On the Second Eigenvalue and Random Walks in Random d-Regular Graphs
August 24, 1988
Revised May 22, 1989

Joel Friedman*
Princeton University

November 26, 1993

Abstract

The main goal of this paper is to estimate the magnitude of the second largest eigenvalue in absolute value, λ_2, of (the adjacency matrix of) a random d-regular graph, G. In order to do so, we study the probability that a random walk on a random graph returns to its originating vertex at the k-th step, for various values of k. Our main theorem about eigenvalues is that

$$\mathrm{E}\left\{|\lambda_2(G)|^m\right\} \le \left(2\sqrt{2d-1}\left(1+\frac{\log d}{\sqrt{2d}}+O\left(\frac{1}{\sqrt{d}}\right)\right)+O\left(\frac{d^{3/2}\log\log n}{\log n}\right)\right)^m$$

for any $m \le 2\left\lfloor \log n \left\lfloor \sqrt{2d-1}/2 \right\rfloor / \log d \right\rfloor$, where $\mathrm{E}\{\,\}$ denotes the expected value over a certain probablity space of $2d$-regular graphs. It follows, for example, that for fixed d the second eigenvalue's magnitude is no more than $2\sqrt{2d-1}+2\log d+C'$ with probability $1-n^{-C}$ for constants C and C' for sufficiently large n.

1 Introduction

Let G be a d-regular (i.e. each vertex has degree d) undirected graph, and let A be its adjacency matrix. We allow G to have multiple edges, in which case the corresponding entry of A is that multiplicity; we allow (possibly multiple) self-loops, in which case the corresponding entry of A is twice the multiplicity of the self-loop. Since A is symmetric it is diagonalizable with (an orthogonal set of eigenvectors and with) real eigenvalues. It is easy to see that d is the largest eigenvalue in absolute value (with the all 1's eigenvector). Let $\lambda_2 = \lambda_2(G)$ be the next largest eigenvalue in absolute value.

*The author wishes to acknowledge the National Science Foundation for supporting this research in part under Grant CCR–8858788, and the Office of Naval Research under Grant N00014–87–K–0467.

The magnitude of $|\lambda_2|$ has received much attention in the literature: for example, it is useful to give some estimate of the expansion properties of G (see [Tan84],[Alo86]) or the rate of convergence of the Markov process on G (with probabilities $1/d$ at each edge) to the stable distribution (see [BS87] for more on this and references). The smaller $|\lambda|$ the better, usually; intuitively it measures the difference (in L_2 operator norm) between A and the all d/n 's matrix, and more generally that between A^m and the all d^m/n 's matrix.

The main goal of this paper is to estimate the magnitude of $|\lambda_2|$ of a random d-regular graph. We view d as fixed and $n \to \infty$; if d is sufficiently large with respect to n, it is well known that $|\lambda_2| = O(\sqrt{d})$ with high probability. In order to estimate $|\lambda_2|$, we study the probability that a random walk on a random graph returns to its originating vertex at the k-th step, for various values of k (for us, $k \approx \log n$ or $(\log n)^2$ are important). We will actually work with $2d$-regular graphs, drawn from a probability space $\mathcal{G}_{n,2d}$ to be defined shortly. Our main theorem is

Theorem A *For $G \in \mathcal{G}_{n,2d}$ we have*

$$E\{|\lambda_2(G)|\} \leq 2\sqrt{2d-1}\left(1 + \frac{\log d}{\sqrt{2d}} + O\left(\frac{1}{\sqrt{d}}\right)\right) + O\left(\frac{d^{3/2}\log\log n}{\log n}\right)$$

(with an absolute constant in the $O(\ )$ notation), where $E\{\}$ denotes the expected value over $\mathcal{G}_{n,2d}$ and where the logarithm is taken base e; more generally we have

$$E\{|\lambda_2(G)|^m\} \leq \left(2\sqrt{2d-1}\left(1 + \frac{\log d}{\sqrt{2d}} + O\left(\frac{1}{\sqrt{d}}\right)\right) + O\left(\frac{d^{3/2}\log\log n}{\log n}\right)\right)^m$$

for any $m \leq 2\lfloor \log n \lfloor \sqrt{2d-1}/2 \rfloor / \log d \rfloor$.

All logarithms in this paper are taken base e. A corollary of this theorem is

Theorem B *For any $\beta > 1$ we have*

$$|\lambda_2(G)| \geq \left(2\sqrt{2d-1}\left(1 + \frac{\log d}{\sqrt{2d}} + O\left(\frac{1}{\sqrt{d}}\right)\right) + O\left(\frac{d^{3/2}\log\log n}{\log n}\right)\right)\beta$$

with probability

$$\leq \frac{\beta^2}{n^{2\lfloor \sqrt{2d-1}/2 \rfloor \log \beta / \log d}} \,.$$

In particular, we have that for any C there is a C' such that for any d there is an n_0 such that if $n \geq n_0$ then a $G \in \mathcal{G}_{n,2d}$ has

$$|\lambda_2(G)| \leq 2\sqrt{2d-1} + 2\log d + C'$$

with probability at least $1 - n^{-C}$.

One can compare this to the well-known lower bound

$$|\lambda_2(G)| \geq 2\sqrt{2d-1} + O\left(\frac{1}{\log_d n}\right)$$

for any $2d$-regular graph, G, with, say, $d \leq \sqrt{n}$ (the bound as stated is easy, and has been discovered by many people; for a proof see section 3; a similar result also holds for the second largest *positive* eigenvector, see [McK81] and Alon and Boppana in[Alo86]). Previously Broder and Shamir, in [BS87] had shown that for fixed d,

$$\mathrm{E}\left\{|\lambda_2(G)|\right\} \leq 2^{5/4} d^{3/4}(1+\epsilon+o(1))$$

as $n \to \infty$, and the analog to the second equation in theorem A with $m \leq (2-\epsilon')\log_{d/2} n$. Our method refines Broder and Shamir's estimate of the probability that a random walk in a random graph returns to its original vertex after some specified number of steps; this therefore improves the estimates on $|\lambda_2(G)|$. Also, Kahn and Szemerédi, in [KS], have independently given a much different approach, which modifies the standard counting argument (the standard counting argument gives no interesting bound), to get $|\lambda_2(G)| = O(\sqrt{d})$ with probability $\to 1$ as $n \to \infty$. The following conjecture of Alon is still unresolved, which states that for *fixed* d, as $n \to \infty$, $|\lambda_2| \leq 2\sqrt{2d-1} + \epsilon$ with high probability for any $epsilon > 0$.

The standard approach to estimating eigenvalues is to estimate the trace of a high power of the adjacency matrix, i.e. estimate the probability that a random walk of a given length will return to the starting vertex on a random graph. Since

$$\mathrm{Trace}\left(A^k\right) = \sum_{i=1}^{n} \lambda_i^k$$

for a matrix, A, with eigenvalues $\lambda_1, \ldots, \lambda_n$, this gives bounds on the magnitude of the largest unknown eigenvalue. This method was used by Wigner in [Wig55] to find the limiting distribution (the "semi-circle law") of the eigenvalues in a random $n \times n$ symmetric matrix with independently chosen entries drawn from a fixed distribution, as $n \to \infty$; implicit in his paper is that when the expected value of the entry is 0 the largest (second largest otherwise) eigenvalue is of magnitude roughly proportional to $\sqrt{n}$. His approach to estimating the trace of high powers of the matrix, and subsequent refinements (see [Gem80], [FK81] for example) do not work when the distribution of the entries vary too much with n, in particular with matrices of graphs whose edge density decreases too fast with n. In [McK81], McKay derived a semi-cirlce law for d regular graphs, i.e. the limiting distribution of the eigenvalues, but the method did not give bounds on $|\lambda_2|$ (the limiting distribution only gives information on $n-o(n)$ of the eigenvalues). In [Mar87] and [LPS86], Margulis and (independently) Lubotzky, Phillips, and Sarnak constructed a class of graphs with $|\lambda_2| \leq 2\sqrt{d-1}$. In [BS87], Broder and Shamir gave a method of estimating traces for random graphs which gave $|\lambda_2| = O(d^{3/4})$ with high probability. In this paper we use Broder and Shamir's framework for trace estimates; we also make use of the cosine substitution and resulting formulas as in [LPS86], and Alon's theorem that says expansion implies "small" second eigenvalue, in [Alo86]. We begin by reviewing the terminology and results in [BS87].

The model we take is for a random $2d$-regular graph on n vertices. Choose independently d permutations of the numbers from 1 to n, each permutation equally likely. We construct a directed graph, $G = (V, E)$ with vertex set $V = \{1, \ldots, n\}$ and edges

$$E = \left\{ (i, \pi_j(i)), (i, \pi_j^{-1}(i)) \mid j = 1, \ldots, d \quad i = 1, \ldots, n \right\} .$$

G is therefore a graph with possibly multiple edges and self-loops, and which is symmetric in the sense that for every edge in the graph the reverse edge is also in the graph. Although G is directed, we can view it as an undirected graph by replacing each pair of edges $(i, \pi(i))$, $(\pi(i), i)$ with one undirected edge. We denote this probability space of random graphs $\mathcal{G}_{n,2d}$.

Let A be the adjacency matrix of a graph $G \in \mathcal{G}_{n,2d}$. Let Π be the alphabet of symbols

$$\Pi = \{\pi_1, \pi_1^{-1}, \pi_2, \ldots, \pi_d^{-1}\} .$$

We think of $\pi_1, \ldots, \pi_d$ as being the d permutations from which G was constructed, and any word, $w = \sigma_1 \ldots \sigma_k$ of Π^* as the permutation which is the composition of the permutations $\sigma_1, \ldots, \sigma_k$. Let

$$i \stackrel{w}{\rightarrow} j \equiv \left\{ \begin{array}{ll} 1 & \text{if } w(i) = j \\ 0 & \text{otherwise} \end{array} \right\} .$$

We have that the i, j-th entry of A^k is

$$\sum_{w \in \Pi^k} i \stackrel{w}{\rightarrow} j .$$

We wish to estimate the expected value of the above sum for $i = j$. In evaluating $i \stackrel{w}{\rightarrow} i$, we can cancel in w any consecutive pair of letters $\pi\pi^{-1}$ with $\pi \in \Pi$. We say that a word $w \in \Pi^k$ is *irreducible* if w has no pair of consecutive letters of the form $\pi\pi^{-1}$. We denote the set of irreducible words of length k by Irred_k. Clearly Irred_k has size $2d(2d-1)^{k-1}$. It turns out that to estimate the second eigenvalue it suffices to get an estimate of the form

$$\sum_{w \in \mathrm{Irred}_k} i \stackrel{w}{\rightarrow} i \; = \; 2d(2d-1)^{k-1} \frac{1}{n} + \text{error} \tag{1.1}$$

for all fixed i with a small error term; to estimate the expected second eigenvalue it suffices to estimate the expected value of the left-hand-side of equation 1.1. Intuitively, for $k \geq 1$ we expect that words in Irred_k will send a fixed vertex to each vertex with more or less equal probability, $1/n$. The corresponding eigenvalue estimate is roughly $O(n^{1/k}(\sqrt{d} + \text{error}^{1/k}))$. In [LPS86], the error terms for a specific family of graphs were each bounded by $(cd)^{k/2}$ for all k, yielding $|\lambda_2| \le 2\sqrt{d-1}$. In [BS87] the error term for the expected value over $\mathcal{G}_{n,2d}$ was bounded as

$$O\left(\frac{k^4}{n^2}\, 2d(2d-1)^{k-1} + \frac{k}{n}\,(cd)^{k/2}\right) ,$$

which for $k = 2\log_d n$ is $\le (cd)^{3k/4}/n$.

In this paper we obtain sharper estimates on the error term for the expected value over $\mathcal{G}_{n,2d}$. One that is useful for information on the eigenvalues is (corollary 3.6)

Theorem C *There is a constant c such that the error term in equation 1.1 is no more than*

$$2d(2d-1)^{k-1}\,(ckd)^c \left(\frac{k^{2\sqrt{2d}}}{n^{1+\lfloor\sqrt{2d-1}/2\rfloor}} + \frac{(2d-1)^{-k/2}}{n}\right) . \tag{1.2}$$

for all k with $k \le n^{1/(c\sqrt{d})}/(cd^{3/2})$ and $d \ge 4$.

(See theorem 2.18 for other estimates on the error term.) From the above theorem we derive estimates on the second eigenvalue.

In section 2 we do the most of the work, which is to prove that the sum in equation 1.1 has an r-th order asymptotic expansion in $1/n$ for any $r \le \sqrt{2d-1}/2$ of the form

$$\mathrm{E}\left\{\sum_{w\in \mathrm{Irred}_k} i \xrightarrow{w} i\right\} \;=\; 2d(2d-1)^{k-1}\left(\frac{1}{n}\,f_0(k) + \frac{1}{n^2}\,f_1(k) + \ldots + \frac{1}{n^r}\,f_{r-1}(k)\right) + \mathrm{error}\ , \tag{1.3}$$

where the error involves a $1/n^{r+1}$ term and a term like the second term in equation 1.2, and where the f_i's are polynomials of degree $5i+2$ whose coefficients are bounded by $(cdr)^{cr^2}$. In section 3 we show why theorem C follows from equation 1.3, i.e. why $f_0(k) = 1$ and all the other f_i's must vanish, and derive information on the eigenvalues, such as theorem A.

2 An Asymptotic Expansion

The goal of this section is to develop the asymptotic expansion of equation 1.3. We first explain the reason that such an expansion should exist, reviewing the ideas in [BS87]. Throughout this section we will bound error terms by expressions involving various absolute constants. Rather than giving each one a distinct name, we shall denote them all by c; in cases where confusion could arise we shall also use c'.

For $w = \sigma_1 \ldots \sigma_k \in \mathrm{Irred}_k$ and i fixed, consider the random variable $i \xrightarrow{w} i$. We wish to estimate $\mathrm{E}\{i \xrightarrow{w} i\}$. In [BS87] the analysis is as follows. It is helpful to consider the random variables $t_1 = \sigma_1(i)$, $t_2 = \sigma_2(t_1)$, etc. which trace i's path through G along w. We simplify the calculation of $\mathrm{E}\{i \xrightarrow{w} i\}$ over $\mathcal{G}_{n,2d}$ by considering only the values of $\pi_1, \ldots, \pi_d$ needed to determine the t_j's and therefore the value of $i \xrightarrow{w} i$. So consider the process of selecting an element of $\mathcal{G}_{n,2d}$ by first selecting t_1, then t_2, until t_k, and then determining the remaining values of the π_j's.

To determine t_1 it suffices to know the value of $\sigma_1(i)$, which clearly takes on the values $\{1, \ldots, n\}$ each with probability $1/n$. Next we determine the value of $\sigma_2(t_1)$; there are a few cases to consider. If $\sigma_2 \neq \sigma_1$, then σ_2 and σ_1 represent different permutations (since w is irreducible, $\sigma_2 \neq \sigma_1^{-1}$), and so t_2 takes on the values $\{1, \ldots, n\}$ each with probability $1/n$. If $\sigma_1 = \sigma_2$, then either $t_1 = i$ and so t_2 is forced to be i, or $t_1 \neq i$, in which case t_2 takes on the values $\{1, \ldots, n\} - \{t_1\}$ each with probability $1/(n-1)$. We can continue to determine $t_3, \ldots, t_k$ in this fashion, each t_j's value conditional on the values of the previous ones. When t_j's value is exactly determined by the previous values, i.e. $\sigma_j(t_{j-1})$'s value is known, we say that t_j is a *forced choice*; otherwise we say that t_j is a *free choice*. If t_j is a free choice, then clearly t_j takes on any one of $n - m$ values with probability $1/(n-m)$ for some $m \leq j - 1$. If a free choice t_j happens to be a previously visited vertex (i.e. $= i$ or $= t_m$ for some $m < j$) we say that t_j is a *coincidence*; given that t_j is a free choice, t_j will be a coincidence with probability $\leq j/(n - j + 1)$. During this paper we will also call an edge (t_{j-1}, t_j) a free choice, forced choice, or coincidence if t_j is respectively a free choice, forced choice, or coincidence.

In [BS87], it is observed that two coincidences occur in any given w and i with probability no greater than

$$\binom{k}{2}\left(\frac{k-1}{n-k+1}\right)^2 = O\left(\frac{k^4}{n^2}\right).$$

This accounts for the first of their error terms. The rest of the work is to analyze the case of one coincidence, since if there are no coincidences then clearly $i \xrightarrow{w} i$ is 0. We will analyze the case of $\leq r$ coincidences for any $r \leq \sqrt{2d-1}/2$, and show that the error term in equation 1.1 actually has an r term asymptotic expansion in $1/n$, whose coefficients are polynomials in k for any fixed d. The key to the proof is to group together walks in the graph which have the same rough shape, which we call the *type* of the walk.

Let

$$i_0 \xrightarrow{w_1} i_1 \xrightarrow{w_2} \cdots \xrightarrow{w_j} i_j \equiv \left\{ \begin{array}{ll} 1 & \text{if } w_k(i_{k-1}) = i_k \text{ for } k = 1, \ldots, j \\ 0 & \text{otherwise} \end{array} \right\}.$$

Consider

$$\mathrm{E}\left\{ \sum_{w \in \mathrm{Irred}_k} i \xrightarrow{w} i \right\} = \sum_{\substack{w=\sigma_1\ldots\sigma_d \in \mathrm{Irred}_k \\ t_j \in \{1,\ldots,n\},\ j=1,\ldots,k-1}} \mathrm{E}\left\{ i \xrightarrow{\sigma_1} t_1 \xrightarrow{\sigma_2} \cdots \xrightarrow{\sigma_{k-1}} t_{k-1} \xrightarrow{\sigma_k} i \right\}.$$

We will group the terms of the right-hand-side summation by their geometric configuration. Fix a word $w = \sigma_1 \ldots \sigma_d$, integer i, and sequence of integers $t = (t_1, \ldots, t_{k-1})$

and let $t_k = i$. With these values we associate a directed graph, $\Gamma_{w,i,t}$, with edges labeled by Π. The graph's vertex set is $V = \{i_1, \ldots, i_m\}$, where m is the number of distinct integers among i and t; we think of i_1 as representing i and, more generally, i_l as representing the l-th distinct member of the sequence $i, t_1, \ldots, t_{k-1}$. The edge set, E, has one edge e_j for each free choice t_j; e_j is an edge (i_a, i_b) where i_a i_b respectively represent the values of t_{j-1} and t_j; in addition e_j is labeled with σ_j. $\Gamma_{w,i,t}$ represents the geometric structure of the intended path from i through t and back to i along w; it contains the amount of conditioning on $\pi_1, \ldots, \pi_d$ given that the intended path is actually taken. We say call $\Gamma_{w,i,t}$ the *generalized form* of w, i, t.

Consider an *abstract* generalized form, $\Gamma = (V_\Gamma, E_\Gamma)$, i.e. a directed graph with vertices named $V_\Gamma = \{i_1, \ldots, i_m\}$ and with edges labelled by Π, which is $= \Gamma_{w,i,t}$ for some (w, i, t) triple, but where we ignore the particular (w, i, t) from which is arises (at least for the time being). Let $a_j(\Gamma)$ denote the total number of occurences of π_j and π_j^{-1} as labels of edges. a_j gives the number of values of π_j determined by Γ. For a fixed w we have

$$\sum_{\{(i,t) | \Gamma_{w,i,t} = \Gamma\}} \mathrm{E}\left\{ i \xrightarrow{\sigma_1} t_1 \xrightarrow{\sigma_2} \cdots \xrightarrow{\sigma_{k-1}} t_{k-1} \xrightarrow{\sigma_k} i \right\} =$$

$$n(n-1)\ldots(n-|V_\Gamma|+1) \prod_{j=1}^{d} \frac{1}{n(n-1)\ldots(n-a_j(\Gamma)+1)} ,$$

where the j-th term in the above product is omitted if $a_j(\Gamma) = 0$. We denote the above by $\Pr(\Gamma)$. We can expand

$$\Pr(\Gamma) = \frac{1}{n^{\mathrm{coin}(\Gamma)-1}} \left(1 + \frac{c_1}{n} + \frac{c_2}{n^2} + \cdots \right) \tag{2.1}$$

where the c_j's are constants depending on Γ and $\mathrm{coin}(\Gamma)$ denotes the number of coincidences in Γ (more precisely in any triple (w, i, t) with $\Gamma_{w,i,t} = \Gamma$), which does not depend on (w, i, t). Clearly

$$\sum_{\substack{w \in \mathrm{Irred}_k \\ i=1,\ldots,n}} \mathrm{E}\left\{ i \xrightarrow{w} i \right\} = \sum_{\Gamma} \Pr(\Gamma)\, \omega(\Gamma, k) , \tag{2.2}$$

where $\omega(\Gamma, k)$ denotes the number of irreducible words, w, of length k compatible with Γ, i.e. for which there are i and t with $\Gamma_{w,i,t} = \Gamma$.

Next we group together generalized forms, Γ, whose underlying graphs have the same rough shape. Consider again, for a fixed word, w, and vertex, i, the process of determining the t_j's. We begin by generating distinct vertices t_1, t_2, ..., adding vertices and edges to Γ, until we encounter a coincidence. We may then encounter forced choices, which is to say walk in Γ without generating new vertices or edges, until we encounter a free choice, t_l. Starting with the edge from t_{l-1} to t_l, we depart from the "old" Γ, walking along new edges and vertices until we encounter our next coincidence. We call a point such as t_{l-1} (more precise, the last vertex in a consecutive sequence of forced choices) a *departure*. We will sometimes also refer to the edge (t_{l-1}, t_l) as a departure, if t_{l-1} is. Clearly departures alternate with coincidences throughout the walk. We now group Γ's by forgetting any vertex which is not a coincidence or a departure. In other words, we forget all degree 2 vertices (except i_1); nothing interesting happens at them.

We say that a vertex i_j of Γ is a coincidence (departure) if one of its corresponding t_l's is a coincidence (departure); as with $\mathrm{coin}(\Gamma)$, these notions are independent of (w, i, t). In fact, i_1 is always a coincidence, i_j for $j > 1$ is a coincidence iff its indegree is > 1, and i_j for all j is a departure iff its outdegree is > 1. Any vertex of Γ which is not a coincidence or a departure has indegree and outdegree 1, and so the edges of Γ are a union of simple (directed) paths from pairs of the subset of vertices, S, consisting of all coincidences and departures. With Γ we associate its *type*, T_Γ, which is an *undirected* graph, possibly with multiple edges and self-loops. Its vertex set is $\{v_1, \ldots, v_l\}$ where $l = |S|$, and as before we think of v_j as representing the j-th vertex of S in the order $i_1, \ldots, i_m$. T_Γ has an edge for each simple directed path from pairs of S.

Given an abstract type, T, which is just an undirected graph (corresponding to at least one Γ), we order any group of multiple edges to distinguish each edge from its copies. In other words, for each edge of multiplicity s, represented by distinct edges $\{e_1, \ldots, e_s\}$, we label $\{e_1, \ldots, e_s\}$ with distinct integers $\{1, \ldots, s\}$ in some fashion. This labelling is equivalent to imposing an order on each group of multiple edges. We call a type together with such an ordering or labelling a *type with ordered multiple edges*, or simply an *ordered type*. The point is that given an abstract generalized form, Γ, any word w compatible with Γ gives rise to a unique *oriented* labelling with labels in Π^* of the edges of the corresponding (abstract) ordered type T_Γ (an oriented labelling being one where we orient the edge and give it a label, assigning the inverse label to the oppositely oriented direction), and to a unique walk in T_Γ. To see this, consider the walk from i_1 to itself along w, which traces out a unique walk in Γ. This gives rise to possibly many walks in T_Γ, but only to one walk such that (1) the edges are labelled and oriented as they are traversed, labelled with the word in Π^* which is simultaneously traced out in Γ, and oriented in the direction traversed, (2) once an edge is labelled it is traversed only when it matches the word in Π^* traced out during the traversal, and (3) multiple edges are traversed in increasing order.

Conversely, let T be an (abstract) ordered type. By a *legal walk* in T we mean a walk from v_1 to itself which traverses every edge at least once, and which traverses multiple edges in increasing order. By a *legal oriented labelling* we mean an oriented labeling of T's edges by irreducible words in Π^* of length ≥ 1 such that for any vertex, v, in T the labels on the outwardly oriented edges from v begin with distinct letters. It is easy to see that we have a one-to-one correspondence

$$\{(\Gamma, w) \mid \text{General forms, } \Gamma,\ w \in \text{Irred compatible with } \Gamma\} \overset{1-1}{\longleftrightarrow}$$
$$\{(T, y, \ell) \mid \text{Ordered types } T, \text{ legal walks } y \text{ and oriented labellings } \ell \text{ for } T\}\ .$$

To calculate the asymptotic expansion, we will sum over types, walks, and oriented labellings, rather than over generalized forms and words.

Given a legal walk, y, on an ordered type, T, we define the *multiplicity of the walk* to be the function which assigns to each edge of T the number of times it is traversed, disregarding orientation. For a legal oriented labelling, ℓ, of T, we define the *lettering of the labelling* to be the function which takes a vertex, v, and an incident edge, e, and returns the first letter of e's label with the outward orientation from v. We define the *weight of the labelling* to be the function which takes an edge and returns the length of the word with which it it labelled. The notion of $\mathrm{coin}\,(T)$ for types T carries over from that of generalized forms (i.e. as equal to $\mathrm{coin}\,(\Gamma)$ for one and therefore all Γ's with $T_\Gamma = T$). Our strategy will be to fix a type, a multiplicity, and a lettering, and then sum the corresponding terms in equation 2.2.

We are now ready for the nitty-gritty. Lemmas which are not followed by proofs are immediate or easy.

Lemma 2.1 *For any type, $T = (V, E)$ we have* $\mathrm{coin}\,(T) = |E| - |V| + 1$.

Lemma 2.2 *For a type, T, of coincidence r we have $|V| \leq 2r$ and $|E| \leq 3r - 1$.*

Proof In any legal walk in a generalized form, the coincidences and departures alternate, starting and ending with coincidences. Any vertex, v_l, of the type, with $l \geq 2$, must be either a coincidence or a departure (or both). Thus $|V| \leq 2r$. This and lemma 2.1 yield $|E| \leq 3r - 1$.

□

Lemma 2.3 *The number of types of coincidence r is less than $(2r)^{6r-2}$. The number of types of coincidence $\leq r$ is less than $(2r)^{6r-1}$.*

Lemma 2.4 *The number of letterings of a type of coincidence r is $\leq (2d)^{6r-2}$.*

Let $\mathcal{T}_r$ be the type with two vertices, v_1, v_2, with one edge from v_1 to v_2 and r edges from v_2 to itself (i.e. self-loops). $\mathcal{T}_r$ is of coincidence r.

Lemma 2.5 *In a type, T, of coincidence r which is not equal to $\mathcal{T}_r$, each vertex has degree at most $2r$.*

Proof Again, consider any legal walk in a generalized form. When a vertex is first encountered it is entered and left, giving it indegree and outdegree 1 up to that point. Then it can be entered by coincidence edges up to r times, and left by departure edges up to $r - 1$ times. The degree of the corresponding vertex in the type will therefore have degree $\leq 2r + 1$. We claim that equality can only occur in the type $\mathcal{T}_r$. For equality implies that the corresponding vertex, v, in the type must have r self-edges and that it is the only coincidence and departure in the type; this leaves room for only (possibly) one other vertex, v_1, which corresponds to i_1. If v is not v_1, then this type must be $\mathcal{T}_r$, and if $v = v_1$, then v's degree is $2r$.

□

Lemma 2.6 *In a fixed type of coincidence r, there are $\leq (2r)^m$ legal walks whose sum of its mutliplicities is m.*

Proof If T is not $\mathcal{T}_r$, then each vertex has degree $\leq 2r$ and so there are $\leq (2r)^m$ legal walks of length m in T. In $\mathcal{T}_r$, a legal walk consists of one step from v_1 to v_2, then $m-2$ steps each taking one of the r self-loops in either of its two directions, and finally returning to v_1, for a total of $\leq (2r)^{m-2}$ legal walks.

□

Consider the function

$$g(x) = (1-b_1x)\dots(1-b_rx)\left(\frac{1}{1-c_1x}\right)\dots\left(\frac{1}{1-c_sx}\right),$$

where $b_1,\dots,b_r,c_1,\dots,c_s$ are positive constants. Its i-th derivative is

$$\sum_{\substack{i_1+\dots+i_r+j_1+\dots+j_s=i\\ i_k\leq 1 \text{ for } k=1,\dots,r}} \binom{i}{i_1,\dots,i_r,j_1,\dots,j_s} (-b_1)^{i_1}(1-b_1x)^{i_1-1}\dots(-b_r)^{i_r}(1-b_rx)^{i_r-1}\times$$
$$j_1!\dots j_s!\,\frac{c_1^{j_1}\dots c_s^{j_s}}{(1-c_1x)^{j_1}\dots(1-c_sx)^{j_s}}.$$

For any $0\leq x\leq 1/c$, where c is an upper bound on the b_k's and c_k's, we have

$$\begin{aligned} |g^{(i)}(x)| &\leq \left(\frac{1}{1-xc}\right)^i i! \sum_{i_1+\dots+i_r+j_1+\dots+j_s=i}\binom{i}{i_1,\dots,i_r,j_1,\dots,j_s} b_1^{i_1}\dots b_r^{i_r}c_1^{j_1}\dots c_s^{j_s} \\ &= \left(\frac{1}{1-xc}\right)^i i!\left(\sum b_k+\sum c_k\right)^i, \end{aligned} \tag{2.3}$$

by the multinomial theorem. As a consequence we have

Lemma 2.7 *For any generalized form Γ of $\leq k$ edges, $\Pr(\Gamma)$ has an expansion*

$$\Pr(\Gamma) = \frac{1}{n^{\mathrm{coin}(\Gamma)-1}}\left(p_0+\frac{p_1}{n}+\frac{p_2}{n^2}+\dots+\frac{p_r}{n^r}+\frac{\epsilon}{n^{r+1}}\right) \tag{2.4}$$

(with $p_0=1$), where ϵ is an error term bounded by

$$e^{(r+1)k/n}k^{2r+2}$$

and the p_i's are constants.

Proof We apply equation 2.3 to equation 2.1, where we have $\sum b_j + \sum c_j \leq k^2$. By Taylor's theorem, ϵ is bounded by

$$\frac{1}{(r+1)!}\left(\frac{1}{1-\ k\xi}\right)^{r+1}(r+1)!\,(k^2)^{r+1}\ ,$$

for some $\xi \in [\,0, 1/n]$ for each n.

□

The p_i of lemma 2.7 are polynomials in the a_j and $|V_\Gamma|$ of equation 2.1, and we wish to bound their size and coefficients.

For any integer a, consider the function

$$g(x) = \frac{1}{1-x}\frac{1}{1-2x}\cdots\frac{1}{1-ax}\ .$$

A power series for g is given in terms of Stirling numbers. We wish to bound the size of their coefficients as polynomials in a.

Lemma 2.8 *$g(x)$'s power series about $x = 0$ is of the form*

$$1 + xR_1(a) + x^2R_2(a) + \cdots + x^rR_r(a) + \cdots\ ,$$

where the R_i's are polynomials of degree $2i$,

$$R_i(a) = \sum_{j=0}^{2i} c_{i,j}\binom{a}{j}\ ,$$

where the $c_{i,j}$'s are non-negative integers with

$$c_{i,0} + \cdots + c_{i,2i} \leq 8^i i!\ .$$

Proof Writing

$$1 + xR_1(a) + x^2R_2(a) + \cdots = (1+ax)(1 + xR_1(a-1) + x^2R_2(a-1) + \cdots)$$

and comparing terms yields

$$R_r(a) = R_r(a-1) + aR_{r-1}(a) \tag{2.5}$$

for $r \geq 1$ (with $R_0(a) \equiv 1$). We get a definition for the R_r's recursive in r. Since $R_0(a)$ is 1 it easily follows that R_r is a polynomial of degree $2r$. Since $R_r(1) = 1$ for all r, setting $a = 1$ in equation 2.5 yields $R_r(0) = 0$, where we identify R_r with the polynomial which agrees with it on all positive integers. Thus $c_{r,0} = 0$ for all r. For the other coefficients we have

$$\begin{aligned} R_r(a) - R_r(a-1) &= \sum_{k=0}^{2r-2} c_{r-1,k}\, a \binom{a}{k} \\ &= \sum_{k=0}^{2r-2} c_{r-1,k} \left(\binom{a}{k+1}(k+1) + \binom{a}{k} k \right) \\ &= \sum_{k=0}^{2r-2} c_{r-1,k} \left(\binom{a-1}{k+1}(k+1) + \binom{a-1}{k}(k+1) + \binom{a-1}{k} k + \binom{a-1}{k-1} k \right) . \end{aligned}$$

Solving the above difference equation it follows by induction on r that all the $c_{r,k}$'s are non-negative integers, and that $\sum_k c_{r,k} \leq (8r-6) \sum_k c_{r-1,k}$.

□

Similarly, we consider

$$g(x) = (1-x)(1-2x)\ldots(1-ax) .$$

Lemma 2.9 *$g(x)$'s power series about $x = 0$ is of the form*

$$1 - xQ_1(a) + x^2 Q_2(a) - \cdots + (-1)^r x^r Q_r(a) + \cdots ,$$

where the Q_i's are polynomials of degree $2i$,

$$Q_i(a) = \sum_{j=0}^{2i} c_{i,j} \binom{a}{j} ,$$

where the $c_{i,j}$'s are non-negative integers with

$$c_{i,0} + \cdots + c_{i,2i} \leq 4^i i! .$$

Proof The Q's satisfy

$$Q_r(a) = Q_r(a-1) + aQ_{r-1}(a-1),$$

and a similar analysis yields the lemma.

Corollary 2.10 *In equation 2.4, we have*

$$p_i(\Gamma) = \sum_{|I| \leq i} c_{i,I} \binom{|V_\Gamma| - 1}{I_0} \binom{a_1(\Gamma) - 1}{I_1} \cdots \binom{a_d(\Gamma) - 1}{I_d} , \tag{2.6}$$

where we have used "tensor" notation with $I = (I_0, \ldots, I_d)$ (i.e. I_k's are non-negative integers, and $|I| = I_0 + \cdots + I_d$), with constants $c_{i,I}$ each of absolute value $\leq 8^i i!$.

Our goal is to estimate the terms of the sum in the right-hand-side of equation 2.6, when summed over all (Γ, w) pairs of a given type. This is easy once a lettering of the type is fixed.

For a fixed $\sigma, \tau \in \Pi$, let $\mathrm{Irred}_{k,\sigma,\tau}$ denote the set of irreducible words of length k beginning with σ and ending with τ.

Lemma 2.11 *For any $\sigma, \tau \in \Pi$ and non-negative integers $s_1, \ldots, s_d$ we have*

$$f(k) \equiv \sum_{w \in Irred_{k,\sigma,\tau}} \binom{a_1(w)}{s_1} \cdots \binom{a_d(w)}{s_d} = (2d-1)^k P(k) + (-1)^k Q(k) + R(k)$$

where P, Q, R are polynomials of degree $s = s_1 + \cdots + s_d$ with coefficients bounded by $(cd)^{cs}$.

Proof Consider the $2d \times 2d$ matrix

$$M = \begin{pmatrix} x_1 & 0 & x_2 & x_2 & \cdots & x_d & x_d \\ 0 & x_1 & x_2 & x_2 & \cdots & x_d & x_d \\ x_1 & x_1 & x_2 & 0 & \cdots & x_d & x_d \\ x_1 & x_1 & 0 & x_2 & \cdots & x_d & x_d \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ x_1 & x_1 & x_2 & x_2 & \cdots & x_d & 0 \\ x_1 & x_1 & x_2 & x_2 & \cdots & 0 & x_d \end{pmatrix}$$

i.e. the matrix M whoose coefficients are

$$\begin{array}{ll} M_{2i-1,2j-1} = M_{2i,2j} = x_j & \\ M_{2i,2j-1} = M_{2i-1,2j} = x_j & \text{if } i \neq j \\ M_{2i,2j-1} = M_{2i-1,2j} = 0 & \text{if } i = j \end{array}$$

over all $0 \leq i \leq d$ and $0 \leq j \leq d$. It is easy to see that

$$(M^k)_{i,j} = \sum_{w \in \mathrm{Irred}_{k,\sigma_i,\sigma_j}} x_1^{a_1(w)} \ldots x_d^{a_d(w)}$$

where σ_{2i-1} is π_i and σ_{2i} is π_i^{-1}. Hence

$$f(k) = \frac{1}{s_1! \ldots s_d!} \left(\left(\frac{\partial}{\partial x_1}\right)^{s_1} \cdots \left(\frac{\partial}{\partial x_d}\right)^{s_d} M^k \Big|_{x_1 = \cdots = x_d = 1} \right)_{a,b}$$

for the a, b with $\sigma_a = \sigma$, $\sigma_b = \tau$. Let $C_i = \frac{\partial}{\partial x_i} M$, which is a matrix with $4d-2$ 1's and the rest 0's. All second derivatives of M vanish, and so

$$\frac{1}{s_1! \ldots s_d!} \left(\frac{\partial}{\partial x_1}\right)^{s_1} \cdots \left(\frac{\partial}{\partial x_d}\right)^{s_d} M^k = \sum_{(i_1,\ldots,i_s)} \sum_{j_0 + \cdots + j_s = k-s} M^{j_0} C_{i_1} M^{j_1} C_{i_2} \ldots C_{i_s} M^{j_s}$$

where the first summation is over all tuples $(i_1, \ldots, i_s)$ which contain s_1 1's, s_2 2's, etc. For fixed $(i_1, \ldots, i_s)$ we have

$$\sum_{j_0+\cdots+j_s=k-s} M^{j_0} C_{i_1} \ldots M^{j_s}|_{x_1=\cdots=x_d=1} = \sum_{j_0+\cdots+j_s=k-s} C^{j_0} C_{i_1} \ldots C^{j_s}$$

where

$$C = M|_{x_1=\cdots=x_d=1} = \begin{pmatrix} 1 & 0 & 1 & 1 & \cdots & 1 & 1 \\ 0 & 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 1 & 1 & 0 & \cdots & 1 & 1 \\ 1 & 1 & 0 & 1 & \cdots & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & 1 & 1 & \cdots & 1 & 0 \\ 1 & 1 & 1 & 1 & \cdots & 0 & 1 \end{pmatrix}.$$

We claim that C has eigenvalues $2d-1$, -1 with multplicity $d-1$, and 1 with multiplicity d. To see this, note that the map $T\colon \mathbf{R}^{2d} \to \mathbf{R}^{2d}$ given by

$$T(y_1, \ldots, y_{2d}) = (y_1 + y_2, y_1 - y_2, y_3 + y_4, y_3 - y_4, \ldots, y_{2d-1} - y_{2d})$$

gives

$$TCT^{-1} = \begin{pmatrix} 1 & 0 & 2 & 0 & \cdots & 2 & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 2 & 0 & 1 & 0 & \cdots & 2 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 2 & 0 & 2 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

By permuting the basis vectors this becomes the block matrix with $d \times d$ blocks

$$\begin{pmatrix} B & 0 \\ 0 & I \end{pmatrix}, \tag{2.7}$$

where I is the identity matrix, and

$$B = \begin{pmatrix} 1 & 2 & 2 & \cdots & 2 \\ 2 & 1 & 2 & \cdots & 2 \\ 2 & 2 & 1 & \cdots & 2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 2 & 2 & 2 & \cdots & 1 \end{pmatrix}.$$

Since $B+I$ is the $d \times d$ matrix of all 2's, whose eigenvalues are $2d$ (simple) and 0 with multiplicity $d-1$, it follows that B's eigenvalues are $2d-1$ and -1 with multiplicity $d-1$. Thus C's eigenvalues, which by equation 2.7 are the union of those of B and those of I, are $2d-1$ (simple), -1 with multiplicity $d-1$, and 1 with multiplicity d.

Thus

$$\sum_{j_0+\cdots+j_s=k-s} C^{j_0} C_{i_1} \ldots C^{j_s} = \sum_j AD^{j_0} A^{-1} C_{i_1} AD^{j_1} A^{-1} \ldots C_{i_s} AD^{j_s} A^{-1}$$

where D is the diagonal matrix with diagonal consisting of one $2d-1$, $d-1$ -1's, and d 1's, and A is an orthogonal matrix diagonalizing C. We can write the sum on the right-hand-side of the above equation as

$$\sum_j AD^{j_0} E_1 D^{j_1} \ldots E_s D_{j_s} A^{-1} \tag{2.8}$$

with

$$E_j = A^{-1} C_{i_j} A \ .$$

Since

$$\|E_j\|_2 = \|A\|_2 \, \|C_{i_j}\|_2 \, \|A^{-1}\|_2 = \|C_{i_j}\|_2 \leq \sqrt{4d-2} \ ,$$

(where $\| \ \|_2$ denotes the L_2 operator norm) we have that each of E_j's entries are $\leq \sqrt{4d-2}$. Clearly each of A and A^{-1}'s entries are ≤ 1 in absolute value. Therefore, the a,b-th entry of the matrix in equation 2.8,

$$\sum_j \sum_{b_0,\ldots,b_s \in [1,\ldots,2d]} (A)_{a,b_1} (D^{j_0})_{b_0,b_0} (E_1)_{b_0,b_1} (D^{j_1})_{b_1,b_1} \cdots (E_s)_{b_{s-1},b_s} (D_{j_s})_{b_s,b_s} (A^{-1})_{b_s,b} \ ,$$

is just

$$\sum_j \sum_{\substack{\epsilon_0,\ldots,\epsilon_s \\ \epsilon_l \in \{2d-1,-1,1\}}} c_{\epsilon_0,\ldots,\epsilon_s} \epsilon_0^{j_0} \ldots \epsilon_s^{j_s} \tag{2.9}$$

where the $c_{\epsilon_0,\ldots,\epsilon_s}$'s are constants. Fixing $\epsilon_0,\ldots,\epsilon_s$ and letting α,β,γ be the number of respective occurences of $2d-1$, -1, 1 among the ϵ's, we see that

$$|c_{\epsilon_0,\ldots,\epsilon_s}| \leq (4d-2)^{s/2} (d-1)^\beta d^\gamma \leq 2^s d^{3s/2} \ . \tag{2.10}$$

Also,

$$\sum_j \epsilon_0^{j_0} \ldots \epsilon_s^{j_s} = \sum_{u+v+w=k-s} (2d-1)^u (-1)^v \binom{u+\alpha-1}{\alpha-1} \binom{v+\beta-1}{\beta-1} \binom{w+\gamma-1}{\gamma-1} \ . \tag{2.11}$$

From

$$\sum_{u=0}^{k} R^u = \frac{R^{k+1}-1}{R-1}$$

we derive

$$\sum_{\substack{u+v\leq k \\ u,v\geq 0}} R^u S^v = \left(\frac{R^{k+1}-1}{R-1}\right) + \left(\frac{R^k-1}{R-1}\right) S + \cdots + \left(\frac{R^1-1}{R-1}\right) S^k$$

$$= \frac{R^{k+1}}{R-1}\left(1+\frac{S}{R}+\cdots+\left(\frac{S}{R}\right)^k\right)-\frac{1}{R-1}(1+S+\cdots+S^k)$$
$$= \frac{S^{k+2}(R-1)+R^{k+2}(1-S)+(S-R)}{(R-1)(S-1)(S-R)}$$

and thus

$$\sum_{\substack{u+v+w=k\\ u,v,w\geq 0}} R^u S^v P^w = \frac{S^{k+2}(R-P)+R^{k+2}(P-S)+P^{k+2}(S-R)}{(R-P)(S-P)(S-R)} .$$

Multiplying the above by $R^{\alpha-1}S^{\beta-1}P^{\gamma-1}$, then differentiating by $(\frac{\partial}{\partial R})^{\alpha-1}(\frac{\partial}{\partial S})^{\beta-1}(\frac{\partial}{\partial P})^{\gamma-1}$, the substituing $k-s$ for k yields

$$\sum_{u+v+w=k-s}(2d-1)^u(-1)^v\binom{u+\alpha-1}{\alpha-1}\binom{v+\beta-1}{\beta-1}\binom{w+\gamma-1}{\gamma-1} \tag{2.12}$$

is equal to a sum of six terms, the first of which is

$$\frac{1}{(\alpha-1)!}\frac{1}{(\beta-1)!}\frac{1}{(\gamma-1)!}\left(\frac{\partial}{\partial R}\right)^{\alpha-1}\left(\frac{\partial}{\partial S}\right)^{\beta-1}\left(\frac{\partial}{\partial P}\right)^{\gamma-1}\left[\frac{R^\alpha S^{k-s+\beta+1}P^{\gamma-1}}{(R-P)(S-P)(S-R)}\right]$$

evaluated at $R=2d-1$, $S=-1$, $P=1$; the other five terms are similar. In the above term, the $\frac{\partial}{\partial P}$'s can be applied to any of the three terms $P^{\gamma-1}$ (in the numerator) and $(R-P)$ and $(S-P)$ (in the denominator), and similarly for the other partials. We get a sum of $3^{\alpha+\beta+\gamma-3}$ terms, each of the form

$$\binom{k-s+\beta+1}{l}S^k\theta ,$$

where l is an integer between 0 and s and θ is a polynomial in R,S,P, independent of k, which satisfies

$$|\theta(R,S,P)|\Big|_{R=2d-1,S=-1,P=1}\leq (2d-1)^s .$$

The coefficients of the polynomial

$$\binom{k-s+\beta+1}{l}$$

are clearly dominated by those of

$$\frac{(k+s)^l}{l!} ,$$

which are clearly less than

$$\max_{j=0,\ldots,l} \frac{s^j\binom{l}{j}}{l!} \leq \max_{j=0,\ldots,l} \frac{s^j}{j!} \leq e^s.$$

After doing a similar analysis for the other five terms whose sum equals the expression in equation 2.12, it follows that the expression in equation 2.11 is of the form

$$(2d-1)^k p(k) + (-1)^k q(k) + r(k)$$

with p, q, r polynomials of degree s whose coefficients are bounded by

$$3^{s-2} e^s (2d-1)^s \ .$$

Summing over the 3^{s+1} possibilities for $\epsilon_0, \ldots, \epsilon_s$ in equation 2.9, using also equation 2.10, yields

$$f(k) \equiv \sum_{w \in \mathrm{Irred}_{k,\sigma,\tau}} \binom{a_1(w)}{s_1} \ldots \binom{a_d(w)}{s_d} = (2d-1)^k P(k) + (-1)^k Q(k) + R(k)$$

where P, Q, R are polynomials in k of degree s whose coefficients are bounded by

$$2^s d^{3s/2} 3^{s+1} 3^{s-2} e^s (2d-1)^s < (36e)^s d^{5s/2} \ .$$

□

Corollary 2.12 *For any $\sigma, \tau \in \Pi$ and non-negative integers $s_1, \ldots, s_d$ we have*

$$f(k) \equiv \sum_{w \in \mathit{Irred}_{k,\sigma,\tau}} (a_1(w))^{s_1} \ldots (a_d(w))^{s_d} = (2d-1)^k P(k) + (-1)^k Q(k) + R(k)$$

where P, Q, R are polynomials of degree $s = s_1 + \cdots + s_d$ with coefficients bounded by $(cds)^{cs}$ for some absolute constant c.

Proof We use the fact that the constants c_m in

$$x^n = \sum_{m=0}^{n} c_m \binom{x}{m}$$

are always positive (see, for example, [Knu73] page 65); it follows that $c_m \leq n^n$ by substituting n for x in the above. Applying this to each $a_i(w)^{s_i}$, expanding the product, and applying the previous lemma yields the desired result.

□

For a type $T = (V_T, E_T)$ with $E_T = \{e_1, \ldots, e_t\}$, let $\mathcal{L}_{T,k_1,\ldots,k_t}$ denote the set of oriented labelling with weights $(k_1, \ldots, k_t)$. For a labelling, ℓ, we define $a_i(\ell)$ to be the number of occurences of π_i and π_i^{-1} in ℓ. Any generalized form, Γ, compatible with T and a labelling in $\mathcal{L}_{T,k_1,\ldots,k_t}$ has

$$|V_\Gamma| = |V_T| + \sum (k_j - 1) = |V_T| + |\ell| - t = |\ell| + 1 - r \ ,$$

by lemma 2.1, where $|\ell| = \sum a_i(\ell)$. We can therefore define the p_i of equation 2.4 as $p_i(T, \ell)$, for T and ℓ determine the value of $p_i(\Gamma)$ for any compatible Γ.

Lemma 2.13 *For any fixed type, T, of coincidence $\leq r$, and $k_1, \ldots, k_t$, let $\mathcal{L} = \mathcal{L}_{T,k_1,\ldots,k_t}$. For any $i \leq r + 1$ we have*

$$\sum_{\ell \in \mathcal{L}} p_i(T, \ell) = \sum_{K_1,K_2,K_3} (2d-1)^{|K_1|}(-1)^{|K_2|} P_{K_1,K_2,K_3}(k_1, \ldots, k_t) \ , \tag{2.13}$$

where the right-hand-side sum is over all partitions of $K = \{k_1, \ldots, k_t\}$ into three disjoint sets K_1, K_2, K_3, where

$$|K_j| = \sum_{k_s \in K_j} k_s \ ,$$

and where P_{K_1,K_2,K_3} are polynomials of degree $\leq 2i$ whose coefficients are bounded by $(cdr)^{cr}$ for some absolute constant c.

Proof By equation 2.6 we have that the sum on the left-hand-side of equation 2.13 is

$$\sum_{|I| \leq i} c_I \binom{k_1 + \cdots + k_t - r}{I_0} \binom{a_1(\ell) - 1}{I_1} \cdots \binom{a_d(\ell) - 1}{I_d} \tag{2.14}$$

with $|c_I| \leq 8^i i!$. Let $a_{jl}(\ell)$ denote the number of π_j, π_j^{-1} occurences in e_l. Since

$$a_j(\ell) = a_{j1}(\ell) + \ldots + a_{jt}(\ell) \ ,$$

we can expand equation 2.14 as a sum of $(t+1)^{I_0} 2^{I_1} \ldots 2^{I_d}$ terms of the form

$$c_s \left(\prod_{l=1}^{t} k_l^{s_l} \right) \left(\prod_{j=1}^{d} \prod_{l=1}^{t} (a_{jl}(\ell))^{s_{jl}} \right)$$

with coefficients c_s bounded by $(cdr)^{cr}$ and with

$$\sum_l s_l + \sum_{j,l} s_{j,l} \leq i.$$

Fix a lettering, and let $\overline{\mathcal{L}}$ denote those $\ell \in \mathcal{L}$ of that lettering. We can write

$$\sum_{\ell \in \overline{\mathcal{L}}} \prod_{j,l} (a_{jl}(\ell))^{s_{jl}} = \prod_{l=1}^{t} \sum_{w \in \mathrm{Irred}_{k_l,\sigma_l,\tau_l}} \prod_{j=1}^{d} (a_{jl}(w))^{s_{jl}} \ ,$$

with $\sigma_1, \ldots, \sigma_t, \tau_1, \ldots, \tau_t$ given by the lettering. By corollary 2.12, the above is

$$= \prod_{l=1}^{t} [(2d-1)^{k_l} P_l(k_l) + (-1)^{k_l} Q_l(k_l) + R_l(k_l)] ,$$

with P_l, Q_l, R_l polynomials of degree $\leq 2\sum_l s_{j,l}$ and coefficients bounded by $(cdr)^{cr}$ for some absolute constant c. Hence the above is

$$\sum_{\substack{K_1,K_2,K_3 \\ \text{partition } \{k_1,\ldots,k_t\}}} (2d-1)^{|K_1|}(-1)^{|K_2|} Q_{K_1,K_2,K_3}(k_1,\ldots,k_t) ,$$

for polynomials Q_{K_1,K_2,K_3} of degree $\leq 2\sum_{j,l} s_{j,l}$ with coefficients bounded by $(cdr)^{cr}$ for some absolute constant c. Summing the above over the $\leq (2d)^{6r-2}$ letterings yields the lemma.

□

Our goal is to sum $p_i(T,\ell)$ over all triples (T,y,ℓ) corresponding to pairs (Γ,w) with $w \in \mathrm{Irred}_k$. So far we can estimate

$$\sum_{\ell \in \mathcal{L}} p_i(T,\ell)$$

with $\mathcal{L} = \mathcal{L}_{T,k_1,\ldots,k_t}$. We start summing over walks. So fix multiplicitites $m_1,\ldots,m_t$, and let $k = k_1 m_1 + \cdots + k_t m_t$. Clearly k is the length of w in any (Γ,w) corresponding to a (T,y,ℓ) with multiplicity $m_1,\ldots,m_t$ and weights $k_1,\ldots k_t$.

Lemma 2.14 *For fixed type, T, with t edges and of coincidence $\leq r$, fixed multiplicities m_l, and $i \leq r+1$ we have for any $k \geq m = m_1 + \cdots + m_t$*

$$\sum_{\substack{k_1,\ldots,k_t \geq 1 \text{ with} \\ k_1 m_1 + \cdots + k_t m_t = k}} \sum_{\ell \in \mathcal{L}_{T,k_1,\ldots,k_t}} p_i(T,\ell) = (2d-1)^{k+t-m} P_i(k) + \epsilon$$

where ϵ vanishes if $m_1 = \ldots = m_t = 1$ and otherwise

$$|\epsilon| \leq (2d-1)^{(k-m)/2} k^{t+2i} (crd+m)^{cr^2} ,$$

and P_i a polynomial of degree $t+2i$ with coefficient bounded by $(crd+m)^{cr^2}$ for some absolute constant c.

Proof Applying lemma 2.13 and exchanging summations yields

$$\sum_{k_1,\ldots,k_t} \sum_{\ell} p_i(T,\ell) = \sum_{K_1,K_2,K_3} \sum_{\substack{k_1,\ldots,k_t \geq 1 \text{ with} \\ k_1 m_1 + \cdots + k_t m_t = k}} (2d-1)^{|K_1|}(-1)^{|K_2|} P_{K_1,K_2,K_3}(k_1,\ldots,k_t) .$$

Fix a partition, K_1, K_2, K_3. We shall need some sublemmas.

Sublemma 2.15 *For any integer r, integer $t \leq 3r-1$, and polynomial P of degree $\leq r+t$, we have for any $k \geq t$*

$$\sum_{\substack{k_1+k_2+\cdots+k_t=k \\ k_i \geq 1}} P(k_1,\ldots,k_t) = Q(k)$$

with Q of degree $\deg(P)+t-1$ with

$$|Q| \leq |P|(cdr)^{crt} ,$$

where $|\;|$ for a polynomial denotes the largest absolute value among its coefficients.

Proof We proceed by induction on t. For $t=1$ there is nothing to prove. We claim that if the sublemma holds for t replaced by $t-1$, then it also holds for t. To see this, we write

$$P(k_1,\ldots,k_t) = \sum_{l=0}^{\deg(P)} \overline{P}_l(k_1,\ldots,k_{t-1})k_t^l,$$

with $\deg(\overline{P}_l) \leq \deg(P)-l$, and $|\overline{P}_l| \leq |P|$. Applying the sublemma to each $\overline{P}_l$ we get

$$\sum_{\substack{k_1+k_2+\cdots+k_t=k \\ k_i \geq 1}} P(k_1,\ldots,k_t) = \sum_{l=0}^{\deg(P)} \sum_{j=1}^{k-t+1} Q_l(k-j)j^l ,$$

with Q_l of degree $\leq \deg(P)-l+(t-2)$ and with coefficients bounded by $|P|(cdr)^{cr(t-1)}$. For each fixed l, the sum over j on the left-hand-side of the above is clearly a polynomial of degree $1+l+\deg(Q_l) \leq \deg(P)+t-1$ and coefficients bounded by $|P|(cdr)^{crt}$, assuming c is sufficiently large. Summing over l, using $\deg(P) \leq r+t \leq 4r-1$ proves the inductive step.

□

Sublemma 2.16 *For any $m_1=1, m_2, \ldots, m_t$, partition K_1, K_2, K_3 of $\{k_1,\ldots,k_t\}$ with $k_1 \in K_1$, $t \leq 3r-1$, and polynomial P of degree $\leq r+t$, we have for any $k \geq m = 1+m_2+\cdots+m_t$*

$$\sum_{\substack{k_1m_1+k_2m_2+\cdots+k_tm_t=k \\ k_i \geq 1}} (2d-1)^{|K_1|}(-1)^{|K_2|}P(k_1,\ldots,k_t) = (2d-1)^{k+t-m}Q(k)+\epsilon \tag{2.15}$$

where ϵ vanishes if $m_i=1$ and $k_i \in K_1$ for all i, and otherwise

$$|\epsilon| \leq (2d-1)^{(k-m)/2}k^{\deg(P)+t-1}|P|(cdr+m)^{crt} , \tag{2.16}$$

and with Q of degree $\deg(P)+t-1$ with

$$|Q| \leq |P|(cdr+m)^{crt} . \tag{2.17}$$

Proof We can assume $k_1, \dots, k_s$ are all in K_1, that $m_1 = \dots = m_s = 1$, and that $K_1 = \{k_1, \dots, k_u\}$ with $m_{s+1}, \dots, m_u$ all greater than 1. Writing

$$P(k_1, \dots, k_t) = \sum_{l=0}^{\deg(P)} \overline{P}_l(k_1, \dots, k_s)\overline{R}_l(k_{s+1}, \dots, k_t)$$

with $\overline{R}_l$ of degree l, by sublemma 2.15 we can write the sum in equation 2.15 as

$$\sum_{l=0}^{\deg(P)} \sum_{j=m-s}^{k-s} \sum_{\substack{k_{s+1}m_{s+1}+\dots+k_t m_t = j \\ k_l \geq 1}} (2d-1)^{k_{s+1}+\dots+k_u}(-1)^{|K_2|}(2d-1)^{k-j} Q_l(k-j) R_l(k_{s+1}, \dots, k_t) \, .$$

For each fixed l we write the above sum over j as

$$\sum_{j=m-s}^{k-s} \Big(\Big) = \sum_{j=m-s}^{\infty} \Big(\Big) - \sum_{k-s+1}^{\infty} \Big(\Big)$$

where ( ) denotes the summand (for j) in the preceding equation. Note that the number of solutions to $k_{s+1}m_{s+1} + \dots + k_t m_t = j$ is clearly $\leq \binom{j+t-s-1}{t-s-1}$, and that

$$\begin{aligned}
k_{s+1} + \dots + k_u &\leq \frac{1}{2}(2k_{s+1} + \dots + 2k_u) \\
&\leq \frac{1}{2}(j - m_{u+1} - \dots - m_t - (m_{s+1} - 2)k_{s+1} - \dots - (m_u - 2)k_u) \\
&\leq \frac{1}{2}(j - m_{u+1} - \dots - m_t - (m_{s+1} - 2) - \dots - (m_u - 2)) \\
&\leq \frac{1}{2}(j - m + 2(u - s) + s) \\
&\leq \frac{1}{2}(j - m - s) + t \, .
\end{aligned}$$

It follows from the identity

$$\sum_{n=0}^{\infty} \binom{n}{\tau} \rho^n = \frac{\rho^{\tau+1}}{(1-\rho)^{\tau}}$$

that

$$\sum_{j=m-s}^{\infty} \Big(\Big) = (2d-1)^{k-m} \widetilde{Q}_l(k)$$

where $\widetilde{Q}$ is a polynomial of degree $\leq \deg(Q_l)$, and $|\widetilde{Q}_l| \leq |Q_l||R_l|(crd+m)^{crt} \leq |P|(crd+m)^{crt}$, and that

$$\left| \sum_{k-s+1}^{\infty} \Big(\Big) \right| \leq (2d-1)^{(k-m)/2} k^{\deg(P)+t-s-1} |P|(cdr+m)^{crt}.$$

Summing over l completes the proof.

□

Sublemma 2.17 *For any* $m_1, m_2, \ldots, m_t$*, partition* K_1, K_2, K_3 *of* $\{k_1, \ldots, k_t\}$ *with* $k_1 \in K_1$*,* $t \le 3r-1$*, and polynomial* P *of degree* $\le r+t$*, we have for any* $k \ge m = m_1 + m_2 + \cdots + m_t$

$$\sum_{\substack{k_1m_1+k_2m_2+\cdots+k_tm_t=k \\ k_i \ge 1}} (2d-1)^{|K_1|}(-1)^{|K_2|}P(k_1, \ldots, k_t) = \epsilon$$

where

$$|\epsilon| \le (2d-1)^{(k-m)/\mu} k^{\deg(P)+t-1}|P|(cdr+m)^{crt} ,$$

where μ *is the smallest integer among the* m_i *with* $k_i \in K_1$*.*

Proof We do the same calculation as in the last sublemma, the difference being that this time we have

$$k_1 + \cdots + k_u \le \frac{1}{\mu}(\mu k_1 + \cdots + \mu k_u) \le \frac{1}{\mu}(j-m) + u$$

and therefore get the $1/\mu$ factor in the exponent of $(2d-1)$.

□

To finish the proof of lemma 2.14, we sum over the 3^t partitions of $\{k_1, \ldots, k_t\}$ into three sets, K_1, K_2, K_3, applying one of the sublemmas (noting that $rt \le 3r^2$).

□

Let

$$P_{i,T,\vec{m}}(k)$$

denote the polynomial P_i in lemma 2.14, which depends (only) on T and $m_1, \ldots, m_t$, where we use $\vec{m}$ to abbreviate the sequence $(m_1, \ldots, m_t)$. For an ordered type T and multiplicities $\vec{m}$, let $W(T; \vec{m})$ be the number of legal walks in T with multiplicities $\vec{m} = (m_1, \ldots, m_t)$. By lemma 2.6,

$$\sum_{|\vec{m}|=m} W(T; \vec{m}) \le (2r)^m$$

for types of coincidence $\le r$, where $|\vec{m}| = m_1 + \ldots + m_t$. It follows from this and lemma 2.14 that for any type of coincidence $\le d-1$ the infinite sum

$$\sum_{m_1=1}^{\infty} \cdots \sum_{m_t=1}^{\infty} W(T; \vec{m})(2d-1)^{t-m} P_{i,T,\vec{m}}(k)$$

converges for each i to a polynomial of degree $2i+t$, which we denote $f_{i,T}$. For $0 \le i \le d-2$, let

$$f_i(k) = \sum_{\mathrm{coin}(T) \le i+1} f_{i+1-\mathrm{coin}(T)\,,\,T}(k) ,$$

which is a polynomial in k of degree $\le 2i + (3i+2) = 5i+2$, since $t \le 3i+2$ for any T of coincidence $\le i+1$.

Theorem 2.18 *Let d and $r \le \sqrt{2d-1}/2$ be fixed positive integers. For any fixed $v \in \{1,\ldots,n\}$ we have*

$$E\left\{\sum_{w\in Irred_k} v \xrightarrow{w} v\right\} = 2d(2d-1)^{k-1}\left(\frac{1}{n}f_0(k) + \frac{1}{n^2}f_1(k) + \ldots + \frac{1}{n^r}f_{r-1}(k) + error\right),$$

where

$$error \le \frac{k^{2r+2}}{n^{r+1}}(1+2k^{2r}e^{(r+1)k/n}) + (2d-1)^{-k/2}\sum_{i=0}^{r-1}\frac{(ckid)^{ci^2}}{n^{i+1}}$$

and where the f_i are polynomials of degree $\le 5i+2$ whose coefficients are bounded by $(cdr)^{cr^2}$, where c is an absolute constant (independent of d and r). For any

$$k \le \frac{1}{cd^{3/2}}n^{1/(c\sqrt{d})}$$

we have

$$error \le \frac{k^{4r+2}}{n^{r+1}}c + (2d-1)^{-k/2}\frac{(ckd)^c}{n}.$$

Proof By virtue of the one-to-one correspondence between pairs (Γ, w) and triples (T, y, ℓ) we have

$$\mathrm{E}\left\{\sum_{w\in \mathrm{Irred}_k} i \xrightarrow{w} i\right\} = \frac{1}{n}\sum_T \sum_{m_1,\ldots,m_t}\left[W(T;\vec{m}) \sum_{\substack{k_1,\ldots,k_t \ \ge 1 \text{ with} \\ k_1m_1+\cdots+k_tm_t=k}} \sum_{\ell\in\mathcal{L}_{T,k_1,\ldots,k_t}} \Pr(T,\ell)\right],$$

where $\Pr()$, as with p_i, extends from a function on general forms, Γ, to a function on pairs (T,ℓ). The probability that the walk along v starting at i has more than r coincidneces is no more than the number of ways of choosing $r+1$ points of the k unknown points times the probability that each of these points is indeed a coincidence. Thus we can replace the summation over all types, T, above, with a summation over all types T of coincidence $\le r$ while incurring an error of no more than

$$2d(2d-1)^{k-1}\binom{k}{r+1}\left(\frac{k}{n-k}\right)^{r+1} \le 2d(2d-1)^{k-1}\frac{k^{2r+2}}{n^{r+1}}$$

if $k \le n/2$. Obviously the second summation has non-zero terms only when $m = m_1 + \cdots + m_t$ is $\le k$; we shall restrict the second sum to being over such m.

The total number of $\Pr(T,\ell)$'s occuring in the above quadruple summation is the same as the total number of (Γ, w) pairs (and less when we restrict ourselves to types of coincidence $\le r-1$). For each $w \in \mathrm{Irred}_k$ there are no more than

$$\sum_{j=0}^{r}\binom{k}{j}k^j \le 2k^{2r}$$

compatible Γ's of coincidence $\leq r$, because Γ there is at most one Γ for any $j \leq r$ specified points of coincidence (among the k unknown points $t_1, \ldots, t_k$ of the walk) once we specify at which of the previous vertices each of the j coincidences arrive. By lemma 2.7 we have for each T of coincidence $\leq r$

$$\Pr(T,\ell) \;=\; \frac{p_0}{n^{\mathrm{coin}(T)-1}} + \frac{p_1}{n^{\mathrm{coin}(T)}} + \cdots + \frac{p_{r-1}}{n^{r-1}} + \frac{\epsilon}{n^r} \; ,$$

where ϵ is bounded by

$$e^{(r+1)k/n} k^{2r+2}.$$

It follows that

$$\frac{1}{n} \sum_{\mathrm{coin}(T)\leq r} \sum_{m\leq k} W(T;\vec{m}) \sum_{\vec{k},\ell} \Pr(T,\ell) = \frac{1}{n}\Big(\tilde{f}_0(k) + \frac{\tilde{f}_1(k)}{n} + \ldots + \frac{\tilde{f}_{r-1}(k)}{n^{r-1}} + \frac{\delta}{n^r}\Big) \tag{2.18}$$

(we have abbreviated the inner two summations by $\sum_{\vec{k},\ell}$) where

$$\delta \leq k^{2r+2}(1 + 2k^{2r} e^{(r+1)k/n}) \; , \tag{2.19}$$

and

$$\tilde{f}_i(k) = \sum_{\mathrm{coin}(T)\leq i+1} \sum_{|\vec{m}|\leq k} W(T;\vec{m}) \sum_{\vec{k},\ell} p_{i+1-\mathrm{coin}(T)}(T,\ell) \; .$$

We claim that $\tilde{f}_i(k)$ do not differ much from the infinite sum

$$f_i(k) = \sum_{\mathrm{coin}(T)\leq i+1} \sum_{\vec{m}} W(T;\vec{m})(2d-1)^{t-m} P_{i+1-\mathrm{coin}(T),T,\vec{m}}(k) \; . \tag{2.20}$$

By lemma 2.14 we have for any type of coincidence $\leq i+1$, for $i \leq r-1$,

$$\left|(2d-1)^{t-m} P_{i+1-\mathrm{coin}(T),T,\vec{m}}(k) - \sum_{\vec{k},\ell} p_{i+1-\mathrm{coin}(T)}(T,\ell)\right| \leq (2d-1)^{(k-m)/2} k^{(3(i+1)-2)+2(i+1)} (crd+m)^{ci^2} \; .$$

Since there are less than $(2i+2)^{6i+5}$ types of coincidnece $\leq i+1$, and since for each such type there are $\leq (2i+2)^m \leq (2d-1)^{m/2}$ legal walks with $|\vec{m}| = m$, we have

$$\sum_{\mathrm{coin}(T)\leq i+1} \sum_{m\leq k} W(T,\vec{m}) \left|(2d-1)^{t-m} P_{i+1-\mathrm{coin}(T),T,\vec{m}}(k) - \sum_{k_l} \sum_{\ell} p_i(T,\ell)\right| \leq (2d-1)^{k/2}(ckid)^{ci^2}. \tag{2.21}$$

It remains to estimate

$$\left| f_i(k) - \sum_{\mathrm{coin}(T)\leq i+1} \sum_{|\vec{m}|\leq k} W(T,\vec{m})(2d-1)^{t-m} P_{i+1-\mathrm{coin}(T),T}(k) \right| \leq$$
$$\sum_{\mathrm{coin}(T)\leq i+1} \sum_{|\vec{m}|>k} W(T,|\vec{m}|)(2d-1)^{t-m} \left| P_{i+1-\mathrm{coin}(T),T}(k) \right| \; . \tag{2.22}$$

By lemma 2.14, each $P_{i,T,\vec{m}}$ has degree $\le t+2i$, and coefficients bounded by $(cid+m)^{ci^2}$. Thus we have

$$\left|P_{i+1-\mathrm{coin}(T),T,\vec{m}}(k)\right| \le k^{t+2i}(cid+m)^{ci^2}$$

and thus for a type, T, of coincidence $\le i+1$ we have that the expression in equation 2.22 is bounded by

$$k^{5i+2}(2d-1)^{3i+2}\sum_{m=k+1}^{\infty}\left(\frac{2i+2}{2d-1}\right)^m (cid+m)^{ci^2} \le (ckid)^{ci^2}\left(\frac{2i+2}{2d-1}\right)^k$$

(for $i \le d-2$). Combining the above with equations 2.18, 2.19, and 2.21 yields the first estimate on the error term. For

$$k \le \frac{1}{c' d^{3/2}} n^{1/(c'\sqrt{d})} ,$$

assuming $i \le \sqrt{2d-1}/2$, the sum

$$\sum_{i=0}^{r-1}\frac{(ckid)^{ci^2}}{n^{i+1}}$$

is clearly bounded by the first term. Finally, f_i's coefficients are bounded by

$$(2i+2)^{6i+5}\sum_{m=1}^{\infty}(cid+m)^{ci^2}\left(\frac{2i+2}{2d-1}\right)^m \le (c'id)^{c'i^2}$$

(for $i \le d-2$).

□

3 Consequences of the Expansion

We start by computing $f_0, f_1, \ldots$ in theorem 2.18. Without too much work, one can use methods akin to those of section 2 to argue that $f_0(k) = 1$ and if $f_1(k)$ is non-zero, then its leading coefficient is negative (which is as good as proving $f_1(k) = 0$ as far as eigenvalue estimates are concerned). Direct arguments for other f_l's seem harder to come by. We will argue by using previously known facts about the second eigenvalue, which will give the desired values for the f_l's. These values for f_l will in turn give much sharper information on the second eigenvalues, as we will show later in this section. We warn the reader that other literature on this subject often work with d regular, not $2d$ regular, graphs; to maintain consistency when quoting results or using techniques from previous work, we will sometimes state theorems in terms of d regular graphs.

Theorem 3.1 *For every $d > 2$ there is a constant $\alpha > 0$ such that for a random graph $G \in \mathcal{G}_{n,2d}$ we have*

$$\begin{aligned} \Pr\{|\lambda_2(G)| \le 2d - \alpha\} &= 1 - \frac{1}{n^{d-1}} + O\Big(\frac{1}{n^{2d-2}}\Big), \\ \Pr\{\lambda_2(G) = 2d\} &= \frac{1}{n^{d-1}} + O\Big(\frac{1}{n^{2d-2}}\Big). \end{aligned}$$

Proof A graph, G, with n vertices is said to be a γ-magnifier if for all subsets of vertices, A, of size $\le n/2$ we have

$$|\Gamma(A) - A| \ge \gamma |A|,$$

where $\Gamma(A)$ denotes those vertices connected to some member of A by an edge. It is known that any d regular γ-magnifier has

$$\lambda_2(G) \le d - \frac{\gamma^2}{4 + 2\gamma^2}$$

(see [Alo86]). We say that G is a γ-expander if for all subsets of vertices, A, of size $\le n/2$ we have

$$|\Gamma(A)| \ge (1+\gamma)|A|.$$

A standard counting argument gives the following.

Lemma 3.2 *$G \in \mathcal{G}_{n,2d}$ is not a γ-expander with probability*

$$\le \sum_{m=1}^{n/2} \binom{n}{m} \binom{n}{\lfloor (1+\gamma)m \rfloor} \left(\frac{m + \lfloor m\gamma \rfloor}{n}\right)^{md},$$

and more precisely

$$\le \frac{1}{n^{d-1}} + O\Big(\frac{1}{n^{2d-2}}\Big) + \sum_{m=3}^{n/2} \binom{n}{m} \binom{n}{\lfloor (1+\gamma)m \rfloor} \left(\frac{m + \lfloor m\gamma \rfloor}{n}\right)^{md},$$

for $\gamma < 1/3$.

Proof For fixed subsets of V, A and B, we have

$$\begin{aligned} \Pr\{\Gamma(A) \subset B\} &\le \Pr\{\pi_i(a) \in B \quad \forall a \in A, \quad \forall i \in \{1, \ldots, d\}\} \\ &\le \left(\frac{|B|}{n} \frac{|B|-1}{n-1} \cdots \frac{|B| - |A| + 1}{n - |A| + 1}\right)^d \le \left(\frac{|B|}{n}\right)^{d|A|}. \end{aligned}$$

If G is not a γ-expander, then there exist some $m \in \{1, \ldots, n/2\}$, $|A| = m$, and $|B| = \lfloor (1+\gamma)m \rfloor$ with $\Gamma(A) \subset B$. This yields the first estimate. To refine this estimate, notice that the bound given for $\Pr\{\Gamma(A) \subset B\}$ is only tight when $A = B$. Indeed, if A and B are disjoint, then for $\Gamma(A)$ to be contained in B we need both π_i *and* π_i^{-1} to map A into B for each π_i, which roughly doubles the amount of conditioning on each π_i. In particular, for $|A| = |B| = 1$, $A \neq B$ we have

$$\Pr\{\Gamma(A) \subset B\} \leq \left(\frac{1}{n(n-1)}\right)^d,$$

and a similar analysis for the case $|A| = 2$ and $|A \cap B| = 1$ or 0, giving the second estimate.

□

Assuming $\gamma < 1/3$ and using the estimate $\binom{n}{m} \leq (ne/m)^m$ we can bound the probability that $G \in \mathcal{G}_{n,2d}$ is not a γ-expander by

$$\frac{1}{n^{d-1}} + O\left(\frac{1}{n^{2d-2}}\right) + \sum_{m=3}^{n/2} \left(\frac{m}{n}\right)^{m(d-2)} \left((1+\gamma)^{2d+1+\gamma} e^{2+\gamma}\right)^m = \frac{1}{n^{d-1}} + O\left(\frac{1}{n^{2d-2}}\right)$$

if γ is sufficiently close to 0 and $d > 2$. If G is a γ-expander, then clearly G^2, the graph whose adjacency matrix is the square of G's, is also a γ-expander, and thus also a γ-magnifier. Therefore

$$(\lambda_2(G))^2 \leq 4d^2 - \frac{\gamma^2}{4 + 2\gamma^2}$$

with probability at least

$$1 - \frac{1}{n^{d-1}} + O\left(\frac{1}{n^{2d-2}}\right).$$

On the other hand, $\lambda_2(G) = d$ if G has an isolated vertex. By the inclusion-exclusion principle, this happens with probability

$$\leq n\left(\frac{1}{n^d}\right)$$

and

$$\geq n\left(\frac{1}{n^d}\right) - \binom{n}{2}\left(\frac{1}{n^d}\right)^2.$$

□

Let $G = (V, E)$ be an undirected d-regular graph. By a *non-backtracking walk* in G we mean a walk that at no point in the walk traverses an edge and then on the next step traverses the same edge in the reverse direction. For vertices v, w, and integer k, let $F_k(v, w)$ be the number of non-backtracking walks of length k from v to w. Let

$$W_G(k) = \sum_{v \in V} F_k(v, v).$$

Theorem 2.18 gives us the expected value of $W_G(k)$ for a random $G \in \mathcal{G}_{n,2d}$. In the spirit of [LPS86] we have:

Lemma 3.3 *Let the eigenvalues of G's adjacency matrix be $\lambda_1, \cdots, \lambda_n$. Then*

$$W_G(k) = \sum_{i=1}^{n} q_k(\lambda_i) \ ,$$

where q_k is a polynomial of degree k. It is given by

$$q_k(2\sqrt{d-1}\cos\theta) = (\sqrt{d-1})^k \left(\frac{2}{d-1} \cos k\theta + \frac{d-2}{d-1} \frac{\sin(k+1)\theta}{\sin\theta} \right) .$$

Proof Regarding F_k as an $n \times n$ matrix, we easily see that $F_0 = I$, $F_1 = A$, where A is the adjacency matrix of G,

$$F_2 = A^2 - dI \ ,$$

and for any $t \geq 3$ we have

$$F_t = AF_{t-1} - (d-1)F_{t-2} \ .$$

It follows that

$$F_k = q_k(A)$$

where q_k is some k degree polynomial independent of A, and that

$$W_G(k) = \text{Trace}\,(F_k) = \text{Trace}\,(q_k(A)) = \sum_{i=1}^{n} q_k(\lambda_i) \ .$$

One can solve for q_k explicitly by noting that for fixed λ, $q_k = q_k(\lambda)$ satisfies the simple recurrence

$$q_k = \lambda q_{k-1} - (d-1)q_{k-2} \ .$$

One therefore has

$$q_k = c_1(r_1)^k + c_2(r_2)^k$$

for some c_1, c_2, where r_1, r_2 are the roots of

$$r^2 - \lambda r + (d-1) = 0 \ .$$

One can directly solve for r_1, r_2, c_1, c_2, but it is easier to substitute $\lambda = 2\sqrt{d-1}\cos\theta$. One gets

$$r_{1,2} = \sqrt{d-1}e^{\pm i\theta} \ ,$$

and then after a few more calculations and simplifications one gets

$$q_k(2\sqrt{d-1}\cos\theta) = (\sqrt{d-1})^k \left(\frac{2}{d-1} \cos k\theta + \frac{d-2}{d-1} \frac{\sin(k+1)\theta}{\sin\theta} \right) .$$

□

The $q_k(\lambda)$'s are quite easy to use. For example, if $\lambda = d$, the first eigenvalue, then clearly

$$q_k(d) = d(d-1)^{k-1} .$$

For $|\lambda| \le 2\sqrt{d-1}$ one has $\lambda = 2\sqrt{d-1}\cos\theta$ for some real θ, and therefore

$$|q_k(\lambda)| \le (\sqrt{d-1})^k (k+2) \tag{3.1}$$

since $|\sin(k+1)\theta/\sin\theta|$ is always $\le k+1$. Otherwise, say for $\lambda > 2\sqrt{d-1}$, one can rewrite the identity in lemma 3.3 as

$$q_k(2\sqrt{d-1}\cosh x) = (\sqrt{d-1})^k \left(\frac{2}{d-1} \cosh kx + \frac{d-2}{d-1} \frac{\sinh(k+1)x}{\sinh x} \right)$$

and solve $\lambda = 2\sqrt{d-1}\cosh x$. For $\lambda < -2\sqrt{d-1}$ one solves for $\lambda = -2\sqrt{d-1}\cosh x$ and proceeds similarly. For $|\lambda|$ near $2\sqrt{d-1}$, $\sin\theta$ or $\sinh x$ is very close to 0, and it is convenient to write

$$\frac{\sin(k+1)\theta}{\sin\theta} = \cos k\theta + \cos\theta\cos(k-1)\theta + \cos^2\theta\cos(k-2)\theta + \cdots + \cos^k\theta$$

for estimating. In particular, since as $x \ge 0$ increases $\lambda = 2\sqrt{d-1}\cosh x$ increases, we have that for all λ with $2\sqrt{d-1} \le \lambda \le d-\alpha$,

$$x = \cosh^{-1}\left(\frac{\lambda}{2\sqrt{d-1}}\right)$$

must be

$$\le \cosh^{-1}\left(\frac{d}{2\sqrt{d-1}}\right) - \epsilon$$

for some positive ϵ. Doing the same for λ negative we therefore get

Lemma 3.4 *For any $\alpha > 0$ there is a $\delta > 0$ such that if $|\lambda| \le d-\alpha$, then*

$$|q_k(\lambda)| \le (k+2)(d-1-\delta)^k .$$

Proof Using

$$\cosh(x+y) = \cosh x \cosh y + \sinh x \sinh y ,$$

it follows that

$$\cosh kx \ge \cosh x \cosh(k-1)x \ge \cosh^2 x \cosh(k-2)x \ge \dots$$

Also, $\cosh kx \le e^{kx} \le (d-1-\delta)^k$ for all x corresponding to $|\lambda| \le d-\alpha$, $|\lambda| \ge 2\sqrt{d-1}$. For $|\lambda| \le 2\sqrt{d-1}$, the lemma follows from equation 3.1

□

Theorem 3.5 *In theorem 2.18, with* $r \le \lfloor\sqrt{2d-1}/2\rfloor$*, we have* $f_0(k) = 1$ *and* $f_j(k) = 0$ *for* $j \ge 1$*, if* $d - 2 > \sqrt{2d-1}/2$ *(i.e.* $d \ge 4$*).*

Proof By theorem 3.1 we have that (remember, the degree is now $2d$)

$$\mathrm{E}\left\{\sum_{w\in\mathrm{Irred}_k} v \xrightarrow{w} v\right\} = \left(1 - \frac{1}{n^{d-1}} + O\Big(\frac{1}{n^{2d-2}}\Big)\right)\Big(2d(2d-1)^{k-1} + O((k+2)(2d-1-\delta)^k)\Big) + \left(\frac{1}{n^{d-1}} + O\Big(\frac{1}{n^{2d-2}}\Big)\right)O(n2d(2d-1)^{k-1}) .$$

Taking k even and $= \lfloor(\log n)^2\rfloor$ or $= \lfloor(\log n)^2\rfloor + 1$, letting $n \to \infty$, and noticing that $(2d-1-\delta)^k/(2d-1)^k$ and $(2d-1)^{-k/2}$ are less than any polynomial in n as $n \to \infty$ yields $f_0(k) = 1$ and $f_j(k) = 0$ for any $j \ge 1$.

□

To restate the results so far:

Corollary 3.6 *For any fixed* v*,* $k \ge 1$*, and* $d - 2 > \sqrt{2d-1}/2$ *(i.e.* $d \ge 4$*) we have*

$$E\left\{\sum_{w\in Irred_k} v \xrightarrow{w} v\right\} = 2d(2d-1)^{k-1}\left(\frac{1}{n} + error_{n,k}\right) ,$$

where

$$error_{n,k} \le (ckd)^c\left(\frac{k^{2\sqrt{2d}}}{n^{1+\lfloor\sqrt{2d-1}/2\rfloor}} + \frac{(2d-1)^{-k/2}}{n}\right) .$$

□

Now we use this corollary to estimate the expected sum of the k-th powers of the eigenvalues. Any word in Π^k can be *reduced* by repeatedly cancelling all consecutive occurences of π, π^{-1} in the word, until we get an irreducible word; this irreducible word is independent of how the reducing was done. Notice that

$$\frac{1}{(2d)^k}\mathrm{E}\left\{\sum_{w\in\Pi^k} v \xrightarrow{w} v\right\} = p_{k,0} + \sum_{s=1}^{k} p_{k,s}\frac{1}{2d(2d-1)^{s-1}}\mathrm{E}\left\{\sum_{w\in\mathrm{Irred}_s} v \xrightarrow{w} v\right\} ,$$

where $p_{k,s}$ is the probability that a random word in Π^k reduces to an irreducible word of size s. Since $\sum_s p_{k,s} = 1$, we have

$$\frac{1}{(2d)^k}\sum_v \mathrm{E}\left\{\sum_{w\in\Pi^k} v \xrightarrow{w} v\right\} = 1 + (n-1)p_{k,0} + \sum_{s=1}^{k} np_{k,s}\,\mathrm{error}_{n,s}$$

and therefore

$$\mathrm{E}\left\{\sum_{i=2}^{n}\lambda_i^k\right\} = (2d)^k(n-1)p_{k,0} + (2d)^k\sum_{s=1}^{k} np_{k,s}\,\mathrm{error}_{n,s}\ . \tag{3.2}$$

To estimate the above, first notice that $p_{k,s} = 0$ if k and s have different parity. In [BS87], the following estimate is given:

Lemma 3.7

$$p_{2k,2s} \leq \frac{2s+1}{2k+1}\binom{2k+1}{k-s}\left(\frac{1}{2d}\right)^{k-s}\left(1-\frac{1}{2d}\right)^{2s-1}$$

Proof See [BS87]; for an exact formula and sharper bounds, see [McK81].

□

Incidentally, from the proof of the above lemma in [BS87] it is clear that also

$$p_{2k,0} \geq \frac{1}{2k+1}\binom{2k+1}{k}\frac{(2d-1)^k}{(2d)^{2k}}\ .$$

It follows that for *any* graph of degree $2d$,

$$\sum_{i=1}^{n}\lambda_i^{2k} \geq (2d)^{2k}(n-1)p_{k,0} \approx (n-1)2^{2k}(2d-1)^k,$$

so that taking $2k$ slightly less than $2\log_d n$ yields the lower bound mentioned in the introduction,

$$|\lambda_2| \geq 2\sqrt{2d-1} + O\left(\frac{1}{\log_d n}\right).$$

Now we take $k = 2\lfloor \log n\, \lfloor\sqrt{2d-1}/2\rfloor / \log d\rfloor$, so that k is even, and calculate using the simplified bound

$$p_{2k,2s} \leq 2^{2k}\left(\frac{1}{2d}\right)^{k-s}\ .$$

It is easy to see that the dominant terms of the summation over s in equation 3.2 are $s = 1$ and $s = k$, and therefore

$$\mathrm{E}\left\{\sum_{i=2}^{n}\lambda_i^k\right\} \leq n^{1+\frac{\log 2}{\log d}}(ckd)^c k^{2\sqrt{2d}}\left(2\sqrt{2d}\sqrt{\frac{2d}{2d-1}}\right)^k\ .$$

Taking k-th roots, applying Hölder's (or Jensen's) inequality, and noticing that

$$\left(n^{1+\frac{\log 2}{\log d}}\right)^{1/k} = 1 + \frac{\log d}{\sqrt{2d}} + O\left(\frac{1}{\sqrt{d}}\right)\ ,$$

that

$$(ckd)^{c/k}k^{2\sqrt{2d}/k} = 1 + O\left(\frac{\log d \, \log\log n}{\log n}\right),$$

and that

$$k \leq \frac{1}{cd^{3/2}} n^{\frac{1}{c\sqrt{d}}}$$

for

$$\frac{\log n}{\log\log n} \geq c'\sqrt{d}$$

yields:

Theorem 3.8 *For $G \in \mathcal{G}_{n,2d}$ we have*

$$E\{|\lambda_2(G)|\} \leq 2\sqrt{2d-1}\left(1 + \frac{\log d}{\sqrt{2d}} + O\left(\frac{1}{\sqrt{d}}\right)\right) + O\left(\frac{d^{3/2}\log\log n}{\log n}\right)$$

(with an absolute constant in the $O(\ )$ notation), and more generally

$$E\{|\lambda_2(G)|^m\} \leq \left(2\sqrt{2d-1}\left(1 + \frac{\log d}{\sqrt{2d}} + O\left(\frac{1}{\sqrt{d}}\right)\right) + O\left(\frac{d^{3/2}\log\log n}{\log n}\right)\right)^m$$

for any $m \leq 2\lfloor \log n \lfloor \sqrt{2d-1}/2 \rfloor / \log d \rfloor$.

□

As a corollary we also get:

Theorem 3.9 *For any $\beta > 1$ we have*

$$|\lambda_2(G)| \geq \left(2\sqrt{2d-1}\left(1 + \frac{\log d}{\sqrt{2d}} + O\left(\frac{1}{\sqrt{d}}\right)\right) + O\left(\frac{d^{3/2}\log\log n}{\log n}\right)\right)\beta$$

with probability

$$\leq \frac{\beta^2}{n^{2\lfloor \sqrt{2d-1}/2 \rfloor \log\beta/\log d}} \,.$$

4 Concluding Remarks

An interesting question would be to see if an expansion in theorem 2.18 exists for $r > \sqrt{2d-1}/2$. In section 2 we defined $f_i(k)$ really for any i such that

$$\sum_{m=0}^{\infty} \left(\frac{2i+2}{2d-1}\right)^m$$

converges, which is to say any $i \le d-2$. We only used $f_i(k)$ for $i \le (\sqrt{2d-1}/2) - 1$, because in equation 2.21, we needed to use

$$(2i+2)^m \le (2d-1)^{m/2}$$

in order to get a reasonable bound on the error term. If there were some way to bound this error term by a reasonable quantity for larger i, one could extend the expansion, with the n^{-j} coefficient being $f_{j-1}(k)$. Whether or not $f_i(k)$ for $i \ge (\sqrt{2d-1}/2) - 1$ is involved in an asymptotic expansion, it might be interesting to evaluate them, since they seem to be a fairly naturally defined quantity.

If one were able to get an expansion up to r-th order for r close to d, one might notice $\mathcal{G}_{n,2d}$'s weakness, that according to theorem 3.1 a graph in $\mathcal{G}_{n,2d}$ has a d as a multiple eigenvalue with probability on the order of $n^{-(d-1)}$. From the arguments in section 3, it would follow that the expansion, if it existed, would fail to vanish by the n^{-d+1} term. However, one might hope for a better second eigenvalue estimate through various approaches. First of all, one might be able to find the graphs which give the non-vanishing terms, such as graphs with isolated vertices for the n^{-d+1} term, and thus show that upon removing them (which does not affect the probability measure very much) one has a better moment estimate. Secondly, one might do better with a different probability distribution, say $\mathcal{H}_{n,2d}$, which is constructed like $\mathcal{G}_{n,2d}$ but only allows permutations, π, which consist of one cycle. There are $(n-1)!$ such permutations, and once a values of such a π have been determined (which do not already give π a cycle), there are

$$\frac{(n-1)!}{(n-1)(n-2)\dots(n-a)}$$

ways of completing π to be a permutation of one cycle. It follows that over $\mathcal{H}_{n,2d}$ one has

$$\Pr(\Gamma) = n(n-1)\dots(n-|V_\Gamma|+1)\prod_{j=1}^{d}\frac{1}{(n-1)(n-2)\dots(n-a_j(\Gamma))},$$

and therefore the asymptotic expansion and second eigenvalue theorems hold for $\mathcal{H}_{n,2d}$ as well. However, in $\mathcal{H}_{n,2d}$ one can never have an isolated vertex or similar phenomena, and so one would have a much better theorem 3.1 for $\mathcal{H}_{n,2d}$, and a possibility of having a further than d-th order expansion for which all lower order terms vanish.

References

[Alo86] N. Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986.

[BS87] Andrei Broder and Eli Shamir. On the second eigenvalue of random regular graphs. In *28th Annual Symposium on Foundations of Computer Science*, pages 286–294, 1987.

[FK81] Z. Füredi and J. Komlós. The eigenvalues of random symmetric matrices. *Combinatorica*, 1(3):233–241, 1981.

[Gem80] S. Geman. A limit theorem for the norm of random matrices. *Ann. of Prob.*, 8(2):252–261, 1980.

[Knu73] Donald E. Knuth. *The Art of Computer Programming*, volume 1: Fundamental Algorithms. Addison-Wesley, Reading, Massachusetts, second edition, 1973.

[KS] J. Kahn and E. Szemeredi. Personal communication, and to appear.

[LPS86] A. Lubotzky, R. Phillips, and P. Sarnak. Explicit expanders and the ramanujan conjectures. In *18th Annual ACM Symposium on Theory of Computing*, pages 240–246, 1986.

[Mar87] G. Margulis. Manuscript in Russian on graphs with large girth, 1987.

[McK81] B. McKay. The expected eigenvalue distribution of a large regular graph. *Lin. Alg. Appl.*, 40:203–216, 1981.

[Tan84] R.M. Tanner. Explicit concentrators from generalized N-gons. *SIAM J. Alg. Disc. Methods*, 5:287–293, 1984.

[Wig55] E. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Annals of Math.*, 62(3):548–564, 1955.