Random walks, directed cycles and Markov chains

Kate Gingell and Franklin Mendivil

November 22, 2021

Abstract

A Markov chain is a random process which iteratively travels around in its state space with each transition only depending on the current position and not on the past. When the state space is discrete, we can think of a Markov chain as a special type of random walk on a directed graph. Although a Markov chain normally never settles down but keeps moving around, it does usually have a well-defined limiting behaviour in a statistical sense.

A given finite directed graph can potentially support many different random walks or Markov chains and each one could have one or more invariant (stationary) distributions. In this paper we explore the question of characterizing the set of all possible invariant distributions. The answer turns out to be quite simple and very natural and involves the cycles on the graph.

1 Introduction.

Suppose we take a random walk on the vertices of the graph shown in the left part of Figure 1. We start at vertex "a" and each time we reach a vertex with a choice of outgoing edges, we choose one according to the probabilities that label these outgoing edges. As is well-known, this generates a (discrete time) Markov chain $x_1, x_2, x_3, \ldots$ with the vertices of the graph as its state space. Since we start at a we have $x_1 = a$.

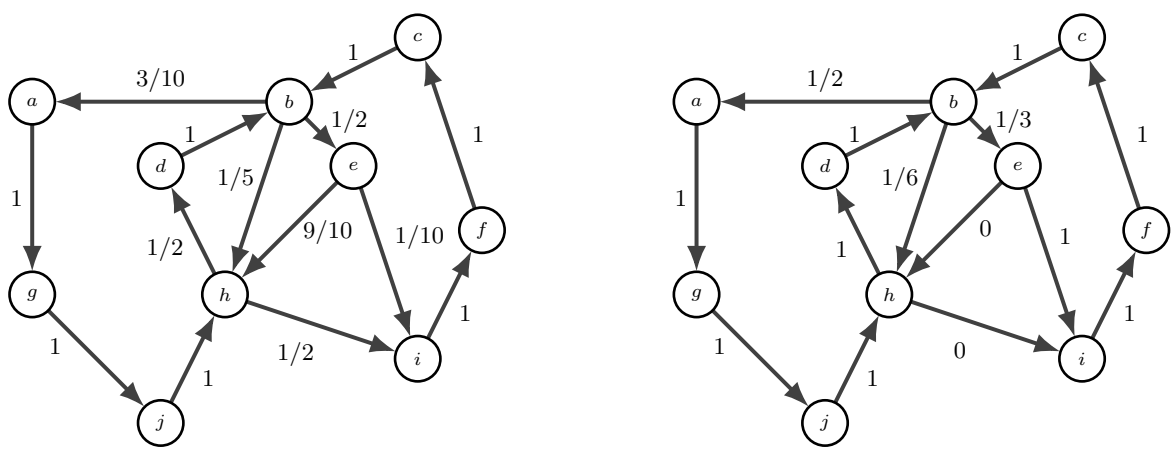

Figure 1: Directed graph with two different probability edge weights.

Because the chain satisfies the hypotheses of the Perron-Frobenius theorem there is a unique (attractive) invariant distribution

$$\pi = (1/18,\ 5/27,\ 7/72,\ 19/216,\ 5/54,\ 7/72,\ 1/18,\ 19/108,\ 7/72,\ 1/18)$$

which satisfies $\pi P = \pi$ where P is the transition matrix for the Markov chain. We notice that $\pi_a = \pi_g = \pi_j = 1/18$ and $\pi_c = \pi_f = \pi_i = 7/72$. Moreover, the ergodic theorem also tells us that (almost surely) for

any vertex v,

$$\pi_v = \lim_m \frac{1}{m} \left|\{1 \le i \le m : x_i = v\}\right|, \tag{1}$$

where we use $|S|$ to denote the size of a finite set S. The quantity $\frac{1}{m}\left|\{1 \le i \le m : x_i = v\}\right|$, as we vary the vertex v, is the empirical occupation distribution, up to iteration m, generated by the walk. We think of both the empirical occupation distribution and the limit π either as functions on the vertices or as vectors with one component for each vertex.

We can imagine changing some of the edge probabilities (at least those for which there is a choice) and running another random walk on the same directed graph. The right-hand image in Figure 1 shows the same graph with different edge probabilities. Again there is a unique attractive invariant distribution, this time given by

$$\pi = (3/31,\ 6/31,\ 2/31,\ 4/31,\ 2/31,\ 2/31,\ 3/31,\ 4/31,\ 2/31,\ 3/31)$$

and again we have that $\pi_a = \pi_g = \pi_j$ and $\pi_c = \pi_f = \pi_i$, though this time we also have $\pi_e = \pi_c$.

Clearly as we change these edge weights we will get different invariant distributions. The main question we explore in this paper is:

Question: *What does the set of possible invariant distributions "look" like?*

Certainly the structure of the given directed graph influences the possible invariant distribution, but in what way? We noticed that for both examples above we have the same limiting probabilities for certain subsets of the vertices. Why these particular vertices? Is this a general feature for every directed graph (perhaps of some special form)?

We expect the invariant distribution to be a continuous function of the edge probabilities, at least for the open set of parameter values where there is a unique such invariant distribution. For each vertex, the set of all possible probability edge weights for its outgoing edges is a compact and convex set. Does this mean that the set of possible invariant distributions is also a compact and convex set? If so, what is the dimension of this set and what are the extreme points?

For our example graph, we only have three vertices where there is a choice of probabilities on the outgoing edges; the general situation is illustrated in Figure 2. The natural constraints on the parameters are

$$\alpha, \beta, \gamma, \delta \in [0, 1] \text{ with } \alpha + \beta \le 1$$

and these ensure that at each vertex the labels on the outgoing edges all correspond to probabilities. With this choice of parameters, the vector

$$\frac{1}{C}(\alpha,\ 1,\ 1 - \gamma\beta\delta + \beta\delta - \delta,\ \gamma\beta\delta - \beta\delta + \delta,\ \beta,\ 1 - \gamma\beta\delta + \beta\delta - \delta,\ \alpha,\ 1 + \beta\gamma - \beta,\ 1 - \gamma\beta\delta + \beta\delta - \delta,\ \alpha) \tag{2}$$

is an invariant distribution, with $C = (2(1-\gamma)\delta + \gamma)\beta + 3\alpha - 2\delta + 5$ being a normalizing factor. For this particular graph all (appropriate) choices of the parameters result in a unique invariant distribution. We notice that $\pi_a = \pi_g = \pi_j$ as well as $\pi_c = \pi_f = \pi_i$, as occurred in the two examples above where we made specific choices of the parameters.

We will divide our discussion into two natural subtopics. First we imagine that we have a somewhat arbitrary walk on $\mathcal{G}$ and find a property that must hold for any "limiting distribution" of the walk. Then, having this property in hand, we will ask if all the distributions with this property are actually possible invariant distributions for some Markov chain on $\mathcal{G}$. Amazingly we will discover that the set of "limiting distributions" for ANY walk on $\mathcal{G}$ is no larger than the set of invariant distributions for Markov chains supported on $\mathcal{G}$ (see Theorem 7 and Theorem 8). Thus even though walks generated by Markov chains form a very small subclass of all possible walks, their sets of "limiting" distributions are the same.

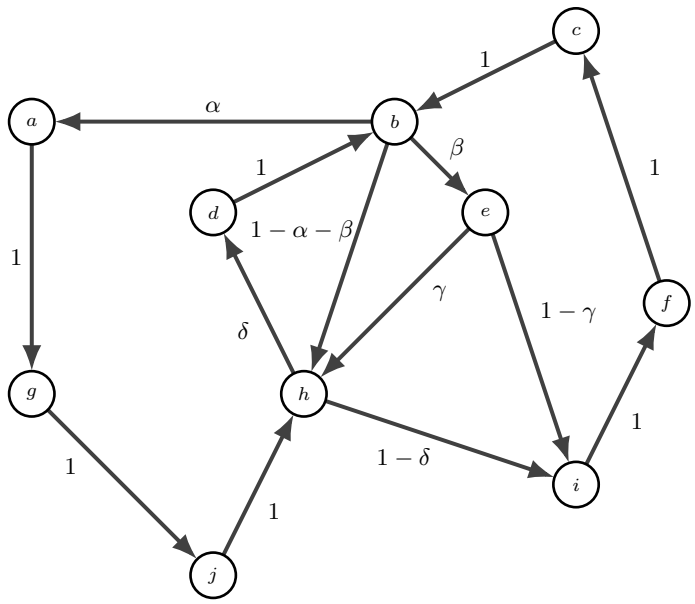

Figure 2: The example directed graph with general probability edge weights.

Notation and definitions.

All our graphs are finite and directed. An edge is thought of as an ordered pair of vertices with a *source* and *target* and we will write such an edge as $u \to v$. We do allow an edge to go from a vertex to itself but do not allow more than one edge with a given source and target. A *directed walk* is a list of vertices $v_0, v_1, \ldots, v_k$ where each $v_{i-1} \to v_i$ is in $\mathcal{G}$ for $i = 1, 2, \ldots, k$. Note that there is no requirement that the vertices are not repeated. A directed walk is *closed* if the walk starts and ends at the same vertex. A *(directed) cycle* is a closed walk in which no vertex is repeated. If $C = v_0, v_1, \ldots, v_k$ is a cycle then the *length of* C, denoted $|C|$, is equal to k, the number of distinct vertices. (We note that this is not the standard definition, as the length is usually defined to be the number of edges.)

The directed graph $\mathcal{G}$ is *strongly connected* if for any two vertices u and v, there is a path from u to v. We note that if $\mathcal{G}$ is strongly connected, then every vertex u is contained in a cycle. Roughly we just piece together the path from u to v with the path from v to u, removing the repeated "side excursions" to get a set of distinct vertices. Thinking about why this is true is a good warm-up to what we will discuss in Section 2. The strongly connected *components* of $\mathcal{G}$ are the maximal strongly connected subgraphs of $\mathcal{G}$.

Not every directed graph $\mathcal{G}$ can support a Markov chain. In order to be able to do so, each vertex in $\mathcal{G}$ has to have at least one outgoing edge (perhaps to itself). Of course a graph where each vertex has only one outgoing edge is rather simple! As a point of nomenclature, we say that a Markov chain with state space the vertices of $\mathcal{G}$ and transition matrix P is *compatible* with $\mathcal{G}$ if $P_{u,v} > 0$ implies $u \to v$ is an edge in $\mathcal{G}$.

For the reader who would like to think ahead on the main questions of this discussion, the graph in Figure 3 is a good place to start. What are the possible limiting distributions for this graph? What happens if we add another chordal edge across the outer cycle? How would this change the possible limiting distributions?

2 From random walks to directed cycles.

In this section we will answer one half of our main question by showing that any "limiting distribution" of an infinite walk on $\mathcal{G}$ has to be of a very special form. For our purposes we say that the walk (x_n) on $\mathcal{G}$ has

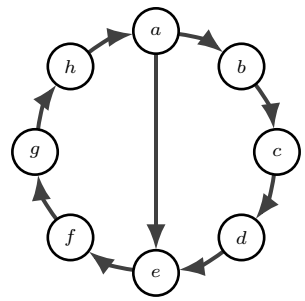

Figure 3: What are the possible limiting distributions for this graph?

a limiting distribution if the limit

$$\pi_v = \lim_{m\to\infty} \frac{1}{m} \left|\{1 \le i \le m : x_i = v\}\right|$$

exists for each vertex v in $\mathcal{G}$; that is, any walk for which the empirical occupation distribution has a limit.

We start with the simplest possible example in order to set the stage for the more formal results which will follow.

Example 1. Suppose we consider the graph $\mathcal{G}$ with two cycles as in Figure 3. The only vertex at which there is a choice of outgoing edges is the vertex "a", where we can either take the edge $a \to b$ or the edge $a \to e$. Set $Pr(a \to b) = p$ so that $Pr(a \to e) = 1 - p$, where $0 \le p \le 1$. This Markov chain has the transition matrix

$$P = \begin{pmatrix} 0 & p & 0 & 0 & 1-p & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

and we can solve that the (unique) invariant distribution is

$$\pi = \left(\frac{1}{3p+5}, \frac{p}{3p+5}, \frac{p}{3p+5}, \frac{p}{3p+5}, \frac{1}{3p+5}, \frac{1}{3p+5}, \frac{1}{3p+5}, \frac{1}{3p+5}\right). \tag{3}$$

In this case it is not so hard to understand exactly how this limiting behavior arises. For a very large number m of visits to a, by the law of large numbers we will take the edge $a \to b$ approximately $p \cdot m$ times and the edge $a \to e$ approximately $(1-p) \cdot m$ times. When we take the $a \to b$ edge, we continue around on the cycle

$$C_1 = a,\ b,\ c,\ d,\ e,\ f,\ g,\ h,\ a$$

while each time we take the $a \to e$ edge we take the cycle

$$C_2 = a,\ e,\ f,\ g,\ h, a.$$

These are obviously the only two cycles in the graph.

On the mth return visit to vertex a, we will have taken cycle C_1 approximately pm times and cycle C_2 approximately $(1-p)m$ times. This gives a rough total of $8pm$ steps corresponding to cycle C_1 and $5(1-p)m$ steps corresponding to C_2 for a total of (approximately) $N = (3p+5)m$ steps since the first visit to vertex a. Using this, we have

$$\frac{1}{N} \left|\{1 \le i \le N : x_i = v\}\right| \approx \begin{cases} \frac{p}{3p+5}, & \text{if } v \in \{b, c, d\} \\ \frac{1}{3p+5}, & \text{if } v \in \{a, e, f, g, h\}, \end{cases} \tag{4}$$

which is consistent with (3).

Now let $\mathcal{L}_i$ count the number of times we go around the cycle C_i during the first $N \approx (3p+5)m$ steps of the walk. The proportion of these N steps that are associated with C_1 is

$$\frac{\mathcal{L}_1|C_1|}{N} \approx \frac{8pm}{(3p+5)m} = \frac{8p}{3p+5} \tag{5}$$

while the proportion for C_2 is

$$\frac{\mathcal{L}_2|C_2|}{N} \approx \frac{5(1-p)m}{(3p+5)m} = \frac{5(1-p)}{3p+5}. \tag{6}$$

We note that

$$\left(\frac{1}{3p+5}, \frac{p}{3p+5}, \frac{p}{3p+5}, \frac{p}{3p+5}, \frac{1}{3p+5}, \frac{1}{3p+5}, \frac{1}{3p+5}, \frac{1}{3p+5}\right) = \tag{7}$$

$$\frac{8p}{3p+5}\left(\frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}\right) + \frac{5(1-p)}{3p+5}\left(\frac{1}{5}, 0, 0, 0, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right). \tag{8}$$

Examining the two parts of this decomposition we see that

$$\left(\frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}\right) \quad \text{and} \quad \left(\frac{1}{5}, 0, 0, 0, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right) \tag{9}$$

are two "uniform" distributions which are supported on C_1 and C_2 (respectively). The weights in this combination (from (7)) are given by the proportions from equations 5 and 6. Note that these weights are rational functions of p but are not as simple as just being p and $1-p$. Of course for this graph the existence of a relationship like that in (7) is not all that surprising since any walk on $\mathcal{G}$ has to "go around" these two cycles and once on a cycle it has to continue until the end of that cycle.

The two distributions in equation (9) are examples of an important class of distributions on $\mathcal{G}$. Given a cycle C, the *uniform cycle distribution based on* C is the probability distribution

$$\theta(v) = \begin{cases} \frac{1}{|C|}, & \text{if } v \in C, \\ 0, & \text{otherwise} \end{cases} \quad = \quad \frac{1}{|C|}\mathbb{1}_C(v), \tag{10}$$

where we use $\mathbb{1}_S$ to denote the indicator function of a set S. The uniform cycle distributions for a given directed graph $\mathcal{G}$ play a central role in our discussions.

The general situation is, of course, more complicated than what we have just seen in Example 1. However, the "feel" is the same in that any closed walk must involve cycles in the graph. The next lemma contains the essential combinatorial insight and is a well-known folklore result. Roughly the lemma states that we cut-and-paste the cycles together to form any closed walk. The cycles in the decomposition are not disjoint (the endpoints of one cycle necessarily have to intersect another cycle) but they are almost disjoint. To describe this situation, given a cycle $v_0, v_1, \ldots, v_k$ we say that its *truncation* is the path $v_1, v_2, \ldots, v_k$.

Lemma 2. *Any finite closed directed walk $x_1, x_2, x_3, \ldots, x_m$ can be decomposed into a collection of directed cycles so that each x_i is in exactly one truncation for each $i > 1$.*

Proof. Since the walk is closed we know that $x_1 = x_m$. If these are the only repeated vertices then the walk is already a directed cycle and there is nothing to prove.

Suppose then that $x_1 = x_m$ is repeated more than twice and let $x_1, x_{n_1}, x_{n_2}, \ldots, x_{n_\ell}, x_m$ be these repeats. Then each of $x_1, x_2, \ldots, x_{n_1}$, and $x_{n_1}, x_{n_1+1}, \ldots, x_{n_2}$, and $x_{n_2}, x_{n_2+1}, \ldots, x_{n_3}$, etc., are finite closed directed walks with the further property that the first vertex (which is the same as the last) is not repeated internal to the walk. Each of these walks is strictly shorter than the original and so we (recursively) repeat the argument with these.

Thus we can assume that we have arrived at a single walk $x_1, x_2, \ldots, x_m$ where $x_1 = x_m$ is not repeated again but some other vertex v is repeated. Letting $x_{n_1}, x_{n_2}, \ldots, x_{n_\ell}$ be the occurrences of v in sequence along the walk, we see that each of

$$\begin{array}{l} x_{n_1}, x_{n_1+1}, x_{n_1+2}, \ldots, x_{n_2}, \text{ and} \\ x_{n_2}, x_{n_2+1}, x_{n_2+2}, \ldots, x_{n_3}, \text{ and} \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ x_{n_{\ell-1}}, x_{n_{\ell-1}+1}, x_{n_{\ell-1}+2}, \ldots, x_{n_\ell}, \text{ and} \\ x_1, x_2, \ldots, x_{n_1}, x_{n_\ell+1}, x_{n_\ell+2}, \ldots, x_m \end{array}$$

is a closed finite directed walk where the first vertex is not repeated internal to the walk (see Figure 4 for an illustration of this case with the initial and final segments joined as one of the cycles). Again each of these walks is strictly shorter than the original and we can recursively repeat the argument with each of these.

After a finite number of these steps we have arrived at a complete decomposition as desired. □

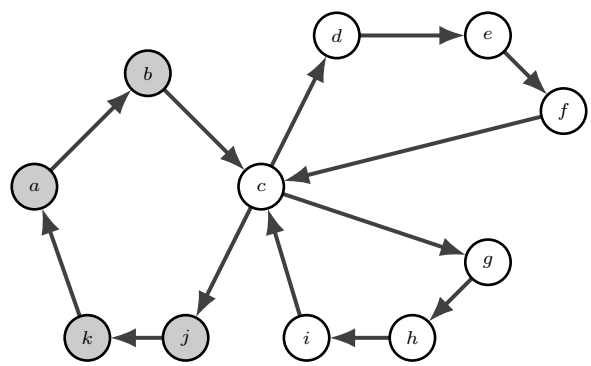

Figure 4: Decomposition of the closed walk $\underline{a}\ \underline{b}\ c\ d\ e\ f\ c\ g\ h\ i\ c\ \underline{j}\ \underline{k}\ \underline{a}$ into cycles.

The decomposition given in the Lemma need not be unique. As an example, consider the closed walk

$$a,\ b,\ c,\ \underline{d},\ e,\ f,\ g,\ \underline{\underline{h}},\ i,\ j,\ \underline{\underline{k}},\ \underline{d},\ l,\ m,\ \underline{\underline{h}},\ n,\ o,\ \underline{\underline{k}},\ p,\ q,\ \underline{\underline{h}},\ r,\ s,\ \underline{d},\ a$$

which can be decomposed in the two different ways illustrated in Figure 5. These two different decompositions can be found by switching which repeated vertex is used to break the overall walk into segments. For the decomposition in the left-side of Figure 5, we first break into segments at the vertex d, giving the closed walks

$$a,\ b,\ c,\ d,\ a \quad \text{and} \quad d,\ e,\ f,\ g,\ h,\ i,\ j,\ k,\ d \quad \text{and} \quad d,\ l,\ m,\ h,\ n,\ o,\ k,\ p,\ q,\ h,\ r,\ s,\ d$$

and then decompose each of these as necessary. For the decomposition in the right-side of Figure 5, we first break into segments at the vertex h, giving the closed walks

$$a,\ b,\ c,\ d,\ e,\ f,\ g,\ h,\ r,\ s,\ d,\ a \quad \text{and} \quad h,\ i,\ j,\ k,\ d,\ l,\ m,\ h \quad \text{and} \quad h,\ n,\ o,\ k,\ p,\ q,\ h$$

then decompose each of these as necessary.

This potential lack of uniqueness will be something that we need to account for but, fortunately, will not cause any major difficulties and is simple enough to deal with.

Our next step is to use this decomposition of a (finite and closed) walk into cycles to break our infinite walk into cycles. Of course the way we do this is by using a sequence of decompositions, but we need to make

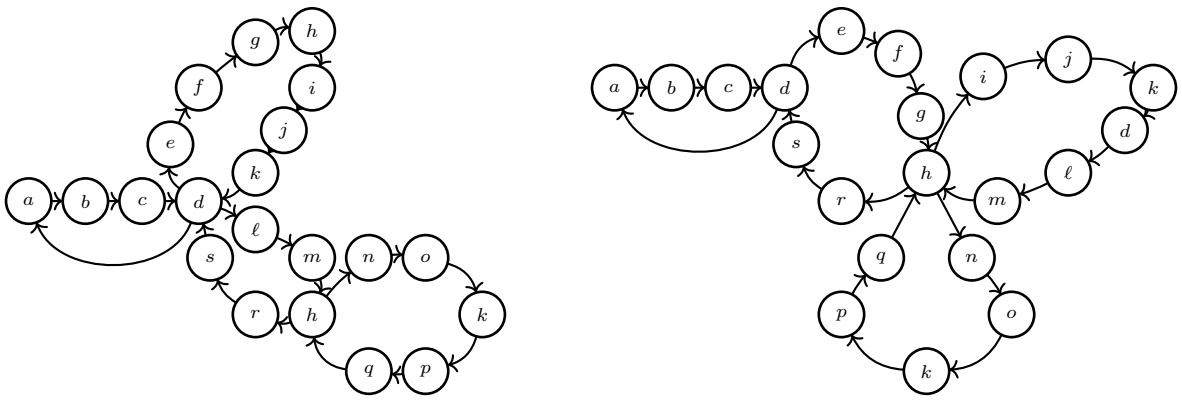

Figure 5: Two different cycle decompositions of the same closed walk.

sure that we have convergence of the relative proportion of "time" that is used for each of the cycles. After this we argue that these limiting proportions (one for each cycle) are related to the limiting distribution for the walk.

Theorem 3. *Let $\mathcal{G}$ be a directed graph and $x_1, x_2, \ldots, x_n, \ldots$ be an infinite walk on $\mathcal{G}$ so that*

$$\pi_v = \lim_{m\to\infty} \frac{1}{m} \left|\{1 \le i \le m : x_i = v\}\right| \tag{11}$$

exists for each vertex v of $\mathcal{G}$. Then there are directed cycles $C_1, C_2, \ldots, C_N$ on G and $\lambda_j \in [0,1]$ with $\sum_{j=1}^N \lambda_j = 1$ so that $\pi = \sum_{j=1}^N \lambda_j \frac{1}{|C_j|} \mathbb{1}_{C_j}$.

Proof. Let the list of all of the possible directed cycles on $\mathcal{G}$ be $C_1, C_2, \ldots, C_N$.

First imagine that the vertex x_1 is visited infinitely many times and let $n_1 = 1 < n_2 < n_3 < \ldots$ be the "times" at which this vertex is visited. From Lemma 1, we can decompose each closed walk $x_{n_{i-1}}, x_{n_{i-1}+1}, \cdots, x_{n_i}$ into "almost" disjoint cycles, which necessarily must come from the list $C_1, C_2, \ldots, C_N$. Let $L_j^{(i)}$ be the number of times cycle C_j occurs in this decomposition. Furthermore, let $\mathcal{L}_j^{(i)} = L_j^{(1)} + L_j^{(2)} + \cdots + L_j^{(i)}$, so that $\mathcal{L}_j^{(i)}$ is the total number of times cycle C_j occurs in the chosen decomposition of the closed walk $x_1, x_2, \ldots, x_{n_i}$.

By counting the number of vertices visited, we can see that

$$\mathcal{L}_1^{(i)}|C_1| + \mathcal{L}_2^{(i)}|C_2| + \cdots + \mathcal{L}_N^{(i)}|C_N| = n_i - 1$$

and thus

$$\frac{\mathcal{L}_1^{(i)}|C_1| + \mathcal{L}_2^{(i)}|C_2| + \cdots + \mathcal{L}_N^{(i)}|C_N|}{n_i - 1} = 1$$

In addition, for each vertex v of the graph $\mathcal{G}$,

$$\left|\{1 \le k < n_i : x_k = v\}\right| = \sum_{j=1}^N \mathcal{L}_j^{(i)} \mathbb{1}_{C_j}(v).$$

Now by assumption we know that

$$\pi_v = \lim_{i\to\infty} \frac{1}{n_i - 1} \sum_{j=1}^N \mathcal{L}_j^{(i)} \mathbb{1}_{C_j}(v) = \lim_{i\to\infty} \sum_{j=1}^N \left(\frac{\mathcal{L}_j^{(i)}|C_j|}{n_i - 1} \right) \frac{1}{|C_j|} \mathbb{1}_{C_j}(v) \tag{12}$$

exists. Thus we wish to take

$$\lambda_j = \lim \frac{\mathcal{L}_j^{(i)}|C_j|}{n_i - 1} \tag{13}$$

and so we have to argue that some form of a suitable limit does exist. Unfortunately the limit as written in (13) with $i \to \infty$ does not have to exist. One reason for this is the amount of choice (because of non-uniqueness) that is allowed when using Lemma 2. To deal with this issue we notice that the sequence of vectors

$$\mathcal{V}_i = \left(\frac{\mathcal{L}_1^{(i)}|C_1|}{n_i - 1}, \ \frac{\mathcal{L}_2^{(i)}|C_2|}{n_i - 1}, \ \cdots, \ \frac{\mathcal{L}_N^{(i)}|C_N|}{n_i - 1} \right)$$

is always contained in $[0,1]^N$, which is a compact set. Thus there is a subsequence which converges to some point $\mathcal{V} \in [0,1]^N$. Since the sum of the components of each $\mathcal{V}_i$ is equal to one, the same is true for the limit $\mathcal{V}$. We take λ_j to be the jth component of $\mathcal{V}$.

With this choice of λ_j, we see that $\pi = \sum_{j=1}^N \lambda_j \frac{1}{|C_j|} \mathbb{1}_{C_j}$ by using (12) along the appropriate subsequence.

If the vertex x_1 is not visited infinitely many times, let x_k be the first vertex which is visited infinitely many times. Clearly the limit in (12) does not depend on $x_1, x_2, \ldots, x_{k-1}$ and thus we can ignore these and repeat the previous argument with a re-indexing starting at x_k. □

The decomposition given in this theorem is valid no matter the source of the walk, as long as the limit exists. Of course, "most" walks on $\mathcal{G}$ would not have a limiting distribution in this sense. However, if the walk (x_n) is generated by a time homogeneous Markov chain on $\mathcal{G}$, then we know that the limit (11) does exist and will give a stationary distribution for the chain. Of course, depending on the structure of $\mathcal{G}$ and the (fixed) transition probabilities, it is possible that running the walk again might result in a different limit (and thus another stationary distribution). However, any such limiting distribution will be a convex combination of uniform cycle distributions.

Other types of random walks on the vertices of $\mathcal{G}$ can also result in a walk with a limiting distribution. This includes time inhomogeneous Markov chains as well as more general random processes where each step might depend on the entire history.

The support of any walk on $\mathcal{G}$ is obviously always connected and "eventually" (that is, after visiting all those vertices which are visited only finitely many times) it is always strongly connected. Thus we might guess that the subgraph of $\mathcal{G}$ given by $\mathrm{supp}(\pi) = \{v \in \mathcal{G} : \pi_v > 0\}$ is always a strongly connected subgraph. Interestingly this is not necessarily the case!

Example 4. As an example, consider the graph shown in Figure 6. We now construct a walk whose limiting distribution is uniformly distributed on all the vertices except for the middle one, vertex T. Take the two cycles

$$C_1 = a,\ b,\ c,\ d,\ e,\ f,\ g,\ h,\ a \quad \text{and} \quad C_2 = A,\ B,\ C,\ D,\ E,\ F,\ G,\ H,\ A.$$

We start the walk at the vertex a. Now we go around C_1 once then go through T to A and go around C_2 twice and then go back through T to a. This is the first "epoch". On the next epoch we go around C_1 three times and C_2 four times and, in general, on the nth such epoch we go $2n-1$ times around C_1 and $2n$ times around C_2. This defines the walk $x_1, x_2, \ldots, x_m, \ldots$.

A couple of things are useful to note. First, during each epoch we visit T twice and each of a and A exactly one more time than any other vertex in each of their respective cycles (either C_1 or C_2). Second, at the end of the nth epoch we have moved a total of n^2 times around C_1 and a total of $n^2 + n$ times around C_2. These observations mean that at the end of the nth epoch we have taken a total of $16n^2 + 12n + 1$ steps. Since

$$\frac{8n^2}{16n^2 + 12n + 1} \to 1/2, \quad \frac{8n^2 + 8n}{16n^2 + 12n + 1} \to 1/2, \quad \text{and} \quad \frac{2n}{16n^2 + 12n + 1} \to 0$$

we know that in the limit the two cycles C_1 and C_2 are visited the same proportion of the time and T is visited a vanishingly small proportion of the total time. Furthermore, even though both a and A are visited more frequently than any of the other vertices, their proportions limit to the same value of $1/16$.

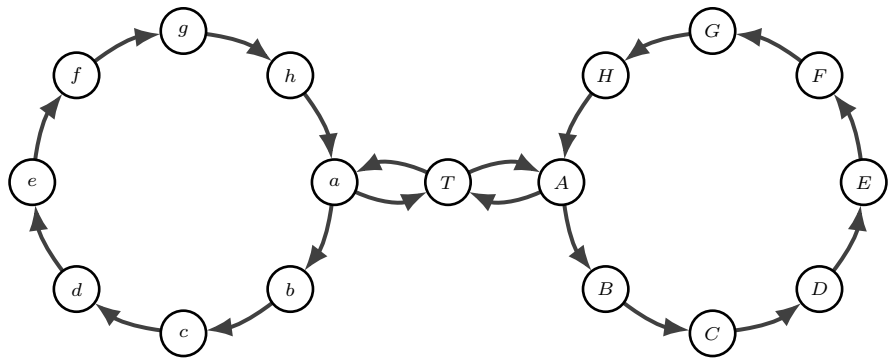

Figure 6: A graph which allows a limiting distribution with disconnected support.

Of course the walk that we have defined for this example cannot be driven by a time homogeneous Markov chain (though it obviously can be driven by a time inhomogeneous chain). For a walk given by a time homogeneous chain, the support of the limiting distribution π has to be strongly connected.

3 From directed cycles to Markov chains.

In the previous section we learned that any "limiting distribution" of a walk on $\mathcal{G}$ has to be a convex combination of uniform cycle distributions. Of course this does not mean that every such linear combination is a limiting distribution of some walk and, in fact, there are usually many of these linear combinations which cannot be limiting distributions for any walk since their support is disconnected and there is no "bi-directional bridge" between the components (imagine the graph in Figure 6 but with the edges $a \to T$ and $A \to T$ removed). However, as we will explain in this section, every such linear combination is an invariant distribution for some Markov chain on $\mathcal{G}$. Our idea for constructing the transition matrix P for the chain comes directly from the theory of *cycle processes* as initiated (as far as we can tell) in [7]. This view gives a very different approach to Markov chains and we encourage the reader to explore it more fully than we can here with our very small application. A fairly comprehensive discussion can be found in the book [5].

Theorem 5. *Suppose that the directed graph $\mathcal{G}$ supports a Markov chain. Let $\pi = \sum_j \lambda_j \frac{1}{|C_j|} \mathbb{1}_{C_j}$ for some directed cycles C_j on $\mathcal{G}$ and $\lambda_j \in [0,1]$ with $\sum_j \lambda_j = 1$. Then there is a Markov transition matrix P with state space the vertices of $\mathcal{G}$ so that $\pi P = \pi$.*

Proof. We use the idea from Section 2 of [7] to define the transition matrix P. With no loss of generality we assume that the cycles C which are used are all the cycles in $\mathcal{G}$ (by setting the appropriate $\lambda_C = 0$ if necessary). For each cycle C, define the weight $w_C = \lambda_C / |C|$. Notice that w_C is the amount of probability that the cycle C contributes (through the uniform cycle distribution $\frac{1}{|C|}\mathbb{1}_C$) to each of its vertices.

If u is a vertex in one of the cycles C_j with $\lambda_j > 0$ and v is any other vertex, we define

$$P_{u,v} = \frac{\displaystyle\sum_{C: u \to v \in C} w_C}{\displaystyle\sum_{D: u \in D} w_D} \tag{14}$$

(the sums are over all the cycles which satisfy the given condition). If u is not in any cycle C with $\lambda_C > 0$, then we choose any vertex u' for which $u \to u'$ is in $\mathcal{G}$ (we know there is at least one such vertex since $\mathcal{G}$ supports a Markov chain) and define

$$P_{u,v} = \begin{cases} 1, & \text{if } v = u' \\ 0, & \text{if otherwise} \end{cases}.$$

It is worthwhile to pause and give some intuition behind equation (14). Since each weight w_C is a probability, the fraction in (14) is a conditional probability. The denominator is the total probability assigned to vertex u while the numerator is the probability assigned to the edge $u \to v$, and so $P_{u,v}$ is set to be the probability of taking the edge $u \to v$ given that one is in vertex u.

We first argue that P is in fact a transition matrix. It is clear that $P_{u,v} \geq 0$ from the definition. Fix a vertex u. If u is not in any cycle with a positive weight, then by definition $P_{u,u'} = 1$ and $P_{u,v} = 0$ for $v \neq u'$ and so $\sum_v P_{u,v} = 1$. Now suppose that $u \in D$ and $w_D > 0$. Then

$$\sum_v P_{u,v} = \sum_v \frac{\sum\limits_{C:u\to v\in C} w_C}{\sum\limits_{D:u\in D} w_D} = \left(\frac{1}{\sum\limits_{D:u\in D} w_D}\right) \sum_v \sum_{C:u\to v\in C} w_C = \frac{\sum\limits_{C:u\in C} w_C}{\sum\limits_{D:u\in D} w_D} = 1$$

and so again the sum is equal to one. Notice that if $P_{u,v} > 0$ then the edge $u \to v$ is in $\mathcal{G}$ and thus the transition matrix P is compatible with the directed graph $\mathcal{G}$.

Next we show that $\pi P = \pi$. If v is a vertex with $\pi_v > 0$, then there is some cycle C which includes v and for which $\lambda_C > 0$. In this case there is clearly some vertex u where $u \to v \in C$ and thus $P_{u,v} > 0$ and $\pi_u > 0$. Then for any such a vertex u we have

$$\pi_u P_{u,v} = \left(\sum_C \frac{\lambda_C}{|C|} \mathbb{1}_C(u)\right) \frac{\sum\limits_{D:u\to v\in D} w_D}{\sum\limits_{E:u\in E} w_E} = \left(\sum_{C:u\in C} w_C\right) \frac{\sum\limits_{D:u\to v\in D} w_D}{\sum\limits_{E:u\in E} w_E} = \sum_{D:u\to v\in D} w_D.$$

Summing over all the vertices u we get

$$(\pi P)_v = \sum_u \pi_u P_{u,v} = \sum_{u:\pi_u>0} \sum_{D:u\to v\in D} w_D = \sum_{D:v\in D} w_D = \sum_D \frac{\lambda_D}{|D|} \mathbb{1}_D(v) = \pi_v.$$

On the other hand, suppose that $\pi_v = 0$ and consider another (not necessarily distinct) vertex u. If $\pi_u > 0$ then there is some cycle C with $\lambda_C > 0$ and $u \in C$. However, in this case $P_{u,v} = 0$ since v is not in a cycle with a positive weight and therefore $\pi_u P_{u,v} = 0$. If $\pi_u = 0$ then $\pi_u P_{u,v} = 0$ regardless of the value of $P_{u,v}$. Thus in any case if $\pi_v = 0$ we have $\pi_u P_{u,v} = 0$ for all vertices u and so again

$$(\pi P)_v = \sum_u \pi_u P_{u,v} = 0 = \pi_v.$$

Thus no matter which vertex we choose, we have that $(\pi P)_v = \pi_v$ and so $\pi P = \pi$. □

The transition matrix defined in the proof does not have to be irreducible and could often have many absorbing states. This means that it could have many different linearly independent invariant distributions including some supported on a single vertex. A simple example of this is illustrated in the left-hand image in Figure 7. The thick edges indicate the support of the chosen cycles, the thin black edges indicate all the $u \to u'$ transitions which are used, and the gray edges are those edges which exist in $\mathcal{G}$ but are not used in constructing P. We take

$$\pi = \frac{\lambda_1}{3}\mathbb{1}_{C_1} + \frac{\lambda_2}{3}\mathbb{1}_{C_2} + \frac{\lambda_3}{4}\mathbb{1}_{C_3}$$

with $\lambda_1 = \lambda_2 = \lambda_3 = 1/3$ and the three cycles

$$C_1 = a,\ b,\ c \quad \text{and} \quad C_2 = g,\ h,\ i \quad \text{and} \quad C_3 = g,\ h,\ j,\ i.$$

This results in the invariant distribution

$$\pi = (\frac{1}{9},\ \frac{1}{9},\ \frac{1}{9},\ 0,\ 0,\ 0,\ \frac{7}{36},\ \frac{7}{36},\ \frac{7}{36},\ \frac{1}{12},\ 0,\ 0)$$

and π having a disconnected support. In addition by using the construction of P from the theorem and the choice of "extra" edges as indicated in the left-hand image in Figure 7, this also results in P being equal to (notice that we labelled the rows and columns to help interpret the entries)

$$
\begin{array}{c}
\\ a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ i \\ j \\ k \\ \ell
\end{array}
\begin{array}{c}
\begin{array}{cccccccccccc} a & b & c & d & e & f & g & h & i & j & k & \ell \end{array} \\
\left(\begin{array}{ccc|cc|c|ccccc|cc}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{4}{7} & \frac{3}{7} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{array}\right)
\end{array}
. \tag{15}
$$

The block diagonal structure of P means that it has several different invariant subspaces; these subspaces correspond to the components of $\mathcal{G}$ (and their unions). Further, P restricted to any of these is also a Markov transition matrix and thus P also has several different invariant distributions (in fact 5 linearly independent ones). Among these are $\frac{1}{3}\mathbb{1}_{C_1}$, $\mathbb{1}_{\{e\}}$, and $\mathbb{1}_{\{f\}}$ and all of these are distinct from π.

The right-hand image in Figure 7 illustrates a different choice for the "extra" (i.e., $u \to u'$) transitions. Here instead of self-loops at some vertices, each vertex that is not in $S = C_1 \cup C_2 \cup C_3$ is connected to S by a directed path. From the figure we can see that after at most two steps any walk (no matter its starting position) will be on S and subsequently remain on S. In terms of the matrix P, this means that αP^2 is supported on S for any starting distribution α. This choice results in fewer linearly independent invariant distributions for P. Since $\mathrm{supp}(\pi)$ has two components there are still two linearly independent invariant distributions, one supported on $\{a, b, c\}$ and the other supported on $\{g, h, i, j\}$.

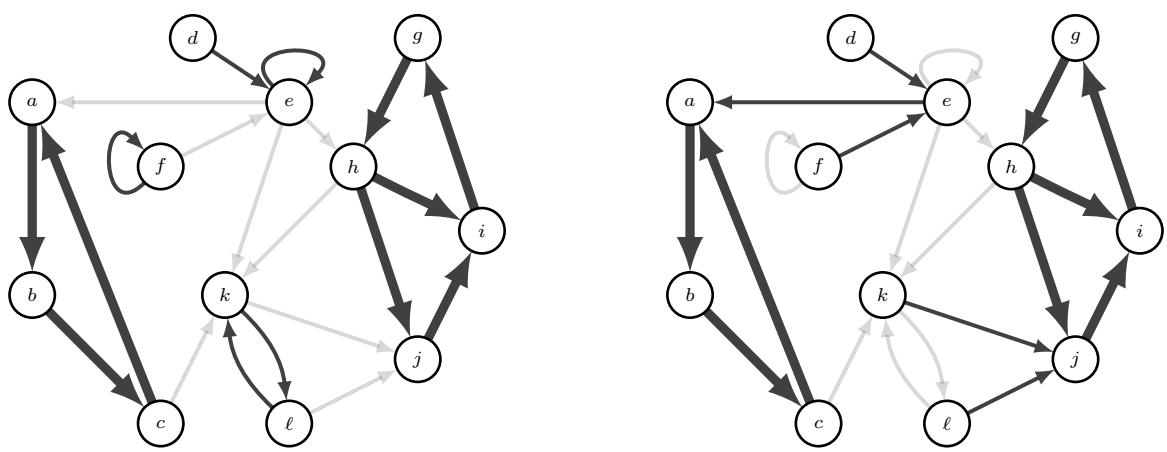

Figure 7: Two different choices of the "extra" transitions in constructing P in Theorem 5.

This discussion naturally leads us to the question of what is the best situation that we can hope for? That is, when can we specify a particular π (as a linear combination of uniform cycle distributions) to be the unique invariant distribution for P? Well, certainly if $\mathrm{supp}(\pi)$ is disconnected then the components of this support will be different ergodic components of any chain with π as an invariant distribution and thus π cannot be unique (as we saw from the simple examples). So, if we want π to be the unique invariant

distribution for P, the support of π must be connected (which immediately implies that it will be strongly connected since it is composed of a collection of cycles). In addition, we need some connectivity property of the graph $\mathcal{G}$ to eliminate "islands" which yield more invariant distributions.

Theorem 6. *Suppose that $\mathcal{G}$ is a directed graph such that for any cycle C and vertex u that is not in any cycle (including C), there is a directed path from u to C. Let $C_1, C_2, \ldots, C_k$ be cycles such that $C_1 \cup C_2 \cup \cdots \cup C_k$ is connected. Then for any distribution $\pi = \sum_i \lambda_i \frac{1}{|C_i|} \mathbb{1}_{C_i}$ with $\lambda_i \in (0,1)$ and $\sum_i \lambda_i = 1$, there is a Markov transition matrix P with state space the vertices of $\mathcal{G}$ so that $\pi P = \pi$ and π is the unique such invariant distribution.*

Proof. First we note that $\pi_u > 0$ for any vertex in any C_i since $\lambda_i > 0$. Thus $\mathrm{supp}(\pi)$ is always $C_1 \cup C_2 \cup \cdots \cup C_k$ for all values of the λ_i.

We define the transition matrix P in a similar way as we did previously, only modifying the definition for vertices not in $\mathrm{supp}(\pi)$. So if u is in one of the cycles and v is any other vertex, we define

$$P_{u,v} = \frac{\displaystyle\sum_{C: u \to v \in C} w_C}{\displaystyle\sum_{D: u \in D} w_D} \tag{16}$$

where again $w_C = \lambda_C / |C|$ and we sum over the chosen cycles.

If every vertex is in $\mathrm{supp}(\pi)$ then we are done. Otherwise, we need to define P for the rest of $\mathcal{G}$. We do this by inductively assembling all the vertices outside of $\mathrm{supp}(\pi)$ into a forest of directed trees with each tree leading to some vertex in $\mathrm{supp}(\pi)$ (see the right-hand image in Figure 7 for a small illustration of this). Once we have this forest the next stage in the proof is to use it to define P.

Thus let $u \in \mathcal{G} \setminus \mathrm{supp}(\pi)$. Then by assumption there is some directed path from u to $\mathrm{supp}(\pi)$. Let this path be $u \to u_1 \to u_2 \to \ldots \to u_m \to v \in \mathrm{supp}(\pi)$ and set $S_1 = \{u, u_1, u_2, \ldots, u_m\}$. If $\mathrm{supp}(\pi) \cup S_1 = \mathcal{G}$ we move on to the stage of defining P, otherwise we continue inductively.

Assuming that $\mathrm{supp}(\pi) \cup S_1$ does not exhaust all of the vertices in $\mathcal{G}$, let $u \in \mathcal{G} \setminus (\mathrm{supp}(\pi) \cup S_1)$. Again there is a path from u to $\mathrm{supp}(\pi)$ and let this path be $u \to u_1 \to u_2 \to \ldots \to u_n \to v \in \mathrm{supp}(\pi)$. Now, it might happen that some $u_i \in S_1$, in which case we don't want to take the entire path to v. So, if $u_i \notin S_1$ for all $i = 1, 2, \ldots, n$, then we set $S_2 = \{u, u_1, \ldots, u_n\}$ and note that $S_1 \cup S_2$ is the union of two disjoint paths. Otherwise suppose that u_i is the first vertex on this path that is in S_1. In this case we set $S_2 = \{u, u_1, u_2, \ldots u_{i-1}\}$ and note that $S_1 \cup S_2$ forms a tree rather than a path.

The general induction step starts with having constructed $S_1, S_2, \ldots, S_\ell$ as a set of disjoint paths. Either $\mathcal{G} = \mathrm{supp}(\pi) \cup S_1 \cup \cdots \cup S_\ell$ and we move on to the next stage of the proof or we can find a vertex $u \in \mathcal{G}$ not in this union. In the later case, there is a path $u \to u_1 \to \ldots \to u_n \to v$ from u to $\mathrm{supp}(\pi)$. Then either this path only meets $\mathrm{supp}(\pi) \cup S_1 \cup \cdots \cup S_\ell$ at the vertex v, in which case we set

$$S_{\ell+1} = \{u, u_1, \ldots, u_n\},$$

or we set

$$S_{\ell+1} = \{u, u_1, \ldots, u_i\},$$

where u_{i+1} is the first vertex in this path to meet $\mathrm{supp}(\pi) \cup S_1 \cup \cdots \cup S_\ell$.

Now we are ready to define P for vertices $u \notin \mathrm{supp}(\pi)$. Given such a vertex u, we know that $u \in S_i$ for some i. Let w be the next vertex along the path that we constructed above (which includes S_i as an initial segment) so that either $w \in S_i$ or $w \in S_j$ for some $j < i$ or $w \in \mathrm{supp}(\pi)$. Then for any vertex $v \in \mathcal{G}$, we define

$$P_{u,v} = \begin{cases} 1, & \text{if } v = w \\ 0 & \text{otherwise} \end{cases}.$$

Showing that P is a Markov transition matrix is just the same as before and so we leave this step to the reader's imagination.

The fact that $\pi P = \pi$ is also shown in the same way as in Theorem 5 and so all that is left to argue is that π is the unique stationary distribution. To see this we first note that for any initial distribution α we have, by our construction of P, that $\alpha P^{|\mathcal{G}|}$ is supported on $\text{supp}(\pi)$ (in fact it would normally take many fewer than $|\mathcal{G}|$ steps for the Markov chain given by P to arrive in one of the given cycles). Thus any invariant distribution for P must be supported on $\text{supp}(\pi)$. Now since $\text{supp}(\pi)$ is a union of cycles and P respects the directed edges of $\mathcal{G}$ (i.e., P is compatible with $\mathcal{G}$), this means that we can restrict P to the vertices of these cycles and get a Markov chain. Since $\text{supp}(\pi)$ is strongly connected and finite, this Markov chain is irreducible and (positively) recurrent and thus has a unique invariant distribution (Section 3.2 of [4] has a discussion of these points). Thus π is the unique invariant distribution for P. □

Summarizing some of what we have discussed we get a rather simple characterization of possible invariant distributions for Markov chains on a graph.

Theorem 7. *Let $\mathcal{G}$ be a directed graph. The set of invariant distributions for all possible compatible time homogeneous Markov chains on subgraphs of $\mathcal{G}$ is equal to the set of convex combinations of the uniform cycle distributions of $\mathcal{G}$.*

The story regarding limiting distributions for walks on $\mathcal{G}$ can also be easily summarized. The next theorem follows both from our other results and also from the techniques which we have used (for instance, modifying Example 4 allows us to construct the desired walk for the converse in part 1). Filling in the details for the proof of this theorem is also a good exercise for any energetic reader.

Theorem 8. *Let $\mathcal{G}$ be a directed graph.*

1. *Any limiting distribution of a walk on $\mathcal{G}$ is a convex combination of uniform cycle distributions from $\mathcal{G}$. Conversely, for any strongly connected subgraph $\mathcal{H}$ and any convex combination π of uniform cycle distributions from $\mathcal{H}$, there is a random walk driven by a time inhomogeneous Markov chain whose limiting distribution is π.*

2. *Any limiting distribution of a walk on $\mathcal{G}$ driven by a time homogeneous Markov chain has support which is (almost surely) a strongly connected subgraph. Conversely, if the support of π is strongly connected and π is a convex combination of uniform cycle distributions on $\mathcal{G}$ then there is a random walk on $\mathcal{G}$ driven by a time homogeneous Markov chain whose limiting distribution is π.*

Example 9. As a final example, we circle back around to the graph $\mathcal{G}$ (from Figure 2) that we started with. For this graph, it is not so hard to find that the possible cycles are:

$$\begin{array}{lll} C_1 = a,\ g,\ j,\ h,\ d,\ b,\ a & C_2 = a,\ g,\ j,\ h,\ i,\ f,\ c,\ b,\ a & C_3 = b,\ e,\ h,\ d,\ b \\ C_4 = b,\ e,\ h,\ i,\ f,\ c,\ b & C_5 = b,\ e,\ i,\ f,\ c,\ b & C_6 = b,\ h,\ d,\ b \\ C_7 = b,\ h,\ i,\ f,\ c,\ b. & & \end{array}$$

Using $w_i = \lambda_i / |C_i|$ (as in the proof of Theorem 5), we can see that $\pi = \sum_i \lambda_i \mathbb{1}_{C_i}$ becomes

$$\begin{aligned} \pi = (&w_1 + w_2,\ w_1 + w_2 + \cdots + w_7,\ w_2 + w_4 + w_5 + w_7,\ w_1 + w_3 + w_6, w_3 + w_4 + w_5, \\ &w_2 + w_4 + w_5 + w_7,\ w_1 + w_2,\ w_1 + w_2 + w_3 + w_4 + w_6 + w_7,\ w_2 + w_4 + w_5 + w_7,\ w_1 + w_2). \end{aligned} \tag{17}$$

Each w_i occurs exactly in the entries corresponding to the vertices from the cycle C_i. For instance, w_1 occurs exactly in the entries corresponding to vertices a, b, d, g, h, and j which are precisely the vertices involved in the cycle C_1. Since $1 = \lambda_1 + \cdots + \lambda_7 = |C_1|w_1 + |C_2|w_2 + \cdots |C_7|w_7$, π is a probability distribution. The expression given in (17) for the invariant distribution π is simpler to understand than the one given in (2)!

What about the transition matrix P that we construct (using the same generic choice of the λ_i) to have π as an invariant distribution? Rather than give the entire 10×10 matrix, we instead just look at one row,

the row corresponding to the vertex b. Since π_b from (17) is equal to $w_1 + w_2 + \cdots + w_7$ and hence not zero, we use equation (16) to construct the entries of this row. Doing this we get

$$\begin{array}{c} \begin{array}{ccccccccccc} b\to a & b\to b & b\to c & b\to d & b\to e & b\to f & b\to g & b\to h & b\to i & b\to j \end{array} \\ \left(\begin{array}{ccccccccccc} \dfrac{w_1+w_2}{w_1+\cdots+w_7} & 0 & 0 & 0 & \dfrac{w_3+w_4+w_5}{w_1+\cdots+w_7} & 0 & 0 & \dfrac{w_6+w_7}{w_1+\cdots+w_7} & 0 & 0 \end{array}\right), \end{array} \tag{18}$$

(we have labelled each entry by its corresponding outgoing edge). The first thing to note is that the denominators for the non-zero terms are all the same and all equal to $\pi_b = w_1 + \cdots + w_7$. This means that the entry in the product πP which corresponds to vertex b will just be the sum of the numerators, which we note is also equal to π_b. Looking at the entire matrix and the product πP term-by-term should make it crystal clear why $\pi P = \pi$ and we encourage the energetic reader to do just this!

In terms of the parameters α, β, γ and δ which are used in Figure 2, we have

$$\alpha = \frac{w_1 + w_2}{w_1 + w_2 + \cdots + w_7}, \qquad \beta = \frac{w_3 + w_4 + w_5}{w_1 + w_2 + \cdots + w_7},$$

$$\gamma = \frac{w_3 + w_4}{w_3 + w_4 + w_5}, \qquad \delta = \frac{w_1 + w_3 + w_6}{w_1 + w_2 + w_3 + w_4 + w_6 + w_7},$$

in agreement with Theorem 5. The expressions for α and β are also in agreement with the $b \to a$ and $b \to e$ entries of P as given in (18). Of course since there are seven w_is and only four parameters α, β, γ, and δ for the Markov chain in Figure 2, we cannot express all the w_is only in terms of α, β, γ, and δ. This is not unexpected since the seven cycles are extreme points of the set of possible invariant distributions and this set is only four dimensional (more precisely, the affine hull of the set is a four-dimensional hyperplane).

We have two different parameterizations of the Markov chain on $\mathcal{G}$, one using α, β, γ, and δ and the other using the w_is. The first gives a conceptually simple description of the outgoing probabilities at each vertex while the second gives a more intuitive description of the invariant distribution.

There are many more questions that can be explored. For instance, in Example 9 the parameters α, β, γ, and δ are from the compact and convex set

$$\mathcal{S} = \{(x_1, x_2) \in [0,1]^2 : x_1 + x_2 \leq 1\} \times [0,1]^2$$

which has twelve extreme points but there are only seven cycles in $\mathcal{G}$ and thus only seven extreme points in the set of possible invariant distributions. Is there a relationship between these different types of "corners" in this specific case? Is there always a relationship between them? Is there a useful mapping between these two different compact and convex sets? In general, the extreme points in the set of possible Markov transition matrices correspond to choosing exactly one outgoing edge at each vertex and so these are also possible to understand in terms of the structure of the graph, but not necessarily in terms of the cycles.

We could also examine the question of how fast the empirical distribution is approaching the limiting distribution, particularly when there is only one possible limit. The construction in Theorem 6 gives one construction of a transition matrix P for a given choice of weights on the uniform cycle distributions. However, normally several different choices of these weights will give the same convex combination and thus the same limit. Which choice results in the fastest convergence? Is there an alternate construction which gives an even faster convergence? The papers [1, 2, 3] all explore variations of this question for special classes of undirected graphs while [6] instead considers more general directed graphs but using a different measure of the rate of convergence.

In our particular considerations we start with $\mathcal{G}$ and then construct chains based on $\mathcal{G}$ and ask about the possible invariant distributions π. Alternatively, we could start with a desired π and build a chain (via the transition matrix P) which will converge to π (of course this implicitly results in some associated graph $\mathcal{G}$). This is what is done in Markov chain Monte Carlo (MCMC) methods which are used to sample from a specified π on a given state space Ω (usually Ω is enormous). Often in MCMC the associated graph $\mathcal{G}$ is undirected (since the resulting chain is *reversible*) and with self-loops at every vertex. Since the uniform

cycle distribution for a self-loop at a vertex v is a point mass at v, this means that the set of achievable distributions is as large as possible. But again we can ask, how can we construct "natural" chains (meaning perhaps having a simple definition of P) with fast convergence? The Metropolis-Hastings algorithm is one framework which can often lead to "natural" chains on the state space.

Markov chains on finite graphs have a large number of connections to other areas of mathematics and an equally large number of applications in computational science. The authors have spent many enjoyable hours exploring this area of mathematics and welcome others to join them.

Acknowledgment.

This paper grew out of KG's undergraduate Honours thesis at Acadia University. The authors would like to thank the anonymous referees and the editor for their careful reading of our paper and their helpful suggestions. FM would also like to thank Ed Vrscay for the conversation in a hallway (many years ago) that started him on the path towards this paper. FM was partially supported by NSERC 2019:05237.

References

[1] Boyd, S., Diaconis, P., Xiao, L. (2004). Fastest mixing Markov chain on a graph. *SIAM Rev.* 46(4): 667–689.

[2] Boyd, S., Diaconis, P., Sun, J., Xiao, L. (2006). Fastest mixing Markov chain on a path. *Amer. Math. Monthly* 113(1): 70–74.

[3] Boyd, S., Diaconis, P., Parrilo, P.. (2009). Fastest mixing Markov chain on graphs with symmetries. *SIAM J. Optim.* 20(2): 792–819.

[4] Brémaud, P. (1999). *Markov chains: Gibbs fields, Monte Carlo simulation, and queues.* New York: Springer-Verlag.

[5] Kalpazidou, S. (1995). *Cycle representations of Markov processes.* New York: Springer-Verlag.

[6] Kirkland, S. (2010). Fastest expected time to mixing for a Markov chain on a directed graph. *Linear Algebra Appl.* 433(11-12): 1988–1996.

[7] MacQueen, J. (1981). Circuit processes. *Ann. Probab.* 9(4): 604–610.