum_{p'' \in N(p')} Q(p'')$ and $Q(q') = \frac{1}{4}\sum_{q'' \in N(q')} Q(q'')$ by proposition (2). We know again by definition (5), that p' and q' have two neighbors in L_k, one in L_{k-1} and one in layer L_{k+1} which we already know is p and q respectively. Then by induction hypothesis, we have $Q(p') = Q_{k-1} + 2 * Q_k + Q(p)$ and $Q(q') = Q_{k-1} + 2 * Q_k + Q(q)$. Furthermore, $Q(p') = Q(q') = Q_k$ which implies $Q(p) = Q(q)$ thus proving by mathematical induction that, for all layers with layer-value $k \geq 0$, the escape probability only depends upon the layer-value k. □

With this result, for further simplicity, in this subsection, we will denote $Q(p)$ as Q_k where the point p is in layer L_k for some k, and next we establish that this Q_k is an arithmetic progression in k.

Lemma 4. *In a layer-symmetric region, the escape probability Q_k can be recursively defined as $\frac{1}{2}(Q_{k-1} + Q_{k+1})$. Moreover, Q_k is an arithmetic progression, i.e. for some fixed r, $Q_k = 1 + r \cdot k$.*

Proof. By proposition (2), we know for any point p, $Q(p) = \frac{1}{4}\sum_{p' \in N(p)} Q(p') \forall p \in D$. Note that by definition (5), one of its 4 neighbors is in layer k-1, two in layer k and one in layer k+1 given that p is in layer k.

Using lemma (3), this implies $Q_k = \frac{1}{4}Q_{k-1} + \frac{2}{4}Q_k + \frac{1}{4}Q_{k+1}$ which simplifies to $Q_k = \frac{1}{2}(Q_{k-1} + Q_{k+1})$. This implies $Q_k - Q_{k-1} = Q_{k+1} - Q_k$ and hence, Q_k is an arithmetic progression with common difference r(say). This gives us $Q_k = Q_0 + r \cdot k$. Furthermore, $Q_0 = 1$ by proposition as layer 0 is the boundary by definition. Thus, $Q_k = 1 + r \cdot k$ for some r independent of k. □

Now, we attempt to establish the following theorem about escape probability in layer-symmetric regions and we will prove it using two cases of having infinitely many layers and finitely many layers, which we will show is not possible.

Theorem 4. *In any layer-symmetric region, the escape probability from any point is 1.*

*Proof - **Infinitely many Layers Case**.* We begin by using lemma (3) and lemma (4), which tell us, for each layer value k, $Q_k = 1 + r \cdot k$. Clearly, r cannot be greater than 0 otherwise $Q_1 = 1 + r$ which violates the axiom of probability but r cannot be negative either as can be seen by the following argument. Suppose $r < 0$ then for all $k > \frac{1}{|r|}$, $Q_k < 0$ as $r \cdot k < \frac{r}{|r|} = -1$. The existence of such k is guaranteed by the layers being infinitely many i.e. k ranges up to infinity. This leaves us with only a choice of $r = 0$ which implies $Q_k = 1$ for all k. □

Corollary 1. *This proves the first half of the theorem* (3)*. It is easy to observe that in a half-plane, the layer value is one of the coordinates itself, so it is a layer-symmetric region with infinitely many layers.*

Note that by definition (5), we know one of the neighbors of every point in layer L_k, lies in layer L_{k-1} and one in layer L_{k+1} but this cannot be followed by the points in layer L_m since there does not exist the layer L_{m+1}, given we have finitely many $m + 1$ (with L_0 being boundary) layers. Technically, a layer-symmetric region cannot have a finite number of layers and we are done with the proof of the theorem (4). It can also be noted that except for their last layer, some finitely many layered regions can still satisfy all of the lemmas above. We can use this situation to establish escape probability properties for such finitely many layered by introducing null points where the escape probability is fixed to be 0. It is presented below in the proof of the first half of theorem (2).

*Proof - **First Half of Theorem*** (2)*.* Here, we know one of the coordinates Y (say) is bounded by 0 and some fixed n, i.e. $D := 0 < y < n$ and boundaries are $y = 0$ and $y = n$. We can apply the following trick, let U_k be the probability that a particle at layer k escapes to the boundary $y = 0$ and let V_{n-k} be the probability that the particle escapes to boundary $y = n$. Now, $U_0 = V_0 = 1$ and $U_n = V_n = 0$ as a particle cannot travel from one boundary to another. Now, take a modified layer-symmetric region S with the condition that the nth layer of S is null which means the escape probability there is fixed to be 0. It can then be easily observed that both U_k and V_k can model the escape probability of layer $k \leq n$ of S as for U_k and V_k, the infinite rectangle is basically $n - 1$ symmetric layers i.e. satisfy the definition (5) as the layer value is dependent on one of the coordinates itself.
Thus, by using lemma (3) and lemma (4), we obtain $U_k = 1 + r_u \cdot k$ and similarly, $V_k = 1 + r_v \cdot k$ and then using $U_n = V_n = 0$, we get $r_u = r_v = \frac{-1}{n}$. Thus, a point $p = (x, k)$ escaping to the boundary $y = 0$ is $U_k = \frac{n-k}{n}$ and to the boundary $y = n$ is $V_{n-k} = \frac{k}{n}$ and hence, the overall escape probability is $\frac{n-k+k}{n} = 1$ as desired. □

3 Markov Chains on $\mathbb{Z}^2$

In this section, we introduce Markov Chains and use it to prove Theorem 1 and discover the escape probability distribution over a closed region's boundary points. Markov Chains are a framework used to evaluate dynamic systems in probability, and they will assist us in formalizing the properties of random walks on $\mathbb{Z}^2$.
To introduce the concept of a Markov Chain, imagine a frog hopping from one lily pad to another on a pond. At each hop, the frog's choice of the next lily pad might depend on various factors: maybe the size of the pad, or its distance from the frog. However, what if the frog only cares about its current lily pad and makes the jump solely based on where it is right now, without any regard for its hopping history? This is the essence of a Markov Chain. At each step, the system's next state depends solely on the present state, and not on the sequence of states that preceded it. This memoryless property can be incredibly powerful because

it simplifies complex systems. Markov chains find applications in numerous fields: from economics, where they can represent the transitions between different economic states, to biology, for modeling the changes in a population or the spread of a disease. In our context, a Markov Chain will be used to model a random walker moving along the two-dimensional lattice in discrete time.
For our purposes, understanding the Markov property will be crucial for rigorously exploring questions like: How often does a random walker return to its starting point? Can it escape a bounded region? To address these, we employ Markov chains as our guiding principle.
Let's begin with a formal introduction:

Definition 6. *A (discrete time) stochastic process is a sequence of random variables $(X_t)_{t\in\mathbb{N}}$ taking values in some set S.*

Now, making our frog analogy rigorous:

Definition 7. *A (discrete time) Markov chain is a stochastic process X_t taking values in S, and satisfying the condition*

$$P(X_t = z_t \mid X_0 = z_0, \ldots, X_{t-1} = z_{t-1}) = P(X_t = z_t \mid X_{t-1} = z_{t-1}).$$

Essentially, the probability of moving to a particular state only depends on the previous state and not the entire history of states. It's as if our frog has a short-term memory.

To cement this understanding, let's examine a basic example:

Example 1. *Consider a simple 2-state Markov chain with the following transition matrix:*

$$P = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix}$$

This matrix is akin to our frog's decision-making process. If the frog is on lily pad 1, there's a 70% chance it'll stay put and a 30% chance it'll hop to lily pad 2. Similarly, if it's on lily pad 2, there's a 60% chance it'll stay and a 40% chance it'll leap back to lily pad 1.

3.1 Properties of a Random Walk on $\mathbb{Z}^2$

Example 2. *Consider a Markov chain representing a random walk in $\mathbb{Z}^2$. This Markov chain, labeled X_t, moves to one of its horizontal or vertical neighbors with probability $1/4$. Let $P^{(t)}(x, y)$ represent the probability that X_t moves from x to y in t steps. Let $V(x, y) = \sum_{t=0}^{\infty} P^{(t)}(x, y)$ be the expected number of visits X_t takes to y starting from x in infinite steps.*

Proposition 3. *X_t is expected to return to its starting point infinitely often almost surely.*

Proof. Without loss of generality, let the origin be the starting point for X_t. Then, X_t returns to its starting point infinitely often if and only if $V(0, 0) = \infty$. We first observe that X_t cannot return to the origin in an odd number of steps, since it is guaranteed to leave the origin at step one. Thus, $V(0, 0) = \sum_{t=0}^{\infty} P^{(2t)}(0, 0)$. For X_t to return to the origin, it must have an equal number of up and down steps, and an equal number of left and right steps. Let the total number of steps be t, and the number of vertical steps as $0 \leq k \leq t$. The total number of paths back to the origin is then given by the multinomial coefficient:

$$\binom{2t}{k, k, t-k, t-k}$$

The probability of any particular path of length l is $(1/4)^l$, so $P^{(2t)}(0, 0) = (1/4)^{2t}\binom{2t}{k,k,t-k,t-k}$.
We can express $\binom{2t}{k,k,t-k,t-k}$ as $\frac{(2t)!}{(k!)^2((t-k)!)^2}$, which gives us $P^{(2t)}(0, 0) = (1/4)^{2t}\sum_{k=0}^{t}\frac{(2t)!}{(k!)^2((t-k)!)^2}$. Therefore, $V(0, 0)$ diverges to infinity, implying X_t is expected to return to the origin infinitely often. [DRAFT NOTE: Still working on proving divergence of series] $\square$

Theorem 5. *X_t is expected to travel to every point on the lattice infinitely often almost surely.*

Proof. A random walk on the integer lattice, by its definition, is a Markov chain on a countable, connected graph. This implies the existence of an integer s such that $P^s(0, y) > 0$ for all y in the lattice.
[DRAFT NOTE: The proof for this theorem is under construction. The primary idea is to demonstrate that the expected number of visits to the origin is infinite, as shown by the proposition. Given that the probability of transitioning from the origin to any other lattice point in a finite number of steps is always positive, it follows that the expected number of visits to any lattice point is infinite.] ☐

Corollary 2. *If X_t starts in some region D with boundary B, then the probability of escaping the boundary is 1.*

Proof. This follows directly from Theorem 2, where we proved that X_t is expected to travel to every point infinitely often. Thus, this corollary holds for any finite or infinite D and any shape of boundary B. [DRAFT NOTE: This proof is not sufficient. We might go back to examples instead] ☐

4 Exit Times in Infinite Regions

Given that the random walk is guaranteed to exit the region, we now want to investigate how long it takes to do so. For arbitrary regions, the exit times are often difficult to compute, so in this section we instead focus on exit times of points inside highly symmetric shapes. This maintains most of the intuition for computing exit times, while simplifying algebraic computation. Specifically, we consider two cases: infinite rectangles and the half plane, where the latter can be considered a limiting case of the former as the height approaches infinity.

4.1 Infinite Rectangle

For the infinite rectangle, without loss of generality, we will consider the region D defined by $0 \leq y \leq n$, where the particle is allowed to move arbitrarily far in either horizontal direction (otherwise, translate the boundaries accordingly). We aim to find the exit time from any point within this region. Observe that the exit time depends only on the vertical coordinate, because the rectangle extends infinitely in the horizontal direction. Therefore, we can define $T_{n,k}$ to be the exit time for a point (x, k) in this region.

Theorem (Second half of Theorem (2))**.** *In an infinite rectangle $D = 0 \leq y \leq n$, the exit time from a point (x, k) is $T_{n,k} = 2k(n-k)$.*

Proof. We have $T_{n,0} = T_{n,n} = 0$, because the particle is considered to have already exited.
Then, for any point (x, k) with $0 < k < n$, we consider the possible moves on the next step of the random walk; we have the following recurrence:

$$T_{n,k} = 1 + \frac{T_{n,k+1} + T_{n,k-1} + 2T_{n,k}}{4}.$$

This is because at any given step of the random walk, the number of moves taken so far increments by 1, and then the random walk moves up, down, left, and right with equal probability.
We simplify the above equation to

$$T_{n,k} = 2 + \frac{1}{2}T_{n,k+1} + \frac{1}{2}T_{n,k-1}.$$

In total, we have the following recurrences for the exit time:

$$T_{n,k} = \begin{cases} 0 & k = 0 \\ 2 + \frac{1}{2}T_{n,k+1} + \frac{1}{2}T_{n,k-1} & 0 < k < n \\ 0 & k = n \end{cases}$$

This set of recurrences has a unique solution. We can observe this by taking all the recurrences except $k = n$; substituting them into each other yields equations that express $T_{n,k}$ for $0 \leq k < n$ in terms of $T_{n,n}$. Plugging in $T_{n,n} = 0$ then uniquely determines the values of all the $T_{n,k}$.

This unique solution is given by $T_{n,k} = 2k(n-k)$, and we can verify it satisfies all the recurrence equations. For $k = 0$ and $k = n$ obviously $2k(n-k) = 0$, and then for the remaining values of k we have:

$$T_{n,k} = 2+\frac{1}{2}T_{n,k+1}+\frac{1}{2}T_{n,k-1} = 2+\frac{1}{2}2(k+1)(n-k-1)+\frac{1}{2}2(k-1)(n-k+1) = 2-2k^2+2kn-2 = 2k(n-k)$$

as desired. □

4.2 Half Plane

The upper half-plane is defined as all points such that $y \geq 0$, and we want to find the exit time of the particle reaching the boundary $y = 0$. We can consider this as the limiting case of the infinite rectangle case as $n \to \infty$. Using the exit time formula for that case of $2k(n-k)$, as $n \to \infty$ we would expect the exit time of the half-plane to also be infinite, and indeed this is the case.

Theorem (Second half of Theorem (3))**.** *The exit time from any point in the half plane is infinite.*

Proof. Define T_k to be the exit time from a point at y-coordinate k. Consider the following recurrence: the expected time to go from $y = 2$ to $y = 0$ is the sum of the time to go from $y = 2$ to $y = 1$, and the time to go from $y = 1$ to $y = 0$. Both of these times are given by T_1, so we have $T_2 = 2T_1$.
Also, note that while at $y = 1$, we have $\frac{1}{4}$ probability of going to $y = 2$, $\frac{1}{4}$ probability of going to $y = 0$, and $\frac{1}{2}$ probability of staying at $y = 1$, so we have $T_1 = 1 + \frac{1}{4}T_2 + \frac{1}{4}T_0 + \frac{1}{2}T_1$, which simplifies to $T_1 = 2 + \frac{1}{2}T_2$.
Plugging in $T_2 = 2T_1$ yields $T_1 = 2 + T_1$, which is a contradiction for any finite T_1. Therefore T_1 must be infinite, and all other T_k for $k \neq 0$ are also infinite.

□

5 Conclusion

In this paper, we've achieved significant progress in exploring random walks on integer lattices. Our introduction of layered lattice structures has led to the establishment of Theorem (1), revealing universal escape probabilities within finite regions, consistently equal to 1. This insight has been extended to encompass special infinite regions that satisfy the criteria outlined in (5), encompassing infinite rectangles and half-planes.

Furthermore, we've unveiled concise formulas for expected exit times within various region scenarios (including infinite rectangles and half-planes), capitalizing on the symmetry of one-dimensional movements to simplify two-dimensional random walks. These findings provide deeper insights into the dynamics of random walks, illuminating the expected exit times.

Finally, it's worth noting that random walks find application in numerous fields, ranging from genetic drift in biology to economic stock market trends and the realm of Brownian motion in physics. By delving into the study of random walks, we aim to establish a comprehensive framework for comprehending such processes across diverse disciplines.

6 Future Work

We can also investigate related problems that arise from various generalizations, constraints, or specific cases of the theorems we studied here.

Problem 1. *We proved that for finite regions, the escape probability is 1. We can extend this to the following question: when a particle exits a certain region, what is the probability that it exits at a particular boundary point?*

Problem 2. *For finite regions, the escape probability is 1 if given infinite time. Now, what is the probability that the particle will escape the region within k steps?*

Problem 3. *For finite regions, how does the exit time vary as a function of the area and perimeter of the region?*

Problem 4. *If the random walk is weighted, i.e. the probabilities of moving in each direction are not necessarily 1/4, how does this affect the escape probabilities of points in the half-plane?*

7 Acknowledgements

This paper is the collaborative result of the efforts of Kartikesh Mishra, Bryce Hancock, and Darren Yao.

Kartikesh began the paper by writing the abstract and, introduction while also drawing all the figures using Python script. He then introduced the concept of a layered lattice structure in the section (2), proved Theorem (1), and extended the theorem to the special infinite regions of the infinite rectangle and half-plane, proving the first half of both theorems (2) and (3). In the end, he wrote the conclusions and the future works and ended the paper with this section and Appendix where all the Python scripts and his unfinished work lie.

Darren derived and proved closed-form formulas for exit times of all points within the infinite rectangle and the half-plane, in section (4). He also added a few questions in the future works section. Bryce analyzed Markov Chains, for a proof of theorem (1), and uncovered the distribution of escape probabilities along the boundary in section (3).

We express our heartfelt gratitude to our esteemed mentors, Wei Zhang and Thomas Pickering, whose unwavering support and guidance, encompassing both the technical and non-technical aspects of this research, have been instrumental in shaping the trajectory of this paper. Their expertise and encouragement have been invaluable in our pursuit of knowledge and discovery.

8 Appendix

8.1 Method of Difference Proof Technique for first half of Theorem (2) and (3)

Lemma 5. *For both of these regions, Q_k can be recursively defined as $\frac{1}{2}(Q_{k-1} + Q_{k+1})$. Furthermore, the dependency on Q_{k-1} can be removed as $Q_k = \frac{1}{k+1} + \frac{k}{k+1}Q_{k+1}$*

Proof. By proposition (2), we know for any point p, $Q(p) = \frac{1}{4}\sum_{p' \in N(p)} Q(p') \forall p \in D$. Note that in both cases, one of its 4 neighbors is in layer k-1, two in layer k and one in layer k+1 given that p is in layer k. This implies $Q_k = \frac{1}{4}Q_{k-1} + \frac{2}{4}Q_k + \frac{1}{4}Q_{k+1}$ which simplifies to $Q_k = \frac{1}{2}(Q_{k-1} + Q_{k+1})$. Furthermore, $Q_0 = 1$ in both cases as layer 0 is the boundary by definition. Thus, $Q_1 = \frac{1}{2} + \frac{1}{2}Q_2$ which satisfies the base case of the claim. Suppose for some valid u i.e. u is not a boundary (in such case particle cannot travel to layer u+1), $Q_u = \frac{1}{u+1} + \frac{u}{u+1}Q_{u+1}$ then,

$$Q_{u+1} = \frac{1}{2}(Q_u + Qu + 2) \implies Q_{u+1} = \frac{1}{2(u+1)} + \frac{u}{2(u+1)}Q_{u+1} + \frac{1}{2}Q_{u+2}$$

$$Q_{u+1}(1 - \frac{u}{2(u+1)}) = \frac{1}{2(u+1)} + \frac{1}{2}Q_{u+2} \implies Q_{u+1} = \frac{1}{2(u+1)} \cdot \frac{2(u+1)}{u+2} + \frac{1}{2} \cdot \frac{2(u+1)}{u+2}Q_{u+2}$$

which establishes the induction step as $Q_{u+1} = \frac{1}{u+2} + \frac{u+1}{u+2}Q_{u+2}$ thus, by principle of mathematical induction for all possible layer value k, $Q_k = \frac{1}{k+1} + \frac{k}{k+1}Q_{k+1}$. □

Theorem 6. *The escape probability from any point in a half-plane is 1.*

Proof. Using lemma (3) and lemma (5), we know, for each layer value k, $Q_k = \frac{1}{k+1} + \frac{k}{k+1}Q_{k+1}$ and in this case, $Q_0 = 1$. Solving recursively for Q_1,

$$Q_1 = \frac{1}{2} + \frac{1}{2}Q_2 = \frac{1}{2} + \frac{1}{6} + \frac{1}{3}Q_3 = ... = \frac{1}{1*2} + \frac{1}{2*3} + ... + \frac{1}{r*(r+1)} + \frac{1}{r+1}Q_{r+1}$$

Now, we know $\frac{1}{i*(i+1)} = \frac{1}{i} - \frac{1}{i+1}$, using this,

$$Q_1 = \sum_{i=1}^{r} \frac{1}{i*(i+1)} + \frac{1}{r+1}Q_{r+1} = 1 - \frac{1}{r+1} + \frac{1}{r+1}Q_{r+1}$$

The last equality holds by the method of difference and now, as r tends to infinity, Q_1 tends to 1, and since for any r, $Q_r \geq 0$ simply by probability axiom, we get the desired $Q_1 = 1$. Since $Q_1 = 1$, similar argument can be applied to Q_2 and henceforth making $Q_k = 1$ for all k $\in \mathbb{N}$. ☐

Theorem 7. *The escape probability from any point in an infinite rectangle is 1.*

Proof. Now, using lemma (3), let U_k be the probability that a particle at layer k escapes to the boundary $y = 0$ and let V_{n-k} be the probability that the particle escapes to boundary $y = n$. Now, $U_0 = V_0 = 1$ and $U_n = V_n = 0$ as a particle cannot travel from one boundary to another. Since both U and V are escape probabilities, by lemma (5), $U_k = \frac{1}{k+1} + \frac{k}{k+1}U_{k+1}$ and $V_k = \frac{1}{k+1} + \frac{k}{k+1}V_{k+1}$. This implies with $U_k = V_k$ for k = 0 and n, $U_k = V_k \forall k \in [n]$ by induction. Now, U_k can be recursively solved by,

$$U_k = \frac{1}{k+1} + \frac{k}{k+1}U_{k+1} = \frac{1}{k+1} + \frac{k}{(k+1)*(k+2)} + \frac{k}{k+2}U_{k+2} = ... = \frac{1}{k+1} + \frac{k}{(k+1)*(k+2)} + ... + \frac{k}{(n-1)*n} + \frac{k}{n}U_n$$

Since $U_n = 0$,

$$U_k = V_k = \frac{1}{k+1} + \sum_{i=k+1}^{n-1} \frac{k}{i*(i+1)}$$

Using the method of difference, similarly, as in theorem (6),

$$U_k = V_k = \frac{1}{k+1} + \frac{k}{k+1} - \frac{k}{n} = 1 - \frac{k}{n}$$

By countable additivity of probability axiom,

$$Q_k = U_k + V_{n-k} = 1 - \frac{k}{n} + 1 - \frac{n-k}{n} = 1 \text{ as required.}$$

☐

8.2 Layer Partition Lemma

Lemma 6. *Consider a countable sequence of a mutually disjoint sets $S_1, S_2, S_3, ...$ and any function $f_n : D_n \to \mathbb{R}$ where $D_n = \bigcup_{i \leq n} S_i$. Let $F_k^n = (f_n(x) : x \in S_k)$ be a functional vector for set S_k. If F_k^n satisfies the following properties,*

- $\forall n, F_k^n = B_{-1}(k)F_{k-1}^n + B_0(k)F_k^n + B_1(k)F_{k+1}^n$ *for all $k \geq 1$, for some real-valued matrices, $B_0(k), B_{-1}(k)$ and $B_1(k)$ independent of n*
- $B_{-1}(1) = B_0(1) = 0$

Then $\forall k, F_k^n = A(k+1)F_{k+1}^n$ for some real-valued matrix A(k+1) iff $I - B_{-1}(k)A(k) + B_0(k)$ is invertible for all k.

Proof. We use proof by induction. The base case is trivial from the given statement i.e. $\forall n, F_1^n = B_1(2)P_2^n$. Thus, $A(2) = B_1(2)$ exist. Now suppose it is true for some u-1 then $\forall n, F_u^n = B_{-1}(u)F_{u-1}^n + B_0(u)F_u^n + B_1(u)F_{u+1}^n = (B_{-1}(u)A(u) + B_0(u))F_u^n + B_1(u)F_{u+1}^n \implies F_u^n = (I - B_{-1}(u)A(u) - B_0(u))^{-1}B_1(u)F_{u+1}^n$. This completes the proof with $A(u+1) = (I - B_{-1}(u)A(u) - B_0(u))^{-1}B_1(u)$.
Now suppose instead that $\forall k, F_k^n = A(k+1)F_{k+1}^n$ then, equating with $\forall n, F_k^n = B_{-1}(k)F_{k-1}^n + B_0(k)F_k^n + B_1(k)F_{k+1}^n$, we get $(I - B_{-1}(k)A(k) - B_0(k))A(k+1)F_{k+1}^n = B_1(k)F_{k+1}^n$. Since this holds for arbitrary F_{k+1}^n, we must have $(I - B_{-1}(k)A(k) - B_0(k))A(k+1) = B_1(k)$. Multiplying both sides by the inverse of $I - B_{-1}(k)A(k) - B_0(k)$ gives $A(k+1) = (I - B_{-1}(k)A(k) - B_0(k))^{-1}B_1(k)$. Therefore, $I - B_{-1}(k)A(k) - B_0(k)$ must be invertible.
Note This proof and lemma are still under work and may progress as this document progresses. $\square$

For any layer L_k, we know by the proposition (2) that its neighbor belongs to its neighbor layers, the next definition establishes a more formal relation between the layers.

Definition 8. *For any layer L_k of a graph G if there exist layer L_{k+1}, the inter-layer relation matrix $V_k \in \mathbb{R}^{|L_k| \times |L_{k+1}|}$ is defined as following using indexing,*

- $V_k[i, j] = 1/4$ *if* $L_{k+1}[j] \in N(L_k[i])$
- $V_k[i, j] = 0$ *otherwise*

Also, some of the neighbors may be actually in the same layer. We define intra-layer relation matrix $U_k \in \mathbb{R}^{|L_k| \times |L_k|}$ as following in a similar way,

- $U_k[i, j] = 1/4$ *if* $L_k[j] \in N(L_k[i])$
- $U_k[i, j] = 0$ *otherwise*

Following this definition, an interesting property can be seen for escape probability, let F_k be the vector $(Q(p) : p \in L_k)$ then by definitions, $F_k = V_{k-1}^T F_{k-1} + U_k F_k + V_k F_{k+1}$. Furthermore, by theorem (1), $F_k = A(k-1)F_{k-1}$ for some real-valued matrix $A(k+1)$ with the condition that for any row in $A(k+1)$, the sum of the row is 1. Then by lemma (6), we know that the inverse of $I - V_k A(k) - U_k$ exists.

8.3 Code for Generating random walks

```
import numpy as np
import matplotlib.pyplot as plt
from matplotlib.patches import FancyArrowPatch

# Function to generate a random strongly connected region
def generate_connected_region(width, height):
    lattice = np.zeros((height, width), dtype=float)
    x, y = np.random.randint(1, width-1), np.random.randint(1, height-1)
    stack = [(x, y)]

    while stack:
        x, y = stack.pop()
        lattice[y, x] = 1.0
        neighbors = [(x+1, y), (x-1, y), (x, y+1), (x, y-1)]
        np.random.shuffle(neighbors)
        for nx, ny in neighbors:
            if 0 < nx < width - 1 and 0 < ny < height - 1 and lattice[ny, nx] == 0.0:
                stack.append((nx, ny))
                break

    for y in range(height):
        for x in range(width):
```

```
            if lattice[y, x] == 0.0:
                is_boundary = False
                for ny, nx in [(y+1, x), (y-1, x), (y, x+1), (y, x-1)]:
                    if 0 <= nx < width and 0 <= ny < height and lattice[ny, nx] == 1.0:
                        is_boundary = True
                        break
                if is_boundary:
                    lattice[y, x] = 0.5
    # For square region comment above and uncomment below
    # lattice = np.ones((width,height))
    # lattice[0,:] = 0.5
    # lattice[width-1,:] = 0.5
    # lattice[:,0] = 0.5
    # lattice[:,height-1] = 0.5

    return lattice

# Generate a connected region
width, height = 10, 10  # Adjust the size of the lattice as needed
connected_region = generate_connected_region(width, height)

# Prepare data for plotting
x_points, y_points = np.where(connected_region == 1)
boundary_points = np.where(connected_region == 0.5)

# Choose a random starting point from the region (red point)
start_x, start_y = np.random.choice(x_points), np.random.choice(y_points)


# Generate the path with length strictly less than 4
path_x = [start_x]
path_y = [start_y]
while len(path_x) < 20  and connected_region[path_x[-1], path_y[-1]]!=0.5:
    lx, ly = path_x[-1], path_y[-1]
    neighbors = [(lx + 1, ly), (lx - 1, ly),
                 (lx, ly + 1), (lx, ly - 1)]
    np.random.shuffle(neighbors)
    for nx, ny in neighbors:
        if len(path_x)>15 or connected_region[nx, ny] == 1:
            path_x.append(nx)
            path_y.append(ny)
            break

# Create a function to add arrows to the plot
def add_arrow(ax, p1, p2, color):
    arrow = FancyArrowPatch(p1, p2, arrowstyle='->', mutation_scale=15, color=color)
    ax.add_patch(arrow)

# Plot the lattice, path with lines and arrows, and labels
plt.figure(figsize=(8, 8))  # Adjust the figure size as needed
plt.scatter(y_points, x_points, c='red', s=80, label='Region Points : D')
plt.scatter(boundary_points[1], boundary_points[0], c='blue', s=80, label='Boundary Points
for i in range(1, len(path_x)):
    plt.annotate("1/4", ((path_y[i] + path_y[i-1]) / 2, (path_x[i] + path_x[i-1]) / 2), fo
```

```
for i in range(1, len(path_x)):
    add_arrow(plt.gca(), (path_y[i-1], path_x[i-1]), (path_y[i], path_x[i]), 'green')
plt.annotate("start", (start_y, start_x), fontsize=20, ha='center')
# for i in range(1, len(path_x2)):
#     plt.annotate("1/4", ((path_y2[i] + path_y2[i-1]) / 2, (path_x2[i] + path_x2[i-1]) / 
# for i in range(1, len(path_x2)):
#     add_arrow(plt.gca(), (path_y2[i-1], path_x2[i-1]), (path_y2[i], path_x2[i]), 'black')

plt.xlabel('X Coordinate')
plt.ylabel('Y Coordinate')
plt.title('A Random Walk Path on a Square Graph')
plt.legend()
plt.gca().set_aspect('equal', adjustable='box')
plt.grid()

# Save the plot with higher DPI for better quality
plt.savefig('random_path.png', dpi=300)

# Show the plot
plt.show()
def m(k):
    if k==1: return 1
    return 8*(k-1)

def I(k):
    return np.eye(m(k))
def B(k):
    n = m(k)
    res = np.zeros((n, n))
    for i in range(n):
        res[i, (i+1)%n] = 0.25
        res[i, (i-1)%n] = 0.25
    return res
def Bp(k):
    n1 = m(k)
    c = int(n1/4)
    n2 = m(k-1)
    res = np.zeros((n1, n2))
    for d in range(4):
        for i in range(d*c+1, (d+1)*c):
            j = (i - (2*d + 1))%n2
            res[i, j] = 0.25
    return res
def Bn(k):
    return Bp(k+1).T
def A(k):
    if k==2:
        return Bn(1)
    else:
        return np.linalg.inv(I(k-1) - Bp(k-1)@A(k-1) - B(k-1))@Bn(k-1)
```