NEAR-EQUILIBRIUM CLASSES OF DISTRIBUTIONS FOR RANDOM WALK NETWORKS

TREVOR MCCARTEN

Abstract. In this paper, we look at the distributions that a random walk network may tend to as time is allowed to go to infinity. Although there are networks where there is a single steady state that all distributions will tend towards and it is what one might expect in the continuous case, there are many networks where we do not end up with a simple steady state but instead finite sets of distributions that cycle between each other at each tick of the clock. We will examine when this is the case and what maximum near-equilibrium class length we can expect for a given graph.

1. Introduction

Imagine a graph (a set of vertices and edges connecting them) in which a particle dropped into it at a particular vertex has some probability for each edge attached to the vertex of moving to the vertex on the other side of the edge after some unit amount of time. Given this we could find the probability of the particle reaching any other vertex after a particular amount of time. This data can be represented as a vector $\vec{v}^{(m)} = (v_1^{(m)}, v_2^{(m)}, \ldots, v_n^{(m)})$, where the vertices are numbered 1 through n and $v_k^{(m)}$ is the probability that, after time m, the particle will be at the k^{th} vertex.

After enough time has passed, it would be reasonable to expect the probability vector $\vec{v}^{(m)}$ to settle down to an equilibrium, like a sloshing pool of water eventually coming to rest. The first simple examples that were looked at with regard to this line of thought gave an indication that the distribution of probabilities does indeed settle down but perhaps not as simply as would be hoped (that is, not a simple stationary distribution).

Example 1. Suppose we have a simple line segment with 5 vertices and 4 edges as below and we say the at each vertex, a particle on the vertex has an equal chance of leaving the vertex via any one the edges attached to it.

Figure 1. Line Segment Graph

If we drop a particle into the node on the far left, the probability for each vertex that the particle will have reached it after a given time is given in Figure 2.

If enough time passes, this will converge to an alternation between these two states:

Date: August 7, 2009.

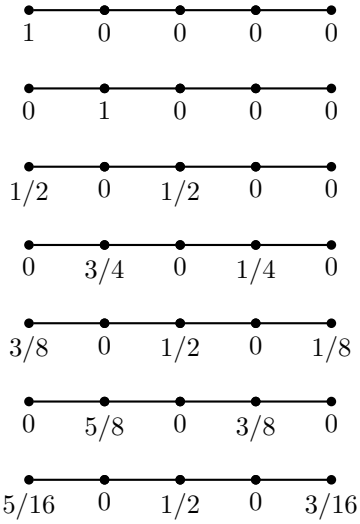
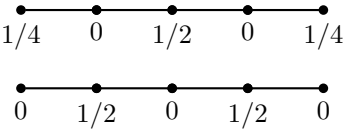


FIGURE 2. Times 1 Through 7



As you can see this is not a simple equilibrium distribution for the network, but instead a pair of states that together comprise a kind of equilibrium. Though this network also has a simple steady state distribution, where $\vec{v} = (1/8, 1/4, 1/4, 1/4, 1/8)$, the fact that when we try to find an equilibrium induced by a particular initial state we end up with something that is analogous but more complicated leads us to the question, what can we say about this phenomenon more generally?

This last example showed that there are cases where a distribution in a graph can converge as time goes on to something that is not a simple equilibrium. However, the next example shows that we can find very similar graphs that may only converge to simple equilibriums.

Example 2. Consider the case where we have a graph shaped like a circle with five nodes (5 vertices, 5 edges) such that a particle at any given vertex has an equal chance of moving to either of its neighboring vertices at the next step (see figure 3). In this case, unlike the case of the line segment, the only type of end-state that a distribution may tend toward is a simple steady state in which all vertices have the same probability assigned to them.

In the case of the line segment, if we assigned by vertices the labels 1, 2, 3, 4, and 5 from left to right, it is reasonably clear that if a particle started at vertex 1, then after an even number of steps, the vertices it could have arrived at would all have the same parity (that is they would all have odd labels). It is easy enough to

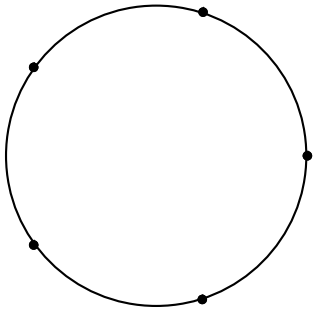


FIGURE 3. Circle Graph

see that this parity-preserving property of that network was precisely what made a set of two states possible as a limiting sequence of states for the network. In the case of the pentagon, however, if a particle started at the vertex on the right side of the circle and followed a path clockwise around to get back to the top, it could reach the same vertex after an odd number of steps instead of an even number. That means that after enough steps have gone by, the probability that the particle will have reached any vertex will be nonzero because not only even-labeled vertices but odd-labeled vertices will be reachable.

This difference between the line segment and the pentagon can be expressed by saying that the only cycles of vertices (paths from a vertex back to itself) are of even length for the line segment, whereas the cirle graph has cycles of both length 2 (moving to a neigboring vertex and right back) and length 5 (travelling all the way around the pentagon), and the difference in parity between 2 and 5 in this case keeps the circle graph from having limiting sequences of states other than the one simple equilibrium. We will see later on (Corollary 3) that this idea can be generalized to give a characterization of the possible limiting sequences of states for random walk networks.

2. Setting the Stage

So far we have only thought intuitively about random walk networks and the near-equilibrium classes of states. The goal of this section is to lay out the definitions and theorems necessary to speak precisely about these ideas.

2.1. **Random Walk Networks as Markov Chains.**

To make the most of our study of random walks, we must first formalize the notion and then think of it in terms of the structures that we have the most tools for working with – in this case, Markov Chains. First things first, definitions.

Definition 1. A *Random Walk Network* is a pair (G, ρ), where G = (V, E) is a finite directed graph and $\rho : V \times V \longrightarrow [0, 1]$ is a function such that if there is no directed edge from vertex v_i to vertex v_j, then $\rho(v_i, v_j) = 0$, if there is a directed edge from v_i to v_j, then $\rho(v_i, v_j) > 0$, and for any vertex $v \in V$ we have that $\sum_{w \in V} \rho(v, w) = 1$

The number $\rho(v, w)$ is viewed as the probability of a particle at vertex v jumping to vertex w at the next step. The function ρ naturally induces a very useful matrix for the random walk network.

Definition 2. The *Transition Matrix* for a random walk network (G, ρ) where V $= \{v_1, v_2, \ldots, v_n\}$ is the matrix $P = [\rho(v_i, v_j)]$.

The i-jth entry of the transtion matrix P is the probability of a particle at v_i jumping to v_j. It is easily verifiable that the i-jth entry of P^n is the probability after n steps of a particle at v_i arriving at v_j. This means that the matrix P completely describes the random walk network (except perhaps for edges that do not allow flow of particles in either direction) and that P^n completely describes where a particle may be after n steps.

Definition 3. Let (G, ρ) be a random walk network and let $n = |V|$. Then a *Distribution* for G is a vector $\vec{w} = (w_1, w_2, \ldots, w_n) \in \mathbb{R}^n$, where $w_k \geq 0$ for $k = 1, \ldots, n$ and $\sum_{k=1}^{n} w_k = 1$.

The interpretation of a distribution $\vec{w}$ is that w_k is the probability of a particle being at vertex v_k for $k = 1, \ldots, n$, where $V = \{v_1, \ldots, v_m)$. It can also be thought of as the amount of water, say, at a particular vertex, and in that case we would think of $\rho(v_i, v_j)$ as the fraction of the water at vertex v_i that will move to vertex v_j at the next step.

Distributions for a random walk network naturally interact with the corresponding transition matrices. If P is the transition matrix for the random walk network (G, ρ) and $\vec{w}$ is a distribution for (G, ρ), then the distribution we get by pushing $\vec{w}$ one step forward in time is $\vec{w}P$. This is most easily seen by verifying it for the standard basis vectors: $\vec{e_k}P$ is just the kth row of P, i.e. the vector $(\rho(k, 1), \rho(k, 2), \ldots, \rho(k, n))$, where $n = |V|$. This row represents the probabilities of a particle at vertex v_k moving to the other vertices – that is, the distribution you get by pushing that particle forward one step.

Remark 1. We will find that it is very important to look at the matrix P^n for different n, rather than P. As noted before, the matrix P^n has a natural interpretation as the matrix representing where the $i - j^{th}$ entry represents the probability of a particle starting at v_i reaching v_j exactly n steps later. And in this sense, the matrix P^n is a transition matrix for a new random walk network with the same vertices as before. The difference is that instead of taking the same edges as before, we say that there is an edge from v_i to v_j if and only if there is a path from v_i to v_j through the graph that is exactly n steps long.

Now that we have random walk networks defined, we will define the objects and the properties that we will use as tools for proving things about random walk networks.

Definition 4. A *Markov Chain* is a sequence of random variables $X_1, X_2, \ldots$ with an at most countable state space S along with a (potentially infinite) $|S| \times |S|$ matrix $P = [p_{i,j}]$ called the transition matrix such that

$Pr(X_{n+1} = i_{n+1} | X_n = i_n) = Pr(X_{n+1} = i_{n+1} | X_1 = i_1, X_2 = i_2, \ldots, X_n = i_n) = p_{i_n, i_{n+1}}$.

In other words, if we are dealing with only at most countably many possible states in discrete time, if we have the probability of those states at the given time, and if the next state's probabilities depend only on the current state and not on any earlier states or the current time, then we have a Markov chain. Note that the notion of a distribution, which we have defined for a random walk network, is

adaptable to Markov chains: it is a (potentially infinite) vector of length $|S|$, where S is the state space, with entries that are nonnegative real numbers and with the entries summing to 1. It is important to remember that S may be countable for Markov chains in general.

It is easy, then, to verify that a random walk network forms a Markov chain. For instance, say that we have a random walk network (G, ρ). Then let the random variables $X_1, X_2, \ldots$ have state space V and be such that $Pr(X_n = v_j | X_{n-1} = v_i) = \rho(v_i, v_j)$. Since the next step in a random walk inside a graph depends only on the vertex the particle is currently on (i.e. the current state of the particle in Markov chain terms) and not on the vertices that the particle visited earlier (i.e. the earlier states of the particle) or the number of steps that have already been traversed and the graph is finite, these random variables make the random walk network a Markov chain. Note that the measure space that we would be computing these probabilities on would essentially be the space of countably long paths through the graph. For a rigorous exposition of this idea, see [3]

Also the notion of a transition matrix for a Markov chain fits the definition of a transition matrix for a random walk. Having this structure with which to look at random walk networks allows us to use the tools that have been developed already for it.

2.2. **Important Facts.** In this section, we will present the basic theorems that allow us to work with random walk networks using Markov chain tools and allow us to apply those tools to the problem of determining the classes of near-equilibrium distributions. First, we'll give a few definitions to be able to talk about the theorems sensibly.

Definition 5. If we have a Markov chain with transition matrix $P = [p_{i,j}]$ and state space S, then a set $I \subset S$ is a *Communicating Class* for the chain if, given $i \in I$, we have for all $j \in S$ that $j \in I$ if and only if $p_{i,j}^{(m)} > 0$ and $p_{j,i}^{(n)} > 0$ for some $m, n \in \mathbb{N}$, where $p_{i,j}^{(n)}$ is the $i - j^{th}$ entry of the matrix P^n.

Note that P^n in this case has the same intrepretation that we mentioned for the transition matrix for a random walk network in Remark 1 – giving us the probability that n steps in the future we go from one state to another. With this interpretation, upon inspection it is clear that a communicating class for a Markov chain is an equivalence class and the set of communicating classes for a Markov chain forms a partition of the subset of the state space of states that can return to themselves.

Definition 6. A Markov chain is said to be *Irreducible* if it has only one communicating class.

For instance, if every state in the state space for the Markov chain can return to itself with nonzero probability, then the Markov Chain is irreducible if and only if S itself is a communicating class.

Definition 7. A Markov chain with random variables $X_1, X_2, \ldots$ and state space S is said to be *Positive Recurrent* if $Pr(X_n = i \text{ for infinitely many } n \in \mathbb{N} | X_1 = i) = 1$ for all $i \in S$ and if the expected return time for every state in S is finite – that is, if $p_n = Pr(X_n = i | X_{n-1} \neq i, X_{n-2} \neq i, \ldots, X_2 \neq i, X_1 = i)$, then $\sum_{n=1}^{\infty} p_n n < \infty$.

Definition 8. If a $\vec{w}$ is a distribution for a Markov chain with transition matrix P, then $\vec{w}$ is said to be a *Stationary Distribution* if $\vec{w}P = \vec{w}$.

Now that we have these definitions describing nice properties a Markov chain can have (which we will find random walk networks share), we can state an important theorem regarding our problem.

Theorem 1. *Suppose a Markov chain is both irreducible and positive recurrent. Then it has a unique stationary distribution λ.*

Proof. See Theorem 1.7.7 in [1]. □

The key idea in this paper is that, although this theorem only applies to chains where there is only one communicating class, each separate communicating class may be viewed as the state space for its own Markov chain, meaning that we can find a stationary distribution for each communicating class and use this to find near-equilibrium classes of distributions (see Definition 11).

Now we want to relate some of these ideas to random walk networks, so that we can get some results in the next section. First a definition.

Definition 9. A random walk network with transition matrix P and vertex set $V$is *Connected* if for every $v_i, v_j \in V$, there is and n such that $p_{i,j}^{(n)} > 0$, where $p_{i,j}^{(n)}$ is the i-jth entry of the matrix P^n.

This can be interpreted as saying that the random walk network is connected if a particle at any one vertex can eventually reach any other. Of particular note is that this just says that the graph G for the network should have a directed path from any vertex to any other (this particular fact will be used as an expedient in a later definition). These random walk networks are special in that they allow us to talk intelligently about their communicating classes.

Remark 2. Note that given a transition matrix P for a random walk network (G, ρ), the matrix P^n is also a transition matrix for some random walk network that shares the vertices of G. Thus, we may talk about the communicating classes for P^n, viewing P^n as a transition matrix.

Lemma 1. *Let (G, ρ) be a random walk network and let P be its transition matrix. For any vertex $v_i \in V$, the communicating class for P^n containing v_i (here, we are using the convention adopted in Remark 2), is $\{v_j : p_{i,j}^{(mn)} > 0$ for some $m \in \mathbb{N}\}$, where $p_{i,j}^{(mn)}$ is the i-jth entry of P^{mn}.*

Proof. It is easiest to prove this if we understand what these things stand for. First, the i-jth entry of P^n is nonzero if and only if there is a path of length n through the graph G from v_i to v_j. So the communicating class for v_i with respect to P^n is the set of vertices that a particle starting at v_i can reach via a path of kn steps for some $k \in \mathbb{N}$ and get back to v_i from there in mn steps for some $m \in \mathbb{N}$.

With this interpretation, we can prove it without much trouble. If $v_k \notin \{v_j : p_{i,j}^{(mn)} > 0$ for some $m \in \mathbb{N}\}$, then there is not even a path of length mn from v_i to v_k, so v_k cannot be in the communicating class for v_i

If, on the other hand, $v_k \in \{v_j : p_{i,j}^{(mn)} > 0$ for some $m \in \mathbb{N}\}$, then there is a path of length mn from v_i to v_k for some $m \in \mathbb{N}$. And since (G, ρ) is assumed to be connected, there is a path of length α from v_k to v_i. Therefore, since mn is the

length of a path from v_i to v_k and α of a path from v_k to v_i, there is a path of length $\alpha + mn$ steps from v_i to v_i. And by repeating that path n times, we get a path of length $(\alpha + mn)n$ steps from v_i back to itself that reaches v_k after mn steps. That just means there is a path from v_k to v_i of length $(\alpha + mn - 1)n$ – that is, a particle can get from v_i to v_k and back in steps of length n. So v_k is in the communicating class of v_i with respect to P^n.

Thus, v_k is in the same communicating class as v_i with respect to P^n if and only if $v_k \in \{v_j : p_{i,j}^{(mn)} > 0 \text{ for some } m \in \mathbb{N}\}$. □

This lemma essentially says that in a connected random walk network with transition matrix P, if we can go from one vertex to another in steps of length n, then the two vertices are in the same communicating class with respect to the transition matrix P^n. So we do not need to worry about checking that we can go in one direction and then get back again – one direction is enough.

Remark 3. The fact that, for a connected random walk network (G, ρ), we only need to check that we can get from v to u in steps of length n to know they are in the same communicating class for P^n has a more significant implication. Since any vertex that can be reached from a vertex in the communicating class $I \subset V$ in the random walk network induced by P^n is also in I, we know that the entire subnetwork of points connected to I via P^n is in I. This means that if there is an edge coming out of $v \in V$ in the random walk network induced by P^n, then that edge lead to an element of I, and from that element of I we can get back to v. On the one hand, having all edges coming out of v leading to vertices in I means that $\sum_{u \in I} \rho^{(n)}(u) = 1$, where $\rho^{(n)}$ is the probability function induced by P^n, so the set I is the vertex set for a random walk network induced by P^n. And since from any element of I we can come back to v, that random walk network is connected.

In short, the communicating class I for P^n is in fact the vertex set for a connected random walk network with transition matrix induced by P^n.

Theorem 2. *Let (G, ρ) be a random walk network and let P be its transition matrix. If G is connected, then (G, ρ) is irreducible and positive recurrent.*

Proof. First off, (G, ρ) is irreducible by Lemma 1 for the case $n = 1$ simply because (G, ρ) is connected.

So let's prove that it is positive recurrent. This means proving that if a particle starts at a particular vertex $v \in V$, then the probability is 1 that it will return to v an infinite number of times as the number of steps is allowed to go to infinity.

To prove this, note that the connectedness of (G, ρ) implies that for every vertex $u \in V$, there is a path A_u from u to v. That means that if $|A_u|$ is the number of steps in A_u, then there is a nonzero probability p_u of a particle at u moving to v in $|A_u|$ steps. Let $p = min_{u \in V, u \neq v} p_u$. Then the probability that a particle at a vertex u does not return to v at some point in $A = max_{u \in V, u \neq v} |A_u|$ steps is always less than or equal to $1 - p \in [0, 1)$. Therefore, if a particle starts at v, the probability that it does not return to v any time after kA steps is less than or equal to $(1-p)^k$. And $(1-p)^k \longrightarrow 0$ as $k \longrightarrow \infty$. So the probability that the particle never returns to v is 0. So the probability that the particle returns to v is 1.

Furthermore, if α_n is the probability that a particle leaving v returns to v after *exactly* n steps and β_m is the probability that a particle starting at v does not return to v any time when m steps have past, then whenever $kA < n$, we get that $\alpha_n \leq \beta_{kA} \leq (1-p)^k$. Therefore, $\sum_{n=1}^{\infty} \alpha_n n \leq 1 + 2 + \ldots + (1-p)^k(kA+1) + (1-$

$p)^k(kA+2)\ldots$ Since $\sqrt[kA+m]{(1-p)^k(kA+m)} \longrightarrow (1-p)^{(1/A)} < 1$ as $k \longrightarrow \infty$, the root test from calculus implies that series converges and hence the expected return time for v is finite.

Repeating the argument from the first paragraph for each of the countably many paths that return to v exactly n times after some number of steps, we get that the probability that the particle returns to v $n+1$ times given that it begins with a particular path that returns n times is 0 for any $n \in \mathbb{N}$. Thus, by the subadditivity of measures and the countability of the set of finite paths, the probability that the particle returns to v a finite number of times is less than or equal to $\sum_{n=1}^{\infty} 0 = 0$, by the subadditivity of probability measures. Thus, the probability that the particle returns to v an infinite number of times is 1. So (G, ρ) is positive recurrent. □

3. Near-Equilibrium Distributions

We begin with the definition of an object that is intuitive enough and will be very important in our proofs (e.g. Lemma 2 and the results that follow from it).

Definition 10. If G is a directed graph, then an *n-cycle* is a list of vertices $v_0v_1\ldots v_n$ such that $v_0 = v_n$ and for $k = 0, \ldots, n-1$ there is a directed edge from v_k to v_{k+1}.

In other words, an n-cycle is a directed path from a vertex back to itself through G. Next, we define the objects that have the starring role: near-equilibrium classes of distributions.

Definition 11. If (G, ρ) is a random walk network with transition matrix P, then a *Near-Equilibrium Class of Distributions* is a finite set $X = \{\vec{w_1}, \vec{w_2}, \ldots, \vec{w_n}\}$ of distributions for (G, ρ) such that $\vec{w_1}P = \vec{w_2}, \vec{w_2}P = \vec{w_3}, \ldots, \vec{w_n}P = \vec{w_1}$ and $\vec{w_i} \neq \vec{w_j}$ for $i \neq j$. In this case, we say the class has length n. And we say that $\vec{w_j}$ is a *Near-Equilibrium Distribution* for $j = 1, 2, \ldots, n$.

In other words a near-equilibrium class of distributions is a set of distributions that cycle through each other as time goes along. The elements of the class can also be seen as left eigenvectors of eigenvalue 1 for the transition matrix P^n, where n is the length of the class. Note that this means that the elements of the class are stationary distributions for P^n.

Lemma 2. *Let (G, ρ) be a connected random walk network with transition matrix P. Let $m_1, m_2, \ldots, m_k, \ldots$ be some lengths of cycles in G. Then the communicating class I for the transition matrix P^n containing the vertex $v \in V$ contains all points within $n, m_1, m_2, \ldots, m_k, \ldots$ steps of v.*

Proof. First, we know by definition that I contains all vertices in G that are n steps away from v.

Second, either G is the collection of directed edges that form a cycle in which no vertices in the cycle are repeated except the initial and terminal vertices, or the cycle of length m_j in G can be exited – that is, if a particle is in the cycle $v_0v_1\ldots v_{m_j}$, there is an $i \in \{0, 1, \ldots, m_j - 1\}$ such that if the particle is at v_i, then at the next step, it can move to a vertex that is not v_{i+1} if $i \neq m_j$ or v_1 if $i = m_j$. If G is such a cycle, then I trivially contains every vertex m_j steps away, since that is simply v.

On the other hand, suppose G is not such a cycle. Since G is connected, a particle starting at v may reach the vertex v_0 in the cycle $v_0\ldots v_{m_j}$ in, say, α_j

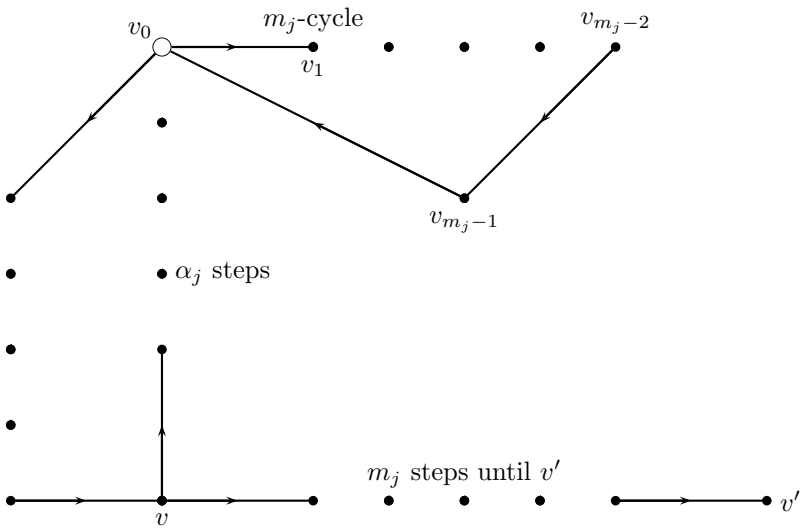


FIGURE 4. After α_j Steps

steps, where we assume the particle may exit the cycle $v_0 \dots v_{m_j}$ at v_0. This is pictured in Figure 4, where the open node represents the location of the particle. Then after a total of $\alpha_j + m_j$ steps, the particle may either have left the cycle, continuing for m_j more steps, or gone around the cycle once more, ending up back at v_0, m_j steps behind its other possible location. Then if β is the number of steps needed for a particle at v to come back to v given that it travels to v_0 and through the cycle with length m_j once along the way, we may repeat this cycle n times to get such a cycle of length βn. This means that there is another cycle of length βn starting at v, going to v_0 in α_j steps, skipping the cycle and going to v in $\beta n - m_j$ steps total and ending m_j steps ahead of v. These two paths of length βn are pictured in Figure 5. Thus, all vertices m_j steps forward from v are reachable with the transition matrix P^n, so Lemma 1 says that they are in I.

This implies that if I is a communicating class for P^n and $v \in I$, then all vertices $n, m_1, m_2, \dots, m_k, \dots$ steps ahead of v are in I. □

Remark 4. Note that in a connected random walk network (G, ρ) with $v \in V$ there is always a cycle either from v directly to itself, or a cycle obtained by leaving v and then returning to v. If either of these cycles is repeated n times, we get a cycle of length mn for some $m \in \mathbb{N}$. That is, in a connected random walk network, each vertex is in fact contained in a communicating class. This is a technical detail, but it is noteworthy since it is implicit in the statement of the next corollary.

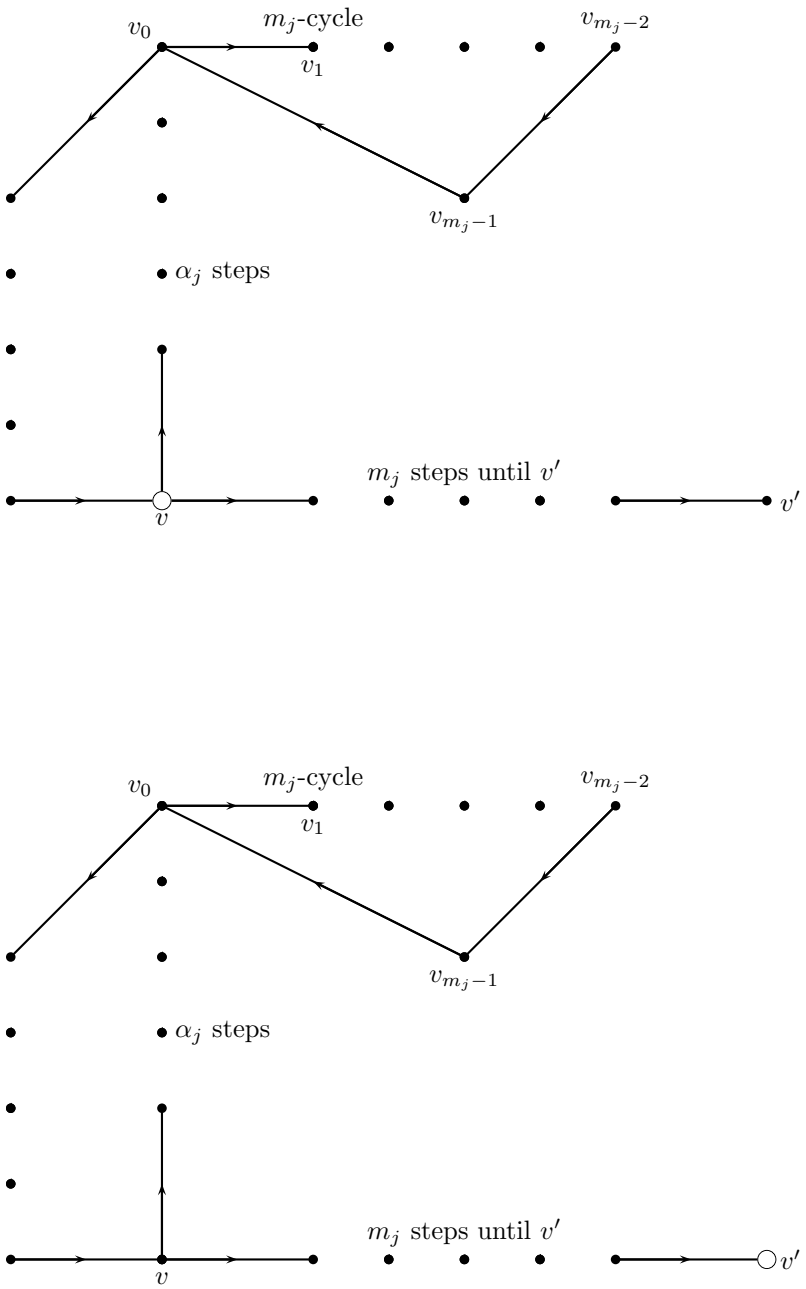


FIGURE 5. At v after βn Steps and m_j Steps Ahead of v After βn Steps

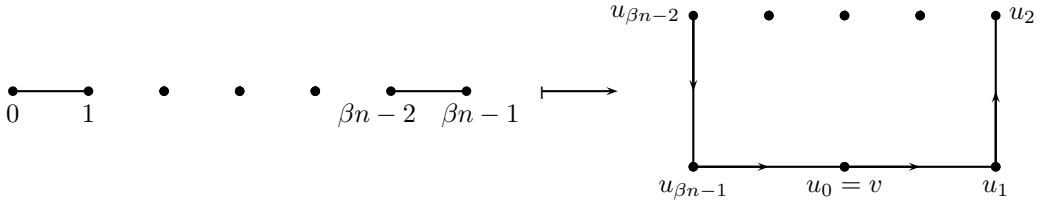


FIGURE 6. $\frac{\mathbb{Z}}{\beta n\mathbb{Z}}$ Mapping To The Cycle C

Corollary 1. *Let (G, ρ) be a connected random walk network with transition matrix P. Let $m_1, m_2, \ldots, m_k, \ldots$ be lengths of cycles in G. Then the communicating class I for the transition matrix P^n containing the vertex $v \in V$ contains all points within $gcd(\{n, m_1, m_2, \ldots, m_k, \ldots\})$ steps of v.*

Proof. Let $v' \in V$ be a vertex that can be reached in $gcd(\{n, m_1, m_2, \ldots, m_k, \ldots\})$ steps from v. Then there is a cycle of length β that starts at v, passes through v' after $gcd(\{n, m_1, m_2, \ldots, m_k, \ldots\})$ steps, and returns to v. Repeating this cycle n times gives us a cycle of length βn, call this cycle C.

If $C = u_0 u_1 \ldots u_{\beta n}$, where we may have $u_i = u_j$ for $i \neq j$, then we have a map $\phi : \frac{\mathbb{Z}}{\beta n\mathbb{Z}} \longrightarrow \{u_0, \ldots, u_{\beta n-1}\}$, defined by $\overline{j} \mapsto u_j$ (this can be thought of as "wrapping" $\frac{\mathbb{Z}}{\beta n\mathbb{Z}}$ around the cycle C as in Figure 6). We know that since at any point of I on this cycle C we may move forward $n, m_1, \ldots, m_k, \ldots$ steps in C to reach another point in I, from Lemma 2, the subgroup $< \overline{n}, \overline{m_1}, \overline{m_2}, \ldots, \overline{m_k}, \ldots >$ of $\frac{\mathbb{Z}}{\beta n\mathbb{Z}}$ is in $\phi^{-1}(I)$. Since $\overline{gcd(\{n, m_1, m_2, \ldots, m_k, \ldots\})}$ is in that subgroup, we know that $v' \in I$, since that is $\phi(\overline{gcd(\{n, m_1, m_2, \ldots, m_k, \ldots\})}) \in I$. □

By putting a cap on the how far apart points in a communicating class for P^n are placed, Corollary 1 effectively allows us to figure out how many communicating classes there are, based on their "concentration." We shall see later in the paper that this, combined with a close connection between the communicating classes and near-equilibrium classes of distributions, allows us to determine information about the near-equilibrium classes.

Next, we want to get an understanding of how the distance between the points in a communicating class relates to the number of communicating classes.

Definition 12. If (G, ρ) is a connected random walk network with transition matrix P and I is a communicating class for P^n, let $k((G, \rho), n, I)$ be the length of the smallest path from one vertex in I to another. We will simply write $k(I)$ when the context is understood.

Definition 13. If (G, ρ) is a connected random walk network with transition matrix P, let $\alpha((G, \rho), n)$ be the number of communicating classes for P^n. We will simply write $\alpha(n)$ when the context is understood.

Lemma 3. *Let (G, ρ) be a connected random walk network with transition matrix P. Let $v \in V$ be a vertex and let $I \subset V$ be the communicating class for P^n containing v. Let $u_1, u_2 \in I$ be such that u_2 can be reached in $k(I)$ steps from u_1. Then if v' is $k(I)$ steps forward from v, then $v' \in I$.*

Proof. First, by Lemma 1, we know that a vertex is in I if and only if it is reachable after mn steps from a vertex in I, for some $m \in \mathbb{N}$. Let $v_0'' \in V$ be a vertex that is n steps ahead of v, where the path connecting them includes v'. Then, since v_0''

is n steps away from $v \in I$, we know that $v_0'' \in I$. Therefore, since $u_2 \in I$ as well, there is a path A of length mn from u_2 to v_0''. Then the path of length mn that starts at u_1, goes $k(I)$ steps to u_2 and then follows A ends at the vertex v_1'' $k(I)$ steps behind v_0''. So $v_1'' \in I$ by Lemma 1.

We may repeat this process for $v_1'', v_2'', etc.$, where after the j^{th} repetition, v_j'' will be $n - jk(I)$ steps ahead of v on the path containing v'. There is, then, a j such that $n - jk > 0$ but $n - (j+1)k(I) \leq 0$. Since k is defined to be the smallest distance from one vertex in I to another, $n - jk(I) \geq k(I)$ but $n - (j+1)k(I) \leq 0$ implies that $n - jk(I) \leq k(I)$. So $n - jk(I) = k(I)$. So $v' = v_j'' \in I$. □

Essentially, this theorem is talking about things similar to what Corollary 1 talks about, except that instead of taking some particular distance between elements of I, it takes the minimum possible distance. The idea is to use that minimality as an exact way of looking at the communicating classes and then use the distance mentioned in Corollary 1 as a bound for that exact number, so that we can estimate it.

Theorem 3. *If (G, ρ) is a connected random walk network with transition matrix P and I and I' are communicating classes for P^n, then $k(I) = k(I')$.*

Proof. Let's assume $I \neq I'$.

First, by Lemma 3, we know that if $v_1' \in I'$, then there is a $v_1 \in I$ such that v_1' is $k(I) - j$ steps forward from v_1, where $0 < j < k(I)$. This follows because the connectedness of G implies there is a path from any vertex in I to v_1', so we can just follow some path from I to v_1' in increments of $k(I)$, by Lemma 3, until we arrive at a vertex $v_1 \in I$ fewer than $k(I)$ steps behind v_1' (it must be fewer and not equal because $I \neq I'$ and if it were equal to $k(I)$ we would get $I = I'$.

Let $v_2 \in I$ be a vertex that is $k(I)$ steps forward from v_1 (which is guaranteed to exist by the connectedness assumption), where the path from v_1 to v_2 goes $k(I) - j$ steps to v_1' and j more steps to v_2. Then, since $v_1, v_2 \in I$, Lemma 1 says that there is a path of length mn from v_2 to v_1, for some $m \in \mathbb{N}$. Then the path of length mn that starts at v_1', goes j steps to v_2 and continues along the path from v_2 to v_1 ends at a vertex v_2' j steps behind v_1. And since v_2' is mn steps from v_1', Lemma 1 says that $v_2' \in I'$. And since v_2' is j steps behind v_1, which is $k(I) - j$ steps behind v_1', we know that v_2' is $k(I)$ steps behind v_1'. Therefore, since $k(I')$ is defined as the minimum number of steps from a vertex in I' to another, $k(I') \leq k(I)$.

The same argument in reverse shows that $k(I) \leq k(I')$. So $k(I) = k(I')$. □

By telling us that the minimum distances between points within communicating classes are all the same, Theorem 3 let's us know that the each communicating class is just a slight perturbation of another, with all the minimum distances preserved. And that brings us to the following corollary.

Corollary 2. *If (G, ρ) is a connected random walk network with transition matrix P and $\{I_1, \ldots, I_{\alpha(n)}\}$ is the set of communicating classes for P^n, then $k(I_1) = \alpha(n)$.*

Proof. Let I_j' be the set containing all vertices that are one step forward from some vertex in I_j. Let $v \in I_j'$. Since I_j is simply all vertices that can be reached in $mk(I_j)$ steps from some particular vertex of I_j for some $m \in \mathbb{N}$, I_j' is similarly just the set of vertices that can be reached after $mk(I_j)$ steps from v for some $m \in \mathbb{N}$.

But since $k(I_j) = k(I_i)$ for all $i = 1, \ldots, \alpha(n)$, by Theorem 3, we know that I'_j is simply the communicating class containing v.

Let $\sigma : \{1, 2, \ldots, \alpha(n)\} \longrightarrow \{1, 2, \ldots, \alpha(n)\}$ be the function defined by $I'_j = I_{\sigma(j)}$. Then, since the connectedness of G implies that any vertex in I_j may reach any vertex of I_i eventually, for any $i, j \in \{1, 2, \ldots, \alpha(n)\}$, we know that σ must in fact be a permutation (that is, we know that it is surjective, so by order considerations, it must also be bijective).

But after $k(I_j)$ steps from a vertex in I_j, we are back in I_j, by Lemma 3 meaning that $I_{\sigma^{k(I_j)}(j)} = I_j$. Therefore, since σ is a permutation and $k(I_j) \leq \alpha(n)$ (this is because between any two vertices of I_j that are $k(I_j)$ steps apart, every pair of vertices is less than $k(I_j)$ steps apart and hence they are from different communicating classes), this implies that $\{I_1, I_2, \ldots, I_{\alpha(n)}\} = \{I_1, I_2, \ldots, I_{k(I_j)}\}$. Therefore, $\alpha(n) = k(I_j)$. □

Remark 5. Not only does the proof of Corollary 2 prove that $k(I) = \alpha(n)$, but it proves that if γ_j is the stationary distribution on I_j (which exists and is unique by Theorems 1 and 2 and the observation that P^n is a transition matrix) for P^n, then $\gamma_j P$ has nonzero values only on $I_{\sigma(j)}$ and $(\gamma_j P)P^n = \gamma_j P$, i.e. $\gamma_j P = \gamma_{\sigma(j)}$.

So the transition matrix P acts on the set $\{\gamma_1, \gamma_2, \ldots, \gamma_{\alpha(n)}\}$ on the right as a permutation.

Lemma 4. *Let (G, ρ) be a random walk network, P its transition matrix, $\{I_1, \ldots, I_{\alpha(n)}\}$ the communicating classes for P^n, and $\gamma_1, \ldots,$ and $\gamma_{\alpha(n)}$ be the stationary distributions on $I_1, \ldots,$ and $I_{\alpha(n)}$ respectively. Let $\vec{\omega}$ be a stationary distribution for P^n. Then $\vec{\omega} = \sum_{i=1}^{\alpha(n)} \beta_i \gamma_i$ for some $\beta_1, \ldots, \beta_{\alpha(n)} \geq 0$ such that $\sum_{i=1}^{\alpha n} \beta_i = 1$.*

Proof. Suppose that $\vec{\omega} = (\omega^1, \omega^2, \ldots, \omega^m\}$ has nonzero values on I_j, where $m = |V|$. Then let $\vec{\omega_j} = (\omega_j^1, \omega_j^2, \ldots, \omega_j^m)$ be the distribution with $\omega_j^i = 0$ for $v_i \notin I_j$ and $\omega_j^i = \frac{\omega^i}{\sum_{v_l \in I_j} \omega^l}$ for $v_i \in I_j$, so that the entries of $\vec{\omega_j}$ in this case still sum to 1. Since under action on the right by P^n, entries for communicating classes do not affect entries for other communicating classes (after n steps, a particle only has a nonzero probability of ending up at particles that are in its communicating class), we know that $\vec{\omega_j} P^n$ only has nonzero entries for I_j.

And if we let $\vec{\omega_j} = \vec{0}$, when $\vec{\omega}$ has no nonzero values on I_j, then we get $\vec{\omega} = \sum_{j=1}^{\alpha(n)} \beta_j \vec{\omega_j}$ for appropriate β_j, where we let $\beta_j = 0$ when $\vec{\omega_j} = \vec{0}$. But since $\vec{\omega} = \vec{\omega} P = \sum_{j=1}^{\alpha(n)} \vec{\omega_j} P$ and both $\vec{\omega_j}$ and $\vec{\omega_j} P$ can only be nonzero on entries for I_j and $I_1, \ldots, I_{\alpha(n)}$ are pairwise disjoint, we must have that $\vec{\omega_j} = \vec{\omega_j} P$. That is, $\vec{\omega_j} = \gamma_j$ when $\vec{\omega_j} \neq \vec{0}$.

Thus, $\vec{\omega} = \sum_{j=1}^{\alpha(n)} \beta_j \vec{\gamma_j}$ and by the definition of β_j and $\vec{\omega_j}$, we have $\beta_1, \ldots, \beta_{\alpha(n)} \geq 0$ and $\sum_{i=1}^{\alpha n} \beta_i = 1$. □

This lemma reduces the problem of describing near-equilibrium distributions and their classes to describing the stationary distributions on the communicating classes. So, for instance, in Example 1, the longest near-equilibrium distribution we could find was of length two, which corresponds to the number of stationary distributions and hence communicating classes for that network for any n: two. Having such a reduction tells us that we can connect near-equilibrium classes in general with the number of communicating classes, which we have already proved much about. The next lemma takes advantage of this.

Lemma 5. *Let (G, ρ) be a random walk network, and P its transition matrix. Then the largest near-equilibrium class of distributions for P^n has as its length the number of communicating classes for P^n: $\alpha(n)$.*

Proof. First, as noted in Remark 5, if $\{\gamma_1, \ldots, \gamma_{\alpha(n)}\}$ is the set of stationary distributions on the communicating classes for P^n, then P act on the right of this set as a permutation. Therefore, $\{\gamma_1, \ldots, \gamma_{\alpha(n)}\}$ is in fact a near-equilibrium class of distributions, meaning that $\alpha(n)$ is less than or equal to the size of the largest near-equilibrium class of distributions.

On the other hand, if $\{\vec{\omega_1}, \vec{\omega_2}, \ldots, \vec{\omega_m}\}$ is a near-equilibrium class of distributions for P^n, then Lemma 4 and what we have just said about $\gamma_1, \ldots, \gamma_{\alpha(n)}$ imply that $\vec{\omega_1}P^{\alpha(n)} = \vec{\omega_1}$, so that $m \leq \alpha(n)$. So the size of the largest near-equilibrium distribution is also less than or equal to $\alpha(n)$.

So $\alpha(n)$ is the size of the largest near-equilibrium distribution. □

Corollary 3. *Let (G, ρ) be a random walk network, and P its transition matrix. If $m_1, m_2, \ldots, m_j, \ldots$ are some lengths of cycles in G and $\{\vec{\omega_1}, \vec{\omega_2}, \ldots, \vec{\omega_l}\}$ is a near-equilibrium class of distributions, then $l \leq gcd(\{m_1, m_2, \ldots, m_j, \ldots\})$ and $l | gcd(\{m_1, m_2, \ldots, m_j, \ldots\})$.*

Proof. Let I be a communicating class for P^l. By Corollary 1 and the definition of $k(I)$, we know that $k(I) \leq gcd(\{m_1, m_2, \ldots, m_j, \ldots\})$. By Corollary 2, $k(I) = \alpha(l)$. By Lemma 5, $l \leq \alpha(l)$ (in fact, we know that $l = \alpha(l)$ in this case). Thus, $l \leq gcd(\{m_1, m_2, \ldots, m_j, \ldots\})$.

And since by Corollary 1, $l = \alpha(l) = k(I) \leq gcd(\{l, m_1, m_2, \ldots, m_j, \ldots\}) \leq l$, we get that $l = gcd(\{l, m_1, m_2, \ldots, m_j, \ldots\}) | gcd(\{m_1, m_2, \ldots, m_j, \ldots\})$.

So $l | gcd(\{m_1, m_2, \ldots, m_j, \ldots\})$. □

Now that we have a good, perhaps even exact, upper bound on the size of near-equilibrium classes of distributions for connected random walk networks, we may generalize it to the rest of the random walk networks.

Definition 14. If (G, ρ) is a random walk network, then a subgraph H of G is a *Connected Component* if H is connected in the sense described immediately following Definition 9, that is, every vertex in H is connected to every other vertex in H via a directed path in H, if H is maximal in this regard (there is no larger connected subgraph containing H), and if there are no directed paths from a vertex of H to a vertex in $G - H$.

The idea now is to break a given graph up into connected components, apply our previous results to those components, and use that to get a result for graphs in general.

Theorem 4. *If (G, ρ) is a random walk network that has connected components $C_1, C_2, \ldots, C_k$ and $f(C_j)$ is the maximum length of a near-equilibrium class of distributions on C_j when it is viewed as a random network, then the maximum length of a near-equilibrium class of distributions for (G, ρ) is $lcm(f(C_1), f(C_2), \ldots, f(C_k))$.*

Proof. First, suppose $\vec{\omega}$ is a near-equilibrium distribution for (G, ρ).

If the vertex $v_i \notin C_1, C_2, \ldots, C_k$, then either v_i has no paths back to itself, or it has some maximal connected subgraph of G containing it.

In the first case, Let B_{v_i} be the set of vertices in G that have paths leading to v_i along with v_i itself, in which case there must be an edge from v_i to a vetex not in B_{v_i} or we would have a path from v_i back to itself.

In the second case, let S_{v_i} be the set of vertices u such that there is a path from u to v_i and from v_i to u. Then S_{v_i} is the vertex set for the largest connected subgraph of G that contains v_i. Therefore, there must be a directed path from a vertex of S_{v_i} to a vertex in $G - S_{v_i}$, since otherwise, S_{v_i} would be a connected component containing v_i. So if, in this case B_{v_i} is the set of all vertices in G with paths leading into S_{v_i}, then this implies that, for every vertex $u \in B_{v_i}$ there is a nonzero probability p_u of a particle starting at u ending at a vertex outside of S_{v_i} after n_u steps for some $n_u \in \mathbb{N}$ (this is also true for the B_{v_i} defined in the first case, so from now on the argument will apply to both). Therefore, if $g(\vec{w})$ is the sum of the entries of $w \in \mathbb{R}^{|V|}$ corresponding to vertices in B_{v_i}, then $g(\vec{\omega}P^{max_{u\in B_{v_i}} n_u}) \leq (1-min_{u\in B_{v_i}} p_u)g(\vec{\omega})$ if $\vec{\omega}$ has nonzero entries corresponding to vertices in S_{v_i}. Therefore, we also get that if m is the length of the near-equilibrium class containing $\vec{\omega}$, then $g(\vec{\omega}) = g(\vec{\omega}P^{m(max_{u\in S_{v_i}} n_u)}) < (1 - min_{u\in S_{v_i}} p_u)g(\vec{\omega}) < g(\vec{\omega})$, a contradiction. Thus, $\vec{\omega}$ is zero on vertices in S_{v_i}. So we know that $\vec{\omega}$ has nonzero values only for vertices of the connected components.

If, then, $v \in C_j$ for some j, then let ρ_j be the restriction of ρ to vertices of C_j. Since the definition of connected component requires that there are no edges leading out of C_j, we still get $\sum_w \rho_j(v, w) = 1$, so (C_j, ρ_j) is in fact a connected random walk network. If $\vec{\omega}$ is nonzero on any vertices of C_j, let $h_j(\vec{\omega})$ be the distribution on (C_j, ρ_j) obtained by giving $h_j(\vec{\omega})$ the same values as $\vec{\omega}$ for the same vertices, only scaled so that the sum of the entries of $h_j(\vec{\omega})$ is 1. If $\vec{\omega}$ is zero on all vertices of C_j, then let $h_j(\vec{\omega}) = \vec{0}$. Since there are no edges from any vertices of C_j to any vertices outside of C_j and any edges leading into C_j must be from vertices that do not belong to connected components and for which $\vec{\omega}$ hence has a zero entry, we get that $h_j(\vec{\omega})P_j = h_j(\vec{\omega}P)$, where P_j is the transition matrix for (C_j, ρ_j). Thus, $h_j(\vec{\omega})P_j^m = h_j(\vec{\omega}P^m) = h_j(\vec{\omega})$, that is, $h_j(\vec{\omega})$ is a near-equilibrium distribution on (C_j, ρ_j).

Therefore, since $h_j(\vec{\omega})$ is a near-equilibrium distribution, the length of the near-equilibrium class containing $h_j(\vec{\omega})$ divides $f(C_j)$ (if l is the length of the class containing $h_j(\vec{\omega})$, then $gcd(l, f(C_j))$ must be a multiple of l, i.e. $l|f(C_j)$). Therefore, if m_j is the length of the near-equilibrium class containing $h_j(\vec{\omega})$, then we get that $lcm(m_1, m_2, \ldots, m_k) \leq lcm(f(C_1), \ldots, f(C_k))$. And if $H_j(h_j(\vec{\omega}))$ is the distribution on (G, ρ) with the same entries as $h_j(\vec{\omega})$ for vertices of C_j and zero entries otherwise, then there are $\beta_1, \ldots, \beta_k \geq 0$ such that $\sum \beta_j = 1$ and $\vec{\omega} = \sum \beta_j H_j(h_j(\vec{\omega}))$, by definition.

Since $H_j(h_j(\vec{\omega}))$ is a near-equilibrium distribution of length m_j, we know that $H_j(h_j(\vec{\omega}))P^{lcm(m_1,m_2,\ldots,m_k)}$, so that $\vec{\omega}P^{lcm(m_1,m_2,\ldots,m_k)} = \vec{\omega}$, meaning that m divides $lcm(m_1, m_2, \ldots, m_k)$. So $m \leq lcm(m_1, m_2, \ldots, m_k) \leq lcm(f(C_1), \ldots, f(C_k))$.

On the other hand, if $\vec{\omega}$ is a near-equilibrium distribution such that $h_j(\vec{\omega})$ is a near-equilibrium distribution of length $f(C_j)$ for each j, then we know that $H_j(h_j(\vec{\omega}))$ is a near-equilibrium distribution with class length $f(C_j)$, so we know that $f(C_j)$ divides m because $H_j(h_j(\vec{\omega}))P^m = H_j(h_j(\vec{\omega}))$, so $lcm(f(C_1), \ldots, f(C_k))$ divides m. Thus, $lcm(f(C_1), \ldots, f(C_k))$ is in fact the maximum length for a near-equilibrium class of distributions for (G, ρ). □

These culminating results gives us a bound on the size of near-equilibrium classes for arbitrary random walk networks, based solely on their directed graphs. The fact that the network in Example 2 did not have a parity-preserving property in the face of the fact that the only near-equilibrium distribution it had was a simple stationary distribution was an example of these results in action, where the fact that $gcd(2,5) = 1$ implied that only a stationary distribution was possible. Having such results allows us to analyze particular types of random walk networks and describe their near-equilibium classes of distributions, at least in number.

For instance, consider electrical networks (for definitions and theorems concerning electrical networks, see [2]).

Definition 15. An *Electrical Network* is a pair (G, γ), where G is a graph with vertex set V and edge set E, $V = \partial V \cup intV$, where $\partial V \cap intV = \emptyset$ and ∂V is called the set of boundary vertices and $intV$ the set of interior vertices, and $\gamma : E \longrightarrow [0, \infty)$ is the *conductivity function* for the network.

From an electrical network, we may induce a random walk network based on the conductivities of the edges.

Definition 16. If (G, γ) is an electrical network, then the *Random Walk Network Induced by* (G, γ) is the random walk network (G', ρ), where the vertex set for G' is the vertex set for G, V, there is a directed edge from an interior vertex of V to any other vertex exactly when there is an edge in G connecting the two, there is exactly one directed edge out of any boundary vertex, which is from itself to itself. And we define $\rho : V \times V \longrightarrow [0, 1]$ like so: if $v \in intV$ and $v' \in V$ and there is an edge between v and v' in G, then $\rho(v, v') = \frac{\gamma(e_{v,v'})}{\sum_{u \in N(v)} \gamma(e_{v,u})}$, where we say that $N(v)$ is the set of vertices that are connected via an edge to v in G and $e_{v,u}$ is the edge from v to u in G. And if $v \in \partial V$, then $\rho(v, u) = 0$ if $u \neq v$ and $\rho(v, v) = 1$.

It is clear enough that ρ in the above definition does in fact make (G', ρ) a random walk network. Given this structure for the random network, we immediately get a result for electrical networks.

Corollary 4. *Let (G, γ) be an electrical network and let (G', ρ) be the random walk network induced by (G, γ). Then any near-equilibrium class of distributions for (G', ρ) has length no more than* 2.

Proof. Suppose C is a connected component of (G', ρ). Then, either C is a single vertex with a directed edge to itself (a cycle of length 1) or C contains vertices v and v' such that there is a directed edge from v to v' and another from v' to v (a cycle of length 2) because if there is a directed edge form one v to v', that means $\rho(v, v') \neq 0$, which means $\frac{\gamma(e_{v,v'})}{\sum_{u \in N(v)} \gamma(e_{v,u})} \neq 0$ and hence $\frac{\gamma(e_{v',v})}{\sum_{u \in N(v')} \gamma(e_{v',u})} \neq 0$, in the notation of Definition 16. The latter inequality says that there is a directed edge form v' to v.

Thus, all connected components of (G', ρ) contain either 1-cycles or 2-cycles. So Corollary 3 says that the maximum length for a near-equilibrium class of distributions on C when it is viewed as a random walk network is 1 or 2. Thus, Theorem 4 says that the maximum length of a near-equilbrium class of distributions for (G', ρ) is the least common multiple of 1's and 2's, so it is either 1 or 2, i.e. it is less than or equal to 2. □

References

[1] J.R. Norris, *Markov Chains*, New York, Cambridge University Press, 2007.
[2] Edward B. Curtis and James A. Morrow, *Inverse Problems for Electrical Networks*, Singapore, World Scientific Publishing, 2000.
[3] Andrew Lewis, *Random Walk Networks*, University of Washington, 2004.