Probability on Trees and Networks

Bronwyn Walsh, Catherine Fontaine

June 2025

Abstract

In this paper, we provide an introductory guide of Markov chains, and a streamlined exposition of electrical network theory and Polya's theorem. We aim to provide a clearer and more concise guide to these topics than is found in existing works. This paper considers a narrow view of electrical networks, so further exploration of the many other proofs of Pólya's Theorem is encouraged.

1 Introduction

Markov chains are a mathematical object that can encode many real-world problems. For example, they can be used in forecasting chemical reactions and windpower trends, and are instrumental in information theory and Bayesian statistics. Markov chains can also be used to represent electrical networks. In this paper, we explore how electrical networks can in fact be used to understand general properties of Markov chains. In particular we use them to prove Pólya's theorem for random walks on a graph, following the exposition of Lyons and Peres' *Probability on Trees and Networks* [LP17b]. This paper is organized as follows. We first provide background on Markov chains, largely adapted from Levin and Peres [LP17a], before discussing simple random walks, and then giving an overview of electrical network theory. We then state and prove Pólya's theorem.

2 Markov chains

In this section, we will introduce Markov chains and prove some interesting properties of them. This section closely follows the introduction to Markov chains found in [LP17a].

Intuitively, (time-homogeneous) Markov chains are a stochastic process valued in a fixed $\mathcal{X}$, where the next position at time t only depends on the position at time $t-1$. Formally, we have

Definition 2.1. *Let $\mathcal{X}$ be a set. Let $(X_0, X_1, \dots)$ be a sequence of random variables. $(X_0, X_1, \dots)$ is a Markov chain with state space $\mathcal{X}$ and transition matrix P if for all $x, y \in \mathcal{X}$ and all $t \geq 1$ and all events $H_{t-1} = \bigcap_{s=0}^{t-1}\{X_s = x_s\}$ (the previous positions), we have that*

$$\mathbb{P}(X_{t+1} = y | H_{t-1} \cap \{X_t = x\}) = \mathbb{P}(X_{t+1} = y | X_t = x) = P(x, y). \tag{2.1}$$

Equation 2.1 is called the Markov property.

The Markov property implies that the conditional probability of going from state x to y is the same, no matter what sequence of $x_0, x_1, \dots, x_{t-1}$ was observed before. Thus, the matrix P of dimension $|\mathcal{X}| \times |\mathcal{X}|$ is enough to describe the transition probability. Also, we will have $\sum_{x \in \mathcal{X}} P(x, y)$ for all $y \in \mathcal{X}$.

We will now look at a nice example. Consider a frog in a pond with 2 lily pads, east and west. The frog jumps from east to west with probability p and from west to east with probability q. Then,

$$P = \begin{pmatrix} P(\text{east}, \text{east}) & P(\text{east}, \text{west}) \\ P(\text{west}, \text{east}) & P(\text{west}, \text{west}) \end{pmatrix} = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}.$$

Let $(X_0, X_1, \dots)$ be the sequence of lily pads occupied by the frog at time $0, 1, \dots$. Then, $(X_0, X_1, \dots)$ is a Markov chain with state space $\{\text{east}, \text{west}\}$ and transition matrix P.

Let $\mu_t = (\mathbb{P}(X_t = \text{east} \,|\, X_0 = \text{east}), \mathbb{P}(X_t = \text{west} \,|\, X_0 = \text{east}))$ and $\mu_0 = (1, 0)$. By the Markov property, $\mu_t = \mu_{t-1}P$ for all $t \geq 1$, and thus $\mu_t = \mu_0 P^t$. Thus, we can consider the limit of μ_t as $t \to \infty$, which we will call π.

Proposition 2.2. *$\pi(east) = \frac{q}{p+q}$ and $\pi(west) = \frac{p}{p+q}$.*

This completes our example.

Definition 2.3. *A chain is irreducible if for any two states $x, y \in \mathcal{X}$, there exists an integer t such that $P^t(x, y) > 0$.*

In other words, starting from any state, there is a positive probability of reaching any other state. Any limiting distribution π of such Markov chain must satisfy that $\pi P = \pi$ for P our transition matrix, which motivates our next definition.

Definition 2.4. *A stationary distribution for a transition matrix P is a distribution π such that $\pi = \pi P$ or, equivalently, $\pi(y) = \sum_{x \in \mathcal{X}} \pi(x) P(x, y)$ for all $y \in \mathcal{X}$.*

We will now see a sufficient condition for a distribution to be stationary.

Definition 2.5. *The detailed balance equations for a probability distribution π on $\mathcal{X}$ and transition matrix P are $\pi(x)P(x, y) = \pi(y)P(y, x)$ for all $x, y \in \mathcal{X}$. A chain $(X_t)_{t \geq 0}$ with a distribution π satisfying these equations is called reversible.*

Proposition 2.6. *Any probability distribution π satisfying the detailed balanced equations is a stationary distribution for P.*

Finally, we will define a type of random variable often used in Markov chains which are hitting times.

Definition 2.7. *For $x \in \mathcal{X}$, the hitting time of x is*

$$\tau_x = \min\{t : X_t = x\}.$$

3 Simple random walk

The rest of the paper follows closely the book [LP17b].

We now define simple random walks, which is the object analyzed in Pólya's theorem.

Simple random walks are examples of irreducible and reversible Markov chains. We will first define the random walks. We start with a weighted graph $G = (V, E, c)$ where the weights are defined as a function $c : E \to [0, \infty)$. For $x, y \in V$, we denote $x \sim y$ if x and y are adjacent. Our state space is $V \times V$. Our transition matrix $P(x, y)$ is

$$P(x, y) = \frac{c(x, y)\mathbf{1}_{y \sim x}}{\sum_{z \sim x} c(x, z)}.$$

This defines our random walk, and its associated stationary distribution π respects the detailed balanced equations and

$$c(x, y) = \pi(x)P(x, y), \tag{3.1}$$

for all $x, y \in V$. Notice, thanks to reversibility, that $c(x, y) = c(y, x)$.

We define a random walk as a walk on G, i.e. when the walk is at vertex x, it chooses to move to an adjacent vertex with probability proportional to the weights. A random walk is a simple random walk when the weights are all equal. We may define a random walk starting with a weighted graph (with positive weights) instead of a Markov chain by the converse process. The simple random walk we are interested in is the simple random walk on $\mathbb{Z}^d$ for $d \geq 1$.

We will now define two key concepts which are of importance for Pólya's theorem.

Definition 3.1. *A Markov chain is said to be recurrent if it returns to its starting point infinitely many times with probability 1.*

Definition 3.2. *A Markov chain is said to be transient if it returns to its starting point only finitely many times with probability 1.*

We are now ready to state Pólya's theorem.

Theorem 3.3 (Pólya)**.** *Simple random walk on $\mathbb{Z}^d$ is recurrent if $d \leq 2$, otherwise it is transient.*

4 Electrical network theory

Before presenting the proof of Pólya's theorem, we must first define some key concepts taken from electrical network theory. Simply put, an electrical network is a weighted graph, where the weights are called conductances and their reciprocals resistances. These are denoted for $x, y \in V$, $c(x, y)$ and $r(x, y)$ respectively.

Definition 4.1. *Given a graph $G = (V, E)$, a function $f : V \to \mathbb{R}$ is harmonic at $x \in V$ if $f(x) = \sum_{y \sim x} P(x, y)f(y)$. If f is harmonic at each point of a set $W \subseteq V$, we say f is harmonic on W.*

A harmonic function is a weighted average of the values of the adjacent vertices.

Definition 4.2. *Given a graph $G = (V, E)$ and two sets of vertices A and Z in V, a function $v : V \to [0, \infty)$ is a voltage if it is a harmonic function on every $x \notin A \cup Z$*

Definition 4.3. *Given a graph $G = (V, E)$ and a voltage function v, the current is a function $i : E \to \mathbb{R}$ such that $i(x, y) = c(x, y)[v(x) - v(y)]$ for $x, y \in V$*

Definition 4.4. *Given a graph $G = (V, E)$, a function $\theta : E \to \mathbb{R}$ is a flow between sets A and Z in V if $\theta(x, y) = -\theta(y, x)$ for all $(x, y) \in E$ and $\sum_{y \sim x} \theta(x, y) = 0$ for all $x \notin A \cup Z$.*

We will now define energy for a graph $G = (V, E)$. We will use the standard inner product for the Hilbert space of functions on the vertex set, i.e. $(f, g) := \sum_{x \in V} f(x)\, g(x)$. We will now assume each edge has both orientations, thus the flow in one direction is negative flow for the other direction. We may give a more telling notation to the endpoints of an edge e with e^+ and e^- and thus positive flow goes from e^- to e^+. Also, for $e = (e^-, e^+)$, we define $-e$ to be (e^+, e^-).

Definition 4.5. *An antisymmetric function $\theta : E \to \mathbb{R}$ is such that $\theta(-e) = -\theta(e)$ for all $e \in E$.*

We will also write for functions f, g, h on our edge set, $(f,g)_h$ to mean (f, gh). In particular, our norm $\|f\|_h$ will be $\sqrt{(f,f)_h} = \sqrt{\sum_{x\in V} f^2(x)h(x)}$.

Definition 4.6. *Given the graph $G = (V, E)$, let the function $r : E \to [0, \infty)$ give the value of the resistance for each edge $e \in E$. For any antisymmetric function $\theta : E \to \mathbb{R}$, we define its energy as $\mathscr{E}(\theta) = \|\theta\|_r$.*

We may define some useful operators.

Definition 4.7. *Given the graph $G = (V, E)$, the coboundary operator d from functions on the vertex set V to antisymmetric functions on the edge set E is defined by*

$$(df)(e) := f(e^-) - f(e^+)$$

Definition 4.8. *Given the graph $G = (V, E)$, the boundary operator d^* from antisymmetric functions on the edge set E to functions on the vertex set V is defined by*

$$(d^*\theta)(x) := \sum_{e^- = x} \theta(e).$$

We can notice that the boundary and coboundary operator are adjoint of each other. We can now redefine the notion of flow using these operators.

Definition 4.9. *Given the graph $G = (V, E)$, let A and Z be two disjoint set of vertices in V. An antisymmetric function $\theta : E \to \mathbb{R}$ is called a flow if $d^*\theta$ is 0 off of A and Z. If it is nonnegative on A and non-positive on Z, then we call it a flow from A to Z.*

Definition 4.10. *The strength of a flow $\theta : E \to \mathbb{R}$ is*

$$Strength(\theta) := \sum_{a\in A} d^*\theta(a),$$

which the amount flowing into the network.

Definition 4.11. *A flow $\theta : E \to \mathbb{R}$ of strength 1 is called a unit flow.*

Finally, we will state two theorems which are essential in the proof of Pólya's theorem.

Theorem 4.12. *(Part of Rayleigh's Monotonicity Principle) Let G be an infinite and connected graph with two assignments, c and c', of conductances such that $c \leq c'$ everywhere. Then, if (G, c) is transient, then so is (G, c')*

Definition 4.13.

- *Given a finite graph $G = (V, E)$, let A and Z be two disjoint sets of vertices in V. A cutset is a set $\Pi \subseteq E$ such that it separates A and Z, i.e. every path with one endpoint in A and the other in Z must contain at least one edge in Π.*
- *Given an infinite graph $G = (V, E)$, a cutset is a set $\Pi \subseteq E$ which separates A and ∞. That is, Π is a cutset iff every simple path from A includes an edge in Π.*

Definition 4.14. *Fix a finite graph G and consider a simple random walk on G. The effective conductance between a and Z is $\mathscr{C}_{eff} = \pi(a)\mathbb{P}(a \to Z) = \mathscr{C}(a \leftrightarrow Z)$. The reciprocal is called the effective resistance and is denoted $\mathscr{R}(a \leftrightarrow Z)$.*

Theorem 4.15. *(The Nash-Williams Inequality) Let a and z be two distinct vertices in a finite graph that are separated by cutsets $\Pi_1, \ldots, \Pi_n$, then*

$$\mathscr{R}(a \leftrightarrow z) \geq \sum_{k=1}^{n} \left(\sum_{e\in\Pi_k} c(e) \right)^{-1}$$

From the Nash-Williams Inequality, we can derive the *Nash-Williams Criterion.*

Corollary 4.16. *(Nash-Williams Criterion) If $\langle \Pi_n \rangle$ is a sequence of pairwise disjoint, finite cutsets in a locally finite infinite network G, each of which separates a from ∞, then*

$$\mathscr{R}(a \leftrightarrow \infty) \geq \sum_{n}^{n} \left(\sum_{e \in \Pi_n} c(e) \right)^{-1}$$

Further, if $\sum_{k=1}^{n} \left(\sum_{e \in \Pi_k} c(e) \right)^{-1}$ *is infinite, then G is recurrent.*

5 Pólya's theorem

Recall Theorem 3.3. Many proofs of Pólya's theorem exist, from combinatoric arguments to direct proofs which calculate the Greens function. We will follow the proof outlined by Lyons & Peres [LP17b].

In order to prove Pólya's theorem, we will require Theorem 5.2, which requires the following definition.

Definition 5.1. *For an infinite network G, a sequence $\langle G_n \rangle$ of finite subgraphs of G* exhausts *G if $G_n \subseteq G_{n+1}$ and $G = \cup G_n$.*

Theorem 5.2. *A random walk on an infinite, connected network is transient iff the effective conductance (4.14) from any of its vertices to infinity is positive.*

Proof. Let $\langle G_n \rangle$ exhaust G. For every n, each edge in G_n is an edge in G, so we give it the same conductance as in G. We also assume that G_n is the graph induced in G by the vertices of G_n, $V(G_n)$.

Let $Z_n := V(G \setminus G_n)$. Let G_n^W be the graph obtained from G by identifying Z_n to a single vertex z_n, and removing loops but keeping duplicate edges. This graph will have finitely many vertices but may have infinitely many edges. even when loops are deleted. This is because some vertex of G_n may have infinite degree.

Given a network random walk on G, if we stop the first time it reaches Z_n, the we obtain a network random walk on G_n^W until it reaches z_n. Now for every $a \in G$, the events $[a \to Z_n]$ are decreasing in n, so $\lim_n \mathbf{P}[a \to Z_n] =$ the *escape* probability from a, that is the probability of never returning to a in G. This is positive iff the random walk on G is transient.

Observe that for a fixed vertex a in G, $\lim_n \mathcal{C}(a \leftrightarrow Z_n)$ is the same for every exhaustive sequence $\langle G_n \rangle$ of induced subgraphs of G.

By this observation, the effective conductance from a to ∞, $\mathcal{C}(a \leftrightarrow \infty) := \lim_{n \to \infty} \mathcal{C}(a \leftrightarrow z_n; G_n^W)$, is positive iff the random walk on G is transient. □

We can now prove Pólya's theorem.

Proof. We first show that for $d = 1, 2$, a simple random walk on $\mathbb{R}^d$ is recurrent.

Indeed, let $\mathbf{o}$ be the origin, and $d(\cdot, \cdot)$ be the graph distance. We define the cutsets

$$\Pi_n := \{e \mid d(\mathbf{o}, e^-) = n - 1,\ d(\mathbf{o}, e^+) = n\}$$

Then, $\sum_n \left(\sum_{e \in \Pi_n} c(e) \right)^{-1} = \sum_n (|\Pi_n| \cdot c)^{-1}$ since all edges have the same conductance, and this sum is infinite for $d = 1, 2$. Thus, by the Nash-Williams criterion, the simple random walk on $\mathbb{R}^d$ is recurrent.

We now prove that for $d \geq 3$, the simple random walk on $\mathbb{R}^d$ is transient.
It is sufficient to prove this for $d = 3$ by Rayleigh's monotonicity principle.

Let L be a uniformly distributed (i.e. uniform intersection on the unit sphere) ray from the origin $\mathbf{o} = (0, 0, 0)$ to ∞. Let $\mathcal{P}(L)$ be a simple path in $\mathbb{Z}^3$ from $\mathbf{o}$ to ∞ such that the distance between $\mathcal{P}(L)$ and L is at most 4, and such that $\mathcal{P}(L)$ is chosen measurably.

Define a flow θ by

$$\theta(e) := \sum_{n \geq 0} (\mathbb{P}[e_n = e] - \mathbb{P}[e_n = -e])$$

where $\mathbb{P}$ is the probability measure on the path $\mathcal{P}(L) := \langle e_n \mid n \geq 0 \rangle$. Note that θ is a unit flow from $\mathbf{o}$ to ∞.

Claim: θ has finite energy.
Observe that there exists some constant $A \in \mathbb{R}$ such that if e is an edge with midpoint at Euclidean distance R from $\mathbf{o}$, then $\mathbb{P}[e \in \mathcal{P}(L)] \leq A/R^2$, since $\mathcal{P}(L)$ goes from $\mathbf{o}$ to ∞, thus there is a nonempty set of edges E_R such that each edge in E_R has midpoint at Euclidean distance R from $\mathbf{o}$ in $\mathcal{P}(L)$. Thus, there exists some $A \in \mathbb{R}$ such that $|E_R| \leq AR^2$ so the probability that $e \in \mathcal{P}(L)$ is less than A/R^2.

Note: A is approximately 4π.

Now note that since the Euclidean distance between the midpoints of any two edges in $\mathbb{Z}^3$ is at least $1/\sqrt{2}$, there exists another constant B such that there are at most Bn^2 edges with midpoints at distance from the origin between n and $n + 1$.

Therefore, the energy of θ is at most $\sum_n A^2 B n^2 n^{-4} < \infty$. Then, by Theorem 5.2, G is transient.

□

Acknowledgements

Thank you to our mentor Tasmin Chu for leading a wonderful Directed Reading Program. Thank you as well to Sasha Bell, Yanees Dobberstein, and Hazem Hassan for facilitating the Winter 2025 Directed Reading Program.

References

[LP17a] David A. Levin and Yuval Peres. *Markov Chains and Mixing Times, Second Edition.* Providence. American Mathematical Society, Rhode Island, 2nd edition, 2017.

[LP17b] Russell Lyons and Yuval Peres. *Probability on Trees and Networks.* Cambridge University Press, 2017.