# MA 351 LECTURE NOTES: RANDOM WALKS ON GRAPHS

ANURAG SAHAY

## 1. Introduction

In these lecture notes[1], we will discuss how linear algebra (specifically: diagonalization) can be applied to study random walks on finite graphs. These notes are peppered with problems – these problems are designed to help understand the material, so I recommend doing them as you read, instead of at the end. The problems are *not* for credit. Pay special attention to problems which are starred, since I am likely to test them in the exam.

Without further ado, let us describe an example of a random walk on a graph.

A dog (named Bella) is in a triangular park. The park has three open spaces which are connected by small roads. Consider[2] Figure 1 for a schematic representation of the park where we have labeled the open spaces. Bella is experiencing zoomies. At time 0, Bella starts in space 1. Then, after each second, she gets excited and randomly runs down one of the paths available to her to the open space at the other end. You should imagine she flipped an unbiased coin: if she flips heads, she runs to the space which is counter-clockwise from her current position, while if she flips tails, she runs to the space which is clock-wise. In Figure 2, we have drawn the possible configurations that show up after two seconds.

This is a very simple example of a random walk on a graph – as you shall see below, it is a walk on the graph $K_3$. Our goal in this note is to systematically student walks like these. While this theory also has purely aesthetic value, it is an extremely useful tool with many diverse applications – for e.g., it is the basis of Google's PageRank algorithm, it is used in modeling the spread of disease in epidemiology, it can be used to study shuffling of a deck of cards, and it underlies many Monte Carlo simulation techniques.

*Date*: 29th April, 2025.

**Acknowledgments:** We thank Abhibhav Garg for useful discussions.

[1]These notes are currently incomplete: §8 will be updated later to include a proof sketch for Theorem 7.4.

[2]All figures for this note are hand-drawn and can be found at:

https://www.math.purdue.edu/~sahay5/spring2025/ma35100/notes2_fig.pdf

## 2. Graphs

Before getting into the thick of things, we should spend some time with our basic objects of study, graphs[3], since they are, on the face of it, somewhat disconnected with the topics we have discussed in class till date.

Formally, a graph $G$ is a pair of sets $(V, E)$. In this note, the set $V$ will always be taken to be $\{1, \dots, n\}$, the first $n$ natural numbers for $n \geqslant 1$. This is called the set of vertices (sing. vertex). The set $E$ is a collection "edges" on $V$; an edge is formally defined as a pair of vertices.

When thinking about graphs, however, one should not be too attached to the formalities. It is usually to think about the picture that a graph represents. To draw a graph, one represents vertices by dots (or points) and edges by lines connecting the corresponding vertices. In Figure 3, we depict the graph with $V = \{1, 2, 3\}$ and $E = \{12, 23, 13\}$. This graph is usually called $K_3$ (K for "complete", since all possible edges between vertices are included), though it is sometimes called $C_3$ (C for "cycle", since it looks like a cycle). In Figure 4, we depict the graph with $V = \{1, 2, 3, 4\}$ and $E = \{12, 23, 34, 14\}$. This graph is called $C_4$.

For us, graphs will always be simple and undirected. This means that:

- The edges are unordered: 12 represents the same edge as 21.
- The graph is loopless – the vertices in each edge are distinct. Thus, the edge 11 or 22 is not allowed.
- Multiple edges are not allowed. Thus, 12 may only appear once in $E$.

In Figure 5, we draw some examples of disallowed configurations.

With these exceptions in mind, we can try drawing all the possible graphs with small $n$. A crucial point here is that the labels of the vertices is relatively unimportant; if we desire, we can always relabel the vertices in the beginning. With this in mind, convince yourself that Figure 6 includes all the possible graphs (up to relabeling) with $n \leqslant 3$. In Figure 6, we also include some (but not all) graphs with $n = 4$, namely $C_4$, $K_4$ and $K_2 \oplus K_2$.

**Problem 1.** Draw all possible graphs on 4 vertices.

We say that two vertices $i, j \in V$ are *adjacent* if $ij \in E$. We say that an edge $e$ is *incident* on a vertex $j$ if $e$ contains $j$. That is, $e = ij = ji$ for some other vertex $i \in V$.

If $j \in V$ is a vertex in $G$, a crucial number associated with it is its degree.

**Definition 2.1.** The *degree* of a vertex $j \in V$ is the number of edges in $E$ that are incident on it. This is denoted by $\deg(j)$.

[3] This is the notion of graphs from combinatorics – nothing to do with graphs of a function from calculus!

Let us start with some computation.

**Problem 2.** Compute the degrees of all the vertices in graphs depicted in Figure 6.

For the rest of this note, we will restrict our attention to graphs having constant degree, viz. the regular graphs.

**Definition 2.2.** For fixed $d \geqslant 2$, a graph $G = (V, E)$ is called *$d$-regular* if for every $j \in V$

$$\deg(j) = d.$$

Further, if $G$ is $d$-regular for some $d \geqslant 2$, then we call $G$ a *regular* graph and call $d$ its *regularity*.

Note that $d = 1$ could also be allowed, in principle, but that does not lead to any interesting graphs, especially when one insists that the graph also be connected (see §8 for a definition of connected).

**Problem 3.** Verify that $K_3$ and $C_4$ are 2-regular and $K_4$ is 3-regular. Are any there any other regular graphs in Figure 6? What about in Problem 1? If so, determine their regularity.

While much of what we will discuss now can also be achieved for irregular graphs, the formulae and definitions get more involved. Thus, for simplicity, we will only discuss random walks on regular graphs.

## 3. Random walks on regular graphs

We will now formally describe what a random walk on a regular graph is. We will fix $d \geqslant 2$ and $G$ a $d$-regular graph throughout this discussion. We now introduce a discrete variable $t$ which takes values in $\mathbb{W} = \mathbb{N} \cup \{0\}$. This should be interpreted as a (discrete) *time* variable. The random walker (such as Bella from the introduction; hereinafter, walker) will do a random walk on the graph which transitions as $t$ runs through $\mathbb{W}$.

At time 0, starts at a vertex of the graph $G$. Without loss of generality, we can assume the walker starts at the vertex 1, since we may relabel the vertices to achieve this.

At time[4] $t \in \mathbb{N}$, the walker moves from the vertex they were at at time $t - 1$ to an adjacent vertex uniformly with probability $1/d$ (i.e., at time $t$, the walker is equally likely to be at any vertex adjacent to their position at time $t - 1$ and is definitely not at any of the other vertices).

This is a simple example of a *Markov chain*. In this context, Markov means memoryless: at any given point in time, the history of the walk is irrelevant when

[4]For me, the natural numbers $\mathbb{N} = \{1, 2, 3, \dots\}$ does not include 0, so in this line it is implicit that $t \geqslant 1$.

deciding the next step of the walk. Bella's zoomies from the introduction are now manifestly seen to be modeled by a random walk on $K_3$.

Now, we seek to formally represent the state of the random walk at time $t$.

**Definition 3.1.** A vector $\vec{p} \in \mathbb{R}^n$ is called a *state vector* or a *probability distribution* if

$$p_j \geqslant 0 \qquad\qquad (1 \leqslant j \leqslant n)$$

and

$$\sum_{j=1}^{n} p_j = p_1 + p_2 + \cdots + p_n = 1.$$

Another(!) alternate name for this is *probability vector*.

**Problem 4.** Show that every probability distribution satisfies

$$p_j \leqslant 1 \qquad\qquad (1 \leqslant j \leqslant n).$$

Since the $p_j \in [0, 1]$ and they sum to 1, they can be interpreted as probabilities of disjoint[5] events. In particular, at each time instant $t$, we can describe the state of the random walk using a probability distribution $\vec{p}(t)$ whose components are given by

$$p_j(t) = \Pr[\text{walker is at state } j \text{ at time } t]. \tag{1}$$

Here $\Pr[\cdot]$ stands for the probability that the event in brackets occurs.

For warm-up, you should try the following problems.

**Problem 5.** Check that for every $t \in \mathbb{W}$, $\vec{p}(t)$ is a state vector.

**Problem 6.** Verify that at $t = 0$, $p_1(0) = 1$ and $p_j(0) = 0$ for all $2 \leqslant j \leqslant n$.

**Problem 7.** Verify that at $t = 1$,

$$p_j(1) = \begin{cases} \frac{1}{d} & \text{if } j \text{ is adjacent to } 1, \\ 0 & \text{otherwise.} \end{cases}$$

In Figure 7, we calculate the state vectors at times $t = 0, 1, 2$ for Bella's random walk.

**Problem 8.** Extend Figure 7 to the next step and hence calculate the state vector $\vec{p}(3)$.

Our next goal is to understand how to systematically obtain $\vec{p}(t + 1)$ from $\vec{p}(t)$.

[5] Two or more events are disjoint if no pair of them can happen at the same time.

## 4. Enter the matrix: adjacency and walk

To achieve this goal, we will finally need to get back to linear algebra. In particular, we will associate to each graph two matrices – the adjacency matrix and the walk matrix – which will help us understand the evolution of the state $\vec{p}(t)$.

**Definition 4.1.** The *adjacency matrix* $A$ of a graph $G$ on $n$ vertices is an $n \times n$ matrix whose entries are given by

$$a_{ij} = \begin{cases} 1 & ij \in E, \\ 0 & \text{otherwise.} \end{cases}$$

In other words, the $ij$th entry of $A$ is 1 if $ij$ is an edge and otherwise it is 0.

Applying this definition to $G = K_3$, we see that

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

Similarly, for $G = C_4$, we get

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

**Problem 9*.** Compute the adjacency matrices for $G = K_4$ and $G = K_2 \oplus K_2$.

Note that $ij$ and $ji$ are the same edge. This implies that $a_{ij} = a_{ji}$ for every $i, j \in V$, whence $A$ is symmetric. Further, $jj$ is never an edge (why?) and hence $a_{jj} = 0$, thus the diagonal entries are always zero.

**Definition 4.2.** The *walk matrix* (also called the *transition matrix*) of a $d$-regular graph $G$ is the $n \times n$ matrix defined by the equation

$$M = (1/d)A.$$

In other words,

$$m_{ij} = \begin{cases} 1/d & ij \in E, \\ 0 & \text{otherwise.} \end{cases}$$

The letter $M$ is used because it is the transition matrix of a Markov chain.

**Problem 10*.** Compute the walk matrices for $G \in \{K_3, C_4, K_4\}$.

An extremely important point is that $m_{ij}$ has a probabilistic interpretation:

$$m_{ij} = \Pr[j \longrightarrow i] = \Pr[\text{walker presently at vertex } j \text{ moves to vertex } i]. \qquad (2)$$

This is because if $i$ and $j$ are adjacent, then both $m_{ij}$ and $\Pr[j \longrightarrow i]$ are $1/d$ while if $i$ and $j$ are not adjacent, they are both 0.

## 5. Deriving the recurrence: the law of total probability

Using (2), we will now derive a recurrence relation for $\vec{p}(t)$ that will let us compute it. To do this, we will need a fact from probability theory called the law of total probability. This states that if $E$ is an event and $\{A_j\}_{j=1}^n$ a collection of mutually exclusive but cumulatively exhaustive events, then

$$\Pr[E] = \sum_{j=1}^{n} \Pr[E|A_j] \Pr(A_j).$$

Here $\Pr[E|A]$ is the (conditional) probability of the event $E$ occurring given that the event $A_j$ has occurred. You may have seen this before in a course on probability; in any case, we will simply assume this.

Taking $E$ to be the event that the walker is at vertex $k \in V$ at time $t+1$ and $A_j$ to be the event that the walker is at vertex $j \in V$ at time $t$, we see that

$$\Pr[E|A_j] = \Pr[j \longrightarrow k] = m_{kj}.$$

This is because if the walker is at vertex $j$ at time $t$ and then at vertex $k$ at time $t+1$, then the walker must have walked along an edge from $j$ to $k$. Further, from (1), we get that

$$\Pr[E] = p_k(t+1), \qquad \Pr[A_j] = p_j(t).$$

Thus, the law of total probability implies that

$$p_k(t+1) = \sum_{j=1}^{n} m_{kj} p_j(t). \qquad (3)$$

In principle, we have achieved our mini-goal: this describes $\vec{p}(t+1)$ in terms of $\vec{p}(t)$. However, this description seems quite involved. Thankfully, this is exactly the kind of complication that matrix multiplication was designed to solve! Using the index formula for matrix multiplication, we see that (3) is equivalent to the equation

$$\vec{p}(t+1) = M\vec{p}(t). \qquad (4)$$

## 6. Solving the recurrence

By recursively using (4), we see that

$$\vec{p}(t) = M\vec{p}(t-1) = M(M\vec{p}(t-2)) = M^2\vec{p}(t-2) = M^3\vec{p}(t-3) = \cdots = M^t\vec{p}(0),$$

for every $t \in \mathbb{N}$. Take care not to confuse $M^t$ here (which is $M$ raised to the power $t$) with $M^T$, the transpose of $M$. Simplifying, we get that the recurrence (4) is solved by

$$\vec{p}(t) = M^t\vec{p}(0). \tag{5}$$

Thus, to understand the time evolution of the random walk, we need to compute high powers of the matrix – a task designed for diagonalization!

### 6.1. Generalities.

Before we do some examples, let us recall the general principles of using diagonalization in such a circumstance. Suppose an $n \times n$ matrix $B$ is diagonalizable and we want to compute $B^t\vec{v}$ for some $\vec{v} \in \mathbb{R}^n$. Thus,

$$B = PDP^{-1},$$

where $P = \begin{bmatrix}\vec{b}_1 & \cdots & \vec{b}_n\end{bmatrix}$ is a matrix whose columns are the eigenvectors of $B$ and $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ is a diagonal matrix whose corresponding entries $\lambda_j$ is the eigenvalue corresponding to $\vec{b}_j$. By a recursive calculation,

$$B^t = PD^tP^{-1}.$$

Thus, we see that

$$B^t\vec{v} = (PD^tP^{-1})\vec{v} = PD^t(P^{-1}\vec{v}) = PD^t[\vec{v}]_{\mathcal{B}}, \tag{6}$$

where we have used the fact that $P^{-1} = C_{\mathcal{B}}$, the coordinate matrix of the ordered basis $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$. Further, $D^t = \operatorname{diag}(\lambda_1^t, \ldots, \lambda_n^t)$. Hence, it suffices to diagonalize $B$ and then write $\vec{v}$ in the basis $\mathcal{B}$.

### 6.2. The case $G = K_3$.

We will now turn to solving Bella's random walk in the introduction, viz., the walk on the graph $K_3$. We wish to diagonalize $M$. Before we do this, we note that this is essentially the same as diagonalizing $A$.

**Problem 11*.** Show that

$$(\lambda, \vec{b}) \text{ is an eigenpair for } A \iff (\lambda/d, \vec{b}) \text{ is an eigenpair for } M.$$

Conclude that if $A = PDP^{-1}$ is the diagonalization of $A$, then $M = P\Delta P^{-1}$ is the diagonalization of $M$ where $\Delta = (1/d)D$.

Let us now diagonalize $A$.

6.2.1. *Eigenvalues.* We have that

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \qquad A - \lambda I = \begin{bmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{bmatrix}.$$

A cofactor calculation (that you should check!) now shows that the characteristic polynomial is given by

$$\det(A - \lambda I) = -\lambda^3 + 3\lambda + 2.$$

By some trial and error, we see that $\lambda = -1$ and $\lambda = 2$ are the two zeroes and we obtain the factorization,

$$\det(A - \lambda I) = -(\lambda - 2)(\lambda + 1)^2.$$

Thus, $\lambda_1 = 2$ is an eigenvalue with algebraic multiplicity 1 and $\lambda_2 = -1$ is an eigenvalue with algebraic multiplicity 2.

6.2.2. *Eigenspace for* $\lambda_1 = 2$. We need to find the nullspace of

$$A - \lambda_1 I = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}.$$

By row-reduction (check!), we find that $A - \lambda_1 I$ is row-equivalent to the RREF matrix

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

From this, it follows that the nullspace (and hence the eigenspace of $\lambda_1$) is the span of $(1, 1, 1)$.

6.2.3. *Eigenspace for* $\lambda_2 = -1$. We seek now, the nullspace of

$$A - \lambda_2 I = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

It is now evident that the RREF matrix will be given by

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus, the eigenspace of $\lambda_2 = -1$ is the span of $\{(-1, 0, 1), (-1, 1, 0)\}$.

6.2.4. *Diagonalizing.* Thus, we get that

$$A = PDP^{-1}, \qquad M = P\Delta P^{-1}$$

where

$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \qquad \Delta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}, \qquad P = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

6.2.5. *Dénouement.* Thus, on utilizing (6) to simplify (5), we find that

$$\vec{p}(t) = M^t \vec{p}(0) = P\Delta^t [\vec{p}(0)]_{\mathcal{B}}$$

where $\mathcal{B} = \{\vec{b}_1, \vec{b}_2, \vec{b}_3\}$ are the columns of $P$. Recall that $\vec{p}(0) = (1, 0, 0)$ since Bella deterministically starts at vertex 1. A quick calculation shows that

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix},$$

whence $[\vec{p}(0)]_{\mathcal{B}} = (1/3, -1/3, -1/3)$.

Thus,

$$\vec{p}(t) = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & (-\frac{1}{2})^t & 0 \\ 0 & 0 & (-\frac{1}{2})^t \end{bmatrix} \begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} + \frac{2}{3} \cdot (\frac{-1}{2})^t \\ \frac{1}{3} - \frac{1}{3} \cdot (\frac{-1}{2})^t \\ \frac{1}{3} - \frac{1}{3} \cdot (\frac{-1}{2})^t \end{bmatrix} \tag{7}$$

Et voilà!

**Problem 12.** Check that (7) agrees with Figure 7 and Problem 8.

**Problem 13.** Observe that (7) implies that

$$\lim_{t\to\infty} p(\vec{t}) = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}.$$

This last problem implies that after a very long time, Bella is essentially equally likely to be in any of the three vertices. This is an example of a *mixing* walk. We will discuss this in more detail in §7.

6.3. **The case** $G = C_4$**.** We now repeat this calculation for $G = C_4$, since this exhibits a different behaviour. We will be a bit sketchier with the details.

6.3.1. *Eigenvalues.* We have that

$$\det(A-\lambda I)=\begin{vmatrix}-\lambda & 1 & 0 & 1\\ 1 & -\lambda & 1 & 0\\ 0 & 1 & -\lambda & 1\\ 1 & 0 & 1 & -\lambda\end{vmatrix}=\lambda^4-4\lambda^2=\lambda^2(\lambda-2)(\lambda+2).$$

6.3.2. *Eigenvectors.* By row-reduction, we find that

- The eigenvector of $\lambda_1=2$ is $\vec{b}_1=(1,1,1,1)$.
- The eigenvector of $\lambda_2=-2$ is $\vec{b}_2=(-1,1,-1,1)$.
- The eigenvectors of $\lambda_3=\lambda_4=0$ are $\vec{b}_3=(-1,0,1,0)$ and $\vec{b}_4=(0,-1,0,1)$.

6.3.3. *Diagonalization.* Thus

$$A=PDP^{-1},\qquad M=P\Delta P^{-1}$$

where

$$D=\begin{bmatrix}2 & 0 & 0 & 0\\ 0 & -2 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix},\qquad \Delta=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix},\qquad P=\begin{bmatrix}1 & -1 & -1 & 0\\ 1 & 1 & 0 & -1\\ 1 & -1 & 1 & 0\\ 1 & 1 & 0 & 1\end{bmatrix}.$$

6.3.4. *Dénouement.* Thus, by (5) and (6)

$$\vec{p}(t)=P\Delta^t[\vec{p}(0)]_{\mathcal{B}},$$

where $\mathcal{B}=\{\vec{b}_1,\vec{b}_2,\vec{b}_3,\vec{b}_4\}$. Since we start the walk at vertex 1,

$$\vec{p}(0)=(1,0,0,0)=(1/4)\vec{b}_1+(-1/4)\vec{b}_2+(-1/2)\vec{b}_3+0\vec{b}_4.$$

Thus, we find that

$$\vec{p}(t)=\begin{bmatrix}1 & -1 & -1 & 0\\ 1 & 1 & 0 & -1\\ 1 & -1 & 1 & 0\\ 1 & 1 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & (-1)^t & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}1/4\\ -1/4\\ -1/2\\ 0\end{bmatrix}=\begin{bmatrix}1/4+(-1/4)^t\\ 1/4-(-1/4)^t\\ 1/4+(-1/4)^t\\ 1/4-(-1/4)^t\end{bmatrix}$$

Note that this formula is only valid for $t\in\mathbb{N}$; it fails for $t=0$!

**Problem 14*.** Check all the details of the above calculation for $G=C_4$.

**Problem 15.** Let $\vec{v}=(1/2,0,1/2,0)$ and $\vec{w}=(0,1/2,0,1/2)$. Show that for $t\geqslant 1$,

$$\vec{p}(t)=\begin{cases}\vec{v} & \text{if } t \text{ is even,}\\ \vec{w} & \text{if } t \text{ is odd.}\end{cases}$$

Thus, conclude that $\vec{p}(t)$ does not tend to a limit as $t\to\infty$.

A visual representation of how a random walk on $C_4$ behaves is given in Figure 8. This is an example of a *non-mixing* walk, since the probability distribution does not spread around the entire graph as time passes.

## 7. Stationary distributions and mixing walks

Let us now try to understand why the long-range behaviour of the random walks on $K_3$ and $C_4$ were different. The first observation is that the limiting distribution $\vec{u} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ for the walk on $K_3$ has the property that it does not change under a single step of the random walk:

$$M\vec{u} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} = \vec{u}.$$

It's not too hard to Figure out that if a limit to $\vec{p}(t)$ exists as $t \to \infty$, then it must satisfy $M\vec{u} = \vec{u}$. To see this, take limits on both sides of (4),

$$\lim_{t\to\infty} \vec{p}(t+1) = \lim_{t\to\infty} M\vec{p}(t).$$

Since $M$ is constant, the right hand side should converge to $M\vec{u}$, while the left hand side clearly converges to $\vec{u}$. A probability distribution which has this property is called a *stationary distribution.*

**Definition 7.1.** A probability distribution $\vec{u} \in \mathbb{R}^n$ is called a stationary distribution for a random walk on a graph on $n$ vertices if

$$M\vec{u} = \vec{u},$$

where $M$ is the walk matrix of the graph.

**Problem 16*.** Show that $\vec{u} = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ is a stationary distribution for the graph $C_4$.

Note that if $\vec{u}$ is a stationary distribution, then it is an eigenvector of $M$ with eigenvalue 1. It's not clear that such an eigenvector always exists. However, we have the following theorem.

**Theorem 7.2.** *Every random walk on an undirected simple graph has a stationary distribution.*

In the general case, this is a difficult theorem; it uses Perron-Frobenius theory. In the case of regular graphs, however, it is reasonably elementary.

**Problem 17.** Prove Theorem 7.2 for $d$-regular graphs. [**Hint:** try to guess what the answer must be by generalizing the cases $K_3$ and $C_4$. Proving your guess will require the index formula for matrix multiplication.]

Problem 16 and 17 indicated that $C_4$ has a stationary distribution. Yet, random walks on $C_4$ do not converge to the stationary distribution. Let us be precise.

**Definition 7.3.** We say that a random walk on a graph *converges* or *mixes* if

$$\lim_{t\to\infty} \vec{p}(t) = \vec{u},$$

where $\vec{p}(t)$ is state vector of a random walk starting on any vertex and $\vec{u}$ is the stationary distribution. In other words, we have

$$\lim_{t\to\infty} p_j(t) = u_j \qquad\qquad (1 \leqslant j \leqslant n).$$

In adjective form, such a walk is called a *convergent* random walk or a *mixing* random walk. This naturally raises the question of which random walks mix. The following theorem gives a complete answer.

**Theorem 7.4.** *Let $G$ be an undirected, simple graph on $n$ vertices. Then,*

*The random walk on the graph $G$ mixes $\iff$ $G$ is connected and non-bipartite.*

In the next (and final) section, we will describe what connected and bipartite mean and provide a sketch of the proof of Theorem 7.4 in the case of regular graphs.

## 8. Mixing if and only if connected and not bipartite

A *path* in a graph is a sequence of vertices $(v_1, v_2, \cdots, v_k)$ where consecutive vertices form an edge. We say that the graph is *between* $v_1$ and $v_k$. Thus, for example (recall Figure 6), $(1, 2, 3)$ is a path in both $K_3$ and $C_4$, but it is not a path in $K_2 \oplus K_2$. A graph is called *connected* there exists a path between any two vertices. Both $K_3$ and $C_4$ are connected, but $K_2 \oplus K_2$ is not.

A graph $G = (V, E)$ is called *bipartite* if there is a partition $V = A \cup B$ of the vertex set into two disjoint sets such that every edge of the graph $G$ has exactly one vertex in both $A$ and $B$. $C_4$ is bipartite (see Figure 9) but $K_3$ is not.

**Problem 18.** For each of the graphs in Figure 6, determine if they are connected, bipartite, both, or neither.

Thus, Theorem 7.4 correctly predicts that the walk on $K_3$ mixes but the walk on $C_4$ does not: they are both connected, but the latter is bipartite while the former is not.

It should be clear that connectedness is a prerequisite for mixing: if the graph is disconnected, then a random walk that starts in one connected component cannot possibly escape and go to another connected component (see Figure 10 for an example).

It is not so clear why bipartiteness should be a problem. Essentially, this is because (cf. Problem 15) there is a sequence of looping states which go back and forth: there are state vectors $\vec{v} \neq \vec{w}$ with the property that

$$M\vec{v} = w, \qquad M\vec{w} = v.$$

Thus, random walks can get caught in a loop where they alternate between the states $\vec{v}$ and $\vec{w}$ and thus never converge. A schematic example is shown in Figure 11. In the general case, these distributions are the ones which have zero probability in $B$ but are equally likely to be in any vertex of $A$ (and vice-versa).

Department of Mathematics, Purdue University
*Email address*: anuragsahay@purdue.edu