A Study of Hitting Times for Random Walks on Finite, Undirected Graphs

by

Ari Binder

Steven J. Miller, Advisor

A thesis submitted in partial fulfillment
of the requirements for the
Degree of Bachelor of Arts with Honors
in Mathematics

Williams College
Williamstown, Massachusetts

April 27, 2011

ABSTRACT

This thesis applies algebraic graph theory to random walks. Using the concept of a graph's fundamental matrix and the method of spectral decomposition, we derive a formula that calculates expected hitting times for discrete-time random walks on finite, undirected, strongly connected graphs. We arrive at this formula independently of existing literature, and do so in a clearer and more explicit manner than previous works. Additionally we apply primitive roots of unity to the calculation of expected hitting times for random walks on circulant graphs. The thesis ends by discussing the difficulty of generalizing these results to higher moments of hitting time distributions, and using a different approach that makes use of the Catalan numbers to investigate hitting time probabilities for random walks on the integer number line.

ACKNOWLEDGEMENTS

I would like to thank Professor Steven J. Miller for being a truly exceptional advisor and mentor. His energetic support and sharp insights have contributed immeasurably both to this project and to my mathematical education over the past three years. I also thank the second reader, Professor Mihai Stoiciu, for providing helpful comments on earlier drafts. I began this project in the summer of 2008 at the NSF Research Experience for Undergraduates at Canisius College; I thank Professor Terrence Bisson for his supervision and guidance throughout this program. I would also like to thank Professor Frank Morgan, who facilitated the continuation of my REU work by directing me to Professor Miller. Additionally I thank Professor Thomas Garrity for lending me his copy of Hardy's *Divergent Series*. Finally, I would like to thank the Williams Mathematics Department for providing me with this research opportunity, and my friends and family, whose support and camaraderie helped me tremendously throughout this endeavor.

CONTENTS

1. INTRODUCTION

Random walk theory is a widely studied field of mathematics with many applications, some of which include household consumption [Ha], electrical networks [Pa-1], and congestion models [Ka]. Examining random walks on graphs in discrete time, we quantify the expected time it takes for these walks to move from one specified vertex of the graph to another, and attempt to determine the distribution of these so-called *hitting times* (see Definitions 1.9 below) for random walks on the integer number line.

We start by recording basic definitions from graph theory that are relevant to our study, and providing a basic example. The informed reader can skip this subsection, and the reader seeking a more rigorous introduction to algebraic graph theory should consult [Bi].

Definitions 1.1. A *graph* $G = (V, E)$ is a vertex set V combined with an edge set E, where members of E connect members of V. We say G is an n-vertex graph if $|V| = n$. By convention we disallow multiple edges connecting the same two vertices and self-loops; that is, there is at most one element of E connecting i to j when $i \neq j$, and no edges connecting i to j when $i = j$, where $i, j \in V$. Furthermore, we call G *undirected* when an edge connects i to j if an only if an edge connects j to i.

Unless otherwise specified, we assume G is undirected throughout this paper, and thus define an edge as a two-way path from i to j where $i, j \in V$ and $i \neq j$.

Definition 1.2. We say that a graph G is *strongly connected* if for each vertex i, there exists at least one path from i to any other vertex j of G.

Definition 1.3. When an undirected edge connects i and j, we say that i and j are *adjacent,* or that i and j are *neighbors.* We denote adjacency as $i \sim j$.

Definition 1.4. We call an n-vertex graph *k-regular* if there are exactly k edges leaving each vertex, where $1 \leq k < n$ and $n > 1$.

Definition 1.5. The n-vertex graph G is *vertex-transitive* if its group of automorphisms acts transitively on its vertex set V. This simply means that all vertices look the same locally, i.e. we cannot uniquely identify any vertex based on the edges and vertices around it.

Definition 1.6. An n-vertex graph G is *bipartite* if we can divide V into two disjoint sets U and W such that no edge in E connects two vertices from U or two vertices from W. Equivalently, all edges connect a vertex in U to a vertex in W.

The undirected 6-cycle, two representations of which are given below in Figure 1, provides an easy way to visualize the above definitions, and is the example we will be using throughout the thesis.

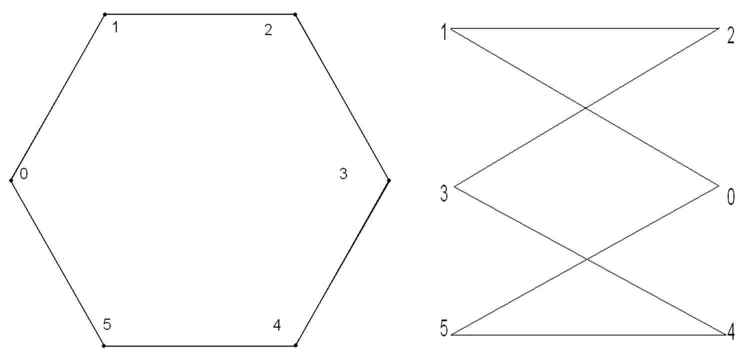

FIGURE 1. Two representations of the undirected 6-cycle.

From Definitions 1.2 and 1.4, we can immediately see that the undirected 6-cycle is strongly connected and 2-regular. Furthermore, Definition 1.5 tells us that the undirected 6-cycle is vertex-transitive, because the labeling of the vertices does not matter. Whatever labeling we choose, each vertex will still be connected to two other vertices by undirected edges, and thus labeling is the only way we can uniquely identify each vertex. In fact, if we assume undirectedness, no multiple edges, and no self-loops, then k-regularity is equivalent to vertex-transitivity. Finally, from Definition 1.6 and from the graph on the right in Figure 1, we can see that the undirected 6-cycle is bipartite, with disjoint vertex sets $U = \{1, 3, 5\}$ and $W = \{2, 0, 4\}$.

Definition 1.7. A *random walk* on an n-vertex graph G is the following process:

- Choose a starting vertex i of G.
- Let $A \subset V = \{j : i \sim j\}$. Choose an element j of A uniformly at random.
- If not at the walk's *stopping time* (see Definition 2.4 below), move to j, and repeat the second step. Otherwise end at j.

Definition 1.8. Consider a random walk on an n-vertex graph G, where G is isomorphic to a bipartite graph. We refer to this phenomenon by saying either that the walk or G is *periodic*. If G is not isomorphic to a bipartite graph, then the walk and G are *aperiodic*.

Thus when referring to G, we use "bipartite" and "periodic" interchangeably. This is a slight abuse of the terminology: processes are periodic, not graphs. However, we are comfortable describing a graph as periodic in this paper because we only discuss graphs in the context of random walks.

Note that our definition of a random walk is a discrete time definition. All of the results in this paper apply to discrete time random walks, but are convertible to continuous time, according to [Al-F1].

Definitions 1.9. The *hitting time* for a random walk from vertex i to a destination is the number of steps it takes the walk to reach the destination starting from i. Define the walk's destination as the first time it visits vertex j; then we denote the walk's *expected hitting time* as $E_i[T_j]$, and the variance of the walk's hitting time distribution as $Var_i[T_j]$.

Just as with any other expected value, we can intuitively define $E_i[T_j]$ as a weighted average. Thus $E_i[T_j] = \sum_{n=0}^{\infty} n \cdot P(\text{walk first reaches j starting from i in n steps})$. We can define $Var_i[T_j]$ in the same probabilistic manner. Because of the infinite number of possible random walks we must take into account, calculating $E_i[T_j]$ and $Var_i[T_j]$ values from the probabilistic definition appears to be a very difficult task to carry out. In this thesis we develop a more tractable method for quantifying hitting time distributions. Our main result is the use of spectral decomposition of the *transition matrix* (see Definitions 2.2 below) to produce a natural formula for the calculation of expected hitting times on n-vertex, undirected, strongly connected graphs. The rest of the thesis is organized as follows: Section 2 constructs a theoretical conceptualization of hitting times for random walks on graphs, Section 3 uses spectral decomposition to develop a hitting time formula and applies primitive roots to the construction of spectra of transition matrices, Section 4 applies our methodology in sample calculations, Section 5 attempts to quantify hitting time distributions on the number line, and Section 6 concludes.

2. A Theoretical Foundation for Random Walks and Hitting Times

In the following subsections, we construct a rigorous framework that both supports natural observations about random walks on graphs and allows us to move toward our goal of calculating expected hitting times for random walks on finite graphs. The reader is instructed to see the literature for proofs of any unproved assertions below.

2.1. **Understanding Random Walks on Graphs in the Context of Markov Chains.**

We appeal to [No] for the following association of random walks on graphs to discrete time Markov chains. The following definitions and results are summaries of relevant results from [No], Sections [1.1] and [1.4]. See [Sa] as well for a background on finite Markov chains.

Definitions 2.1. Let I be a countable set. Denote $i \in I$ as a *state* and I as the *state-space.* Furthermore, if $\rho = \{\rho_i : i \in I\}$ is such that $0 \leq \rho_i \leq 1$ for all $i \in I$, and $\sum_{i \in I} \rho_i = 1$, then ρ is a *distribution* on I.

Consider a random variable X, and let $\rho_i = P(X = i)$; then X has distribution ρ, and takes the value i with probability ρ_i. Let X_n refer to the state at time n; introducing time-dependence allows us to use matrices to conceptualize the process of changing states.

Definitions 2.2. We say that a matrix $P = \{p_{ij} : i, j \in I\}$ is *stochastic* if every row is a distribution, which is to say that each entry is non-negative and each row sums to 1. Furthermore, P is a *transition matrix* if P is stochastic such that $p_{ij} = P(X_{n+1} = j | X_n = i)$, where p_{ij} is independent of $X_0, \ldots, X_{n-1}$. We express $P(X_n = j | X_0 = i)$ as $p_{ij}^{(n)}$.

Definition 2.3. (See Theorem [1.1.1] of [No]). A *Markov chain* is a set $\{X_n\}$ such that X_0 has distribution ρ, and the distribution of X_{n+1} is given by the i^{th} row of the transition matrix P for $n \geq 0$, given that $X_n = i$.

Definition 2.4. A random variable S, where $0 \leq S \leq \infty$, is called a *stopping time* if the events $\{S = n\}$, where $n \geq 0$, depend only on $X_1, \ldots, X_n$.

These definitions allow us to introduce a property that is necessary for our conception of hitting times for random walks on graphs.

Theorem 2.5. (Norrris's strong Markov property; see Theorem [1.4.2] of [No]). Let $\{X_n\}$ be a Markov chain with initial distribution ρ and transition matrix P, and let S be a stopping time of $\{X_n\}$. Furthermore, let $X_S = i$. Then $\{X_{S+n}\}$ is a Markov chain with initial distribution δ and transition matrix P, where $\delta_j = 1$ if $j = i$, and 0 otherwise.

We now apply random walks on graphs to this Markov chain framework.

Proposition 2.6. Consider a random walk starting at vertex i of an n-vertex graph G. Let S be a stopping time such that $X_S = i$. Then the random walk, together with S, exhibits Norris's strong Markov property.

Proof. First of all, define X_t as the position of the walk at time $t,\ t \geq 0$. Trivially, we see that $X_0 = i$. Let $\rho = \{\rho_j : 1 \leq j \leq n\}$, where $\rho_j = 1$ if $j = i$ and 0 otherwise. Then X_0 has distribution ρ. Furthermore, define P as $\{p_{ij} : 1 \leq i, j \leq n\}$ where p_{ij} is the number of edges going from vertex i to j divided by the total number of edges leaving i. From the definition of a random walk given above, we see that $p_{ij} = P(X_{t+1} = j | X_t = i)$. Furthermore, we calculate p_{ij} from the structure of the graph; the positions of the walk have no bearing on the transition probabilities. Thus p_{ij} is independent of $X_0, \ldots, X_{t-1}$. Then $\{X_t\}$ is a Markov chain with initial distribution ρ and transition matrix P.

We now consider S. The event set $\{S = n\}$ refers to the set of occurrences in which after n steps, the walk is again at vertex i, where $n \geq 0$. Clearly these events depend only on the first n steps of the walk; hence S is a stopping time. Thus according to Theorem 2.5, $\{X_{S+t}\}$ is a Markov chain with initial distribution $\delta_i = \rho$ and transition matrix P. □

This result is simply saying that the Markov process "begins afresh" after each stopping time is reached, and is intuitively clear. In the next subsection, we use this framework to develop a theoretical conception of hitting times for random walks on graphs.

2.2. **Interpreting Hitting Times as a Renewal-Reward Process.** In accordance with [Al-F1], we use renewal-reward theory as a way to think about hitting times for random walks on graphs. Doing so is a prerequisite for all the results we obtain regarding hitting times. We appeal to [Co] for the following outline of a renewal-reward process.

Definitions 2.7. Let S_1, S_2, S_3, S_4, $\ldots$ be a sequence of positive i.i.d.r.v.s such that $0 < E[S_i] < \infty$. Denote $J_n = \sum_{i=1}^{n} S_i$ as the n^{th} *jump time* and $[J_{n-1}, J_n]$ as the n^{th} *renewal interval.* Then the random variable $(R_t)_{t \geq 0}$ such that $R_t = \sup\{n : J_n \leq t\}$ is called a *renewal process.* R_t refers to the number of renewals that have occurred by time t. Now let W_1, W_2, W_3, W_4, $\ldots$ be a sequence of i.i.d.r.v.s such that $-\infty < E[W_i] < \infty$. Then $Y_t = \sum_{i=1}^{R_t} W_i$ is a *renewal-reward process.*

We can see a natural relation between Definitions 2.7 and Markov chains, which we prove as the following proposition.

Proposition 2.8. Consider a random walk on an n-vertex graph G starting from vertex i. Consider a sequence of stopping times $S_1, S_2, S_3, \ldots$ such that $X_{S_n} = i$ for all $n \geq 1$. Let Y_t refer to the total number of visits to an arbitrary vertex j that have occurred by the most recent stopping time. Then Y_t is a renewal-reward process.

Proof. Let S_n refer to the number of steps the walk takes between the $(n-1)^{st}$ and n^{th} stopping times, such that $X_{S_1+S_2+\ldots+S_{n-1}} = X_{S_1+S_2+\ldots+S_n} = i$. Note that by Proposition 2.6, the random walk on G together with the stopping time S_n satisfies Theorem 2.5, which is to say that the portion of the walk between the $(n-1)^{st}$ and n^{th} stopping times is independent of all previous portions of the walk. Furthermore, the starting and ending positions of each portion of the walk between stopping times are the same, and are governed by the same transition matrix P, also in accordance with the strong Markov property. Thus $S_1, S_2, S_3, S_4, \ldots$ is a sequence of i.i.d.r.v.s such that $0 < E[S_i] < \infty$.

Now, let $J_n = \sum_{i=1}^{n} S_i$ refer to the number of steps the walk has taken by the time the n^{th} stopping time occurs. Furthermore let $R_t = \sup\{n : J_n \leq t\}$ be the number of stopping times that have been reached by time t. Then according to Definitions 2.7, R_t is a renewal process. Now let V_n be the number of visits to vertex j between the $(n-1)^{st}$ and n^{th} stopping times. Once again, because the portion of the walk between the $(n-1)^{st}$ and n^{th} stopping times is independent of all previous portions of the walk; for all n, the portion of the walk between the $(n-1)^{st}$ and n^{th} stopping times share the same start and endpoints; and the entire walk is governed by transition matrix P, we conclude that $V_1, V_2, V_3, V_4, \ldots$ is a sequence of i.i.d.r.v.s such that $-\infty < E[V_i] < \infty$. Therefore $Y_t = \sum_{i=1}^{R_t} V_i$ is a renewal-reward process, and Y_t is the number of visits to vertex j by the most recent stopping time. $\square$

We now introduce some key insights that allow us to think about how to calculate expected hitting times on strongly connected finite graphs. The reader is instructed to see [No] for proofs.

Theorem 2.9. (The Asymptotic Property of Renewal-Reward Processes). Consider renewal-reward process $Y_t = \sum_{j=1}^{R_t} W_j$ as described in Definitions 2.7. Assume the process begins in state i. Then $\lim_{t\to\infty} \frac{Y_t}{t} = \frac{E_i[W_1]}{E_i[S_1]}$.

Definitions 2.10. (Irreducibility of the Transition Matrix). If an n-vertex graph G is strongly connected, then we say its transition matrix P is *irreducible.* Irreducibility implies the existence of a unique probability distribution π on the n vertices of G such that $\pi_j = \sum_{i=1}^{n} \pi_i p_{ij}$ for $1 \leq j \leq n$. We refer to π as the graph's *stable distribution.*

Theorem 2.11. (The Ergodic Theorem; see Theorem 1 of [Al-F1]). Let $N_j(t)$ be the number of visits to vertex j during times $1, 2, \ldots, t$. Then for any initial distribution, $t^{-1}N_j(t) \to \pi_j$ as $t \to \infty$, where $\pi = (\pi_1, \ldots, \pi_j, \ldots, \pi_n)$ is the stable distribution.

With these insights, we arrive at a crucial proposition.

Proposition 2.12. (See Proposition 3 of [Al-F1]). Consider a random walk starting at vertex i on an n-vertex, strongly connected graph G. Furthermore, let S be a stopping time such that $X_S = i$ and $0 < E_i[S] < \infty$. Then $E_i[\text{number of visits to j before time S}]$ $= \pi_j E_i[S]$ for each vertex j.

Proof. We bring the previous results to bear. Consider a sequence of stopping times $S_1, S_2, S_3, \ldots$, and assume that these stopping times, together with a stopping time S, are all governed by the same distribution. Let Y_t refer to the total number of visits to an arbitrary vertex j that have occurred by the most recent stopping time. By Proposition 2.8, Y_t is a renewal-reward process. Then according to the asymptotic property of renewal-reward processes, $\lim_{t\to\infty} \frac{Y_t}{t} = \frac{E_i[W_1]}{E_i[S_1]}$. Since S is distributed identically with $S_1, S_2, S_3, \ldots$, replace S_1 with S and W_1 with W in the above expression:

$$\lim_{t\to\infty} \frac{Y_t}{t} = \frac{E_i[W]}{E_i[S]}.$$

Now by Definitions 2.10, because G is strongly connected, its corresponding transition matrix P is irreducible. Hence we can use the Ergodic Theorem to express the asymptotic proportion of time the walk spends visiting vertex j:

$$\lim_{t\to\infty} \frac{N_j(t)}{t} = \pi_j,$$

where $N_j(t)$ is the total number of visits to vertex j by the t^{th} step of the walk. Note that $N_j(t) \neq Y_t$; rather, $Y_t = N_j(S_{\text{most recent}})$. However, relative to t, $N_j(t) - Y_t$ becomes arbitrarily small as $t \to \infty$. Hence the two limits are equal, and thus we have

$$\frac{E_i[W]}{E_i[S]} = \pi_j$$

or $E_i[W] = \pi_j E_i[S]$. Noting that W refers to the number of visits to j before stopping time S, we arrive at our result. $\square$

2.3. **Two Basic Results.** Proposition 2.12 is very useful because as long as S is a stopping time such that $X_S = i$, we can choose S to be whatever we want. By smart choices of S and clever manipulation, [Al-F1] uses this proposition to prove many useful properties, some of them regarding expected hitting times. We prove one of these properties below.

Result 2.13. (See Lemma 5 of [Al-F1]). Consider a random walk starting from vertex i on an n-vertex, strongly connected graph G. Define T_i^+ as the *first return time* to i. Then

$$E_i[T_i^+] = \frac{1}{\pi_i}.$$

Proof. Substituting i for j and T_i^+ for S in Proposition 2.12, we get

$$E_i[\text{number of visits to i before time T}_\text{i}^+] = \pi_i E_i[T_i^+].$$

We follow [Al-F1]'s convention, and include time 0 and exclude time t in accounting for the vertices a t-step walk visits. Note that we visit vertex i exactly once by the first return time: the visit occurs exactly at time 0, since we exclude time T_i^+. Substituting 1 in for $E_i[\text{number of visits to i before time T}_\text{i}^+]$, we arrive at our result. □

Applying Result 2.13 to k-regular graphs, we can quantify specific hitting times, as shown below.

Proposition 2.14. Consider an n-vertex, k-regular, undirected, strongly connected graph G. Let π represent the stable distribution on G's vertices. Then π is the uniform distribution.

Proof. First of all, because G is strongly connected, the transition matrix P is irreducible, which implies the existence of a stable distribution π. To show π is uniform, note that since G is k-regular, k edges leave each vertex. Furthermore, because G is undirected, we know that the number of edges leaving vertex i is equal to the number of edges leading to i for each vertex i of G. Thus $P(X_{t+1} = j|X_t = i) = P(X_{t+1} = i|X_t = j)$. This implies that $p_{ij} = p_{ji}$ for $1 \leq i, j \leq n$, which means that P is symmetric. Furthermore, if i is connected to j, then $p_{ij} = p_{ji} = 1/k$, and $p_{ij} = p_{ji} = 0$ otherwise. Because each vertex is connected to k vertices, we see that each row of P, and hence each column as well by symmetry, contains k non-zero terms, each equalling $1/k$. Now, let $\pi_i = 1/n$ for $1 \leq i \leq n$. We see that

$$\begin{aligned}
\pi_i &= \frac{1}{n} \\
&= \pi_j \text{ for } 1 \ \leq \text{j} \leq \text{n} \\
&= k \cdot \pi_j \cdot \frac{1}{k} \\
&= k \cdot \pi_j \cdot \frac{1}{k} + (n-k) \cdot \pi_j \cdot 0 \\
&= \sum_{j=1}^{n} \pi_j p_{ji},
\end{aligned}$$

which is the irreducibility condition. Thus the stable distribution is uniform, and is unique by Definition 2.10. $\square$

From Proposition 2.14 and Result 2.13, we can immediately see that the expected first return time for undirected, k-regular, strongly connected graphs is simply equal to n, the number of vertices on the graph:

$$E_i[T_i^+] = n. \tag{2.14a}$$

From (2.14a) we can infer another result regarding expected hitting times for random walks on this class of graphs.

Result 2.15. Assume vertices i and j are adjacent. Consider a random walk starting at vertex j on an n-vertex k-regular undirected graph G. Then $E_j[T_i] = n - 1$.

Proof. Consider a random walk on G starting at vertex i. After one step, the walk is at j with probability $1/k$ for all j such that $j \sim i$. Thus we see the following:

$$\begin{aligned} E_i[T_i^+] &= n \\ &= 1 + \frac{1}{k} \sum_{i \sim j} E_j[T_i] \\ &= 1 + \frac{1}{k} \cdot k \cdot E_j[T_i] \\ &= 1 + E_j[T_i], \end{aligned}$$

yielding $E_j[T_i] = n - 1$ for $j \sim i$. $\square$

When G is regular, it is possible to construct general formulas for $E_i[T_j]$ for the cases when the shortest path from i to j is of length 2 or 3 (see [Pa-2]). In what follows, we work toward an explicit formula that does not depend on the regularity of G, and yields expected hitting time values for random walks on any finite, undirected, strongly connected graph. We employ spectral decomposition of G's transition matrix to achieve such a formula. This application of the spectral decomposition of P to hitting times is well-known; see [Al, Al-F2, Br, Ta], for instance. However, this literature assumes the background we have just provided, it often gives results in continuous time, and it does not explicitly state a formula for determining $E_i[T_j]$ values. Thus while the following methodology is by no means novel, by fully deriving and stating the formula that it permits, and by remaining in discrete time, we present hitting time results in what we believe to be a clearer, more straightforward, and more intuitive manner than the existing literature.

3. A Methodology for Calculating Expected Hitting Times

3.1. **The Fundamental Matrix and its Role in Determining Expected Hitting Times.** With Proposition 2.12 and a clever choice of S, [Al-F1] defines the *fundamental matrix*, from which we can calculate expected hitting times for random walks on any graph with stable distribution π. We paraphrase their ingenuity in this subsection.

Definition 3.1. Consider a graph G with irreducible transition matrix P and stable distribution π. The graph's *fundamental matrix* Z is defined as the matrix such that

$$Z_{ij} = \sum_{t=0}^{\infty}(p_{ij}^{(t)} - \pi_j).$$

It is verifiable that the fundamental matrix Z has rows summing to 0, and when π is uniform, constant diagonal entries (proved as Proposition 7.3 in the appendix). These properties will come into play later in our discussion, especially in manipulating the following definition.

Definition 3.2. Consider a random walk on an n-vertex graph G in which for each vertex j, $P(\text{walk starts at j}) = \pi_j$. Then we refer to the expected number of steps it takes the walk to reach vertex i as $E_\pi[T_i]$.

Theorem 3.3. (The Convergence Theorem; See Theorem 2 of [Al-F1]). Consider a Markov chain $\{X_n\}$. Then $P(X_t = j) \to \pi_j$ as $t \to \infty$ for all states j, provided the chain is aperiodic.

These definitions set us up to prove our first general result about expected hitting times.

Proposition 3.4. (See Lemma 11 of [Al-F1]). For an n-vertex graph G with fundamental matrix Z and stable distribution π,

$$E_\pi[T_i] = \frac{Z_{ii}}{\pi_i}.$$

Proof. We get at the notion of Z in the following manner. Consider a random walk on G starting at vertex i. Let $t_0 \geq 1$ be some arbitrary stopping time and define S as the time taken by the following process:

- wait time t_0,
- and then wait, if necessary, until the walk next visits vertex i.

Note that after time t_0, the graph has probability distribution δ, where $P(X_{t_0} = j) = \delta_j$ for $1 \leq j \leq n$. Hence $E[S] = t_0 + E_\delta[T_i]$. Furthermore, note that the number of visits to i by time t_0 is geometrically distributed, because we can treat $X_n = i$ as a Bernoulli Trial for all n from 1 to $t_0 - 1$. Also, it easy to see that the walk makes no visits to i after time $t_0 - 1$ (remember that we do not include time S in our accounting). Therefore we can express $E_i[N_i(t_0)]$ as the sum of the i to i transition probabilities:

$$\begin{aligned} E[N_i(S)] &= \sum_{n=1}^{t_0-1} p_{ii}^{(n)} + 1 \\ &= \sum_{n=0}^{t_0-1} p_{ii}^{(n)} \text{ because the walk starts at i.} \end{aligned}$$

Invoking Proposition 2.12 and setting $j = i$, we achieve the following expression:

$$\sum_{n=0}^{t_0-1} p_{ii}^{(n)} = \pi_i(t_0 + E_\delta[T_i]).$$

Rearranging, we get

$$\sum_{n=0}^{t_0-1} (p_{ii}^{(n)} - \pi_i) = \pi_i E_\delta[T_i].$$

Now let $t_0 \to \infty$. By definition, the left hand term converges to Z_{ii}. Furthermore, by the convergence theorem, as t_0 becomes large, δ approaches π if G is aperiodic. Therefore the expression becomes $Z_{ii} = \pi_i E_\pi[T_i]$, and we arrive at our result.

Now let G be periodic. Then by definition X_{t_0} oscillates between disjoint vertex subsets U and W as $t_0 \to \infty$. Without loss of generality assume that when n is odd, $P(X_n \in U) = 0$ and when n is even, $P(X_n \in W) = 0$. Thus $i \in U$, meaning that for odd n, $p_{ii}^{(n)} - \pi_i = -\pi_i$, a constant. Hence the infinite series does not converge. However, if we scale each partial sum by $1/n$, these terms become arbitrarily small as $n \to \infty$. Denoting the m^{th} partial sum of the series as s_m, we see that $\lim_{t_0 \to \infty} \frac{1}{t_0-1} \sum_{m=0}^{t_0-1} s_m$ exists. Thus the series $\sum_{n=0}^{t_0-1} (p_{ii}^{(n)} - \pi_i)$ is Cesàro summable (see [Har], for instance), and by definition of Z, the Cesàro sum is equal to Z_{ii}. Considering the right hand side, note that by the convergence theorem, for odd n, as $n \to \infty$, the distribution of X_n converges to a stable distribution on the elements of W; we call this distribution π_W. Furthermore, for even n, as $n \to \infty$, the distribution of X_n converges to a stable distribution on the elements of U; call this distribution π_U. Define δ_n as the distribution of X_n on the vertices of G; then the odd terms of the sequence $\{\delta_n\}$ approach π_W and the even terms approach π_U. Consider

$\rho_n = \frac{1}{n}\sum_{i=0}^{n}\delta_n$; then for large n,

$$\rho_n \approx \frac{1}{n}\left(\frac{n}{2}\pi_W + \frac{n}{2}\pi_U\right)$$
$$= \frac{1}{2}(\pi_W + \pi_U),$$

or equivalently,

$$\lim_{n\to\infty}\rho_n = \frac{1}{2}(\pi_W + \pi_U).$$

Hence the Cesàro average of this sequence exists.

Now, consider a random walk on G that starts from vertex i, where $i \in U$ with probability $1/2$, and $i \in W$ with probability $1/2$. Then from this starting distribution δ_0, $\{\delta_n\}$ converges to $\pi = \frac{1}{2}(\pi_W + \pi_U)$, as indicated above. Using Cesàro summation as above, we again see that

$$\sum_{n=0}^{t_0-1}(p_{ii}^{(n)} - \pi_i) \to Z_{ii} \text{ for } t_0 \to \infty,$$

where p_{ii} refers to the average of the two i to i transition probabilities. Therefore $Z_{ii} = \pi_i E_\pi[T_i]$, completing the proof for the periodic case. □

Proposition 3.4 allows us to calculate expected hitting times for a random walk to a specific vertex when the starting vertex is determined by the stable distribution. We seek to extend this result so that we can determine expected hitting times for explicit walks from vertex i to vertex j. We use the following lemma to do so.

Lemma 3.5. (See Lemma 7 of [Al-F1]). Consider a random walk on a graph G with stable distribution π. Then, for $j \neq i$,

$$E_j[\text{number of visits to j before time T}_\text{i}] = \pi_j(E_j[T_i] + E_i[T_j]).$$

Proof. Assume the walk starts at j. Define S as "the first return to j after the first visit to i." Then we have $E_j[S] = E_j[T_i] + E_i[T_j]$. Because there are no visits to i before T_i and no visits to j after T_i by our accounting convention, substitution into Proposition 2.12 yields the result. □

We are now ready to associate Z with explicit hitting times, and do so below.

Proposition 3.6. (See Lemma 12 of [Al-F1]). For an n-vertex graph G with fundamental matrix Z and stable distribution π,

$$E_i[T_j] = \frac{Z_{jj} - Z_{ij}}{\pi_j}.$$

Proof. We proceed in the same manner as above. Consider a random walk on G starting at vertex j. Define S as the time taken by the following process:

- wait until the walk first reaches vertex i,
- then wait a further time t_0,
- and finally wait, if necessary, until the walk returns to j.

Define δ as the probability distribution such that $\delta_k = P_k(X_{t_0} = k)$. Then $E_j[S] = E_j[T_i] + t_0 + E_\delta T_j$. Furthermore note that between times T_i and t_0, the number of visits to j is geometrically distributed, so

$$E_j[N_j(t_0)] = E_j[\text{number of visits to j before time T}_\text{i}] + \sum_{t=0}^{t_0-1} p_{ij}^{(t)}.$$

Once again we appeal to Proposition 2.12, and by the same reasoning as above, we achieve

$$E_j[\text{number of visits to j before time T}_\text{i}] + \sum_{t=0}^{t_0-1} p_{ij}^{(t)} = \pi_j(E_j[T_i] + t_0 + E_\delta T_j).$$

Subtracting $E_j[\text{number of visits to j before time T}_\text{i}]$ from both sides and appealing to Lemma 3.5, we obtain

$$\sum_{t=0}^{t_0-1} p_{ij}^{(t)} = \pi_j(-E_i[T_j] + t_0 + E_\delta T_j).$$

Rearranging yields

$$\sum_{t=0}^{t_0-1} (p_{ij}^{(t)} - \pi_j) = \pi_j(E_\delta T_j - E_i[T_j]).$$

Once again, we let $t_0 \to \infty$ and end up with

$$Z_{ij} = \pi_j(E_\pi[T_j] - E_i[T_j]).$$

Finally, Proposition 3.4 allows us to write

$$Z_{ij} = Z_{jj} - \pi_j E_i[T_j],$$

and the proof is complete for the aperiodic case. When G is periodic, we employ the same argument as above, with the addition that if $j \in U$, $p_{ij}^{(n)} = 0$ for odd n, and if $j \in W$, then $p_{ij}^{(n)} = 0$ for even n. $\square$

As Aldous and Fill observe in [Al-F1], because hitting time results are transferable to continuous time, where the periodicity issue does not arise, it is easier to switch to continuous time to evaluate random walks that are periodic in the discrete case. However, we believe that remaining in discrete time is fruitful because it gives us a more tangible sense of the oscillatory nature of discrete time random walks on bipartite graphs.

3.2. **Making Sense of the Fundamental Matrix.** As we can see from the above section and ultimately from Proposition 3.6, the fundamental matrix is a powerful tool for our purposes, so long as we can determine it. However, for most graphs it is essentially impossible to calculate the sum of this infinite series, just as it seems difficult to calculate an expected hitting time from its probabilistic definition. We therefore seek to equate this definition of Z with a more useable one. To do so, we make use of the following lemma. While the results of the next two sections are known, we arrive at them independently of the existing literature.

Lemma 3.7. Let graph G have irreducible transition matrix P. Define P_∞ as the n by n matrix such that $P_{\infty_{ij}} = \pi_j$ for all i. Then $PP_\infty = P_\infty = P_\infty P$. Furthermore, $P_\infty P_\infty = P_\infty$.

Proof. Define J_n as the n column vector of all ones. Thus we have $PP_\infty = PJ_n\pi$. Now, by stochasticity of P, $PJ_n = J_n$, so $PP_\infty = J_n\pi = P_\infty$. Similarly, $P_\infty P = J_n\pi P$. Since P is irreducible, $\pi P = \pi$. So, $P_\infty P = J_n\pi = P_\infty$. To show $P_\infty P_\infty = P_\infty$, we write the following:

$$\begin{aligned}
(P_\infty P_\infty)_{ij} &= \pi_1\pi_j + \pi_2\pi_j + \cdots + \pi_n\pi_j \\
&= \pi_j(\pi_1 + \pi_2 + \cdots + \pi_n) \\
&= \pi_j \;\text{ because } \pi \text{ is a distribution} \\
&= P_{\infty_{ij}}.
\end{aligned}$$

□

The commutativity of P with respect to P_∞ is a crucial property, and we use it to express Z in terms of matrices we can easily calculate.

Proposition 3.8. Consider an n-vertex graph G with irreducible transition matrix P and fundamental matrix Z. Then

$$\sum_{t=0}^{\infty} \left(P^t - P_\infty\right) = Z = (I - (P - P_\infty))^{-1} - P_\infty.$$

Proof. Consider $\sum_{t=0}^{\infty} (P^t - P_\infty)$ and let $t = 0$. In this case, $P^0 - P_\infty = I - P_\infty = (P - P_\infty)^0 - P_\infty$. For $t \geq 1$, we appeal to the binomial theorem, and note that $(1+x)^t = \sum_{i=0}^{t} \binom{t}{i} x^i$. Now, let $x = -1$. Clearly

$$0 = (1 + (-1))^t = \sum_{i=0}^{t} \binom{t}{i} (-1)^i.$$

This means that for $t \geq 1$, if we negate every other term of the expansion of $(P - P_\infty)^t$, the coefficients of the terms will sum to 0. Since P and P_∞ commute, this yields

$$\begin{aligned}(P - P_\infty)^t &= \binom{t}{0} P^t - \binom{t}{1} P^{t-1} P_\infty + \binom{t}{2} P^{t-2} P_\infty^2 - \cdots \pm \binom{t}{t} P_\infty^t \\ &= \binom{t}{0} P^t - \binom{t}{1} P_\infty + \binom{t}{2} P_\infty - \cdots \pm \binom{t}{t} P_\infty \text{ by Lemma 3.7} \\ &= P^t - P_\infty \text{ by the binomial theorem.}\end{aligned}$$

We get from the second line to the third line above by knowing that the coefficients of the terms sum to 0, and the first term, P^t, has coefficient 1. So we now have

$$\begin{aligned}\sum_{t=0}^{\infty} (P^t - P_\infty) &= (P - P_\infty)^0 - P_\infty + \sum_{t=1}^{\infty} (P - P_\infty)^t \\ &= \sum_{t=0}^{\infty} (P - P_\infty)^t - P_\infty.\end{aligned}$$

Consider $\sum_{t=0}^{\infty} (P - P_\infty)^t$. The first term of this geometric series is I, and the common factor is $(P - P_\infty)$. From linear algebra, we know that geometric series of powers of the matrix A converge if and only if A's eigenvalues are strictly less than 1 in absolute value. Now because P and P_∞ are real and symmetric, the eigenvalues of $(P - P_\infty)$ are real; furthermore, because P and P_∞ commute, they are simultaneously diagonalizable, and hence $\lambda_{P_i} - \lambda_{P_{\infty_i}} = \lambda_{(P-P_\infty)_i}$, where $\lambda_{P_i} P = \lambda_{P_i} \vec{v}$ and $\lambda_{P_{\infty_i}} P = \lambda_{P_{\infty_i}} \vec{v}$, and $\vec{v}$ is the eigenvector that corresponds to both λ_{P_i} and $\lambda_{P_{\infty_i}}$. Note that P_∞ has rank 1 and null space of dimension $n - 1$; therefore it has the eigenvalue 1 with multiplicity 1 and 0 with multiplicity $n - 1$. Let G be aperiodic; then P has eigenvalue 1 with multiplicity 1, and all other eigenvalues are strictly less than 1 in absolute value. Because P and P_∞ are doubly stochastic, the only eigenvector $\vec{v}$ such that $P\vec{v} = \vec{v}$ and $P_\infty = P_\infty \vec{v}$ is J_n. Thus $\lambda_{P_1} = \lambda_{P_{\infty_1}} = 1$, and $\lambda_{(P-P_\infty)_1} = 1 - 1 = 0$. Furthermore, because $\lambda_{P_{\infty_i}} = 0$ for $2 \leq i \leq n$ and $|\lambda_{P_{\infty_i}}| < 1$ for $2 \leq i \leq n$, we conclude that the eigenvalues of $A = (P - P_\infty)$ are strictly less than 1 in absolute value. Therefore, from linear algebra, the series converges to a sum satisfying

$$(I - A)(I + A + A^2 + \cdots) = I.$$

Because A has eigenvalues all less than 1 in absolute value, $(I - A)$ has all nonzero eigenvalues, and hence is invertible. Thus the sum of the series is $(I - A)^{-1} = (I - (P - P_\infty))^{-1}$, completing the proof for the aperiodic case:

$$\begin{aligned}Z &= \sum_{t=0}^{\infty}(P^t - P_\infty)\\ &= \sum_{t=0}^{\infty}(P - P_\infty)^t - P_\infty\\ &= (I - (P - P_\infty))^{-1} - P_\infty.\end{aligned}$$

When G is periodic, -1 is an eigenvalue of P, a result which we prove as Proposition 7.2 in the appendix. Because P and P_∞ are simultaneously diagonalizable, this eigenvalue of -1 from P is paired with an eigenvalue of 0 from P_∞, and so $P - P_\infty$ has an eigenvalue of -1. Thus $\sum_{t=0}^{\infty}(P - P_\infty)^t$ oscillates in the limit, which we would expect based on the proof of Proposition 3.4 for the periodic case. We appeal to Abel summation (see [Har]) to evaluate this sum. To do so, let $0 < \alpha < 1$. Write

$$\begin{aligned}&\sum_{t=0}^{\infty}(P^t - P_\infty)\alpha^t\\ = &\sum_{t=0}^{\infty}(P - P_\infty)^t\alpha^t - P_\infty \text{ from above}\\ = &\sum_{t=0}^{\infty}(\alpha(P - P_\infty))^t - P_\infty.\end{aligned}$$

Because $|\alpha| < 1$, $\alpha(P - P_\infty)$ has eigenvalues all strictly less than 1 in absolute value. Once again, because $\alpha(P - P_\infty)$ has no eigenvalues equal to 1, $(I - \alpha(P - P_\infty))$ has all nonzero eigenvalues and hence is invertible. Therefore the sum $\sum_{t=0}^{\infty}(P^t - P_\infty)\alpha^t$ converges to $(I - \alpha(P - P_\infty))^{-1} - P_\infty$. Making use of this equality and noting that the Abel sum of $\sum_{t=0}^{\infty}(P - P_\infty)^t$ is simply the limit of $\sum_{t=0}^{\infty}(P^t - P_\infty)\alpha^t$ as α approaches 1 from below, we find that

$$\sum_{t=0}^{\infty}(P - P_\infty)^t = (I - (P - P_\infty))^{-1}.$$

Note that from the proof of Proposition 3.6, the series $\sum_{t=0}^{\infty}(p_{ij}^{(t)} - P_{\infty_{ij}})$ is Cesàro summable for each i and j, with Cesàro sum equal to Z_{ij}. Hence $\sum_{t=0}^{n}(P^t - P_\infty)$ is Cesàro summable, with Cesàro sum equal to Z. Because the Abel sum exists and equals the Cesàro sum whenever the latter exists, we see that the Abel sum of $\sum_{t=0}^{n}(P^t - P_\infty)$ is equal to Z, and therefore $Z = (I - (P - P_\infty))^{-1} - P_\infty$ in the periodic case as well. $\square$

3.3. **Toward an Explicit Hitting Time Formula.** Using Proposition 3.8 and Proposition 3.6, we can quantify hitting times on any finite, strongly connected graph. However, the fundamental matrix is an abstract concept that is hard to visualize. It is hard to tell exactly where the actual values for the hitting times are coming from. Thus it would be nice if we could determine hitting times straight from the transition matrix, because we obtain the transition matrix directly from the graph itself. We work toward quantifying first $E_\pi[T_j]$ values and then $E_i[T_j]$ values using only the transition matrix and its spectrum in this subsection. Doing so requires associating the spectrum of the fundamental matrix with that of the transition matrix, which we accomplish with the following lemma.

Lemma 3.9. For an n-vertex graph G with irreducible transition matrix P and fundamental matrix Z, $\lambda_{Z_1} = 0$, and for $2 \leq m \leq n$,

$$\lambda_{Z_m} = (1 - \lambda_{P_m})^{-1},$$

where λ_{P_m} refers to the m^{th} largest eigenvalue of P.

Proof. From Proposition 3.8, we write $Z = (I - (P - P_\infty))^{-1} - P_\infty$. From linear algebra, we know that if two matrices commute with each other and are symmetric, then they are simultaneously diagonalizable. To show Z is symmetric, note that P is symmetric by undirectedness of G, P_∞ is symmetric by uniformity of π, and I is symmetric by definition. Because differences of symmetric matrices are symmetric and inverses of symmetric matrices are symmetric, $Z = (I - (P - P_\infty))^{-1} - P_\infty$ is symmetric.

We now must show that P and Z commute. By definition, P commutes with I and with itself, and by Lemma 3.7, P commutes with P_∞. Thus we have

$$\begin{aligned}
PZ &= P((I - (P - P_\infty))^{-1} - P_\infty) && \text{from Proposition 3.8} \\
&= P(I - (P - P_\infty))^{-1} - PP_\infty && \text{by distributivity of matrix multiplication} \\
&= P(I - (P - P_\infty))^{-1} - P_\infty P && \text{by Lemma 3.7.} \\
&= (I - (P - P_\infty))^{-1} P - P_\infty P && \text{because P commutes with I, } \mathrm{P}_\infty\text{, and itself} \\
&= ((I - (P - P_\infty))^{-1} - P_\infty) P && \text{by distributivity of matrix multiplication} \\
&= ZP.
\end{aligned}$$

Thus P commutes with Z, meaning that the two are simultaneously diagonalizable. This implies that for every eigenvalue λ_{Z_m} such that $Z\vec{v} = \lambda_{Z_m}\vec{v}$, $P\vec{v} = \lambda_{P_m}\vec{v}$, where $1 \leq m \leq n$ and $\vec{v}$ is an eigenvector in the shared eigenspace. Now we know that $\lambda_{P_1} = 1$ (see appendix) with eigenvector J_n. Hence λ_{Z_1} is also associated with J_n. Because Z has rows summing to 0 and J_n has constant entries, λ_{Z_1} must be 0.

For $2 \leq m \leq n$, simultaneous diagonalizability of P and Z combined with properties of eigenvalues allows us to write

$$\begin{aligned}
\lambda_{Z_m} &= \lambda_{((I-(P-P_\infty))^{-1}-P_\infty)_m} \\
&= \lambda_{(I-(P-P_\infty))_m^{-1}} - \lambda_{P_{\infty m}} \\
&= (\lambda_{I_m} - (\lambda_{P_m} - \lambda_{P_{\infty m}}))^{-1} - \lambda_{P_{\infty m}}.
\end{aligned}$$

As we observed in the proof of Proposition 3.8, $\lambda_{P_{\infty_k}} = 0$ for $2 \leq k \leq n$, so we get rid of these terms and arrive at our result:

$$\begin{aligned}
&(\lambda_{I_m} - (\lambda_{P_m} - \lambda_{P_{\infty m}}))^{-1} - \lambda_{P_{\infty m}} \\
=&(\lambda_{I_m} - \lambda_{P_m})^{-1} \\
=&(1 - \lambda_{P_m})^{-1}, \text{ which is defined for } 2 \ \leq \ \text{m} \ \leq \ \text{n}.
\end{aligned}$$

□

With this crucial lemma, which provides a natural formula relating the eigenvalues of Z to those of P, we are in position to explicitly quantify expected hitting times. Recall that $E_\pi[T_j]$ refers to the expected number of steps it takes a random walk to reach vertex j in which the starting vertex is determined by the stationary distribution. Thus it makes intuitive sense to think of these values as an average of hitting times with an explicit starting vertex. We prove this a priori intuition below.

Proposition 3.10. For a random walk on an n-vertex graph G with irreducible transition matrix P and uniform stable distribution π, we have

$$E_\pi[T_j] = \frac{1}{n}\sum_{j=1}^{n} E_i[T_j] = \sum_{m=2}^{n}(1-\lambda_{P_m})^{-1} = \tau \text{ for all j}.$$

Proof. By Proposition 3.4, $E_\pi[T_j] = \frac{1}{\pi_j}Z_{jj}$. Because the rows of Z sum to 0, we can write

$$E_\pi[T_j] = \frac{1}{\pi_j}Z_{jj} - \sum_{j=1}^{n} Z_{ij}.$$

Since the diagonal entries of Z are constant when π is uniform, this becomes

$$\begin{aligned} E_\pi[T_j] &= \frac{1}{n} \cdot \frac{1}{\pi_j} \sum_{j=1}^{n} (Z_{jj} - Z_{ij}) \\ &= \frac{1}{n} \sum_{j=1}^{n} \frac{Z_{jj} - Z_{ij}}{\pi_j} \\ &= \frac{1}{n} \sum_{j=1}^{n} E_i[T_j] \ \text{ by Proposition 3.6.} \end{aligned}$$

Now, when π is uniform, Proposition 3.4 gives $E_\pi[T_j] = \frac{1}{\pi_j} Z_{jj} = nZ_{jj}$. Using m in the same way as in Lemma 3.9, we write

$$\begin{aligned} nZ_{jj} &= Tr(Z) \ \text{ because Z has constant diagonal entries for uniform } \pi \\ &= \sum_{m=1}^{n} \lambda_{Z_m} \\ &= \sum_{m=2}^{n} \lambda_{Z_m} \\ &= \sum_{m=2}^{n} (1 - \lambda_{P_m})^{-1} \text{ by Lemma 3.9.} \end{aligned}$$

□

Thus when π is uniform, the average of the expected hitting times for random walks on G is simply equal to the sum of the fundamental matrix's eigenvalues; and because of the natural relation between Z and P, we can easily express this sum in terms of the eigenvalues of P. This is a very nice result. Spectral decomposition of Z and P allows us to determine a similarly nice formula for $E_i[T_j]$ values, and one that does not rely on uniformity of π.

Proposition 3.11. For a random walk on an n-vertex graph G with irreducible transition matrix P, for any vertices i and j,

$$E_i[T_j] = \frac{1}{\pi_j} \sum_{m=2}^{n} (1 - \lambda_{P_m})^{-1} u_{jm}(u_{jm} - u_{im}),$$

where u_{jm} refers to the j^{th} entry of the eigenvector of P that corresponds to λ_{P_m}.

Proof. Recall that Z is symmetric, and from linear algebra we know that symmetric matrices can be diagonalized by an orthogonal transformation of their eigenvectors and eigenvalues. Thus $Z = U \Lambda U^T$, where U is an $n \times n$ matrix whose columns are the unit eigenvectors of Z, and Λ is the diagonal matrix consisting of Z's eigenvalues. Hence by definition, $Z_{ij} = (U\Lambda U^T)_{ij}$. By the rules of matrix multiplication,

$$(U\Lambda U^T)_{ij} = \sum_{m=1}^{n}\sum_{k=1}^{n}(u_{im}\Lambda_{mk}u^T_{kj}).$$

Now, Λ_{mk} is nonzero only where $m = k$, so we have

$$\begin{aligned}
Z_{ij} &= (U\Lambda U^T)_{ij} \\
&= \sum_{m=1}^{n} u_{im}\Lambda_{mm}u^T_{mj} \\
&= \sum_{m=1}^{n} u_{im}\lambda_{Z_m}u^T_{mj} \\
&= \sum_{m=1}^{n} \lambda_{Z_m}u_{im}u_{jm} \\
&= \sum_{m=2}^{n} \lambda_{Z_m}u_{im}u_{jm} \text{ as observed in the proof of Lemma 3.9} \\
&= \lambda_{Z_2}u_{i2}u_{j2} + \cdots + \lambda_{Z_n}u_{in}u_{jn} \\
&= (1-\lambda_{P_2})^{-1}u_{i2}u_{j2} + \cdots + (1-\lambda_{Z_n})^{-1}u_{in}u_{jn} \text{ by Lemma 3.9} \\
&= \sum_{m=2}^{n}(1-\lambda_{P_m})^{-1}u_{im}u_{jm}.
\end{aligned}$$

Appealing to Proposition 3.6 completes the proof:

$$\begin{aligned}
E_i[T_j] &= \frac{1}{\pi_j}(Z_{jj} - Z_{ij}) \\
&= \frac{1}{\pi_j}\left(\sum_{m=2}^{n}(1-\lambda_{P_m})^{-1}u_{jm}u_{jm} - \sum_{m=2}^{n}(1-\lambda_{P_m})^{-1}u_{im}u_{jm}\right) \\
&= \frac{1}{\pi_j}\sum_{m=2}^{n}(1-\lambda_{P_m})^{-1}u_{jm}(u_{jm} - u_{im}).
\end{aligned}$$

□

When G is k-regular and π is uniform, Proposition 3.11 becomes

$$E_i[T_j] = n\sum_{m=2}^{n}(1-\lambda_{P_m})^{-1}u_{jm}(u_{jm} - u_{im}). \tag{3.11a}$$

3.4. **Applying Roots of Unity to Hitting Times.** Almost everything we have done up to this point is generalizable to any undirected, strongly connected graph G. However, when G exhibits *circulancy*, as in the case of the undirected 6-cycle, results from complex analysis give us a unique way to construct the spectrum of P. [Al-F2] alludes to the process of using roots of unity to construct the spectra of transition matrices of n-cycles and hence obtain expected hitting times.

Definition 3.12. We call a graph G *circulant* if each row of its transition matrix P is equal to a cyclic shift of the first row. That is, the i-th row is equal to the first row, except that each entry is shifted, in a cyclic manner, $i - 1$ places to the right.

We can immediately verify that the undirected 6-cycle is circulant from looking back at its transition matrix. The following definitions formalize the notion of circulancy.

Definitions 3.13. A *Cayley graph* is a visual representation of the group $(\mathcal{S}, *)$, where $\mathcal{S}$ is a set and $*$ is an operation on $\mathcal{S}$ that is closed and associative, with exactly one element in $\mathcal{S}$ being the identity element. We label the vertices of Cayley graph G such that each vertex corresponds to exactly one element of the set $\mathcal{S}$. No elements of $\mathcal{S}$ are left out and no elements are represented by multiple vertices; that is, our labeling must result in a map $\phi : \mathcal{S} \to V$ (recall that V is G's vertex set) such that ϕ is both one-to-one and onto. Hence G is an $|\mathcal{S}|$-vertex graph. We choose a set $\mathcal{A} \subset \mathcal{S}$, which we call the alphabet. Then for all $\mathsf{a} \in \mathcal{A}$, a directed edge connects vertex $\phi(\mathsf{x})$ to vertex $\phi(\mathsf{y})$ if and only if $\mathsf{x} * \mathsf{a} = \mathsf{y}$, where $\mathsf{x}, \mathsf{y} \in \mathcal{S}$ and $\phi(\mathsf{x}), \phi(\mathsf{y}) \in V$. Note that G is $|\mathcal{A}|$-regular. By convention, we say that $G = Cay(\mathcal{S}, \mathcal{A})$.

Thus the undirected 6-cycle is a Cayley graph representing $(\mathbb{Z}_6, +_6)$ on generators 1 and 5 (or -1); that is, when G is the undirected 6-cycle, $G = Cay(\mathbb{Z}_6, \{\pm 1\})$. This is why we label the vertices of the undirected 6-cycle from 0 to 5 instead of from 1 to 6. It is easy to show that finite undirected graphs are Cayley graphs if and only if they are circulant (proved as Proposition 7.4 in the appendix). When this is the case, we can use primitive roots of unity to generate orthonormal spectra for their transition matrices. The following theorem and proof is taken from [La]:

Theorem 3.14. If $X = Cay(\mathbb{Z}_n, \mathcal{S})$, then $Spec(X) = \{\lambda_x : x \in \mathbb{Z}_n\}$ where

$$\lambda_x = \sum_{\mathsf{s} \in \mathcal{S}} \exp\left(\frac{2\pi i x s}{n}\right).$$

Proof. (See [La]). Let T be the linear operator corresponding to the adjacency matrix of a circulant graph $X = Cay(\mathbb{Z}_n, \{a_1, a_2, \dots, a_m\})$. If f is any real function on the vertices of X we have

$$T(f)(x) = f(x + a_1) + f(x + a_2) + \cdots + f(x + a_k).$$

Let ω be a primitive n^{th} root of unity and let $g(x) = \omega^{mx}$ for some $m \in \mathbb{Z}_n$. Then

$$\begin{aligned} T(g)(x) &= \omega^{mx+ma_1} + \omega^{mx+ma_2} + \cdots + \omega^{mx+ma_k} \\ &= \omega^{mx}\left(\omega^{ma_1} + \omega^{ma_2} + \cdots + \omega^{ma_k}\right). \end{aligned}$$

Thus g is an eigenfunction and $\omega^{ma_1} + \omega^{ma_2} + \cdots + \omega^{ma_k}$ is an eigenvalue. □

This theorem applies to the spectrum of the adjacency matrix of a Cayley graph. However, the transition matrix is simply the adjacency matrix divided by the graph's regularity. That is, assuming that $|\mathcal{S}| = k$, we divide each entry of a graph's adjacency matrix by k to come up with the graph's transition matrix. The following corollary naturally arises.

Corollary 3.15. If $X = Cay(\mathbb{Z}_n, \mathcal{S})$ has transition matrix P, then $Spec(X) = \{\lambda_{P_x} : x \in \mathbb{Z}_n\}$ where

$$\lambda_x = \frac{1}{k} \sum_{s \in \mathcal{S}} \exp\left(\frac{2\pi i x s}{n}\right).$$

Proof. Let R be a linear operator corresponding to G's transition matrix, and substitute R for T in the above proof:

$$R(f)(x) = \frac{1}{k}\left(f(x + a_1) + f(x + a_2) + \cdots + f(x + a_k)\right).$$

Defining ω and g exactly as above,

$$\begin{aligned} R(g)(x) &= \frac{1}{k}\left(\omega^{mx+ma_1} + \omega^{mx+ma_2} + \cdots + \omega^{mx+ma_k}\right) \\ &= \frac{\omega^{mx}}{k}\left(\omega^{ma_1} + \omega^{ma_2} + \cdots + \omega^{ma_k}\right). \end{aligned}$$

So, this new g is an eigenfunction of R, and $\frac{1}{k}\left(\omega^{ma_1} + \omega^{ma_2} + \cdots + \omega^{ma_k}\right)$ is an eigenvalue of P. □

Hence we can generate hitting times for random walks on circulant graphs directly from primitive roots of unity.

4. Sample Calculations

In the series of calculations that follow, we quantify expected hitting times for random walks on G, where G is the undirected 6-cycle (reproduced below as Figure 2).

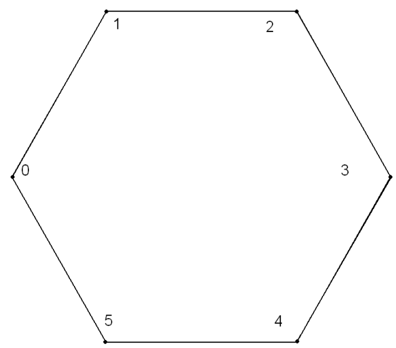

Figure 2. For reference, a reprint of the graph on the left in Figure 1.

Calculation 4.1. Using Mathematica, we approximate expected hitting times by simulating a random walk on G 100,000 times with a given starting vertex and a given stopping vertex, and taking the average of the sample. Using this method, we find that $E_0[T_1] \approx 5,\ E_0[T_2] \approx 8,\ \text{and}\ E_0[T_3] \approx 9.$

As $0 \sim 1$ and G is 2-regular, π is uniform and hence $E_0[T_1] = n - 1 = 5$, which agrees with our Mathematica approximation. Furthermore, due to the undirectedness and regularity of G, $E_0[T_5] = E_0[T_1] = 5$, $E_0[T_4] = E_0[T_2] \approx 8$, and $E_0[T_j] = E_j[T_0]$ for all j. Noting that $E_0[T_0]$ trivially equals 0, we use Proposition 4.4 to calculate $E_\pi[T_0]$:

$$\begin{aligned} E_\pi[T_0] &= \frac{1}{6}\sum_{i=0}^{5} E_i[T_0] \\ &= \frac{1}{6}(0+5+8+9+8+5) \\ &= \frac{35}{6}. \end{aligned}$$

Calculation 4.2. Using Proposition 3.8 combined with Propositions 3.4 and 3.6, we recalculate G's fundamental matrix and use it to determine $E_\pi[T_j]$ and $E_i[T_j]$ values, hoping these values agree with our Mathematica simulations. First, note that G has the following transition matrix P:

$$P = \begin{bmatrix} 0 & 1/2 & 0 & 0 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 0 & 0 & 1/2 & 0 \end{bmatrix}.$$

Using P, the identity matrix, and P_∞, which in this case is the 6×6 matrix where each entry is equal to the uniform stable distribution value $\frac{1}{6}$, we construct Z according to Proposition 3.8:

$$\begin{aligned} Z &= (I - (P - P_\infty))^{-1} - P_\infty \\ &= \begin{bmatrix} 35/36 & 5/36 & -13/36 & -19/36 & -13/36 & 5/36 \\ 5/36 & 35/36 & 5/36 & -13/36 & -19/36 & -13/36 \\ -13/36 & 5/36 & 35/36 & 5/36 & -13/36 & -19/36 \\ -19/36 & -13/36 & 5/36 & 35/36 & 5/36 & -13/36 \\ -13/36 & -19/36 & -13/36 & 5/36 & 35/36 & 5/36 \\ 5/36 & -13/36 & -19/36 & -13/36 & 5/36 & 35/36 \end{bmatrix}. \end{aligned}$$

Appealing to Propositions 3.4 and 3.6, we can calculate the hitting times as follows:

$E_0[T_1] = \frac{1}{\pi_1}(Z_{11} - Z_{01}) = 6\left(\frac{35}{36} - \frac{5}{36}\right) = 5.$

$E_0[T_2] = \frac{1}{\pi_2}(Z_{22} - Z_{02}) = 6\left(\frac{35}{36} + \frac{13}{36}\right) = 8.$

$E_0[T_3] = \frac{1}{\pi_3}(Z_{33} - Z_{03}) = 6\left(\frac{35}{36} + \frac{19}{36}\right) = 9.$

$E_\pi[T_0] = \frac{1}{\pi_0}(Z_{00}) = 6\left(\frac{35}{36}\right) = \frac{35}{6} = E_\pi[T_j] = \tau$ for all j.

Indeed, the Mathematica simulations verify these results.

Calculation 4.3. Because the G has the uniform stable distribution, for all j, $E_\pi[T_j] = \tau$, as verified in the above calculation. Thus we use the spectra of Z and P, as suggested by Proposition 3.10, to calculate τ again. The set of eigenvalues of the transition matrix is $\left\{1, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -1\right\}$, and the set of eigenvalues of the fundamental matrix is

$\{0, 2, 2, \frac{2}{3}, \frac{2}{3}, \frac{1}{2}\}$. In accordance with Proposition 3.10, we write

$$\begin{aligned}
\sum_{m=2}^{n}(1-\lambda_{P_m})^{-1} &= \frac{1}{1-1/2}+\frac{1}{1-1/2}+\frac{1}{1+1/2}+\frac{1}{1+1/2}+\frac{1}{1+1} \\
&= 2+2+\frac{2}{3}+\frac{2}{3}+\frac{1}{2} \\
&= \frac{35}{6} \\
&= \tau \\
&= \sum_{m=2}^{n}\lambda_{Z_m}.
\end{aligned}$$

Hence our Mathematica simulations verify Proposition 3.10 and Lemma 3.9.

Calculation 4.4. Using Proposition 3.11, we recalculate $E_i[T_j]$ values on G using spectral decomposition of P. The eigenvalues and orthonormal eigenvectors of P are as follows:

$\lambda_{P_1} = 1$, with eigenvector $\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^T/\sqrt{6}$;
$\lambda_{P_2} = \frac{1}{2}$, with eigenvector $\begin{bmatrix} 1 & 0 & -1 & -1 & 0 & 1 \end{bmatrix}^T/2$;
$\lambda_{P_3} = \frac{1}{2}$, with eigenvector $\begin{bmatrix} -1/2 & -1 & -1/2 & 1/2 & 1 & 1/2 \end{bmatrix}^T/\sqrt{3}$;
$\lambda_{P_4} = -\frac{1}{2}$, with eigenvector $\begin{bmatrix} -1 & 0 & 1 & -1 & 0 & 1 \end{bmatrix}^T/2$;
$\lambda_{P_5} = -\frac{1}{2}$, with eigenvector $\begin{bmatrix} -1/2 & 1 & -1/2 & -1/2 & 1 & -1/2 \end{bmatrix}^T/\sqrt{3}$;
and $\lambda_{P_6} = -1$, with eigenvector $\begin{bmatrix} -1 & 1 & -1 & 1 & -1 & 1 \end{bmatrix}^T/\sqrt{6}$.

$$\begin{aligned}
E_0(T_1) &= 6\sum_{m=2}^{n}(1-\lambda_m(P))^{-1}u_{1m}(u_{1m}-u_{0m}) \\
&= 6\Big(2\cdot 0(0-1/2)+2\cdot -1/\sqrt{3}(-1/\sqrt{3}+1/2\sqrt{3})+2/3\cdot 0(0+1/2) \\
&\qquad +2/3\cdot 1/\sqrt{3}(1/\sqrt{3}+1/2\sqrt{3})+1/2\cdot 1/\sqrt{6}(1/\sqrt{6}+1/\sqrt{6})\Big) \\
&= 6\Big(2\cdot 0+2\cdot -1/\sqrt{3}\cdot -1/2\sqrt{3}+2/3\cdot 0 \\
&\qquad +2/3\cdot 1/\sqrt{3}\cdot 1/\sqrt{3}+1/2\cdot 1/\sqrt{6}\cdot 1/\sqrt{6}\Big) \\
&= 6\,(0+1/3+0+1/3+1/6) \\
&= 5 \\
&= 6(Z_{11}-Z_{01}).
\end{aligned}$$

$$
\begin{aligned}
E_0(T_2) &= 6\sum_{m=2}^{n}(1-\lambda_m(P))^{-1}\,u_{2m}(u_{2m}-u_{0m}) \\
&= 6\Big(2\cdot -1/2(-1/2-1/2)+2\cdot -1/2\sqrt{3}(-1/2\sqrt{3}+1/2\sqrt{3}) \\
&\qquad +2/3\cdot 1/2(1/2+1/2)+2/3\cdot -1/2\sqrt{3}(-1/2\sqrt{3} \\
&\qquad +1/2\sqrt{3})+1/2\cdot -1/\sqrt{6}(-1/\sqrt{6}+1/\sqrt{6})\Big) \\
&= 6\Big(2\cdot -1/2\cdot -1+2\cdot -1/2\sqrt{3}\cdot 0+ \\
&\qquad 2/3\cdot 1/2\cdot 1+2/3\cdot -1/2\sqrt{3}\cdot 0+1/2\cdot -1/\sqrt{6}\cdot 0\Big) \\
&= 6\,(1+0+1/3+0+0) \\
&= 8 \\
&= 6(Z_{22}-Z_{02}).
\end{aligned}
$$

Once again, Mathematica supports Proposition 3.11.

Calculation 4.5. Finally, we use primitive roots of unity, as suggested by Corollary 3.15, to recalculate the spectrum and expected hitting times for random walks on the undirected 6-cycle. Note that $\omega = \exp(\frac{2\pi i}{6})$, m ranges from 0 to 5, $k = 2$, $a_1 = -1$, and $a_2 = a_k = 1$. Using these parameters we compute eigenvectors and eigenvalues as follows:

$$
\lambda_{P_1} = \frac{\omega^{0\cdot 1}+\omega^{0\cdot -1}}{2} = \frac{1+1}{2} = 1.
$$

We have $u_1 = [\omega^{0x}]^T$ for x ranging from 0 to 5. Thus,

$$
\begin{aligned}
u_1 &= \begin{bmatrix} \omega^0 & \omega^0 & \omega^0 & \omega^0 & \omega^0 & \omega^0 \end{bmatrix}^T \\
&= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^T.
\end{aligned}
$$

We calculate the second eigenvalue and eigenvector in a similar manner:

$$
\begin{aligned}
\lambda_{P_2} &= \frac{\omega^{1\cdot 1}+\omega^{1\cdot -1}}{2} \\
&= \frac{\exp(\frac{2\pi i}{6})+\exp(\frac{-2\pi i}{6})}{2} \\
&= \frac{\cos\frac{\pi}{3}+i\sin\frac{\pi}{3}+\cos\frac{-\pi}{3}+i\sin\frac{-\pi}{3}}{2} \\
&= \frac{1+i\left(\frac{\sqrt{3}}{2}-\frac{\sqrt{3}}{2}\right)}{2} \\
&= \frac{1}{2}.
\end{aligned}
$$

Furthermore, $u_2 = [\omega^{1x}]^T$ for x ranging from 0 to 5. Thus,

$$\begin{aligned} u_2 &= [\ \omega^{1\cdot 0} \quad \omega^{1\cdot 1} \quad \omega^{1\cdot 2} \quad \omega^{1\cdot 3} \quad \omega^{1\cdot 4} \quad \omega^{1\cdot 5}\ ]^T \\ &= [\ \omega^0 \quad \omega^1 \quad \omega^2 \quad \omega^3 \quad \omega^4 \quad \omega^5\ ]^T. \end{aligned}$$

Now, all but one of these entries will be a complex number. To come up with eigenvectors without complex entries, we add this eigenvector to the one generated when $m = 5$:

$$\begin{aligned} u_{2_{\text{Real}}} &= [\ \omega^0 \quad \omega^1 \quad \omega^2 \quad \omega^3 \quad \omega^4 \quad \omega^5\ ]^T + [\ \omega^0 \quad \omega^5 \quad \omega^{10} \quad \omega^{15} \quad \omega^{20} \quad \omega^{25}\ ]^T \\ &= [\ 1 \quad 1/2 \quad -1/2 \quad -1 \quad -1/2 \quad 1/2\ ]^T. \end{aligned}$$

Using this method, we calculate all eigenvalues and eigenvectors, create linear combinations of the eigenvectors with real entries as necessary, normalize these linear combinations, and come up with the following:

$$\lambda_{P_1} = 1, \text{ with unit eigenvector } \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^T / \sqrt{6}.$$

$$\lambda_{P_2} = \frac{1}{2}, \text{ with unit eigenvector } \begin{bmatrix} 1 & 1/2 & -1/2 & -1 & -1/2 & 1/2 \end{bmatrix}^T / \sqrt{3}.$$

$$\lambda_{P_3} = \frac{1}{2}, \text{ with unit eigenvector } \begin{bmatrix} 0 & 1 & 1 & 0 & -1 & -1 \end{bmatrix}^T / 2.$$

$$\lambda_{P_4} = -\frac{1}{2}, \text{ with unit eigenvector } \begin{bmatrix} 1 & -1/2 & -1/2 & 1 & -1/2 & -1/2 \end{bmatrix}^T / \sqrt{3}.$$

$$\lambda_{P_5} = -\frac{1}{2}, \text{ with unit eigenvector } \begin{bmatrix} 0 & 1 & -1 & 0 & 1 & -1 \end{bmatrix}^T / 2.$$

$$\lambda_{P_6} = -1, \text{ with unit eigenvector } \begin{bmatrix} 1 & -1 & 1 & -1 & 1 & -1 \end{bmatrix}^T / \sqrt{6}.$$

Note that these eigenvectors are not all the same as they were in Calculation 4.4, yet they are still orthonormal. In fact, we note that this process automatically produces orthonormal eigenvectors, though we do have to correct for complex entries. In Calculation 4.4 we used the Gram-Schmidt process to orthogonalize the eigenvectors. Using Proposition 3.11, we once again compute $E_0(T_1)$.

$$\begin{aligned}
E_0[T_1] &= 6\sum_{m=2}^{n}(1-\lambda_{P_m})^{-1}u_{1m}(u_{1m}-u_{0m}) \\
&= 6\Big(2\cdot 1/2\sqrt{3}(1/2\sqrt{3}-1/\sqrt{3})+2\cdot 1/2(1/2-0)+ \\
&\qquad 2/3\cdot -1/2\sqrt{3}(-1/2\sqrt{3}-1/\sqrt{3})+2/3\cdot 1/2(1/2-0) \\
&\qquad +1/2\cdot -1/\sqrt{6}(-1/\sqrt{6}-1/\sqrt{6})\Big) \\
&= 6\Big(2\cdot 1/2\sqrt{3}\cdot -1/2\sqrt{3}+2\cdot 1/4+2/3\cdot -1/2\sqrt{3}\cdot -3/2\sqrt{3} \\
&\qquad +2/3\cdot 1/4+1/2\cdot -1/\sqrt{6}\cdot -2/\sqrt{6}\Big) \\
&= 6\,(-1/6+1/2+1/6+1/6+1/6) \\
&= 5 \\
&= n-1 \\
&= 6(Z_{11}-Z_{01}).
\end{aligned}$$

Calculations 4.1 through 4.4 provide numerical verification, in the case of the undirected 6-cycle, of the technique we developed in Sections 2 and 3 for calculating expected hitting times for random walks on finite, undirected, strongly connected graphs. Furthermore, Calculation 4.5 verifies that mean hitting times on circulant graphs appear to come directly from primitive roots of unity. While we do need to correct for complex entries, the above process automatically constructs orthogonal spectra, and does not depend on the symmetry of P. Further investigation may reveal that as long as G is circulant, P need not be symmetric in order to use roots of unity to construct hitting times. This completes our analysis of mean hitting times.

5. Using Catalan Numbers to Quantify Hitting Time Probabilities

5.1. **The Question of Variance: A Closer Look at Cycles.** We have developed a methodology for calculating expected hitting times on undirected, strongly connected graphs. The natural question is: can we use this same methodology to calculate higher moments of the distribution of hitting times? In short, the answer appears to be no. The spectra of transition matrices are naturally suited to calculating expected hitting time values, but an analogous general formula appears not to exist for calculating variances of hitting times for random walks on graphs. The theoretical background developed in Section 2 is appropriate and necessary for understanding expected values of hitting times, but it does not seem to offer any insight about how to explicitly calculate their variances.

While our hopes of achieving any result regarding variances in the general case are slim, some literature exists on variances of hitting times in the asymptotic case. We turn our attention to [Al].

Proposition 5.1. (See Proposition 7 of [Al]). For a sequence of n-vertex, vertex-transitive graphs $\{G_n\}$ with $n \to \infty$, the following are equivalent:

- $1/\tau(1-\lambda_{P_2}) \to 0$;
- $\mathcal{D}_\pi(T_j/\tau) \to \mu_1$ and $Var_\pi[T_w/\tau] \to 1$ for all j,

where $\mathcal{D}_\pi(T_j/\tau)$ refers to the distribution $\mathcal{D}$ of stopping times T_j/τ for a random walk on graph G_n in which the starting vertex is determined by the uniform stable distribution π, and μ_1 is the exponential distribution with mean 1.

Using Mathematica, we investigate whether Proposition 5.1 holds for random walks with explicit starting and destination vertices. We consider the sequence $\{C_{5n}\}$, where C_n is the undirected n-cycle. Letting n range from 200 to 1000, we simulate 75 random walks from vertex 0 to vertex 1 and obtain the variance of this sample for each value of n. We display our results below in Figure 3.

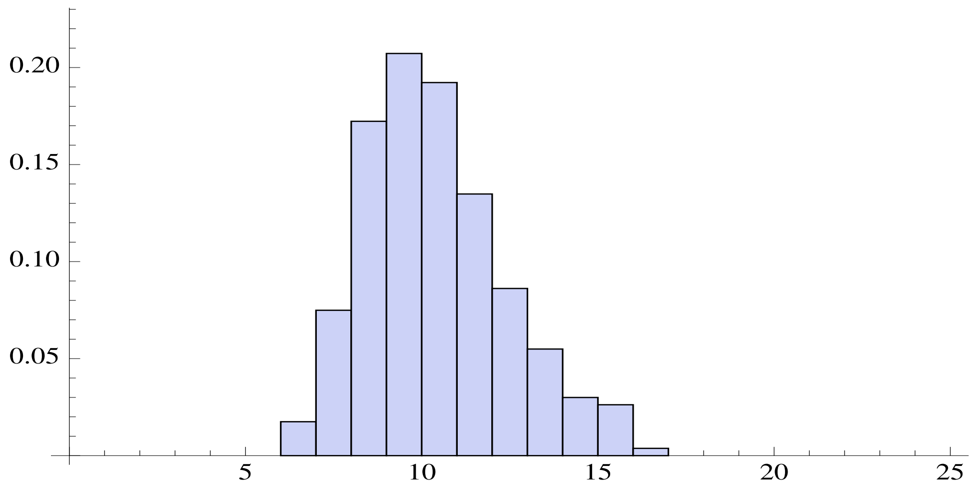

FIGURE 3. A probability density histogram describing the distribution of our results. We plot the probability of occurrence as a function of the base 10 logarithm of $Var_0[T_1/(n-1)]$. We expect a distribution that tails off to the right, with a mode near 0, but instead see a distribution centered at approximately 10. While the histogram does not seem to verify Proposition 5.1, we feel that this is very likely due to the lack of a sufficient number of iterations or a large enough range of n, and believe that Proposition 5.1 does indeed apply to the explicit case.

Vertex-transitive graphs exhibit many nice properties (circulancy, regularity, uniform π), and undirected n-cycles are very simple vertex-transitive graphs. Thus we continue to study cycles in an effort to say something quantitative about hitting time distributions. In particular, we study the asymptotic cycle, which is simply the integer number line. Because the methodology developed in the previous sections appears irrelevant, and we have no suitable alternative, we turn to the probabilistic method given in the introduction for further analysis. When using the probabilistic method we must account for all possible paths of a certain length w from vertex i to vertex j, and iterate this process over all possible lengths w. This is a daunting task. As analyzed below, studying random walks on the number line is most promising in terms of yielding results because there exists a systematic method of accounting for all such paths.

Definition 5.2. $P_i(T_j = w)$ refers to the probability that a random walk starting at vertex i first reaches j in w steps.

We seek to determine first hitting time probabilities as given in Definition 5.2 with the help of the Catalan numbers.

5.2. **A Brief Introduction to the Catalan Numbers.** Catalan numbers are well-studied quantities, and the informed reader may skip this subsection. [Hi] provides a rigorous overview of Catalan numbers, and we appeal there for the following definitions.

Definition 5.3. Consider an integer lattice. A path from point (c, d) to point (a, b) is called *p-good* if it lies entirely below the line $y = (p-1)x$, where each step of the path is in either an eastward or a northward direction.

Definition 5.4. (See (1.2) of [Hi]). The k^{th} *Catalan number,* C_k, is the number of 2-good paths from $(0, -1)$ to $(k, k-1)$. Moreover,

$$C_k = \frac{1}{k+1}\binom{2k}{k}.$$

For our purposes, we introduce and use the following definitions:

Definition 5.5. A *p-fine path* is a path from (c, d) to (a, b) that is never above the line $y = (p-1)x$ and contains only eastward and northward steps.

Definition 5.6. (Equivalent definition of the k^{th} Catalan Number). C_k is the number of 2-fine paths from $(0, 0)$ to (k, k).

We prefer Definition 5.6 because it allows for a more tangible conceptualization of the Catalans. In particular, this definition lets us view the k^{th} Catalan number as the number of paths of k "eastward" steps and k "northward" steps on the integer lattice, such that the number of northward steps never exceeds the number of eastward steps. Figure 4 below uses this concept to illustrate C_4:

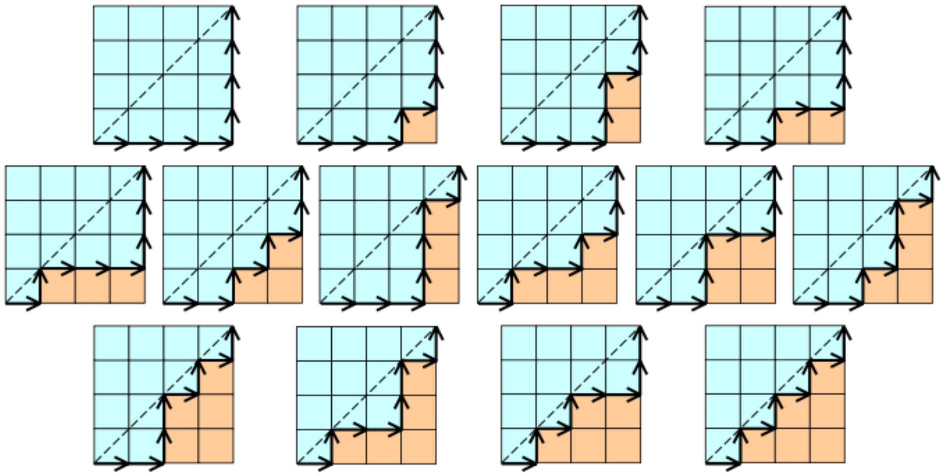

FIGURE 4. (Source: Wikimedia Commons). A demonstration of the 14 possible ways to move from $(0,0)$ to $(4,4)$ such that the number of northward steps never exceeds the number of eastward steps. Thus $C_4 = 14$.

An equivalent conceptualization of C_k is the number of "legal" arrangements of k pairs of parentheses, meaning that when the sequence of parentheses is scanned sequentially, the number of right parentheses never exceeds the number of left parentheses. In this vein, we see that

$$\begin{aligned} &C_1 = 1 = |\{()\}| \\ &C_2 = 2 = |\{()(), (())\}| \\ &C_3 = 5 = |\{()()(), (())(), (()()), ()(()), ((()))\}| \\ &\vdots \end{aligned}$$

The following well-known lemma provides another natural observation about the Catalan numbers.

Lemma 5.7. (See (1.4) of [Hi]). Catalan numbers satisfy the following recurrence relation:

$$C_k = \sum_{i+j=k-1} C_i C_j, \; k \geq 1; \; C_0 = 1.$$

5.3. **Application to Hitting Time Probabilities.** We now apply the Catalan numbers to our purposes. We start by proving the following proposition.

Proposition 5.8. Consider a random walk on the integer number line starting at 0 and stopping upon the first visit to 1. Then,

$$P_0(T_1 = 2k+1) = \frac{C_k}{2^{2k+1}},$$

where k is a non-negative integer.

Proof. Let us consider a $2k+1$ step random walk on the number line, starting at 0, ending at 1, and not visiting 1 until the walk's last step. Let $l(n)$ represent the number of "lefts" after the n^{th} step of the walk and $r(n)$ be the number of "rights" after the n^{th} step of the walk. We immediately know the following:

- $l(2k+1) = k = r(2k+1) - 1$.
- $l(n) \geq r(n)$ for $n < 2k+1$.

That is, in a $2k+1$ step walk from 0 to 1, there can never be more rights than lefts until $n = 2k+1$, i.e. until the last step of the walk. This is because as soon as the walk ventures to the right of 0, it hits 1 and stops. So, the walk ventures to the right of 0 exactly at its last step. Thus all the path variation in the $2k+1$ step walks from 0 to 1 occurs in the first $2k$ steps; when trying to account for all the $2k+1$ step walks from 0 to 1, we need only consider the first $2k$ steps. Consider a transformation of the first $2k$ steps of the walk, which maps the number line to the integer lattice, the starting position of the walk to $(0,0)$, the ending position of the walk to (k,k), "left" to "east", and "right" to "north." Note that in the partial walk, which has length $2k$, the number of rights never exceeds the number of lefts. This means that the transformed partial walk's number of northward steps never exceeds its number of eastward steps, which is to say that the path is never above the line $y = x$. Hence, the partial walk is isomorphic to a 2-fine path from $(0,0)$ to (k,k). From Definition 5.6, we know that C_k such paths exist. Because the last step of the walk is fixed, there are C_k different $2k+1$ step random walks starting at 0 and ending at 1, with no intermediary visits to 1. Noting that each step of the walk occurs with probability $1/2$, we see that $P_0(T_1 = 2k+1) = \frac{C_k}{2^{2k}} \cdot \frac{1}{2} = \frac{C_k}{2^{2k+1}}$. $\square$

We devote the next proposition to the calculation of $P_0(T_2 = 2m)$.

Proposition 5.9. Consider a random walk on the number line starting at 0 and stopping upon the first visit to 2. Then

$$P_0(T_2 = 2m) = \frac{C_m}{2^{2m}},$$

where k is a non-negative integer.

Proof. Note that a $2m$ step walk from 0 to 2 must visit 1 at least once. Let $k \leq m$. Then,

$$\begin{aligned}
P_0(T_2 = 2m) &= P_0(T_1 = 2k+1)P_1(T_2 = 2(m-k)-1) \\
&= P_0(T_1 = 2k+1)P_0(T_1 = 2(m-k)-1) \text{ by vertex transitivity} \\
&= \sum_{k=0}^{m} P_0(T_1 = 2k+1)P_0(T_1 = 2(m-k-1)+1) \\
&= \sum_{k=0}^{m} \frac{C_k}{2^{2k+1}} \cdot \frac{C_{m-k-1}}{2^{2(m-k-1)+1}} \text{ by Proposition 5.8} \\
&= \frac{1}{2^{2m}} \sum_{k=0}^{m} C_k C_{m-k-1} \\
&= \frac{C_m}{2^{2m}} \text{ by Lemma 5.7.}
\end{aligned}$$

□

In a similar manner, we find that

- $P_0(T_3 = 2m+1) = \frac{1}{2^{2m+1}}(C_{m+1} - C_m)$, and
- $P_0(T_4 = 2m) = \frac{1}{2^{2m}}(C_{m+1} - C_m)$.

These derivations set us up to make a generalization about hitting times on the number line.

Theorem 5.10. Consider a random walk on the number line starting at 0. We can express $P_0(T_{2^n} = 2m)$ as a linear combination of Catalan numbers, ranging from $C_{m-2^{n-1}+n}$ to $C_{m+2^{n-1}-1}$, with $m \geq 0$ and $n \geq 1$. That is,

$$\begin{aligned}
P_0(T_{2^n} = 2m) &= \alpha_{n,m+2^{n-1}-1}C_{m+2^{n-1}-1} + \alpha_{n,m+2^{n-1}-2}C_{m+2^{n-1}-2} + \cdots \\
&\quad + \alpha_{n,m-2^{n-1}+n}C_{m-2^{n-1}+n} \\
&= [C_{m+2^{n-1}-1}, C_{m+2^{n-1}-2}, \cdots, C_{m-2^{n-1}+n}]\vec{\alpha}^T,
\end{aligned}$$

where $\vec{\alpha}^T$ is the column vector of the α_n coefficients, and has length $2^n - n$.

Proof. We proceed by induction. As given by Proposition 5.9, when $n = 1$,

$$\begin{aligned}
P_0(T_{2^1} = 2m) &= \frac{C_m}{2^{2m}} \\
&= \alpha_{1,m+2^{1-1}-1}C_{m+2^{1-1}-1}.
\end{aligned}$$

For the inductive assumption, we assume the result holds for all $n = k$; that is,

$$\begin{aligned}
P_0(T_{2^k} = 2m) &= \alpha_{k,m+2^{k-1}-1}C_{m+2^{k-1}-1} + \alpha_{k,m+2^{n-1}-2}C_{m+2^{k-1}-2} + \cdots \\
&\quad + \alpha_{k,m-2^{k-1}+n}C_{m-2^{k-1}-1}.
\end{aligned}$$

Let us now calculate $P_0(T_{2^{k+1}}2m)$. To do so, we split the random walk from 0 to 2^{k+1} into two partial walks, one from 0 to 2^k and one from 2^k to 2^{k+1}. Thus we have

$$\begin{aligned}
P_0(T_{2^{k+1}} = 2m) &= P_0(T_{2^k} = 2l)P_{2^k}(T_{2^{k+1}} = 2(m-l)) \\
&= P_0(T_{2^k} = 2l)P_0(T_{2^k} = 2(m-l)) \\
&= \sum_{l=2^{k-1}}^{m-2^{k-1}} \big(\alpha_{k,l+2^{k-1}-1}C_{l+2^{k-1}-1} + \alpha_{k,l+2^{n-1}-2}C_{l+2^{k-1}-2} \\
&\qquad\qquad + \cdots + \alpha_{k,l-2^{k-1}+n}C_{l-2^{k-1}-1}\big) \\
&\qquad \cdot \big(\alpha_{k,m-l+2^{k-1}-1}C_{m-l+2^{k-1}-1} + \alpha_{k,m-l+2^{k-1}-2}C_{m-l+2^{k-1}-2} \\
&\qquad\qquad + \cdots + \alpha_{k,m-l-2^{k-1}+n}C_{m-l-2^{k-1}-1}\big)\,.
\end{aligned}$$

Note that we index from 2^{k-1} to $m - 2^{k-1}$ because each partial walk must contain at least $2 \cdot 2^{k-1} = 2^k$ steps. That is, if l was less than 2^{k-1}, then the walk from 0 to 2^{k-1} would be impossible, and if l was greater than $m - 2^{k-1}$, then the walk from 2^{k-1} to 2^k would be impossible. We now simplify the expression by introducing indices a and b to account for the sums of Catalan numbers corresponding to the first and second partial walks, respectively, and by using $(\star)$ to denote the α coefficients:

$$\begin{aligned}
&= \sum_{l=2^{k-1}}^{m-2^{k-1}} \sum_{a=-2^{k-1}+k}^{2^{k-1}-1} \sum_{b=-2^{k-1}+k}^{2^{k-1}-1} C_{l+a}C_{m-l+b}(\star) \\
&= \sum_a \sum_b \sum_{l=2^{k-1}}^{m-2^{k-1}} C_{l+a}C_{m-l+b}(\star) \\
&= \sum_a \sum_b \sum_{l=2^{k-1}}^{m-2^{k-1}} C_{l+a}C_{m+b+a-(l+a)}(\star).
\end{aligned}$$

Let $u = l + a$; our expression becomes

$$\begin{aligned}
&\sum_a \sum_b \sum_{u=2^{k-1}+a}^{m+a+b-(b+2^{k-1})} C_u C_{m+b+a-u}(\star) \\
= &\sum_a \sum_b \left(\sum_{u=0}^{m+a+b} C_u C_{m+a+b-u} - \sum_{u=0}^{2^{k-1}+a-1} C_u C_{m+a+b-u} \right. \\
&\qquad \left. - \sum_{u=m+a+b-(b+2^{k-1})+1}^{m+a+b} C_u C_{m+a+b-u} \right) (\star).
\end{aligned}$$

Letting $v = m + a + b - u$ and rearranging, we get

$$\begin{aligned} &\sum_a \sum_b \left(C_{m+a+b+1} - \sum_{u=0}^{2^{k-1}+a-1} C_u C_{m+a+b-u} - \sum_{v=0}^{2^{k-1}+b-1} C_v C_{m+a+b-v} \right) (\star) \\ = \quad &\sum_a \sum_b \left(C_{m+a+b+1} - \sum_{u=0}^{2^{k-1}+a-1} C_{m+a+b-u} - \sum_{v=0}^{2^{k-1}+b-1} C_{m+a+b-v} \right) (\star), \end{aligned}$$

because u and v are independent of m, and hence C_u and C_v are absorbed into $(\star)$. Thus we now have 3 terms, and we seek to show that the indices of each must lie between $m+2^k-1$ and $m - 2^k + k + 1$, inclusive. To do so, we recall that $-2^{k-1} + k \leq a, b \leq 2^{k-1} - 1$, and hence $-2^k + 2k \leq a + b \leq 2^k - 2$. We consider each term individually:

- $C_{m+a+b+1}$. This term has the largest index, so if its index satisfies the upper bound, then the indices of all terms must as well. When $a+b$ is at its maximum, $C_{m+a+b+1} = C_{m+2^k-1}$. Hence, all indices satisfy the upper bound, and the upper bound is sharp. When $a + b$ is at its minimum, $C_{m+a+b+1} = C_{m-2^k+2k+1}$, and the lower bound is satisfied.
- $\sum_{u=0}^{2^{k-1}+a-1} C_{m+a+b-u}$. We already know that the upper bound is satisfied. To check the lower bound, we assign u its maximum value of $2^{k-1} + a - 1$: $C_{m+a+b-u} = C_{m+b+1-2^{k-1}}$. Now we assign b its minimum value of $-2^{k-1} + k$: $C_{m+b+1-2^{k-1}} = C_m - 2^k + k + 1$. Hence the lower bound is satisfied.
- $\sum_{v=0}^{2^{k-1}+a-1} C_{m+a+b-v}$. By symmetry, the indices of this term satisfy the upper and lower bounds as well. Hence the lower bound is sharp.

Therefore, the indices of the Catalan numbers that express the probability $P_{0\to 2^{k+1}}(2m)$ lie exactly between $m + 2^k - 1$ and $m - 2^k + k + 1$, inclusive. That is,

$$\begin{aligned} P_0(T_{2^{k+1}} = 2m) \quad &= \quad \alpha_{k+1,m+2^k-1} C_{m+2^k-1} + \alpha_{k+1,m+2^k-2} C_{m+2^k-2} + \cdots \\ &\qquad\qquad + \alpha_{k+1,m-2^k+n} C_{m-2^k-1} \\ &= \quad [C_{m+2^k-1}, C_{m+2^k-2}, \cdots, C_{m-2^k+k+1}] \vec{\alpha}^T, \end{aligned}$$

where $\vec{\alpha}^T$ is the column vector of the α_{k+1} coefficients, and has length $2^{k+1} - (k+1)$. This completes the proof. $\square$

As indicated by the $0 \to 3$ case, sometimes we can express $P_i(T_j = 2m)$ if $j - i$ is even (or $P_i(T_j = 2m + 1)$ if $j - i$ is odd) as a linear combination of Catalan numbers even if $j - i \neq 2^n$ for $n \geq 0$. We believe that following the same inductive process as in the proof of Theorem 5.10 may reveal that all hitting time probabilities on the number can be expressed in terms of the Catalan numbers. Furthermore, we may be able to calculate $\vec{\alpha}^T$

by generalizing from a sufficient number of base cases. If so, we are not far from knowing the probability distribution for all hitting times on the number line.

Would we be able to apply such a result to finite cycles? Assume $i \sim j$ such that j is one vertex clockwise from i. Then we reach j by making a net of one clockwise step *or* a net of $n-1$ counterclockwise steps. The Catalan number method does not take into account the probability of this second occurrence, though this probability of course approaches 0 as $n \to \infty$. Not taking into account this tendency to underestimate $P_i(T_j = 2m)$, the Catalan number method overestimates $P_i(T_j = 2m)$ for large m. This is because there is no limit on the number of net leftward steps we are allowed to make on the number line. However, in the finite case, we cannot make too many counterclockwise steps in a row or we will hit j before having taken $2m$ steps. The prospect of reaching our destination before the specified number of steps constrains the number of legal walks we are allowed to make to a greater extent in the finite case than in the infinite case. Thus we see that the Catalan method both underestimates and overestimates $P_i(T_j = 2m)$ in different ways. In the limit, these errors exactly offset each other, and in the finite case, it appears that the underestimation dominates the overestimation (consider $P_0(T_2 = 2)$ on the square for a simple verification of this). It would be interesting to get a sense of the interaction between the underestimation and the overestimation as $n \to \infty$.

6. Conclusion

By appealing to the literature to show that random walks on graphs are discrete time Markov chains that behave as renewal processes, we define the fundamental matrix, and hence are able to work toward an explicit hitting time formula. Independently of the existing literature, and remaining in discrete time, we relate the fundamental matrix to the transition matrix, and use spectral decomposition to generate a hitting time formula that yields expected hitting time values for random walks on any finite, undirected, strongly connected graph. Assuming G is undirected makes P symmetric, which facilitates the proofs of several of our results. However, we believe that strong connectedness of G is the only necessary condition for the existence of an explicit expected hitting time formula. Symmetrization of P using powers of π in the case of directed graphs looks promising as a way to increase the scope of our results.

Furthermore, we investigate, in the spirit of [La], the role primitive roots of unity play in calculating expected hitting times on circulant graphs. Finally, we discuss the likely impossibility of generalizing the expected hitting time formula to higher moments, and use a probabilistic method involving the Catalan numbers to move closer to quantifying hitting time distributions for random walks on the number line. With some more investigation, we believe we can both quantify hitting time distributions for random walks to any destination vertex on the number line, and analyze how applicable such a distribution is to finite cycles.

7. Appendix

We devote this section to giving our own proofs of known results that are not taken directly from the literature and may be obvious to some, but not all, informed readers.

Proposition 7.1. An n-vertex undirected graph containing at least one edge is strongly connected if and only if its transition matrix P has largest eigenvalue 1 of multiplicity 1.

Proof. First of all, note that P has all non-negative entries. Then from linear algebra, we know that the eigenvalues of a matrix are bounded in absolute value by the largest row sum. Because P is stochastic, we know that the largest eigenvalue of P must be at most 1, and furthermore,

$$P\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Thus, 1 is an eigenvalue with eigenvector J_n. Assume 1 is an eigenvalue again with eigenvector $\vec{v}$ independent of J_n. Consider vertex i of G, and assume that v_i is a maximum value of $\vec{v}$; without loss of generality we rescale $\vec{v}$ so that $v_i = 1$. Because G is strongly connected, i has k neighbors $x_1, \ldots, x_k$, where $1 \le k \le n-1$. This implies that the i^{th} row of P has k non-zero terms, each one corresponding to a neighbor of i, and each one equaling $1/k$. The other $n-k$ terms in the row are 0. Thus we have

$$\begin{aligned} 1 &= P_i \vec{v} \\ &= \frac{1}{k} v_{x_1} + \frac{1}{k} v_{x_2} + \cdots + \frac{1}{k} v_{x_k} \\ &= \frac{1}{k}(v_{x_1} + \cdots + v_{x_k}). \end{aligned}$$

The v_x's must sum to k; since none of them can be greater than 1, all v_x's must equal 1. Thus the k neighbors of i correspond to 1's in $\vec{v}$ as well. Now, choose a neighbor of i: x_1 for instance. We know x_1 has m neighbors $y_1, \ldots, y_m$, where $1 \le m \le n$. Then,

$$\begin{aligned} 1 &= P_{x_1} \vec{v} \\ &= \frac{1}{m} v_{y_1} + \frac{1}{m} v_{y_2} + \cdots + \frac{1}{m} v_{y_m} \\ &= \frac{1}{m}(v_{y_1} + \cdots + v_{y_m}). \end{aligned}$$

Again, all v_y's must equal 1, and so the m neighbors of x_1 correspond to 1's in $\vec{v}$ as well. Now, we select a neighbor of x_1 different from i, and repeat this process. Because G is strongly connected, we find that all its vertices are neighbors to vertices corresponding to 1's in the eigenvector $\vec{v}$. Hence all vertices correspond to 1's in $\vec{v}$; that is, $\vec{v} = J_n$. Employing the spectral theorem once again, we see that because P is real and symmetric (by undirectedness of G), there exists an orthonormal basis of eigenvectors for its eigenvalues. We just showed that $\lambda = 1$ has the one-dimensional basis $\vec{v} = aJ_n$, where a is a constant. Hence any eigenvector with 1 as its eigenvalue cannot be linearly independent of J_n. Therefore we cannot have 1 as a multiple eigenvalue; $\lambda = 1$ has multiplicity 1.

Proving the reverse conditional, assume that G is not strongly connected. Because G contains at least one edge, we can view G as multiple subgraphs, at least one of which is strongly connected. Hence,

$$P = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

where A is an $m \times m$ (irreducible) transition matrix and B is an $n - m \times n - m$ transition matrix. Note that if vertex i has no neighbors, then $p_{ii} = 1$ and $p_{ij} = 0$ for all $j \neq i$. Then we have the following:

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} J_m \\ 0 \end{bmatrix} = 1 \begin{bmatrix} J_m \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} 0 \\ J_{n-m} \end{bmatrix} = 1 \begin{bmatrix} 0 \\ J_{n-m} \end{bmatrix}$$

Thus, we see that $\lambda = 1$ has multiplicity ≥ 2. Therefore if P has eigenvalue 1 with multiplicity exactly 1, G must be strongly connected. This completes the proof. □

Proposition 7.2. An n-vertex, undirected, strongly connected graph G with transition matrix P is periodic if and only if P has an eigenvalue of -1.

Proof. Consider a random walk on periodic graph G starting from vertex i. Since G is periodic, we can view it as a bipartite graph. This is true if and only if for all vertices j, $P_i(T_j = 2n + \phi(i)) = 0$ for all $n \geq 0$, where $\phi(i) = 0$ or 1, and $\phi(i) + \phi(j) = 1$ if $i \sim j$. Because G is strongly connected, j has k neighbors $a_1, \ldots, a_k$ such that $p_{ja_l} = 1/k$ for $1 \leq l \leq k$. Hence $p_{ja} = 1/k$ if and only if $\phi(j) + \phi(a) = 1$; otherwise it takes the value 0. Consider the n-column vector $\vec{v}$ where $\vec{v}_j = 1$ if $\phi(j) = 0$, and $\vec{v}_j = -1$ if $\phi(j) = 1$. This is equivalent to the following:

- Let $\phi(j) = 0$; then

$$\begin{aligned}
\vec{P v_j} &= p_{ja_1} \cdot -1 + \cdots + p_{ja_k} \cdot -1 \\
&= \frac{1}{k} \cdot -1 + \cdots \frac{1}{k} \cdot -1 \\
&= \frac{-1}{k} + \cdots \frac{-1}{k} \\
&= -1 \\
&= -\vec{v}_j.
\end{aligned}$$

- Let $\phi(j) = 1$; then

$$\begin{aligned}
\vec{P v_j} &= p_{ja_1} \cdot 1 + \cdots + p_{ja_k} \cdot 1 \\
&= \frac{1}{k} \cdot 1 + \cdots \frac{1}{k} \cdot 1 \\
&= \frac{1}{k} + \cdots \frac{1}{k} \\
&= 1 \\
&= -\vec{v}_j.
\end{aligned}$$

Finally, the above two conditions hold if and only if $P\vec{v} = -\vec{v}$, which is equivalent to saying that -1 is an eigenvalue of P. This result is what gives random walks on bipartite graphs their oscillatory behavior. □

Proposition 7.3. The fundamental matrix Z has rows summing to 0, and, when π is uniform, constant diagonal entries.

Proof. Using Proposition 3.8, write $Z = (I - (P - P_\infty))^{-1} - P_\infty$. Because we disallow self-loops, by definition $p_{ii} = 0$ for all i. When π is uniform, $P_{\infty_{ij}} = 1/n$ for all i and j, so clearly P_∞ has constant diagonal entries. Finally, I has constant diagonal entries by definition. Because differences of matrices with constant diagonal entries themselves have diagonal entries, and because $I-(P-P_\infty)$ is symmetric, we conclude that $(I-(P-P_\infty))^{-1}$ has constant diagonal entries. Therefore Z has constant diagonal entries.

To show Z has rows summing to 0, define the i^{th} row sum of Z as $|Z_i|$ and note that for all i,

$$\begin{aligned}
|Z_i| &= \sum_j Z_{ij} \\
&= \sum_j \left(\sum_{t=0}^{\infty} p_{ij}^{(t)} - \pi_j \right) \\
&= \sum_j \left((p_{ij}^{(0)} - \pi_j) + (p_{ij} - pi_j) + \cdots \right) \\
&= \sum_j (p_{ij}^{(0)} - \pi_j) + \sum_j (p_{ij} - \pi_j) + \cdots \\
&= (1-1) + (1-1) + \cdots \\
&= 0
\end{aligned}$$

because all powers of P are stochastic and because π is a distribution. □

Proposition 7.4. An undirected n-vertex graph G is a Cayley graph if and only if it is circulant.

Proof. Let G be a Cayley graph; then $G = Cay(\mathcal{S}, \mathcal{A})$. Without loss of generality assume that the group operation is addition modulo n. From Definitions 2.2 and 3.13, we see that for vertices $\phi(x)$ and $\phi(y)$ and for all $\mathsf{a} \in \mathcal{S}$, $p_{\phi(x)\phi(y)} = 1/k$ if and only if $\mathsf{x} +_n \mathsf{a} = \mathsf{y}$, where $\mathsf{x}, \mathsf{y} \in \mathcal{S}$. Otherwise $p_{\phi(x)\phi(y)} = 0$. By associativity of addition, $(1 +_n \mathsf{x}) +_n \mathsf{a} = 1 +_n \mathsf{y}$. Thus by Definitions 3.13, $p_{\phi(x)\phi(y)} = 1/k$ if and only if $p_{1+_n\phi(x),1+_n\phi(y)} = 1/k$, and $p_{\phi(x)\phi(y)} = 0$ if and only if $p_{1+_n\phi(x),1+_n\phi(y)} = 0$. This is equivalent to saying that $p_{\phi(x)\phi(y)} = p_{1+_n\phi(x),1+_n\phi(y)}$, or that G is circulant. □

REFERENCES

[Al] Aldous, D. (1988), "Hitting Times for Random Walks on Vertex-Transitive Graphs", Mathematical Proceedings of the Cambridge Philosophical Society, 106: 179-191.

[Al-F1] Aldous, D. and J. Fill (1999), "General Markov Chains", Chapter 2 of *Reversible Markov Chains and Random Walks on Graphs*, unpublished manuscript: 1-36.

[Al-F2] Aldous, D. and J. Fill (1996), "Examples: Special Graphs and Trees", Chapter 5 of *Reversible Markov Chains and Random Walks on Graphs*, unpublished manuscript: 1-61.

[Bi] Biggs, N. (1993), *Algebraic Graph Theory*, 2^{nd} ed., Cambridge, Cambridge University Press: 16-17.

[Br] Broder, A. and A. Karlin (1989), "Bounds on the Cover Time", *Journal of Theoretical Probability*, 2, 1: 101-120.

[Co] Cox, D.R. (1967), *Renewal Theory*, London, Methuen & Company: 1-142.

[Ha] Hall, R. (1978), "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence", *The Journal of Political Economy*, 86, 6: 971-987.

[Har] Hardy, G. (1991), *Divergent Series*, 2^{nd} ed., New York, Chelsea Publishing Company: 71-73, 96.

[Hi] Hilton, P. and J. Pedersen (1991), "Catalan Numbers, Their Generalization, and Their Uses", *The Mathematical Intelligencer*, 13, 2: 64-75.

[Ka] Kang, M. (2003), "First Hitting Times of Simple Random Walks on Graphs with Congestion Points", *International Journal of Mathematics and Mathematical Science*, 30: 1911-1922.

[La] Lazenby, J. (2008), "Circulant Graphs and Their Spectra", Reed College Undergraduate Thesis: 5-6.

[No] Norris, J. (1997), *Markov Chains*, Cambridge, Cambridge University Press: 1-56.

[Pa-1] Palacios, J. (2009), "On Hitting Times of Random Walks on Trees", *Statistics and Probability Letters*, 79: 234-236.

[Pa-2] Palacios, J., J. Renom, and P. Berrizbeitia (1999), "Random Walks on Edge-Transitive Graphs", *Statistics and Probability Letters*, 43: 25-32.

[Sa] Saloff-Coste, L., E. Ginè, and G. Grimmet (1997), "Lectures on Finite Markov Chains", *Lecture Notes in Mathematics*, Heidelberg, Springer Berlin: 301-413.

[Ta] Takacs, C. (2006), "On the Fundamental Matrix of Finite State Markov Chains, its Eigensystem and its Relation to Hitting Times", *Mathematica Pannonica*, 17, 2: 183-193.