Network Reputation Games

John Hopcroft
Cornell University
jeh@cs.cornell.edu

Daniel Sheldon
Cornell University
dsheldon@cs.cornell.edu

October 3, 2008

Abstract

Originally, hyperlinks on the web were placed for organic reasons, presumably to aid navigation or identify a resource deemed relevant by the human author. However, link-based reputation measures used by search engines (e.g., PageRank) have altered the dynamics of link-placement by introducing new incentives into the system. Strategic authors — spammers and others — now explicitly attempt to boost their own PageRank by careful link-placement. This paper investigates the consequences of such strategic behavior via a network formation game. Our model assumes that authors may place outlinks arbitrarily, but have no control over their inlinks, and their objective is to maximize reputation. What is the best link-placement strategy? What are the equilibrium outcomes? What properties do equilibria possess?

We show that two similar reputation measures — PageRank and hitting time — lead to dramatically different equilibrium outcomes. Since hitting time is immune to strategic placement of outlinks, any directed graph is a Nash equilibrium. On the other hand, equilibria in the PageRank game have a very rich structure: unless the links are delicately balanced, some page can increase its PageRank by dropping all of its links and pointing to just one carefully chosen page. Every equilibrium has a core in which all edges are bidirectional. In a slightly restricted setting, equilibria are characterized exactly by simple properties, the essential of which is a combintorial equivalence among all (bidirectional) edges called *edgewise walk-regularity*. We also demonstrate surprising algebraic properties of equilibria, relating eigenvalues and their multiplicities to graph structure.

1 Introduction

In network reputation systems, participants rate others by placing links. For example, link analysis algorithms for web search interpret a hyperlink as a positive rating, then analyze the global link structure to determine the reputation of each page; reputations are then used to help rank search results and organize indexes. Most link analysis algorithms rely on some notion of transitivity, propagating endorsement along paths so highly connected nodes receive high reputation scores. The popular PageRank [4] reputation system scores pages according to their stationary probability in a random walk on the web that periodically jumps to a random page. Similar reputation systems are also used in peer-to-peer networks [18] and social networks [15].

A common problem in reputation systems is manipulation: strategic users arrange links attempting to boost their own reputation. On the web, this phenomenon is called link spam, and usually targets PageRank. An obvious boosting strategy for PageRank is to obtain inlinks, which is possible by a variety of methods [13], but sites can also boost PageRank significantly by carefully arranging outlinks [2, 6, 12]. Thus PageRank incentivizes certain link placement strategies that undermine its original premise, that links are placed organically and reflect human judgments. In fact, the phenomenon of strategic link-placement is much more widespread than "spam" sites. In 2007, businesses spent over $1 billion on search engine optimization (SEO) [22], the practice of optimizing website content and link structure to be better ranked by search engines.

Previous work has focused on understanding and detecting manipulation [2, 6, 11–13, 19], and designing reputation systems that are robust to manipulation [7, 16, 24]. This paper does not attempt to combat manipulation, but instead investigates its consequences:

How does a reputation system affect the dynamics of link formation and the structure of the underlying network by inducing selfish behavior?

We address this question with tools from game theory, introducing a network formation game called the *network reputation game*, where players are nodes attempting to optimize their reputation by placement of outlinks. Then, a Nash equilibrium reflects the outcome of selfish play — it is a directed graph where no player can further improve her own reputation by choosing different outlinks. This game depends on the choice of reputation system; we show that two qualitatively similar reputation systems lead to dramatically different outcomes. PageRank requires a delicately balanced link structure for a graph to be in equilibrium; otherwise, some node can improve its reputation by dropping all of its links and linking to one carefully chosen target. In this sense, the dynamics are destructive. We also look at a reputation system based on hitting time in a random walk, introduced by the authors in [16] to combat manipulation in PageRank; since a page cannot alter its own hitting time using outlinks, *every* graph is a Nash equilibrium. Though our model is simple — all participants are selfish and their sole objective is to optimize reputation — these differences highlight the nuances of designing a reputation system.

Equilibria in the PageRank game possess surprisingly rich structure, to which we devote considerable attention. We characterize equilibria, and demonstrate fruitful connections to algebraic graph theory, using properties derived from the Nash conditions (in a slightly restricted setting) to prove new spectral results about edge-transitive graphs. Here is a brief highlight of our results:

Best Responses. We describe best responses in the PageRank game. PageRank assigns score $\pi(v)$ to each node v in a directed graph G, where π is the stationary distribution in a random walk on G that is modified to make a *random jump* with probability α at each step, choosing the next node at random instead of following a link. Intuitively, node v has no control over the walk until it reaches v, at which point, barring a random jump, it can direct the next step to a favorable location by choice of outlinks, with the goal of maximizing the expected number of times the walk visits v. We show that v's optimal strategy is to link to nodes from which the probability of returning to v before the first random jump is maximum.

Bidirectionality. We show that all equilibria contain a core in which edges are bidirectional. All nodes outside the core have no inlinks; they may link to the core, but not each other. Furthermore, the core decomposes into connected components that are themselves equilibria when taken as subgraphs. We may then restrict our focus to bidirectional connected graphs.

α-Insensitive Nash Equilibria. Certain equilibria (from which all known equilibria are derived) are α-*insensitive*, meaning they do not rely on the particular setting of the PageRank parameter α controlling the random jump probability. For example, edge-transitive graphs possess enough symmetry that we can show they are Nash equilibria without appealing to the particular value of α. Hence, an edge-transitive graph is an α-insensitive Nash equilibrium. We give an exact characterization of α-insensitive Nash equilibria using a simple graph property called *edgewise walk-regularity*, in which all (bidirectional) edges are equivalent in the following combinatorial sense: for all t, the number of walks of length t connecting the endpoints of an edge is a constant independent of the identity of the edge.

Algebraic Implications. We demonstrate strong algebraic implications for α-insensitive Nash equilibria, borrowing two concepts from algebraic graph theory: graph representations and quotients with respect to equitable partitions. Informally, a representation of G is a drawing of G in an eigenspace of its adjacency matrix A. For α-insensitive Nash equilibria, the geometry of these drawings is highly constrained and tied to the eigenvalues. Two-dimensional representations (corresponding to eigenvalues of multiplicity two) are particularly simple: the drawing of G is a regular k-gon. Furthermore, grouping vertices of G that are drawn at the same location yields an *equitable partition*, whose definition we defer. Then, using the theory of quotients with respect to equitable partitions, we can conclude quite a lot about G — every α-insensitive Nash equilibrium with an eigenvalue of multiplicity two satisfies all of the following:

- All vertex degrees are even,
- The spectrum (set of eigenvalues) of G contains the entire spectrum of a certain weighted k-cycle; the only additional eigenvalue possible is zero,
- If G is edge-transitive, then the automorphism group of G has the cyclic group of order k as a quotient.

Since every edge-transitive graph is an α-insensitive Nash equilibrium, these results are of independent interest when viewed as results about edge-transitive graphs, which have been widely studied.

1.1 Related Work

We briefly mention some additional related work. Since PageRank was introduced [4], it has been modified and adapted to various applications, including personalized web search [20], web spam detection [14], and trust systems in peer-to-peer networks [18]. As mentioned above, manipulation of PageRank has received considerable attention; Friedman et al. [7] provide an overview of manipulation in reputation systems.

Network formation games have been proposed to model various types of strategic network formation, ranging from communication networks to social networks. Tardos and Wexler [23] and Jackson [17] provide two recent surveys. Most games deal with undirected graphs, in which case modelers face a decision: are links formed unilaterally by a single player, or through cooperation by the players at the two endpoints? In our model, graphs are directed and players have unilateral control over their outlinks.

Finally, our work has strong ties to algebraic graph theory, especially the theory of representations of distance-regular graphs. All of the relevant concepts are found in the book of Godsil [8], particularly Chapters 5 and 13. Other useful references are [3, 5, 10]. Our definition of edgewise walk-regularity is analogous to two existing concepts from this literature: *walk-regularity* (introduced in [9]), and *distance-regularity*. Each defines an equivalence among certain pairs of vertices based on the number of walks of length t connecting them. In edgewise walk-regular graphs, all pairs $(u, v) \in E$ are equivalent. In walk-regular graphs, all pairs (v, v) are equivalent; that is, each vertex has the same number of closed walks of length t. In distance-regular graphs, all pairs of vertices at a given distance have the same number of walks of length t connecting them.

2 Preliminaries

PageRank and Hitting Time. Let G be a directed graph. The *α-random walk* on G is a random walk on G modified to restart (also called a jump), with probability α in each step, by choosing the next node from distribution q instead of following a link. Self-loops are ignored, and we assume that $0 < \alpha < 1$. To specify the transitions for nodes with no outlinks (*dangling nodes*), we preprocess the graph as follows: create an artificial node called `wait` which has a self-loop but no other outlinks, and add a link from each dangling node to `wait`. Hence, each time the walk reaches a dangling node, it immediately moves to `wait` and remains there until the next restart. PageRank assigns score $\pi(v)$ to node v, where π is the stationary distribution of the α-random walk; this is guaranteed to be unique by virtue of the random jump.

The *hitting time* of node j is a random variable counting the number of steps to first hit j in a random walk, given a particular distribution over starting nodes. In [16], the authors of this paper proposed ranking nodes by expected hitting time in the α-random walk when the start node is chosen from q, and showed that it is equivalent to ranking nodes by the probability of being reached before the first random jump. Let Φ be the matrix with entries $\phi_{ij} = \Pr_i[\text{hit } j \text{ before first jump}]$, where $\Pr_i[\cdot]$ indicates conditioning on the walk starting at i (below, we use the notation $E_i[\cdot]$ analogously). If not specified, the start node is distributed according to q. Then, the reputation score assigned in [16] is $\text{rep}(j) = \Pr[\text{hit } j \text{ before first jump}] = (q^T\Phi)_j$. This reputation score does not depend on j's outlinks.

The following proposition, proved in the appendix, will allow us to reason about PageRank and hitting time more simply. Let N be the matrix with entries $N_{ij} = E_i[\#\text{ visits to } j \text{ before first jump}]$, and let P be the transition matrix of the *unmodified* random walk on G (after preprocessing to handle dangling nodes).

Proposition 1. *Let N, P and Φ be defined as above. Then (i) the matrix N is equal to $\sum_{t=0}^{\infty} P^t(1-\alpha)^t = (I-(1-\alpha)P)^{-1}$, (ii) the PageRank vector π^T is proportional to $q^T N$, (iii) for all i and j, we have $N_{ij} = \phi_{ij}N_{jj}$, and (iv) for all j, the quantities $\text{rep}(j)$ and ϕ_{ij} do not depend on j's outlinks.*

3 The Hitting Time Game

First, we introduce the reputation game and analyze the simple case when we use the hitting time reputation function $\text{rep}(\cdot)$. Let V be a set of n players, the nodes in a directed graph. A *strategy* for node v is a set of outlinks. An outcome is a directed graph G consiting of the outlinks chosen by each player. A *best response* for player v with respect to G is a set of outlinks E_v, such that, if v deletes its outlinks from G and adds outlinks E_v, then v maximizes its reputation score over all possible choices of E_v. A directed graph G is a *Nash equilibrium* (or *Nash*) if the set of outlinks for each node is a best response: no player can increase its reputation by choosing different outlinks. Since v's outlinks have no effect on $\text{rep}(v)$, no matter what links are placed by others, node v can never improve its reputation. Hence, any directed graph is a Nash equilibrium in the hitting time game.

4 The PageRank Game

The PageRank game is much more sophisticated. In this section, we will begin by characterizing best response strategies, then prove two main theorems describing equilibria. The decomposition theorem describes the strongly connected components of a Nash equilibrium, which consist of (1) bidirectional components and (2) additional nodes that have no inlinks but may link to the bidirectional components. Ignoring the additional nodes, we can restrict our attention to Nash equilibria that are bidirectional and connected. We also make one additional restriction, considering only α-insensitive Nash equilibria. Then, the characterization theorem shows that for a bidirectional and connected graph, the property of being an α-insensitive Nash is essentially equivalent to edgewise walk-regularity.

4.1 Best Responses

Recall that node j is scored according to the stationary probability $\pi(j)$ — in other words, the fraction of time spent at j in a very long random walk. Intuitively, node j should place its outlinks so that upon each visit to j, the links direct the walk to a neighbor from which a quick return to j is likely. Recall that $\phi_{ij} = \Pr_i[\text{hit } j \text{ before first restart}]$. Henceforth, call ϕ_{ij} the *potential*[1] *of i with respect to j*. By our definition, ϕ_{jj} is equal to one. We define ϕ_{jj}^{+} to be the probability of a similar event that excludes the first visit — that is, ϕ_{jj}^{+} is the probability that a walk starting at j *returns* to j before restart. Several of our arguments make use of the simple relationship among potentials obtained by conditioning on the first step of the random walk. Let D_i be the outdegree of node i, and let $\Gamma(i)$ be the set of out-neighbors of i. For $i \neq j$,

$$\phi_{ij} = \frac{1-\alpha}{D_i} \sum_{k \in \Gamma(i)} \phi_{kj}, \quad \phi_{jj}^{+} = \frac{1-\alpha}{D_j} \sum_{k \in \Gamma(j)} \phi_{kj}. \tag{1}$$

The following lemma indicates that j should always link to nodes that maximize potential, matching our intuition about directing the walk to return quickly.

Lemma 1. *In the PageRank game, a best response strategy for node j is any strategy which links to a nonempty subset of nodes that maximize potential with respect to j. If j has no inlinks, then all nodes have zero potential, and the empty set is also a best response.*

Proof of Lemma 1. From Proposition 1, the PageRank vector is proportional to $q^T N$, and $N_{ij} = \phi_{ij} N_{jj} = \frac{\phi_{ij}}{1-\phi_{jj}^{+}}$. In the final equality, we use the fact that N_{jj} is a geometric random variable counting the number of returns to j before jumping; each time, the walk returns with probability ϕ_{jj}^{+} independent of previous returns. Node j seeks the placement of links to maximize

$$(q^T N)_j = \sum_i q_i N_{ij} = \frac{1}{1-\phi_{jj}^{+}} \sum_i q_i \phi_{ij}$$

But in the final expression, all terms inside the sum are independent of j's outlinks, so it suffices for j to maximize ϕ_{jj}^{+}. From (1), we see that ϕ_{jj}^{+} is equal to a constant times the average potential of j's out-neighbors, and, as before, these potentials are independent of j's outlinks. Hence j can only control which terms appear in the average. To maximize the average, node j should link to any nonempty subset of nodes that maximize potential. Placing no outlinks is equivalent to linking to `wait`, which has zero potential. This is never a best response unless j has no inlinks, in which case all nodes have zero potential. □

4.2 Bidirectionality

In this section we prove that every Nash consists of a core of bidirectional connected components, each of which is a Nash when taken as a subgraph, plus additional nodes that have no inlinks but may link to the bidirectional core. The crux of the argument is the following lemma.

Lemma 2. *If j has any inlinks, then j links only to in-neighbors in best response strategies.*

Proof. Suppose j has an in-neighbor i, and that k does *not* link to j. We will show that k cannot maximize potential with respect to j. Since i links to j, we know that $\phi_{ij} > 0$. Suppose for contradiction that k maximizes potential. Then $\phi_{kj} \geq \phi_{ij} > 0$. By conditioning on the first step from k as in Equation (1), we have $\phi_{kj} = \frac{1-\alpha}{D_k} \sum_{\ell \in \Gamma(k)} \phi_{\ell j}$. Since ϕ_{kj} and α are both nonzero, we see that ϕ_{kj} is strictly less than the average potential over k's out-neighbors; one of these must have strictly greater potential. □

Theorem 1 (Decomposition)**.** *Let G be an Nash, and let $C_1, \ldots, C_K$ be the strongly connected components of G consisting of two or more nodes. Let S be the set containing all other nodes. Then (i) for all i, the subgraph induced by C_i is a Nash that is bidirectional and connected, and there are no links leaving C_i, and (ii) nodes in S have no inlinks.*

[1]For an undirected graph G, take the electrical network corresponding to G having unit conductance on each edge, and additional links of conductance α from each node to a sink. Connect a one-volt battery causing potential of 1 volt at j and 0 volts at the sink. Then ϕ_{ij} is the electrical potential, in volts, at node i.

Proof of Decomposition Theorem. Suppose G is a Nash and let v be a vertex in a strongly connected component C_i with two or more nodes. Then v must have at least one in-neighbor, and by Lemma 2, vertex v links only to in-neighbors. Hence, if v links to w, it must be the case that w links to v, and that w is in C_i. This establishes that C_i is bidirectional and connected, and that there are no links leaving C_i.

To see that the subgraph induced by C_i is a Nash, we must consider how potentials with respect to v change when we remove C_i from G, and argue that all of v's neighbors (we no longer distinguish in- and out-neighbors because C_i is bidirectional) continue to maximize potential. Consider a random walk in G that starts in C_i. Since there are no edges leaving C_i, this walk is probabilistically identical to a walk in the subgraph induced by C_i *until the time of the first restart*, at which point the walk may escape to some other component. But potential with respect to v measures the probability of hitting v *before* the first restart, so restricting to C_i does not change the potential of nodes in C_i. Since v's neighbors maximize potential in G, they continue to maximize potential in the subgraph induced by C_i.

Finally, we show that nodes in S have no inlinks. Let $\{v\}$ be a strongly connected component of size one. If v had any inlinks, by Lemmas 1 and 2, v would link to a non-empty set of in-neighbors, hence forming at least one bidirectional edge, contradicting our assumption that $\{v\}$ is a a strongly connected component of size one. □

Remark. With the decomposition theorem in hand, we can restrict our attention to equilibria that are bidirectional and connected. Henceforth, we dispense with directionality, speaking only of degrees, links, and neighbors, instead of their directed equivalents.

Examples. It is instructive to look at some examples. In Figure 1, graph (a) is not an equilibrium because node i has incentive to drop its link to k. The potential ϕ_{ji} is equal to $1-\alpha$, because the walk from j is guaranteed to hit i in one step unless it jumps. But ϕ_{ki} is strictly smaller, because the walk from k could step away from i. Hence, i should not link to k.

However, it is easy to see by symmetry arguments that graphs (b) and (c) are equilibria. Both have the property that, for any node i, all of its neighbors "look identical" in the following sense: if j and k are two neighbors of i, then there is an automorphism mapping j to k that fixes i. Hence $\phi_{ji} = \phi_{ki}$, and all neighbors have the same potential, which must be maximum since non-neighbors cannot maximize potential. This symmetry property is related to edge-transitivity. A graph is *edge-transitive* if, for all edges $e = (e_1, e_2)$ and $f = (f_1, f_2)$, there is a graph automorphism π mapping e to f, setwise (that is, π could map e_1 to either f_1 or f_2). In fact, our symmetry property implies edge-transitivity: for any pair of edges $e = (i, j)$ and $f = (i, k)$ that share an endpoint, there is an automorphism mapping e to f. For arbitrary e and f, because the graph is connected, there is a path containing both e and f, and we can compose automorphisms at each node along the path to map e to f. It turns out that the weaker property of edge-transitivity is also enough to guarantee that a graph is a Nash. Edge-transitive graphs provide most known examples of Nash equilibria, and all that are simple to describe: cycles, cliques, bipartite cliques, and many more combinatorial and algebraic constructions (e.g., the Peterson graph). Our analysis in the following section will reveal additional Nash equilibria.

Since we argued that the examples in Figure 1 are Nash equilibria by symmetry, without appealing to the particular value of α, they are Nash for *all* settings of α, and we call them α*-insensitive*. All currently known equilibria are α-insensitive — it is an interesting open question to determine whether others exist. In the next section, we will study α-insensitive equilibria in much more detail. Requiring G to be a Nash for *all* values of α seems a severe restriction, but we can restate α-insensitivity in a more innocuous form:

Lemma 3. *If G is a Nash for at least n different values of α, then G is α-insensitive.*

For example, if G is a Nash for all α in some tiny interval, reflecting a minimal amount of stability with respect to the specific parameter setting, then G is α-insensitive. Lemma 3 is proved in the appendix; it is best read following the proof of Theorem 2.

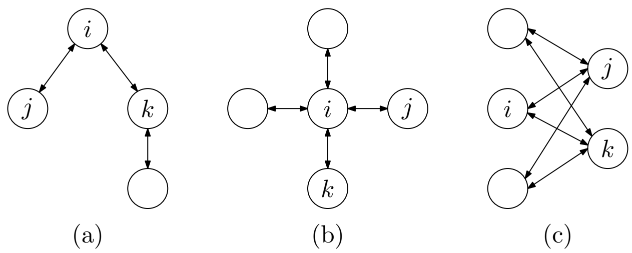

Figure 1: Graph (a) is not a Nash equilibrium; graphs (b) and (c) are Nash equilibria. See text for explanation.

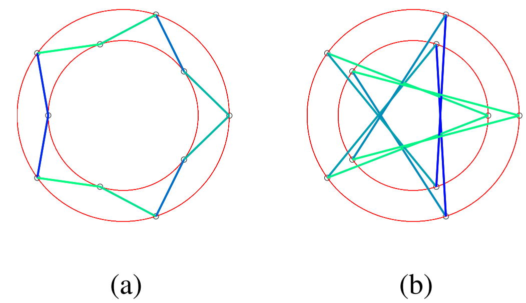

Figure 2: The $(2, 4)$-weighted 10-cycle, plotted (a) using the eigenvectors $x^{(1)}$ and $y^{(1)}$, and (b) using the eigenvectors $x^{(4)}$ and $y^{(4)}$.

4.3 Characterizing α-Insensitive Equilibria

In this section we give a complete combinatorial characterization of α-insensitive Nash equilibria. The essential property is called *edgewise walk-regularity*.

Definition 1. *A graph G is* edgewise walk-regular *if for all $t \geq 0$, the number of walks of length t connecting the two endpoints of an edge does not depend on the identity of the edge. In other words, if A is the adjacency matrix of G, then there exist constants c_t such that $(i, j) \in E$ implies $A^t_{ij} = c_t$.*

We also recall some basic definitions from graph theory. A graph is *regular* if all of its vertices have the same degree. A bipartite graph is *semiregular* if, within each of the two partitions, every vertex has the same degree.

Theorem 2 (Characterization)**.** *A bidirectional, connected graph G is an α-insensitive Nash if and only if the following properties hold: (i) G is edgewise walk-regular, and (ii) G is either regular or bipartite and semiregular.*

Proof. Suppose G is bidirectional and connected. To indicate dependence on α, we now write $\phi_{ij}(\alpha)$ and $N_{ij}(\alpha)$ instead of ϕ_{ij} and N_{ij}. We begin by stating the Nash conditions succinctly and working out a sequence of equivalent formulations. Node i plays a best response if all of its neighbors maximize potential. Since a non-neighbor cannot maximize potential, it is equivalent for all of i's neighbors to have the same potential. Let $\Gamma(i)$ be the neighbors of i. Then, G is an α-insensitive Nash if and only if, for all i,

$$j, k \in \Gamma(i) \implies \forall \alpha,\ \phi_{ji}(\alpha) = \phi_{ki}(\alpha). \tag{2}$$

Since $N_{ji}(\alpha) = \phi_{ji}(\alpha) N_{ii}(\alpha)$, multiplying both sides of the equality in (2) by $N_{ii}(\alpha)$, we get the following equivalent conditions:

$$j, k \in \Gamma(i) \implies \forall \alpha,\ N_{ji}(\alpha) = N_{ki}(\alpha). \tag{3}$$

Recall that $N_{ji}(\alpha) = \sum_{t=0}^{\infty} P^t_{ji}(1-\alpha)^t$, so $N_{ji}(\alpha)$ is the generating function for the sequence $\{P^t_{ji}\}_{t \geq 0}$, evaluated at $(1 - \alpha)$. The two generating functions in (3) are equal if and only if their corresponding sequences are identical, so the following is equivalent to (3):

$$j, k \in \Gamma(i) \implies \forall t \geq 0,\ P^t_{ji} = P^t_{ki}. \tag{4}$$

We have shown that the Nash conditions are equivalent to (4). Now, we will show that (4) holds if and only if the conclusions of the theorem hold. First, we assume that (4) holds; we will show that G is either regular or bipartite and semiregular. Let j and k be neighbors of i. Taking $t = 1$ in (4), we get $(D_j)^{-1} = P^1_{ji} = P^1_{ki} = (D_k)^{-1}$. Hence, j and k have the same degree. But i, j, k were chosen arbitrarily, so for any node, all of its neighbors have the same degree. If G is not regular, then consider a vertex i with degree d_1. All of i's neighbors have degree d_2, and all neighbors of i's neighbors (this set includes i) have degree d_1. It is easy to see that the connected component containing i is bipartite and semiregular. Since G is connected, this is the entire graph.

We will now show that G must be edgewise walk-regular, using (4) and the fact that G is either regular or bipartite and semiregular. We implicitly assume the semiregular case, letting vertices of degree d_1 have neighbors of degree d_2, and vice versa. In all cases, if we let $d_1 = d_2$, then the statements hold for the regular case as well. Write W_{ji}^t for the set of all walks of length t from j to i. Then we can write $P_{ji}^t = \sum_{w \in W_{ji}^t} \Pr[w]$. Note that $|W_{ji}^t| = A_{ji}^t$. Suppose j has degree d_1. All walks in W_{ji}^t have the same probability, because the degrees of vertices visited by the walk alternate between d_1 and d_2, starting with d_1. Specifically, if $w \in W_{ji}^t$, then $\Pr[w] = \left(d_1^{\lceil t/2 \rceil} d_2^{\lfloor t/2 \rfloor}\right)^{-1}$. Hence we have

$$P_{ji}^t = \frac{A_{ji}^t}{d_1^{\lceil t/2 \rceil} d_2^{\lfloor t/2 \rfloor}}. \tag{5}$$

We conclude that if $P_{ji}^t = P_{ki}^t$ as in (4), then $A_{ji}^t = A_{ki}^t$; in other words, any two edges incident on i have the same the number of walks of length t connecting their endpoints. Since i is arbitrary, this is true for all vertices. Since G is connected, any two edges are connected by a path, and we can extend this equality along each vertex in the path to conclude that all edges have the same number of walks of length t connecting their endpoints. Hence G is edgewise walk-regular.

Now we show the converse, that the conclusions of the theorem imply (4). Suppose G is edgewise walk-regular and either regular or bipartite and semiregular, and let j and k be neighbors of i. Nodes j and k have the same degree, say d_1, and by the same argument as before, equation (5) holds for edges (j, i) and (k, i). Because G is edgewise walk-regular, we know that $A_{ji}^t = A_{ki}^t$ for all t, so by (5) we have $P_{ji}^t = P_{ki}^t$ for all t. □

Additional Equilibria. Theorem 2 makes it easy to see that edge-transitive graphs are α-insensitive Nash: there is an automorphism mapping any edge to any other, which guarantees the combinatorial equivalence of edgewise walk-regularity. Also, it is well known that edge-transitive graphs are either regular or bipartite and semiregular. Distance-regular graphs (see Chapters 11 and 13 of Godsil's book [8]) are also α-insensitive Nash equilibria. They are regular, and Lemmas 1.1 and 1.2 of Chapter 13 in [8] imply that in distance-regular graphs, the number of walks of length t between vertices v and w depends only on the distance $d(v, w)$. Taking $d(v, w) = 1$, this implies edgewise walk-regularity. Distance-regular graphs, of which strongly-regular graphs are a special case, provide examples of Nash equilibria that are not edge-transitive: there are strongly-regular graphs based on Latin squares that have no nontrivial automorphisms [21].[2]

5 Algebraic Properties

The α-insensitive Nash conditions have surprising algebraic implications. Our results in this section will use two concepts from algebraic graph theory: graph representations and equitable partitions. We introduce all of the necessary concepts below; more details can be found in Chapters 5 and 13 of [8]. A representation of G is an embedding into a Euclidean space constructed from an eigenspace of the adjacency matrix. If G is an α-insensitive Nash, then the embedding places all vertices on two concentric spheres, and for all edges, the distance separating the endpoints is the same.

The picture is particularly simple in two dimensions: the representation is a regular k-gon. Furthermore, grouping vertices of G that map to the same vertex of the k-gon (the representation need not be one-to-one), yields an *equitable partition*: every vertex in a given cell of the partition has the same number of neighbors in each other cell. Then, using the theory of quotients with respect to equitable partitions, we learn quite

[2] Thanks to Chris Godsil for pointing out that strongly-regular graphs are edgewise walk-regular, and providing this example.

a lot about G, simply from the fact that there exists an eigenvalue of multiplicity two. In particular, all vertex degrees must be even, we can identify all possible eigenvalues (and many eigenvectors), and if G is edge-transitive, then its automorphism group must have a cyclic quotient structure.

5.1 Representations of α-Insensitive Nash

Let A be the adjacency matrix of G. Since A is symmetric, it admits an eigenvalue decomposition $A = \sum_\lambda \lambda U_\lambda U_\lambda^T$. The sum is over distinct eigenvalues. The matrix U_λ has dimension $n \times m_\lambda$, where m_λ is the multiplicity of λ, and its columns form an orthonormal basis for the eigenspace associated with λ. Let $u_\lambda(v)$ be the transpose of the vth row of U_λ; then u_λ is a map from V to $\mathbf{R}^{m_\lambda}$ which we call the *representation belonging to* λ. We call the point $u_\lambda(v)$ the *image* of v under this representation. The Gram matrix $U_\lambda U_\lambda^T$ determines the geometry of the representation by specifying all inner products between points: $(U_\lambda U_\lambda^T)_{vw} = u_\lambda(v)^T u_\lambda(w)$.

The next lemma shows that for α-insensitive Nash equilibria, a great deal of this geometry is determined by the spectrum and simple graph parameters: for any vertex v, the norm of $u_\lambda(v)$ is proportional to $\sqrt{D_v}$, and for any edge, the inner product between the images of its endpoints is a constant.

Lemma 4. *Let $G = (V, E)$ be edgewise walk-regular, let λ be an eigenvalue of A, and let $C_\lambda = m_\lambda/(2|E|)$. For any vertex v, we have $\|u_\lambda(v)\|^2 = D_v C_\lambda$. For any edge (v, w), we have $u_\lambda(v)^T u_\lambda(w) = \lambda C_\lambda$.*

Lemma 4 is proved in the appendix. In a semiregular graph, there are only two possible degrees, so all vertices map onto two spheres centered at the origin. For a vertex v, we can describe the arrangement of the images of v's neighbors: all have the same inner product with $u_\lambda(v)$, hence they lie on an affine hyperplane H. If v has degree d_1, then v's neighbors all have degree d_2, so the images also lie on the sphere S opposite the image of v. The intersection of H and S is a sphere S' of dimension $m_\lambda - 1$ containing the images of v's neighbors. We can also show that the images are "evenly distributed" on S': their centroid coincides with the center of the sphere. Lemma 5, proved in the appendix, makes these statements precise.

Lemma 5. *In the representation belonging to λ, the images of v's neighbors lie on a sphere of dimension $m_\lambda - 1$ centered at the point $y = \frac{\lambda u_\lambda(v)}{D_v}$, and contained in the affine hyperplane orthogonal to $u_\lambda(v)$ through y. For $m_\lambda > 1$, the sphere has nonzero radius r, given by $r^2 = \frac{C_\lambda}{D_v}(d_1 d_2 - \lambda^2)$. Furthermore, the images of v's neighbors are distributed on the sphere so their centroid is also equal to y.*

5.2 Two-Dimensional Representations

Next, we look at the particularly simple geometry of a two-dimensional representation: it is a regular k-gon. Given this observation, the connection between the geometry and the spectrum allows us to draw many conclusions about G whenever there is an eigenvalue of multiplicity two. A key ingredient in the analysis is the theory of equitable partitions and their quotients.

Equitable Partitions. Let $C = (C_1, C_2, \ldots, C_k)$ be a partition of the vertex set of G into k cells. We call C an *equitable partition* if there exist constants c_{ij} such that any vertex v in cell C_i has exactly c_{ij} neighbors in cell C_j. Note that c_{ij} need not equal c_{ji}. The quotient of G with respect to an equitable partition C, denoted G/C, is the weighted directed graph with vertices $\{1, 2, \ldots, k\}$ representing the cells, and having c_{ij} directed edges from cell C_i to cell C_j. Let Q be the $n \times k$ matrix whose jth column indicates membership in cell C_j, so Q_{ij} is equal to one if vertex i is in cell C_j, and zero otherwise. Let B be the adjacency matrix for the quotient G/C. The following lemma states that eigenvectors of B "lift" to eigenvectors of A with the same eigenvalue. The lift operation is simple if we view x as a function on cells, the lift of x is the vector Qx, which assigns the value of a cell to all of its members. If $Bx = \lambda x$, we call (λ, x) an *eigenpair* of B.

Lemma 6 (Godsil [8] Chapter 5, Lemma 2.2)**.** *If (λ, x) is an eigenpair of B, then (λ, Qx) is an eigenpair of A. All nonzero eigenvalues of A are also eigenvalues of B.*

Eigenvalues of (a, b)-Weighted k-Cycles. We will show that partitioning vertices according to their images in a representation of dimension two gives an equitable partition, and the quotient of G with respect to this partition is a cycle described as follows. The (a, b)-*weighted k-cycle* is a bidirectional cycle on vertices $\{1, 2, \ldots, k\}$. For odd i, vertex i has directed edges of weight $a/2$ to its two neighbors: the vertices $i - 1$ (mod k) and $i + 1$ (mod k). Even vertices are similar, but the outgoing edges both have weight $b/2$. Hence the parameters a and b specify the degrees — the total weight of outgoing edges — of odd and even vertices, respectively. If $a \neq b$, we require k to be even so the degrees alternate around the cycle.

Lemma 7. *The eigenvalues of the (a, b)-weighted k-cycle are the values $\lambda_i = \sqrt{ab}\cos\left(\frac{2\pi i}{k}\right)$, for $i = 0, 1, \ldots, \left\lfloor \frac{k}{2} \right\rfloor$, with corresponding eigenvectors $x^{(i)}$ and $y^{(i)}$ given by*

$$x_v^{(i)} = \begin{cases} \sqrt{a}\cos\left(\frac{2\pi i v}{k}\right) & v \text{ odd} \\ \sqrt{b}\cos\left(\frac{2\pi i v}{k}\right) & v \text{ even} \end{cases}, \qquad y_v^{(i)} = \begin{cases} \sqrt{a}\sin\left(\frac{2\pi i v}{k}\right) & v \text{ odd} \\ \sqrt{b}\sin\left(\frac{2\pi i v}{k}\right) & v \text{ even} \end{cases}.$$

If $i = 0$ or $i = \frac{k}{2}$, the vector $y^{(i)}$ is identically zero, so we disregard it as an eigenvector, and the eigenvalue λ_i has multiplicity one instead of two.

Lemma 7 is proved in the appendix. Geometrically, one can plot the (a, b)-weighted k-cycle in two dimensions (see Figure 2) using the vectors $x^{(i)}$ and $y^{(i)}$ as the coordinates. An odd vertex lies on the circle of radius $\sqrt{a}$, and its neighbors are the points on the circle of radius $\sqrt{b}$ that are $\frac{2\pi i}{k}$ radians in either direction.

Main Result. We are ready for the main result of this section. Let $\mathrm{Aut}(G)$ be the automorphism group of G. For a fixed eigenvalue λ, let C_{x_1} be the set of vertices v such that $u_\lambda(v) = x_1$, and let $C = (C_{x_1}, \ldots, C_{x_p})$ be the partition of vertices according to their image under u_λ.

Theorem 3. *Let G be a connected α-insensitive Nash where vertices of degree d_1 have neighbors of degree d_2, and vice versa (allowing the possibility that $d_1 = d_2$). Suppose G has an eigenvalue λ of multiplicity two, and let the partition C be defined as above. Then all of the following are true: (i) the degrees d_1 and d_2 are even, (ii) the partition C is equitable, and the quotient G/C is a (d_1, d_2)-weighted k-cycle, and (iii) if G is edge-transitive, the cyclic group of order k is a quotient of* $\mathrm{Aut}(G)$.

As a consequence, all eigenpairs of the (d_1, d_2)-weighted k-cycle lift to G as described in Lemma 6, and these are the *only* nonzero eigenvalues. For a given eigenvalue, there may be additional eigenvectors.

Proof of Theorem 3. Suppose λ has multiplicity two. For any neighbor w of v, Lemma 5 tells us that $u_\lambda(w)$ lies in a sphere of dimension one orthogonal to $u_\lambda(v)$ with center y and radius r given by the lemma. Such a sphere contains exactly two points: $y \pm z$, where z is a vector of length r orthogonal to $u_\lambda(v)$.

Let a be the number of neighbors of v with image $y + z$, and let b be the number of neighbors with image $y - z$. By Lemma 5, the centroid of these points is also equal to y. Hence

$$y = \frac{a(y+z) + b(y-z)}{D_v} = y + \frac{(a-b)z}{D_v}$$

We conclude that $a = b$, and since $a + b = D_v$, that both are equal to $D_v/2$. Since a and b are integers, D_v is even, proving (i). Furthermore, the values of y and z depend on v only through the image $u_\lambda(v)$, and the degree of v, which itself is determined by the norm of $u_\lambda(v)$ according to to Lemma 4. Hence, any vertex with the same image as v will have neighbors at the same points $y \pm z$.

Let C_x be the cell containing all vertices v such that $u_\lambda(v) = x$. By the above argument, all vertices in cell C_x have the same degree, say d_1, and have $d_1/2$ neighbors in each of the two neighboring cells C_{y+z} and C_{y-z}. Since this number depends only on the cells and not the vertex, the partition is equitable. Each cell has exactly two neighbors, so the cells are arranged in a cycle — it must be connected and not a union of disjoint cycles because G is connected. If a cell contains vertices of degree d_1, the neighboring cells contain vertices of degree d_2, so the degrees alternate. Hence the quotient is a (d_1, d_2)-weighted k-cycle.

Next, we demonstrate a natural group homomorphism from $\mathrm{Aut}(G)$ to the automorphism group of the quotient. Now, we consider the *unweighted* quotient G_λ with vertex set $u_\lambda(V) = \{u_\lambda(v) : v \in V\}$ and edge set $u_\lambda(E) = \{(u_\lambda(v), u_\lambda(w)) : (v, w) \in E\}$. This is the same as the quotient G/C, except all edges have unit weight. Then, by the previous argument, G_λ is a k-cycle. Clearly u_λ is a homomorphism from G to G_λ. We will demonstrate a group homomorphism $\rho : \mathrm{Aut}(G) \longrightarrow \mathrm{Aut}(G_\lambda)$ such that, for any $\pi \in \mathrm{Aut}(G)$, the following diagram commutes

$$\begin{array}{ccc} G & \xrightarrow{u_\lambda} & G_\lambda \\ \Big\downarrow{\scriptstyle \pi} & & \Big\downarrow{\scriptstyle \rho\pi} \\ G & \xrightarrow{u_\lambda} & G_\lambda \end{array}$$

We define ρ in terms of the permutation matrix P associated with π, which satisfies $Pe_v = e_{\pi(v)}$ for each vertex v. For any $x \in u_\lambda(V)$, we define $\rho\pi(x) = U_\lambda^T P U_\lambda x$. Since π is an automorphism, the matrix P commutes with A (that is, $PA = AP$). Equivalently, P and A are simultaneously diagonalizable, so P commutes with $U_\lambda U_\lambda^T$ for all λ. Then, to see that the diagram above commutes, for any vertex v, we have

$$\rho\pi u_\lambda(v) = U_\lambda^T P U_\lambda U_\lambda^T e_v = U_\lambda^T U_\lambda U_\lambda^T P e_v = U_\lambda^T e_{\pi(v)} = u_\lambda \pi(v)$$

To show that $\rho\pi$ is an automorphism, we first show it maps edges onto to edges. Here, we consider edges as unordered pairs; for an edge $e = \{u, v\}$ and a map h, we write $h(e) = \{h(u), h(v)\}$. Let $f \in u_\lambda(E)$ be an edge in the quotient. Then there exists an edge $e \in E$ such that $u_\lambda(e) = f$, and we have $\rho\pi(f) = \rho\pi u_\lambda(e) = u_\lambda\pi(e)$. But $u_\lambda\pi(e)$ is an edge because both π and u_λ are graph homomorphisms. Therefore, $\rho\pi$ is a graph homomorphism. Furthermore, $\rho\pi$ is bijective because the matrix $U_\lambda^T P U_\lambda$ is nonsingular (it is orthogonal). To see that ρ is a group homomorphism, we must argue that $\rho(\pi\sigma) = (\rho\pi)(\rho\sigma)$. Suppose P is the permutation matrix associated with π and Q is the permutation matrix associated with σ. Then for any x,

$$\rho(\pi\sigma)(x) = U_\lambda^T (PQ) U_\lambda x = (U_\lambda^T U_\lambda) U_\lambda^T (PQ) U_\lambda x = (U_\lambda^T P U_\lambda)(U_\lambda^T Q U_\lambda) x = (\rho\pi)(\rho\sigma)(x)$$

Hence ρ is a representation of $\mathrm{Aut}(G)$ as a subgroup of $\mathrm{Aut}(G_\lambda)$. Finally, we show that, if G is edge-transitive, then the $\mathrm{im}(\rho)$ acts transitively on the edges of G_λ. Let f_1 and f_2 be two edges of G_λ. Then there exist edges e_1 and e_2 in G such that $u_\lambda(e_1) = f_1$ and $u_\lambda(e_2) = f_2$. Since G is edge-transitive, there is an automorphism π such that $\pi(e_1) = e_2$. Then

$$\rho\pi(f_1) = \rho\pi u_\lambda(e_1) = u_\lambda\pi(e_1) = u_\lambda(e_2) = f_2.$$

Since $\mathrm{im}(\rho)$ acts transitively on the edges of a k-cycle, it contains the cyclic group of order k as a subgroup; by the properties of group homomorphisms, any subgroup of $\mathrm{im}(\rho)$ is a quotient of $\mathrm{Aut}(G)$. □

Acknowledgments We thank Jon Kleinberg and Robert Kleinberg for useful discussions and suggestions about this work. We thank Chris Godsil for pointing out examples of edgewise walk-regular graphs that are not edge-transitive. This material is based upon work supported by the National Science Foundation (NSF) under Grant No. 0514429, and by the AFOSR under Award No. FA9550-07-1-0124. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF or the AFOSR.

References

[1] David Aldous and James Fill. *Reversible Markov Chains and Random Walks on Graphs*. Monograph in Preparation, `http://www.stat.berkeley.edu/users/aldous/ RWG/book.html`.

[2] Monica Bianchini, Marco Gori, and Franco Scarselli. Inside PageRank. *ACM Trans. Inter. Tech.*, 5(1):92–128, 2005.

[3] N. Biggs. *Algebraic Graph Theory*. Cambridge University Press, 1993.

[4] Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, 30(1–7):107–117, 1998.

[5] A. Chan and C.D. Godsil. Symmetry and eigenvectors. In G. Hahn and G. Sabidussi, editors, *Graph Symmetry: Algebraic Methods and Applications*. Kluwer Academic Pub, 1997.

[6] Alice Cheng and Eric Friedman. Manipulability of PageRank under sybil strategies. In *Proceedings of the First Workshop of Networked Systems (NetEcon06)*, 2006.

[7] Eric Friedman, Paul Resnick, and Rahul Sami. Manipulation-resistant reputation systems. In N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani, editors, *Algorithmic Game Theory*. Cambridge University Press, 2007.

[8] C.D. Godsil. *Algebraic Combinatorics*. Chapman & Hall, 1993.

[9] C.D. Godsil and B.D. McKay. Feasibility conditions for the existence of walk-regular graphs. *Linear Algebra Appl*, 30:51–61, 1980.

[10] C.D. Godsil and G. Royle. *Algebraic Graph Theory*. Springer, 2001.

[11] Zoltán Gyöngyi, Pavel Berkhin, Hector Garcia-Molina, and Jan Pedersen. Link spam detection based on mass estimation. In *Proceedings of the 32nd International Conference on Very Large Databases*. ACM, 2006.

[12] Zoltán Gyöngyi and Hector Garcia-Molina. Link spam alliances. In *Proceedings of the 31st International Conference on Very Large Databases*, pages 517–528. ACM, 2005.

[13] Zoltán Gyöngyi and Hector Garcia-Molina. Web spam taxonomy. In *First International Workshop on Adversarial Information Retrieval on the Web*, 2005.

[14] Zoltán Gyöngyi, Hector Garcia-Molina, and Jan Pedersen. Combating web spam with TrustRank. In *Proceedings of the 30th International Conference on Very Large Databases*, pages 576–587. Morgan Kaufmann, 2004.

[15] T. Hogg and L. Adamic. Enhancing reputation mechanisms via online social networks. *Proceedings of the 5th ACM Conference on Electronic Commerce*, pages 236–237, 2004.

[16] J. Hopcroft and D. Sheldon. Manipulation-Resistant Reputations Using Hitting Time. *Lecture Notes in Computer Science*, 4863:68, 2007.

[17] M.O. Jackson. A Survey of Models of Network Formation: Stability and Efficiency. *Group Formation in Economics: Networks, Clubs and Coalitions*, 2003.

[18] Sepandar D. Kamvar, Mario T. Schlosser, and Hector Garcia-Molina. The eigentrust algorithm for reputation management in P2P networks. In *WWW '03: Proceedings of the 12th international World Wide Web Conference*, pages 640–651, New York, NY, USA, 2003. ACM Press.

[19] Kahn Mason. *Detecting Colluders in PageRank - Finding Slow Mixing States in a Markov Chain*. PhD thesis, Stanford University, 2005.

[20] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The PageRank citation ranking: Bringing order to the web. Technical report, Stanford Digital Library Technologies Project, 1998.

[21] Kevin T. Phelps. Automorphism free latin square graphs. *Discrete Mathematics*, 31(2):193–200, 1980.

[22] Search Engine Marketing Professional Organization (SEMPO). 2007 state of the markey survey (preview news release). `http://www.sempo.org/news/releases/03-17-08`, 2008.

[23] Éva Tardos and Tom Wexler. Network formation games and the potential function method. In N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani, editors, *Algorithmic Game Theory*. Cambridge University Press, 2007.

[24] Hui Zhang, Ashish Goel, Ramesh Govindian, Kahn Mason, and Benjamin Van Roy. Making eigenvector-based reputation systems robust to collusion. In *Proceedings of the 3rd Workshop on Algorithms and Models for the Web-Graph (WAW2004), in conjunction with FOCS 2004*, 2004.

A Additional Proofs

Proof of Proposition 1

Proof. Let $\{X_t\}_{t\geq 0}$ be the sequence of nodes visited by the random walk, let T be the time of the first jump, and let $I\{\cdot\}$ be the indicator variable for the event in braces. Since the decision to jump is independent of location, T is independent of X_t for all t. Then

$$N_{ij} = E_i\left[\sum_{t=0}^{\infty} I\{X_t = j\} I\{T > t\}\right] = \sum_{t=0}^{\infty} \mathrm{Pr}_i[X_t = j]\mathrm{Pr}_i[T > t] = \sum_{t=0}^{\infty} P_{ij}^t (1-\alpha)^t.$$

Hence $N = \sum_{t=0}^{\infty} P^t(1-\alpha)^t$. Since P is stochastic and $\alpha > 0$, this sum converges and is equal to $(I - (1-\alpha)P)^{-1}$. Part (ii) is well-known, and can be verified algebraically. We prefer a probabilistic argument. The value $(q^T N)_j$ is the expected number of visits to j between two restarts. Since the walk is a sequence of probabilistically identical segments delimited by restarts, the number of visits between restarts is proportional to stationary probability. See Proposition 3 in Chapter 2 of Aldous and Fill [1] for details. Part (iii) is a simple probabilistic statement. To count the number of visits to j starting from i, we first decide if the walk hits j, then count the number of visits starting from j. Part (iv) is proved in [16]. The intuition is: to measure the probability that v is hit before the first jump by Monte Carlo simulation, one need never follow a link leaving v. □

Proof of Lemma 3

Proof. We use the Nash conditions in Equation (3) from the proof of Theorem 2. We will show that $N_{ji}(\alpha)$ and $N_{ki}(\alpha)$ can be expressed using polynomials of degree $n-1$ in the variable α. Hence, if they are equal for at least n values of α, they are equal for *all* $\alpha \in (0,1)$. Let $M(\alpha) = I - (1-\alpha)P$, and recall from Proposition 1 that $N(\alpha) = M^{-1}(\alpha)$. Let $M_{[j,i]}(\alpha)$ denote the matrix of size $(n-1)\times(n-1)$ obtained from $M(\alpha)$ by deleting row j and column i. Using the formula for the matrix inverse in terms of cofactors,

$$N_{ji}(\alpha) = M_{ji}^{-1}(\alpha) = \frac{(-1)^{j+i}\det M_{[j,i]}(\alpha)}{\det M(\alpha)}.$$

Hence, $N_{ji}(\alpha) = N_{ki}(\alpha)$ if and only if

$$(-1)^{j+i}\det M_{[j,i]}(\alpha) = (-1)^{k+i}\det M_{[k,i]}(\alpha). \tag{6}$$

Each determinant in (6) is a sum of products of $n-1$ matrix elements, each of which is a monomial in α. Hence both sides of (6) are polynomials of degree $n-1$ in α; if they are equal for at least n distinct values of α, they are equal as polynomials. □

Proof of Lemma 4

Proof. Since G is edgewise walk-regular, there exist constants c_t for all $t \geq 0$ such that $A^t_{vw} = c_t$ for any edge (v, w). Also,

$$A^{t+1}_{vv} = (AA^t)_{vv} = \sum_{w \in V} A_{vw} A^t_{wv} = \sum_{w \in \Gamma(v)} A^t_{wv} = D_v c_t.$$

We can write two different expressions relating the spectrum of G to the constants c_t. First, the trace of A^{t+1} is the sum of the $(t+1)$th powers of eigenvalues, taken according to their multiplicities:

$$\sum_\lambda m_\lambda \lambda^{t+1} = \operatorname{tr}(A^{t+1}) = \sum_v A^{t+1}_{vv} = \sum_v D_v c_t = 2c_t|E|$$

Dividing by $2|E|$, we get

$$c_t = \sum_\lambda \lambda^{t+1} \frac{m_\lambda}{2|E|} = \sum_\lambda \lambda^{t+1} C_\lambda. \tag{7}$$

Second, we can compute c_t using the eigenvalue decomposition. Let $(v, w) \in E$. Then

$$c_t = A^t_{vw} = \sum_\lambda \lambda^t (U_\lambda U_\lambda^T)_{vw}. \tag{8}$$

Combining (7) and (8),

$$\sum_\lambda \lambda^t (U_\lambda U_\lambda^T)_{vw} = \sum_\lambda \lambda^{t+1} C_\lambda \tag{9}$$

Let $z_\lambda = (U_\lambda U_\lambda^T)_{vw} - \lambda C_\lambda$, and rewrite equation (9) as $\sum_\lambda \lambda^t z_\lambda = 0$. This is a linear equation in the variables z_λ. Since it holds for all $t \geq 0$, we can take the equations for $t = 0, 1, \ldots, \ell - 1$, where ℓ is the number of unique eigenvalues, obtaining a linear system whose coefficient matrix is an $\ell \times \ell$ Vandermonde matrix. Such a matrix is nonsingular, so it must be the case that $z_\lambda = 0$ for all λ, giving $(U_\lambda U_\lambda^T)_{vw} = \lambda C_\lambda$. For the diagonal entries, we write

$$c_t = \frac{A^{t+1}_{vv}}{D_v} = \sum_\lambda \lambda^{t+1} \frac{(U_\lambda U_\lambda^T)_{vv}}{D_v}. \tag{10}$$

Combining (10) and (7), we follow a similar argument to conclude that $(U_\lambda U_\lambda^T)_{vv} = D_v C_\lambda$. □

Proof of Lemma 5

Proof. Fix a vertex v. First we show that the centroid of v's neighbors is equal to y. Since the columns of U_λ are eigenvectors, we have $AU_\lambda = \lambda U_\lambda$. Taking the vth row of this matrix equation, we get $\sum_{w \in \Gamma(v)} u_\lambda(w) = \lambda u_\lambda(v)$, and dividing both sides by D_v, we see the centroid is equal to y.

Next we show that for all neighbors w, the vectors $u_\lambda(w) - y$ are orthogonal to $u_\lambda(v)$ and have length r. By Lemma 4, if w is a neighbor of v, then $u_\lambda(w)$ lies in the affine hyperplane $H = \{x : u_\lambda(v)^T x = \lambda C_\lambda\}$. But y is the centroid of points in H, so y also lies in H. Hence $u_\lambda(w) - y$ is orthogonal to the normal vector $u_\lambda(v)$, and also orthogonal to y, which is a scalar multiple of $u_\lambda(v)$. Assume v has degree d_1 and w has degree d_2. Using orthogonality and Lemma 4,

$$\|u_\lambda(w) - y\|^2 \;=\; \|u_\lambda(w)\|^2 - \left\| \frac{\lambda}{d_1} u_\lambda(v) \right\|^2 \;=\; d_2 C_\lambda - \frac{\lambda^2}{d_1^2} \cdot d_1 C_\lambda \;=\; \frac{C_\lambda}{d_1}\left(d_1 d_2 - \lambda^2\right) \;=\; r^2$$

It remains to show that $r > 0$ whenever $m_\lambda > 1$. Suppose $r = 0$. Then $u_\lambda(w)$ is equal to y, which is a scalar multiple of $u_\lambda(v)$. By the same argument, if $r = 0$ then the images of w's neighbors are scalar multiples of $u_\lambda(w)$; continuing this argument, since G is connected, the images of all vertices are scalar multiples of $u_\lambda(v)$. In other words, all rows of U_λ are scalar multiples of the vth row, which cannot happen unless the rank of U_λ is equal to one, i.e., unless $m_\lambda = 1$. □

Proof of Lemma 7

Proof. We verify that $x^{(i)}$ is an eigenvector. Let B be the adjacency matrix of the (a,b)-weighted k-cycle. Suppose v is odd. Then

$$\begin{aligned}
(Bx^{(i)})_v &= \frac{a}{2} \cdot x^{(i)}_{v-1} + \frac{a}{2} \cdot x^{(i)}_{v+1} \\
&= \frac{a\sqrt{b}}{2}\left(\cos\left(\frac{2\pi i(v-1)}{k}\right) + \cos\left(\frac{2\pi i(v+1)}{k}\right)\right) \\
&= \frac{a\sqrt{b}}{2}\left(\cos\left(\frac{2\pi iv}{k} - \frac{2\pi i}{k}\right) + \cos\left(\frac{2\pi iv}{k} + \frac{2\pi i}{k}\right)\right) \\
&= \frac{a\sqrt{b}}{2}\left(2\cos\left(\frac{2\pi iv}{k}\right)\cos\left(\frac{2\pi i}{k}\right)\right) \\
&= \sqrt{ab}\cos\left(\frac{2\pi i}{k}\right)\sqrt{a}\cos\left(\frac{2\pi iv}{k}\right) \\
&= \lambda_i x^{(i)}_v
\end{aligned}$$

The calculation is similar when v is even, and for the eigenvector $y^{(i)}$. □