

# The Second Eigenvalue and Random Walks in Random Regular Graphs with Increasing Girth

by

Alexander Izsak

B.Sc., Harvey Mudd College, 2007

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF  
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate Studies

(Computer Science)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2009

© Alexander Izsak 2009

# Abstract

The goal of this thesis is to upper bound the expected value of the second largest eigenvalue in magnitude of random regular graphs with a given minimum girth. Having a small upper bound implies such random graphs are likely to be expanders and thus have several combinatorial properties useful in various fields of computer science. The best possible upper bound asymptotically on the second eigenvalue has already been proven for random regular graphs without conditions on the girth. Finding this upper bound though required long and complicated analysis due to tangles, which are certain small subgraphs that contain cycles. This thesis thus hypothesizes that specifying a minimum girth large enough will prevent tangles from occurring in random graphs and thus proving an optimal upper bound on the second eigenvalue can avoid the difficult analysis required in order to handle tangles.

To find such an upper bound on random regular graphs with specified minimum girth we consider the probability that a random walk in such a random graph returns to the first vertex of the walk in the  $k$ -th step of the walk. We prove for 2-regular graphs that the random walk is more likely to visit any given vertex not in the walk than the starting vertex of the walk on the  $k$ -th step, and bound how much more likely this event is. We also analyze the  $d$ -regular case and we believe our findings will lead to a similar result in this case.

# Table of Contents

<b>Abstract</b> . . . . .	ii
<b>Table of Contents</b> . . . . .	iii
<b>Acknowledgements</b> . . . . .	iv
<b>Dedication</b> . . . . .	v
<b>1 Introduction</b> . . . . .	1
1.1 Applications of Expanders . . . . .	1
1.2 The Second Eigenvalue of Random Graphs . . . . .	3
1.3 A Random Graph Model That Specifies Girth . . . . .	5
1.4 The Trace Method . . . . .	5
1.5 Open Questions . . . . .	8
<b>2 Random Walks in <math>\mathcal{K}_{n,2,g}</math></b> . . . . .	10
2.1 The Size of $\mathcal{K}_{n,2,g}$ . . . . .	10
2.2 The Probability of a Coincidence in $\mathcal{K}_{n,2,g}$ . . . . .	12
<b>3 Random Walks in <math>\mathcal{K}_{n,d,g}</math></b> . . . . .	14
3.1 The Size of $\mathcal{K}_{n,d,g}$ . . . . .	14
3.2 The Probability of a Coincidence in $\mathcal{K}_{n,d,g}$ . . . . .	17
<b>4 Conclusion</b> . . . . .	20
<b>Bibliography</b> . . . . .	21

# Acknowledgements

My gratitude goes to the faculty, staff and students of UBC whose support have made this thesis possible. In particular I thank my supervisor Prof. Friedman, whose deep insights in the field of expanders inspired me. I thank Prof. Kirkpatrick for helping me clarify the beauty and utility in this field. Craig Weidert, Jeff Hendy and Emily Sors all have my vast appreciation for their advice and support. And of course special recognition goes to my family for their love and support.

# Dedication

*This thesis is dedicated to my parents, my siblings, my partner and my friends.  
You are all always in my heart.*

# Chapter 1

## Introduction

The adjacency matrix of a finite undirected graph  $G$  has only real eigenvalues, which we order by

$$\lambda_1(G) \geq \lambda_2(G) \geq \dots \geq \lambda_n(G)$$

where  $n$  is the size of  $G$  (i.e. number of vertices.) For  $d$ -regular graphs we have  $\lambda_1(G) = d$ . Note for the rest of this thesis the graphs we will be referring to are regular and may contain self-loops and multiple edges. We consider  $\lambda(G)$  where

$$\lambda(G) = \max_{1 < i \leq n} (|\lambda_i(G)|).$$

If the distance between  $\lambda_1(G)$  and  $\lambda(G)$  is large than  $G$  is known as an *expander*. Expanders have various combinatorial properties which lead to their utility in several fields of computer science. Our definition of expanders is intentionally vague since there is no consensus on how small  $\lambda$  should be for a graph to be an expander, and sometimes expanders are defined instead in terms of combinatorial properties which are implied by (and also occasionally imply)  $\lambda$  being small.

### 1.1 Applications of Expanders

In order to convey the fundamental usefulness of expander graphs to computer science, we give several brief overviews of ways in which expander graphs appear in several fields of computer science. We first consider the use of expanders in network theory, and most of the discussion here is explored further in [7]. One example of such a combinatorial property implied by  $\lambda$  being small is edge expansion. If  $S$  is a subset of the vertices in  $G$  and  $\delta(S)$  is the set of edges with only one incident vertex in  $S$  than the *edge expansion ratio* is

$$h(G) = \min_{\{S \mid |S| \leq n/2\}} \frac{\delta(S)}{|S|}.$$

Then a well-known theorem states that

$$\frac{d - \lambda_2(G)}{2} \leq h(G).$$

So if the distance between  $d$  and  $\lambda_2$  is large then the number of edges that must be removed from the graph in order to make some large set  $S$  no longer connected to the rest of the graph is also large. Thus expanders can represent

robust networks of communications even though there would be only  $O(n)$  edges in such a network. Note we consider  $n$  as growing and  $d$  as constant in this thesis. If  $\lambda < cd$  for constant  $c$  then the shortest path between any two vertices in the graph is  $O(\log n)$ . The maximum of the lengths of these shortest path is called the *diameter* of the graph. So if a network is represented by an expander, a message to be transmitted between any two nodes does not have to pass through too many nodes along the way. We note that another important use of expander graphs in network theory is in the development of AKS sorting networks.

Expanders have played an interesting role in complexity theory as well, including the recent proof that  $\mathbf{SL} = \mathbf{L}$ . Recall  $\mathbf{L}$  is the class of problems that can be solved in logarithmic space. The class  $\mathbf{SL}$  is composed of all decision problems that have a logarithmic space reduction to USTCONN, the problem of determining whether there exists a path between two vertices  $s$  and  $t$  in an undirected graph. This problem is fairly trivial if we assume that each component of the graph is an expander. If this is true then there exists a  $O(\log n)$  length path between  $s$  and  $t$  assuming they are in the same component. Thus checking all logarithmically long paths that start at  $s$  for whether or not they reach  $t$  would solve the problem in logarithmic space. Reingold [12] showed one could transform in logarithmic space a graph into a graph where all the components are expanders using clever and repeated use of the zig-zag product. This graph product creates a new graph that combines the expansion properties and general structure of two graphs. Reingold's algorithm uses repeated zig-zag products of the original graph with a good expander, and using the algorithm in practice would depend on being able to find good expanders. Although we only mentioned USTCONN, this algorithm proved that many problems via logarithmic space reduction to USTCONN could be solved in logarithmic space. For more information on zig-zag products and Reingold's algorithm see [12] and [7]. Expanders also feature prominently in studying the approximability of  $\mathbf{NP}$ -hard problems [7].

Expanders are also "random-like" graphs that have uses in pseudorandom sampling. Suppose want to sample uniformly over a huge set. Given an equally large expander graph, we label the vertices of the graph by elements of the set bijectively. We can transform the expander into a Markov chain where every edge has probability  $1/d$ . Then no matter what vertex we start from, after few random steps in this Markov chain we are soon very close to equally likely to be at any vertex in the graph. So selecting the element of the set we are at after a few steps in this Markov chain resembles uniformly sampling from that set. This method is useful since it requires few random bits to approximate a uniform distribution of a massive set. On this application of expander graphs and many more applications I do not detail here such as error-correcting codes more information can be found in [7].

## 1.2 The Second Eigenvalue of Random Graphs

We define  $G$  as *Ramanujan* if  $\lambda(G) \leq 2\sqrt{d-1}$ . In this paper we consider  $d$  as a constant and  $n$  tending towards infinity. In this setting, the constant term  $2\sqrt{d-1}$  is as small as possible due to the Alon-Boppana bound [? ], which implies

$$\liminf_{n \rightarrow \infty} \lambda_2(G_n) \geq 2\sqrt{d-1},$$

where  $G_n$  is any  $d$ -regular undirected graph of size  $n$ . A stronger version of this bound due to Nilli is proven in [11]. For more about this or other basic information about expander graphs see [7] which is a vast overview on the study of expander graphs. Explicit constructions of Ramanujan graphs are known, but not for some  $n$  and  $d$ . On the other hand Alon's Second Eigenvalue Conjecture, recently proven by Friedman [6], states that for 'most'  $d$ -regular graphs,  $\lambda(G) \leq 2\sqrt{d-1} + \epsilon$  with  $\epsilon$  being any positive real. This implies that one can expect to find a nearly Ramanujan graph quickly by selecting a random graph, checking the second eigenvalue and repeating this process until a graph with sufficiently low second eigenvalue is found. This is fine enough for many of the applications of expander graphs mentioned earlier.

The probability space  $\mathcal{G}_{n,d}$  of random  $2d$ -regular graphs is often used in proving an upper bound on the expected second eigenvalue as in [3] and [4] and we describe it here. Given  $n$  and  $d$  with  $d$  even, let  $\pi_1, \dots, \pi_d$  be permutations on  $[n] = \{1, 2, \dots, n\}$  chosen uniformly and independently from the set of all  $n!$  such permutations. We create a graph by setting vertices  $V = [n]$  and edges

$$E = \{(i, \pi_j(i)), (\pi_j(i), i) \mid i = 1, 2, \dots, n, j = 1, 2, \dots, d/2\}.$$

This graph is directed, though may be viewed as undirected by considering each  $(i, \pi_j(i))$  and  $(\pi_j(i), i)$  as a single directed edge. The graph also may contain self-loops and multiple edges.

The general tool used by Friedman to prove most graphs are nearly Ramanujan is known as the trace method. In the most basic form, the trace method bounds the *trace* of a large power of the adjacency matrix of graph by bounding the number of walks of a given length that return to their starting vertex. The trace of a matrix (or respectively graph) is defined as the sum of the diagonal entries of the matrix (or respectively of the adjacency matrix of the graph.) The trace method has been used to examine the expected eigenvalues of graphs since at least 1981 when McKay used it to give the expected distribution the eigenvalues of regular graphs as their size goes to infinity [10]. There also have been several attempts to use this method to prove Alon's second eigenvalue conjecture since then such as [3] and [4].

Friedman's proof of the conjecture not only made improvements on the trace method, but also demonstrated a fundamental reason why previous research using the trace method was not able to prove Alon's second eigenvalue conjecture. Friedman showed that the trace method cannot achieve a proof of Alon's second eigenvalue conjecture unless we discount the contribution of *tangles* to the trace of the graph. Tangles are certain subgraphs which cause the second eigenvalue

of the graph to be large as well as causing certain sums in the trace method's analysis to diverge. We avoid a more technical definition, but one can be found here [6]. Theorem 3.13 of that paper also implies that tangles contain cycles. This can be seen by noting that if  $G$  is acyclic (a tree), then, using the terminology found in that theorem,  $\text{Tree}_d(G)$  is the infinite  $d$ -regular tree and the norm of the infinite  $d$ -regular tree is  $2\sqrt{d-1}$  as is shown in [6]. Not only that, but tangles will need to contain small cycles in order to impede the trace method. This can be seen due since Friedman only needs to deal with tangles of up to a small size in the paper.

Friedman avoided tangles in his analysis by creating a *selective trace*. A power of the selective trace of a graph is equal to the amount of only some of the walks of a certain length that return to the original vertex. The walks the selective trace counts are only those where no small contiguous section of the walk form a graph containing a tangle as a subgraph. This thesis posits that instead of a selective trace, choosing random graphs in a way that avoids tangle would also allow the trace method to show that most of the graphs from the random set are Ramanujan. The selective trace adds quite a bit to the complexity and the length of the proof that most graphs are Ramanujan, and not needing to use a selective trace due to our choice of random graph model could simplify things greatly. We note that Alon did not specify a random graph model in his second eigenvalue conjecture.

This paper introduces a new random model of graphs that avoids tangles by specifying a lower bound on the *girth*, the length of the shortest cycle in the graph. The girth in our random model is increasing monotonically as the size of the graphs increases. We believe such a model will allow us to avoid tangles since tangles are small subgraphs which contain cycles. Thus a given tangle will not appear in our model for  $n$  large enough. Our new model is also similar to the standard random model used in [3], [4] and [6] with regards to the trace method. This similarity is by design so that the analysis in the trace method may extend to this model more readily.

To use the trace method in our new model, we must first consider random walks in our model as [3] and [4] did in the standard model. In particular, we need to consider the following question. Given that a randomly chosen graph contains a specific walk and an edge label  $\pi$ , what is the probability that  $\pi$  maps the final vertex of that walk to any given vertex? In the standard model, this next vertex is just as likely to be a given vertex already in our walk as it is likely to be a given vertex not already in our walk, assuming that we haven't already determined the edge labeled  $\pi$  incident on either of those vertices from the walk so far [3]. This fact follows almost immediately from the definition of the standard model. What this probability is in my new model is not so immediately clear though, and most of my thesis is devoted to understanding this probability. In chapter 2 I show that the probability this final vertex is any given vertex already visited by the walk is smaller than probability the final vertex is any given vertex not along the walk so far, assuming the graph is 2-regular and neither of those given vertices are along the walk so far. I also bound the difference between these two probabilities. Chapter 3 works towards similar

results for  $d$ -regular graphs with  $d > 2$  and finds two different formulas relating the two probabilities. Our analysis so far does not find which probability is larger. We believe though that as in the 2-regular case, the walk will be less likely to return to a previously visited vertex. We justify this belief in the end of Chapter 3.

## 1.3 A Random Graph Model That Specifies Girth

The *girth* of a graph is the length of the shortest cycle in the graph. If the graph contains no cycles, we define the girth as infinite. We also define a self-loop as a cycle of length one and two edges between a pair of vertices as a cycle of length two. Given a size  $n$  of a  $d$ -regular graph it is known that the maximal girth of the graph is  $\Omega(\log n)$  [1]. One commonality between graphs with very small second eigenvalue and graphs with high girth is that they both seem difficult to explicitly construct.

We can use  $\mathcal{G}_{n,d}$  to create a new random graph model that also bounds the girth from below in the following way. Given a desired lower bound  $g$  on the girth, select randomly a graph  $G$  from  $\mathcal{G}_{n,d}$ . If the graph has girth less than  $g$ , chose another one and repeat this process until a suitable graph is chosen. Note that all graphs in  $\mathcal{G}_{n,d}$  with girth at least  $G$  are equally likely to be chosen. We call this new probability space  $\mathcal{K}_{n,d,g}$ . As before, we consider  $n$  as increasing towards infinity and  $d$  as constant, but we also will consider  $g$  as increasing towards infinity with  $g = O(\log(n))$ .

The girth of the graph has several effects on the trace method used to bound the second eigenvalue of the graph. As mentioned earlier, certain subgraphs known as tangles prevent the trace method from giving a proof of Alon's Second Eigenvalue conjecture [6]. Given a particular tangle, if we set the girth high enough we can be sure that tangle will not exist in our graph. This may allow us to give a good bound on the second eigenvalue of most graphs in  $\mathcal{K}_{n,d,g}$ . We also know that the probability that a walk contains a cycle of length less than  $g$  is 0 in our new model.

## 1.4 The Trace Method

In this section we describe the trace method [10] and the analysis of random walks used originally by Broder and Shamir in [3] and subsequently improved in [4] and [6]. We also begin to discuss how girth affects the trace method. The trace of a large power of the adjacency matrix  $A$  of a graph is related to  $A$ 's eigenvalues via

$$\text{Trace}(A^k) = \lambda_1^k + \lambda_2^k + \dots + \lambda_n^k.$$

Thus, if we consider this equality over a graph probability space, the expected value of the trace of  $A^k$  is equal to the sum of the  $k$ th power of the expected eigenvalues of  $A$ . So estimating the expected value of  $\text{Trace}(A^k)$  for some large

#### 1.4. The Trace Method

---

even  $k$  allows us to give a bound on the expected value of  $\max\{\lambda_2, |\lambda_n|\}$  (since  $\lambda_1 = d$ .)

A walk that begins and ends at the same vertex is *closed*. The trace  $A^k$  then is also the number of length  $k$  closed walks. The trace method combinatorially counts the number (or expected number in the random case) of closed walks in order to derive information about the spectrum of the graph (or averaged information in the random case).

We can represent a walk of length  $k$  in a graph from  $\mathcal{G}_{n,d}$  or  $\mathcal{K}_{n,d,g}$  using a starting vertex  $i$  and a word  $w = \sigma_1\sigma_2\dots\sigma_k$  with characters taken from the alphabet

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

The word  $w$  gives the edges of the walk in order, but also is thought of as representing the permutation formed by taking the product of the permutations its characters represent. The probability that  $w$  maps  $i$  back to  $i$  is independent of  $i$  in both  $\mathcal{G}_{n,d}$  and  $\mathcal{K}_{n,d,g}$  since they are closed under relabeling of the vertices. So we may consider the trace in terms of the probability  $w$  maps a vertex back to itself,  $P(w)$ , which gives

$$E(\text{Trace}(A^k)) = n \sum_{w \in \Pi^k} P(w).$$

A word is *irreducible* if no character is immediately followed by its inverse. A walk is said to be *irreducible* if the corresponding word is irreducible, or equivalently if the walk does not cross an edge and then immediately cross that edge again in the opposite direction. We *reduce* a word  $w$  by repeatedly removing all consecutive occurrences of a character and its inverse until we are left with an irreducible word. Note that if we reduce a word  $w$  we arrive at a unique irreducible word  $w'$  and that  $P(w) = P(w')$ .

Let  $\text{Irred}_k$  be the set of all irreducible word in  $\Pi^k$ . The  $k$ -th *irreducible trace* is the number of closed irreducible walks of length  $k$  in the graph, and is denoted by  $\text{IrredTr}(A, k)$  where  $A$  is the graph's adjacency matrix. It is not hard to see that

$$E(\text{IrredTr}(A, k)) = n \sum_{w \in \text{Irred}_k} P(w)$$

and more detail on this can be found in [3]. The irreducible trace is closely related to the usual trace as well as the eigenvalues of the graph, and in fact estimating the irreducible trace will be sufficient for bounding the second eigenvalue. The irreducible trace has been used in [3], [9] and [4]. The relation between the irreducible trace and the eigenvalues of  $A$  is given in Lemma 2.3 of [6], and the discussion immediately after shows how irreducible traces relate to the usual trace.

Given a word  $w = \sigma_1\sigma_2\dots\sigma_k \in \text{Irred}_k$  and a sequence of integers  $I = (i_1, i_2, \dots, i_k)$  with  $i_1 = i_k$  and each integer taken from  $\{1, 2, \dots, n\}$ , we find the probability that the characters of  $w$  take  $i_1$  along the sequence of vertices represented by  $I$ . To do this, we find the probability that  $\sigma_1(i_1) = i_2$ , and then

the probability that  $\sigma_2(i_2) = i_3$  given that  $\sigma_1(i_1) = i_2$ , and then the probability that  $\sigma_3(i_3) = i_4$  given that  $\sigma_1(i_1) = i_2$  and  $\sigma_2(i_2) = i_3$ . This in fact is the method used by [3] to originally bound  $P(w)$  and thus bound the expected second eigenvalue of random graphs.

In  $\mathcal{G}_{n,2d}$ ,  $\sigma_1(i_1)$  takes on any value of  $\{1, 2, \dots, n\}$  with probability  $1/n$ . For any of the next steps of the walk, we will need to consider several cases. Define the random variables  $t_2 = \sigma_1(i_1)$  and  $t_j = \sigma_j(t_{j-1})$  for  $k \geq j > 1$ . Now given that the first  $j - 1$  edges of the walk are along the first  $j$  integers of  $I$ , the value of  $t_j$  may be already determined if we visited the edge that determines  $t_j$  previously in our walk. For example if  $\sigma_1(i) = i$  and  $\sigma_2 = \sigma_1$ , then clearly  $\sigma_2(i) = i$ . We call such a  $t_j$  a *fixed choice*, and if  $t_j$  is not a fixed choice it is called a *free choice*. Similarly, we call the edge that determines  $t_j$  a fixed or a free choice.

If  $t_2$  is a free choice and  $\sigma_1 \neq \sigma_2$  then  $\sigma_2$  is a new permutation and  $t_2$  equals  $i_3$  with probability  $1/n$ . On the other hand if  $t_2$  is a free choice but  $\sigma_1 = \sigma_2$ , then  $\sigma_2$  maps  $i_1$  to  $i_2$  and so  $t_2$  can take on any value from  $\{1, 2, \dots, n\} - \{i_2\}$ . All of these values are also equally likely, so  $\Pr(t_2 = i_3) = \frac{1}{n-1}$  in this case so long  $i_2 \neq i_3$ . In general if  $t_j$  is a free choice then  $\Pr(t_j = i_{j+1}) = \frac{1}{n-m}$  where  $m$  is the number of edges that were free choices that are created due to the permutation  $\sigma_j$ .

Now given the walk  $w$  and the sequence of vertices  $w$  visits we consider the directed graph  $\Gamma_{w,I}$  represented by this walk. The edges of this graph are labelled by characters from  $\Pi$  and the vertices are all the distinct elements of  $I$ . So the number of edges of  $\Gamma_{w,I}$  is the number of free choices in the walk  $w$ . We call  $\Gamma_{w,I}$  the *generalized form* of  $w, I$ .

Noticing that to determine the probability that we walk the path represented by  $\Gamma_{w,I}$ , the exact values of  $I$  did not matter. What did matter was which members of  $I$  were equal to one another, which is determined in  $\Gamma_{w,I}$  by the corresponding vertices being equal. So we define  $\Gamma_w$  to be the same graph but without the labels on the vertices. If  $a_j(w)$  is the number of times that  $\pi_j$  or  $\pi_j^{-1}$  appears in  $w$  and we define  $\Pr(\Gamma_{w,I})$  as the probability that the word  $w$  produces a walk along the sequence of vertices  $I$ , we have

$$\Pr(\Gamma_{w,I}) = \prod_{j=1}^d \frac{1}{n(n-1)\dots(n-a_j(w)+1)}$$

where the  $j$ -term of the product on the right-hand side is 1 if  $a_j(w) = 0$ . Since this probability is independent over the labels on the vertices of  $\Gamma_{w,I}$ , and there are  $n(n-1)\dots(n-|V_\Gamma|)$  possible labels on the vertices, where  $|V_\Gamma|$  is the number of vertices of the generalized form it makes sense to consider

$$\Pr(\Gamma_w) = n(n-1)\dots(n-|V_\Gamma|+1)\Pr(\Gamma_{w,I}).$$

Here  $\Pr(\Gamma_w)$  represents the probability that  $w$  maps any vertex to itself via a walk that has the same form as  $\Gamma_w$ . We can then estimated the irreducible trace

using that

$$E(\text{IrredTr}(A, k)) = \sum_{\Gamma_w | w \in \text{Irred}_k} \Pr(\Gamma_w).$$

Note that the formula above applies to  $\mathcal{K}_{n,2,g}$  as well, although  $\Pr(\Gamma_w)$  will have different values than before. A simple example of this is if  $\Gamma_w$  has girth less than  $g$ . Clearly no such walk could occur in a graph from  $\mathcal{K}_{n,2,g}$ , and so  $\Pr(\Gamma_w) = 0$ . This fact hints that we may be able to achieve similar if not better bounds on the expected second eigenvalue in a random model that conditions on girth.

Now we consider a random walk in  $\mathcal{K}_{n,2,g}$ . The first  $g - 1$  steps of the walk are all free choices, since there are no cycles in our walk of length less than  $g$ . Using the  $t_j$  and  $I$  from before, we see that  $t_1$  can take on any value in  $\{1, 2, 3, \dots, n\} - \{i_1\}$ , and is equally likely to take on any of these values due to the symmetry of  $\mathcal{K}_{n,2,g}$ . More generally  $t_j$  for  $j < g$  can take on any value in  $\{1, 2, 3, \dots, n\} - \{i_1, i_2, \dots, i_j\}$ , and similarly any of these values are equally likely due to symmetry. The bulk of the original research in my thesis is dedicated to describing  $t_j$  for  $j \geq g$ . For  $j$  at least  $g$ , we see that  $t_j$  can take on any value from  $\{1, 2, 3, \dots, n\} - P_g$  where  $P_g$  is the set of vertices that are a distance at most  $g - 1$  from  $i_j$  in  $\Gamma_{\sigma_1 \sigma_2 \dots \sigma_{j-1}, (i_1, i_2, \dots, i_{j-1})}$ , the portion of  $\Gamma_{w,I}$  determined by the walk so far. We call  $t_j$  a *coincidence* if it is equal to a vertex previously visited by the walk. If  $t_j = r$  and  $t_j = s$  each imply  $t_j$  is no coincidence, then both these events are equally likely. The reason once again is symmetry, and this will be discussed in more detail in Chapter 2.

The next few chapters are original research concerning  $t_j$  in the new random model. The second chapter studies the case where  $d = 1$ . Given vertices  $r$  and  $s$  the second chapter shows that if the event  $t_j = r$  is a coincidence and the event  $t_j = s$  is not then  $t_j = r$  is less likely than  $t_j = s$ . We also describe how much more likely the second event is. The chapter following that concerns the  $d$ -regular case for  $d \geq 3$ . Producing similar results to those we found in the  $d = 2$  case appears to be far more difficult for  $d$  larger than 2, but we hope the analysis in that chapter will lead to such results.

## 1.5 Open Questions

One unknown is whether the trace method can prove that  $\lambda \leq 2\sqrt{d-1} + \epsilon$  for  $\epsilon = 0$  or even perhaps  $\epsilon$  being a negative function of  $n$ , tending to 0 as  $n$  approaches infinity. Numerical experiments do indicate the latter but so far the best bound we have is with  $\epsilon$  being any nonzero constant [6]. Even apparently avoiding tangles with a selective trace in that paper is not sufficient if we are to have nonpositive  $\epsilon$ . Being able to use the trace method in the model  $\mathcal{K}_{n,d,g}$  may lead to results with  $\epsilon$  nonpositive although this is completely speculative. We believe the expected trace of graphs from  $\mathcal{K}_{n,d,g}$  is significantly smaller than the expected trace of graphs from  $\mathcal{G}_{n,d}$  since heuristically no small cycles should imply fewer walks that begin and end at the same vertex.

## 1.5. Open Questions

---

There is also an unproven stronger form Alon's second eigenvalue conjecture that continues to drive research in the area. This version asks that given any *base* graph and  $\epsilon > 0$ , can we prove that most random *coverings* of degree  $n$  of the base graph have all new eigenvalues at most  $\rho + \epsilon$ . Here  $\rho$  is the spectral radius of the *universal cover*. This is a stronger version of Alon's second eigenvalue conjecture since every  $d$ -regular graph of size  $n$  can be viewed as a covering of the base graph of a single vertex with  $d/2$  self loops and  $\rho = 2\sqrt{d-1}$ . For more information about this see [5] and [8]. The best result currently along these lines is due to Linial and Puder, which says that the new eigenvalues in most random lifts are in  $O(\lambda_1^{\frac{1}{3}}\rho^{\frac{2}{3}})$  where  $\lambda_1$  is the largest eigenvalue of the base graph.

## Chapter 2

# Random Walks in $\mathcal{K}_{n,2,g}$

We first restrict our discussion to the  $\mathcal{K}_{n,2,g}$ , the 2-regular case of our model that restricts the size of cycles. Note that in the 2-regular case  $\lambda_1 = 2$  and graphs are Ramanujan if  $\lambda_2 \leq 2$ . Although this immediately implies all graphs of  $\mathcal{K}_{n,2,g}$  are Ramanujan, considering random walks in this  $\mathcal{K}_{n,2,g}$  will allow us to develop methods and theorems that will be useful later.

### 2.1 The Size of $\mathcal{K}_{n,2,g}$

Define a *simple cycle* as a closed path with no repeated vertices except for the start and end vertices. Also let  $I$  be a sequence of  $s$  unique integers from  $\{1, 2, \dots, n\}$  with the smallest integer first, and call such sequences *cyclic*. Let  $C_{s,I}$  be the set of graphs from  $\mathcal{G}_{n,4}$  with a simple cycle of length  $s$  occurring along the sequence of vertices represented by  $I$ . We place the smallest integer first in  $I$  so as not to count cycles multiple times by considering the same cycle with different vertices first. Then if  $k_{n,2,g}$  is the number of graphs in  $\mathcal{K}_{n,2,g}$  we have

$$k_{n,2,g} = n! - \left| \bigcup_{s < g, I} C_{s,I} \right|. \quad (2.1)$$

Inclusion-exclusion principle allows us to rewrite the union above. But first note that for  $I \neq J$  and  $I$  and  $J$  share at least one vertex we have  $|C_{s,I} \cap C_{r,J}| = 0$  since a single vertex being in two different simple cycles contradict that these graphs are all 2-regular.

Also intuitively there are just as many graphs that have cycles of length  $s$  over any any possible cyclic sequence of  $s$  integers. We extend this to the number of graphs with  $j$  cycles of lengths  $s_1, s_2, \dots, s_j$ .

Let  $I_1, I_2, \dots, I_j$  be cyclic sequences that share none of the same integers and let the same hold for the cyclic sequences  $J_1, J_2, \dots, J_j$ , then

$$\left| \bigcap_{i=1}^j C_{s_i, I_i} \right| = \left| \bigcap_{i=1}^j C_{s_i, J_i} \right|.$$

So let  $r_i$  be the number of  $s_h = i$  for  $1 \leq i < g$  and  $1 \leq h \leq j$ . We think of  $r_i$  as being the number of cycles of length  $i$  and let  $r = (r_1, r_2, \dots, r_{g-1})$ . We also define  $|r| = r_1 + r_2 + \dots + r_{g-1}$  and  $|r|_* = r_1 + 2r_2 + \dots + (g-1)r_{g-1}$ . Now let's define  $|\bigcap_{i=1}^j C_{s_i, I_i}| = C(r)$ . In other words  $C(r)$  is the number of graphs in the permutation model with  $r_i$  simple cycles of length  $i$  for  $1 \leq i < g$

2.1. The Size of  $\mathcal{K}_{n,2,g}$

---

if the sequences of vertices in the cycles are already given. Clearly  $C(r) = 0$  if  $|r|_* > n$ , since otherwise we'd have disjoint cycles over more than  $n$  vertices. So applying everything we just discussed to 2.1 implies

$$k_{n,2,g} = n! - \sum_{0 < |r|_* \leq n} \omega(r)C(r)(-1)^{|r|} \quad (2.2)$$

where  $\omega(r_1, r_2, \dots, r_{g-1})$  is the number of ways to choose the sequences  $I_1, I_2, \dots, I_j$ .

We can think of  $C(r)$  as the number of ways to order everything besides what is in the specified cycles, and  $\omega(r)$  as the number of ways to choose which vertices to place in the specified cycles and order those vertices. Immediately we find that  $C(r)$  is the factorial of the number of vertices not in the specified cycles and so  $C_r = (n - |r|_*)!$ . Now if we define

$$\binom{n}{r} = \frac{n(n-1)(n-2)\dots(n - |r|_* + 1)}{\prod_{i=1}^{g-1} (i!)^{r_i}}$$

we see that the above definition is equivalent to  $n$  multinomial choose  $r_1$  ones,  $r_2$  twos, and so on. This selects the correct amounts of integers from  $n$  to place in the selected cycles. However this over-counts the number of ways to select integers in the following way. Suppose  $r_2 = 2$  and all the other  $r_i$  are 0. Then  $\binom{n}{r}$  counts the number of ways to select two elements to place in sets  $S_1$  and  $S_2$ . But for our purposes we do not want the sets to be labeled since  $S_1 = \{1, 2\}, S_2 = \{3, 4\}$  and  $S_1 = \{3, 4\}, S_2 = \{1, 2\}$  represent equivalent choices for our purposes. So the number of ways to select the integers to place in the cycles is

$$\frac{\binom{n}{r}}{\prod_{i=1}^{g-1} (r_i)!}$$

Now the number of ways to produce a cycle over  $i$  elements is  $(i-1)!$ . A quick reason for this is you can imagine counting cycles by choosing the smallest element to be the first and last element, and then the number of cycles is simply number of ways to permute all  $i-1$  other elements. So this implies that

$$\omega(r) = \frac{\binom{n}{r}}{\prod_{i=1}^{g-1} (r_i)!} \prod_{i=1}^{g-1} ((i-1)!)^{r_i}$$

We rewrite 2.2 by substituting in for  $\omega(r)$  and  $C(r)$  and simplifying to achieve

$$k_{n,2,g} = n! \left( \sum_{0 \leq |r|_* \leq n} \prod_{i=1}^{g-1} \frac{(-1)^{r_i}}{i^{r_i} r_i!} \right). \quad (2.3)$$

As  $n$  approaches infinity, the right-hand side becomes  $n!$  multiplied by the Maclaurin series for  $e^{x_1+x_2+\dots+x_{g-1}}$  at  $(x_1, x_2, x_3, \dots, x_{g-1}) = (-1, -1/2, -1/3, \dots, -1/g - 1)$ . Notice that if we consider the case  $g = 2$ , then we are counting the number of

derangements, that is permutations with no fixed points. Then our results so far imply the well-known result that the number of derangements quickly approaches  $n!/e$ .

By Taylor's Theorem,

$$\sum_{0 \leq |r| \leq \lfloor n/(g-1) \rfloor} \prod_{i=1}^{g-1} \frac{(-1)^{r_i}}{i^{r_i} r_i!} = e^{-1-1/2-\dots-1/(g-1)} + R_{\lfloor n/(g-1) \rfloor}$$

where  $|R_n| < (g-1)^{n+1}/(n+1)!$ . The bound on the remainder term  $R_n$  follows from Taylor's Theorem since any sequence of partial derivate on  $e^{x_1+x_2+\dots+x_{g-1}}$  has absolute value less than 1 when  $0 \geq x_i \geq -1/i$ .

Now we still need to bound the rest of the terms in 2.3. These are the terms where  $|r| > \lfloor n/(g-1) \rfloor$  and  $|r|_* \leq n$ . Note that the absolute values of all of these terms are included among the terms of Macluarin series for  $e^{x_1+x_2+\dots+x_{g-1}}$  when  $x_i = 1/i$ . So the sum of the leftover terms is bounded in absolute value by the remainder

$$R'_{\lfloor n/(g-1) \rfloor} = e^{1+1/2-\dots+/(g-1)} - \sum_{0 \leq |r| \leq \lfloor n/(g-1) \rfloor} \prod_{i=1}^{g-1} \frac{1}{i^{r_i} r_i!}$$

. Since  $e$  is the largest possible value of any sequence of partial derivatives on  $e^{x_1+x_2+\dots+x_{g-1}}$  with  $0 \leq x_i \leq 1/i$  for any  $i$ , we have that  $|R'_{\lfloor n/(g-1) \rfloor}| < e(g-1)^{\lfloor n/(g-1) \rfloor+1}/(\lfloor n/(g-1) \rfloor+1)!$ .

So finally we have arrived upon the expression

$$k_{n,2,g} = n!(e^{-1-1/2-\dots-1/(g-1)} + E_n) \tag{2.4}$$

where  $|E_n| < (e+1)(g-1)^{\lfloor n/(g-1) \rfloor+1}/(\lfloor n/(g-1) \rfloor+1)!$ . This error term approaches zero as  $n$  goes to infinity. Also interestingly if  $n$  and  $g$  approach infinity then the probability that a randomly selected graph from  $\mathcal{G}_{n,2}$  is in  $\mathcal{K}_{n,2,g}$  goes to zero. This is true regardless of the rate at which  $g$  grows in comparison to the rate at which  $n$  grows.

## 2.2 The Probability of a Coincidence in $\mathcal{K}_{n,2,g}$

Consider a walk in a graph from  $\mathcal{K}_{n,2,g}$  starting at a vertex  $i_1$ . Let  $i_2, i_3, \dots, i_j$  be vertices along the walk up to  $j$ -th vertex in order. Also let  $t_j$  be a random variable denoting the  $j+1$ -th vertex along the walk. Suppose that  $t_j$  is a free choice and let  $\text{coin}(t_j)$  be the event that  $t_j$  is a coincidence. Notice that  $t_j$  can only be a coincidence if the  $j+1$ -th vertex along the walk is  $i_1$ . Then  $t_j$  can be any value from  $\{1, 2, 3, \dots, n\} - \{1_2, 1_3, \dots, 1_j\}$  since the graph is 2-regular.

Suppose  $h_1$  and  $h_2$  are from the set of possible values for  $t_j$ . As mentioned in Chapter 1,  $\Pr(t_j = h_1) = \Pr(t_j = h_2)$  due to symmetry. One way to argue this more explicitly is to consider a function from the graphs of  $\mathcal{K}_{n,2,g}$  that contain a walk along  $i_1, i_2, \dots, i_j, h_1$  to graphs that contain a walk along  $i_1, i_2, \dots, i_j, h_2$

2.2. The Probability of a Coincidence in  $\mathcal{K}_{n,2,g}$

---

by swapping the labels on vertices  $h_1$  and  $h_2$ . This function is a bijection since  $\mathcal{K}_{n,2,g}$  is closed under , implying  $\Pr(t_j = h_1) = \Pr(t_j = h_2)$ . Since  $t_j$  is equally likely to be any vertex that is not a coincidence we have

$$\Pr(\text{coin}(t_j) + (n - j)\Pr(t_j = h_1) = 1.$$

We can use a similar bijective argument to compare the number of graphs where  $t_j = i_1$  and  $t_j = h_1$ . Let function  $f$  map graphs with a simple cycle  $(i_1, i_2, \dots, i_j, h_1, h_2, \dots, h_k)$  to graphs with the same structure except that the cycle is replaced by the simple cycles  $(i_1, i_2, \dots, i_j), (h_1, h_2, \dots, h_k)$ . Now consider the domain of  $f$  as all the graphs of  $\mathcal{K}_{n,2,g}$  that have the walk  $(i_1, i_2, \dots, i_j, h_1)$ . Then the codomain is all the graphs in  $\mathcal{K}_{n,2,g}$  with the simple cycle  $(i_1, i_2, \dots, i_j)$  as well as all graphs in  $\mathcal{G}_{n,2}$  with the simple cycle  $(i_1, i_2, \dots, i_j)$  and no cycles of length less than  $g$  except for the cycle that contains  $h_1$ . Also we see that over this domain and codomain, our function is a bijection.

An immediate result of this is that  $t_j$  has a higher probability of being any specified vertex that is not a coincidence than being a coincidence. We can specify the probability of either of these events. Let  $T(n, j, g)$  be the number of graphs in  $\mathcal{G}_{n,2}$  with the simple cycle  $(i_1, i_2, \dots, i_j)$  and no cycles of length less than  $g$  except for the cycle that contains  $h_1$ . Define

$$s = \frac{k_{n,2,g}}{(n-1)(n-2)\dots(n-j)},$$

the number graphs in  $\mathcal{K}_{n,2,g}$  with the walk  $(i_1, i_2, \dots, i_j)$ . Taking the equality implied by  $f$  being a bijection and dividing by  $s$  implies

$$\Pr(t_j = h_1) = \Pr(\text{coin}(t_j)) + T(n, j, g)/s. \quad (2.5)$$

If we let  $l$  be the length of the cycle of length less than  $g$ , we can count the number of ways to make the cycle shorter than  $g$  and we already know the number of ways to make the rest of the graph is  $k_{n-l-j+1,2,g}$ . So we have

$$T(n, j, g) = \sum_{l=1}^{g-1} (n-j)_{(l-1)} k_{n-l-j+1,2,g} \quad (2.6)$$

where  $n_{(l)}$  is defined as  $n(n-1)(n-2)\dots(n-l+1)$  for  $l > 0$  and  $n_{(0)} = 1$ . Using 2.4 to substitute in for  $k_{n-l-j+1,2,g}$  results in

$$T(n, j, g) = (n-j)! \left( e_g(g-1) + \sum_{l=1}^{g-1} E_{n-l-j+1} \right) \quad (2.7)$$

where  $e_g = e^{-1-1/2-\dots-1/(g-1)}$ . Using 2.7 to substitute in for 2.5 and also substituting out the  $s$  term results in

$$\Pr(t_j = h_1) = \Pr(\text{coin}(t_j)) + \Theta(g). \quad (2.8)$$

Here we assumed that  $j \in O(\log^i)$  for some power  $i$  since we only consider walks of such length for the trace method. This allows the error terms to become negligible.

## Chapter 3

# Random Walks in $\mathcal{K}_{n,d,g}$

Now we no longer restrict our discussion to the 2-regular case. This chapter will show efforts to use methods similar to those in the last chapter in order to find the probability of a coincidence in  $\mathcal{K}_{n,d,g}$ . And similarly to the last chapter we will first consider the size of  $\mathcal{K}_{n,d,g}$  using an inclusion-exclusion argument. Several nice cancellations do not occur in the following inclusion-exclusion argument, and as a result we do not achieve a useful estimate of the size of  $\mathcal{K}_{n,d,g}$ .

### 3.1 The Size of $\mathcal{K}_{n,d,g}$

Previously we found the size of the set of all graphs in  $\mathcal{G}_{n,4}$  with girth less than  $g$  in order to find the size of  $\mathcal{K}_{n,2,g}$ . In the terminology of the last chapter, this set is equivalent to  $\bigcup_{s < g, I} C_{s,I}$ . We recall  $C_{s,I}$  was the set of graphs in  $\mathcal{G}_{n,4}$  which had a simple cycle along the length  $s$  sequence of vertices  $I$ . Finding the size of  $C_{s,I}$  was relatively straightforward, as was finding the size of any intersection of such set. This allowed us to use an inclusion-exclusion argument to find the size of  $\bigcup_{s < g, I} C_{s,I}$ .

For the  $d$ -regular case, instead let us consider the set of graphs in  $\mathcal{G}_{n,d}$  that have length  $s$  simple cycle along a sequence  $I$  of vertices and along the edges represented by a reduced word  $w \in \Pi^s$ . We call this set  $C_{s,I,w}$ . A graph is in  $\mathcal{G}_{n,d}$  but not in  $\mathcal{K}_{n,d,g}$  if and only if it the graph is a member of  $C_{s,I,w}$  for some  $I$  and  $w$  and  $s < g$ . There are  $(n!)^{\frac{d}{2}}$  graphs in  $\mathcal{G}_{n,d}$  since any graph from this probability space is uniquely defined by  $d/2$  permutations of size  $n$ . So if we let  $k_{n,d,g}$  be the number of graphs in  $\mathcal{K}_{n,d,g}$  then we have

$$k_{n,d,g} = (n!)^{\frac{d}{2}} - \left| \bigcup_{s < g, I, w} C_{s,I,w} \right|.$$

The first term of our inclusion-exclusion argument is then  $\sum_{s < g, I, w} |C_{s,I,w}|/2$ . We have to divide by two, since summing over all  $w$  and  $I$  means we would count every cycle twice. This happens because  $w$  and  $I = (i_1, i_2, \dots, i_s)$  represents a cycle traversed in a particular direction and  $w^{-1}$  and  $(i_1, i_s, i_{s-1}, \dots, i_2)$  represent the very same cycle but traversed in the opposite direction.

Once again, we let  $a_j(w)$  be the number of occurrences of  $\pi_j$  and  $\pi_j^{-1}$  in  $w$ .

### 3.1. The Size of $\mathcal{K}_{n,d,g}$

---

Then we have that

$$|C_{s,I,w}| = (n!)^{\frac{d}{2}} \prod_{j=1}^d \frac{1}{n_{(a_j(w))}}.$$

This previous inequality is independent of  $I$  as before. So summing  $|C_{s,I,w}|$  over all possible  $I$  is equal to multiplying  $|C_{s,I,w}|$  by the number of ways to choose  $s$  elements from a set of  $n$  elements and arrange them in a cycle, which is

$$\binom{n}{s} (s-1)! = \frac{n(s)}{s}.$$

If  $w = \pi^s$  for some character  $\pi$ , as it was for for all words in the  $d = 1$  case, then

$$|C_{s,I,w}| \frac{n(s)}{s} = (n!)^{\frac{d}{2}} \frac{1}{s}.$$

If  $w$  is composed of several different characters, then we do not have as much cancelation and instead

$$|C_{s,I,w}| \frac{n(s)}{s} = (n!)^{\frac{d}{2}} \frac{n(s)}{s} \prod_{j=1}^d \frac{1}{n_{(a_j(w))}}.$$

Now we begin to see how finding the size of  $\mathcal{K}_{n,d,g}$  using inclusion-exclusion differs significantly from finding the size of  $\mathcal{K}_{n,2,g}$  similarly. Each term of our inclusion-exclusion argument when finding the size of  $\mathcal{K}_{n,2,g}$  was  $n!$  times a sum of constants with respect to  $n$ . Now instead we have  $n!^{\frac{d}{2}}$  times a sum of rational functions with respect to  $n$  as the first term. So if our inclusion-exclusion argument implies that  $k_{n,d,g}$  approaches a Taylor series as  $n$  goes to infinity, the function which that series represent will have to be a function on  $n$ . This contrasts with the  $k_{n,2,g}$  which approached the Maclaurin series for  $e^{x_1+x_2+\dots+x_{g-1}}$ .

Now let us consider the size of  $\cup_{i=1}^r C_{s_i, I_i, w_i}$ , where the  $I_i$  are length  $s_i$  sequences of integers from  $\{1, 2, \dots, n\}$  and  $w_i$  are length  $s_i$  words from  $\cdot$ . Clearly this set is the set of all graphs in  $\mathcal{K}_{n,d,g}$  that contain all of the simple cycles represented by the  $I_i$  and  $w_i$ . When we considered the  $d = 1$  case, this intersection was the empty set if any pair of the cycles both shared the same vertex. Now this is no longer the case, as cycles may share vertices and edges.

We can implement the previously used notion of a generalized form applied to the new setting of unions of simple cycles. Let  $I = (I_1, I_2, \dots, I_r)$  and  $W = (w_1, w_2, \dots, w_r)$ . Then define the generalized form  $\Gamma_{I,W}$  to be the vertex-labelled and edge-labelled graph that is composed of the simple cycles represented by  $I$  and  $W$ . So now we have that  $|\cup_{i=1}^r C_{s_i, I_i, w_i}|$  is the number of graphs in  $\mathcal{K}_{n,d,g}$  that contain  $\Gamma_{I,W}$  as a subgraph.

Clearly if any vertex in  $\Gamma_{I,W}$  is incident upon two edges with the same label from  $\Pi$  then no graphs in  $\mathcal{K}_{n,d,g}$  contain  $\Gamma_{I,W}$ . Else, we can count the number of graphs in  $\mathcal{K}_{n,d,g}$  containing  $\Gamma_{I,W}$  by considering the probability that every

### 3.1. The Size of $\mathcal{K}_{n,d,g}$

---

edge of  $\Gamma_{I,W}$  occurs in the graph. This yields

$$\left| \bigcup_{i=1}^r C_{s_i, I_i, w_i} \right| = (n!)^{\frac{d}{2}} \prod_{j=1}^{d/2} \frac{1}{n_{(a_j(\Gamma_{I,W}))}}. \quad (3.1)$$

Once again this is independent of the labels on  $\Gamma_{I,W}$ 's vertices. Let  $\Gamma_W$  be the graph  $\Gamma_{I,W}$  without labels on its vertices and let  $C(\Gamma_W) = \left| \bigcup_{i=1}^r C_{s_i, I_i, w_i} \right|$ . Note the definition of  $C(\Gamma_W)$  makes sense since it is independent of any specific labeling on the vertices which is determined by  $I$ . Also define  $X_r$  as the set of graphs that correspond to some  $\Gamma_W$  that is determined by  $r$  simple cycles. Then the  $r$ -th term of the inclusion-exclusion argument is

$$\sum_{\Gamma_W \in X_r} C(\Gamma_W) \omega(\Gamma_W) (-1)^{r+1} \quad (3.2)$$

where  $\omega(\Gamma_W)$  is the number of ways to assign labels to the vertices from  $\{1, 2, \dots, n\}$ .

Now we again consider a fixed  $\Gamma_W$  in order to describe  $\omega(\Gamma_W)$ . If  $v$  is the number of vertices in  $\Gamma_W$  then  $\omega(\Gamma_W)$  is equivalent to the number of ways to select  $v$  vertices from a set of  $n$  vertices,  $\binom{n}{v}$ , times the number of ways to assign labels to the vertices of  $\omega(\Gamma_W)$  from a set of  $v$  labels. So clearly

$$\omega(\Gamma_W) \geq \binom{n}{v}.$$

There are  $v!$  ways to assign  $v$  labels to  $v$  vertices, but some of these labelings of  $G_w$  may result in equivalent graphs. Hence

$$\omega(\Gamma_W) \leq \binom{n}{v} v! = n_{(v)}.$$

There are cases where this upper bound does equal  $\omega(\Gamma_W)$ . Namely this happens if and only if the only automorphism of  $\Gamma_W$  is the identity map.

Define  $e$  as the number of edges in  $\Gamma_W$ . Since  $e = \sum_1^{d/2} a_j(\Gamma_W)$  we can rewrite 3.1 as

$$C(\Gamma_W) = (n!)^{\frac{d}{2}} O(n^{-e}).$$

Using this and our bounds on  $\omega(\Gamma_W)$  to perform gives

$$C(\Gamma_W) \omega(\Gamma_W) = (n!)^{\frac{d}{2}} O(n^{v-e}). \quad (3.3)$$

Note that if the simple cycles that compose  $\Gamma_W$  do not share any vertices then  $v - e = 0$ , else  $v - e < 0$ . So although in the  $d$ -regular case, as opposed to the 2-regular case considered in the previous chapter, our inclusion-exclusion argument must include overlapping simple cycles, the non-overlapping simple cycles contribute far more to the inclusion sum.

### 3.2 The Probability of a Coincidence in $\mathcal{K}_{n,d,g}$

We want to use a similar bijective argument as in the previous chapter in order to estimate the probability that the next vertex in a specified walk in a random graph from  $\mathcal{K}_{n,d,g}$  is equal to a particular vertex. Consider a walk in a graph from  $\mathcal{K}_{n,d,g}$  starting at a vertex  $i_1$  and going along vertices  $i_2, i_3, \dots, i_j$  in that order. These vertices are not all necessarily different. Let  $w \in \Pi^j$  be the word that represents the edges along this walk. Also let  $t_j$  be a random variable denoting the  $j+1$ -th vertex along the walk and suppose that  $t_j$  is a free choice. Let  $\Gamma_{I,w}$  represent the walk so far. For  $1 \leq k < j$  we have  $\Pr(t_j = i_k) = 0$  if  $i_k$  is at a distance in  $\Gamma_{I,w}$  less than  $g-1$  from  $i_j$ . The set of vertices that  $t_j$  may equal if  $t_j$  is a coincidence is not just  $i_1$  as it was in the 2-regular case though.

Let  $\pi$  be the final character of  $w$ , and so  $\pi$  is also the permutation that gives rise to the final edge in our walk  $(i_j, t_j)$ . Also let  $i$  be a vertex in  $\Gamma_{I,w}$  such that the distance in  $\Gamma_{I,w}$  between  $i$  and  $i_j$  is at least  $g-1$ . In this section we are going to compare the probability that  $t_j = i$  with the probability that  $t_j = h$  for some  $h \notin V(\Gamma_{I,w})$ . Given vertices  $h_1$  and  $h_2$  not in  $V(\Gamma_{I,w})$  then  $t_j$  is equally likely to be either of those vertices due to symmetry. So if  $h \notin V(\Gamma_{I,w})$  and  $V = |V(\Gamma_{I,w})|$  then

$$1 = (n - V)\Pr(t_j = h) + \sum_{u \in V(\Gamma_{I,w})} \Pr(t_j = u).$$

One difference from the two regular case is that  $i_j$  and  $i$  are not necessarily in the same cycle of  $\pi$ , in which case clearly  $\Pr(t_j = i) = 0$ . Another difference is that  $t_j = i$  does not imply that  $h$  and  $i$  are in different cycles in  $\pi$ . Because of these two differences, we need to modify the bijective function considered in the previous chapter so that it can map between all the cases where  $t_j = i$  to all cases where  $t_j = h$ .

Our new bijective function  $f$  will map between graphs in  $\mathcal{G}_{n,d}$  with the walk  $\Gamma_{I,w}$  as a subgraph and  $t_j = h$  to the graphs in  $\mathcal{G}_{n,d}$  with the walk  $\Gamma_{I,w}$  as a subgraph and  $t_j = i$ . Since  $t_j$  is determined by  $\pi$  we will have to alter this permutation in some way using  $f$ . Given a  $G$  graph in the domain of  $f$  define  $f(G)$  as the graph created by deleting the edges  $(i_j, h)$  and  $(\pi^{-1}(i), i)$  which are labelled by  $\pi$  and replacing them by edges labeled by edges labeled by  $\pi$   $(i_j, i)$  and  $(\pi^{-1}(i), h)$ . Note that when we delete (resp. replace) an edge  $(x, y)$  labeled by  $\pi$  I mean we delete (resp. replace) the directed edge  $(x, y)$  labeled by  $\pi$  as well as the directed edge  $(x, y)$  labeled by  $\pi^{-1}$ .

The function  $f$  is bijection over the given domain and codomain, but we will need to restrict the domain and codomain to smaller sets so that we can compare the size of the set  $H$  of graphs in  $\mathcal{K}_{n,d,g}$  with the walk  $\Gamma_{I,w}$  as a subgraph and  $t_j = h$  to the size of the set  $J$  of graphs in  $\mathcal{K}_{n,d,g}$  with the walk  $\Gamma_{I,w}$  as a subgraph and  $t_j = i$ . We are interested in the sizes of  $H$  and  $J$  since if  $G(\Gamma_{I,w})$  is the number of graphs in  $\mathcal{K}_{n,d,g}$  that contain the walk  $\Gamma_{I,w}$  then

$$|H|/G(\Gamma_{I,w}) = \Pr(t_j = h) \quad \text{and}$$

### 3.2. The Probability of a Coincidence in $\mathcal{K}_{n,d,g}$

---

$$|J|/G(\Gamma_{I,w}) = \Pr(t_j = i). \quad (3.4)$$

If  $H$  is our domain  $f$  maps to some of  $J$  as well as to some graphs that are not in  $\mathcal{K}_{n,d,g}$ . Due to the definition of  $f$ , graph  $G \in H$  is mapped to a graph not in  $\mathcal{K}_{n,d,g}$  if and only if  $i_j$  and  $i$  or  $\pi^{-1}(i)$  and  $h$  are a distance less than  $g - 1$  apart. The reason this is the one and only way that  $f(G) \notin \mathcal{K}_{n,d,g}$  is that deleting edges does not add any new small cycles, and any new cycles must use the edges that  $f$  adds to  $G$ . Let  $H' \subset H$  be the set of such graphs.

To identify the graphs in  $J$  that  $f$  does not map to it is helpful to first consider  $f^{-1}$ . If we call the set of such graphs  $J'$  then  $J'$  is the set of graphs  $j \in J$  such that  $f^{-1}(j) \notin H$ . Examining  $f$  shows that  $f^{-1}(j)$  is the graph made by deleting edges  $\pi(i_j, i)$  and  $(\pi^{-1}(h), h)$  labeled by  $\pi$  and replacing them with edges  $(i_j, h)$  and  $(\pi^{-1}(h), i)$  labeled by  $\pi$ . So by the same argument that allowed us to describe  $H'$ , we have that  $J'$  is the set of graphs in  $J$  where the distance between  $i_j$  and  $h$  or the distance between  $\pi^{-1}(h)$  and  $i$  is less than  $g - 1$ . We know  $f$  is a bijection between  $H - H'$  and  $J - J'$  and so we have

$$|H| - |J| = |H'| - |J'|.$$

So by 3.4

$$\Pr(t_j = h) - \Pr(t_j = i) = \frac{|H'| - |J'|}{G(\Gamma_{I,w})}.$$

Unfortunately we do not yet know even the sign of  $|H'| - |J'|$  so the above formula is not too useful in comparing  $\Pr(t_j = h)$  and  $\Pr(t_j = i)$ . A more useful comparison might be made if we could somehow modify our function  $f$  and create a new function  $b$  that maps from all of  $J$  to graphs in  $\mathcal{K}_{n,d,g}$ . The reason  $f^{-1}$  did not do this was because some graphs in  $J$  had a short path between a pair of vertices where our function introduces a new edge, thereby creating a cycle of length less than  $g$ . Our function  $b$  will choose which edges to remove and replace in order to avoid this, but we do this at the cost that the graphs in our codomain will not all have  $\pi(i_j)$  equal to the same vertex.

There are at most  $1 + d(d - 1)^{g-1}$  vertices that are at a distance less than  $g + 1$  from any given vertex in a  $d$ -regular graph. So there are at least  $n - (1 + d(d - 1)^{g-1} + \Gamma_{I,w})$  vertices that are at least distance  $g + 1$  away from  $i$  and not in  $\Gamma_{I,w}$ . Call the set of such vertices  $D(G)$ . Suppose  $g$  is small enough so that  $n - (1 + d(d - 1)^{g-1} + \Gamma_{I,w}) > 0$ . Let  $v$  be the vertex in  $D$  with the smallest label. Now given  $G \in J$ ,  $b(G)$  will be  $G$  with the  $\pi$ -labeled edges  $(i_j, i)$  and  $(\pi^{-1}(v), v)$  deleted and replaced by  $\pi$ -labeled edges  $(i_j, v)$  and  $(\pi^{-1}(v), i)$ . Note that  $\pi^{-1}(v)$  and  $i$  are at least distance  $g - 1$  apart, since if not  $i$  and  $v$  would be distance less than  $g$  apart. So both new edges of  $b(G)$  do not result in any cycles of length smaller than  $g$ .

The function  $b$  is an injective map from  $J$  to the set  $N$  of graphs in  $\mathcal{K}_{n,d,g}$  that contain  $\Gamma_{I,w}$  and  $t_j$  is not a coincidence. We can define  $b^{-1}$  as the function on  $N$  that deletes  $\pi$ -labeled edges  $(i_j, v)$  and  $(\pi^{-1}(i), i)$  and replaces them with  $\pi$ -labeled edges  $(i_j, i)$  and  $(\pi^{-1}(i), v)$ . Let  $v_1, v_2, \dots, v_{n-|\Gamma_{I,w}|}$  be the vertices not in  $\Gamma_{I,w}$  in order from smallest to largest. Then  $G \in N$  is not in the range

### 3.2. The Probability of a Coincidence in $\mathcal{K}_{n,d,g}$

---

of  $b$  if  $\pi(i_j) = v_r$  and there exists some  $v_s$  at a distance at least  $g + 1$  from  $i$  in  $g^{-1}(G)$  such that  $s < r$ . Define  $N_r$  as the set of all graphs in  $N$  such that  $\pi(i_j) = v_r$  or  $v_r$  is not a distance at least  $g + 1$  from  $i$  in  $g^{-1}(G)$ . Given an uniformly chosen random graph in  $N_r$  we denote the event that the graph is in the range of  $b$  by  $\text{Range}(r)$ . So since  $b$  is a bijection from  $J$  if we define the codomain as the range we have that

$$|J| = \sum_{r=1}^{n-|V(\Gamma_{I,w})|} |N_r| \Pr(\text{Range}(r)). \quad (3.5)$$

Note that for  $0 < s \leq n - |V(\Gamma_{I,w})|$  we have that  $|N_r| = |N_s|$  since the probability that  $t_j$  is  $v_r$  or  $v_s$  is equally likely as neither of these events are coincidences. So if we divide 3.5 by  $G(\Gamma_{I,w})$  on both sides we find that

$$\Pr(t_j = i) = \Pr(t_j = v_1) \sum_{r=1}^{n-|V(\Gamma_{I,w})|} \Pr(\text{Range}(r)). \quad (3.6)$$

So  $\Pr(\text{Range}(r))$  is at most the probability that a random  $G \in N_r$  has  $\{v_1, v_2, \dots, v_{r-1}\} \subset D(b^{-1}(G))$  and  $v_r \notin D(b^{-1}(G))$  since only such graphs can be in the range. Due to symmetry this equals the probability that a random  $G \in N_1$  has  $\{v_r, v_2, \dots, v_{r-1}\} \subset D(b^{-1}(G))$  and  $v_1 \notin D(b^{-1}(G))$ .

We believe that eventually we will be able to understand  $\Pr(\text{Range}(r))$  well enough to find useful comparisons between  $\Pr(t_j = i)$  and  $\Pr(t_j = v_1)$ . In particular we suspect that  $\Pr(t_j = i) < \Pr(t_j = v_1)$  as it was in the 2-regular case. A heuristic reason for this is that a walk with more coincidences implies more vertices are fairly close to one another and this in turn implies fewer possible paths between vertices and thus fewer graphs containing that walk.

## Chapter 4

# Conclusion

This research examined random walks in a new probability space  $\mathcal{K}_{n,d,g}$ . Understanding how these walks behave is the basis for the trace method, which gives upper bounds on the expected second eigenvalue. We believe that the trace method when applied over a graph probability space with girth suitably high may result in a proof that most graphs in the probability space are Ramanujan even if we do not adjust the trace method as in [6] to account for tangles. The reason for this is that tangles are certain subgraphs that contain cycles, and so for girth sufficiently high any given tangle can not exist in our graph. We were able to find the probabilities that the next vertex in some given walk returns to a previous vertex in the walk or not in the 2-regular case. We only have weak results in this direction though in the  $d$ -regular case. Further research is required on this model still in order to use the trace method on it.

There are different random graph models than  $\mathcal{K}_{n,d,g}$  that still avoids graphs containing tangles and thus would not require a selective trace. Studying such models may be beneficial since in one of them the analysis of random walks might be straight-forward and the trace method could be readily applied. We note that by specifying the girth, our random model  $\mathcal{K}_{n,d,g}$  possibly ignores several graphs that do not contain tangles. For example, a single self-loop does not indicate a graph contains a cycle, although a single vertex with enough self loops is a tangle [6]. Any model that avoids tangles will have to do some conditioning on small cycles in the graph.

# Bibliography

- [1] A. Amit, S. Hoory and N. Linial. A continuous analogue of the girth problem. *Jour. Comb. Th. ser. B*, 84: 340 - 363, 2002.
- [2] Norman Biggs. Girth, valency and excess. *Linear Algebra and Applications*, 31: 55-59, 1980.
- [3] Andrei Broder and Eli Shamir. On the second eigenvalue of random regular graphs. In *28th Annual Symposium on Foundations of Computer Science*, 286294, 1987.
- [4] Joel Friedman. On the second eigenvalue and random walks in  $d$ -regular graphs. *Combinatorica*, 11(4):331-362, 1991.
- [5] Joel Friedman. Relative expanders of weakly relatively Ramanujan graphs. *Duke Math. J.*, 118(1):19-35, 2003.
- [6] Joel Friedman. A proof of alon's second eigenvalue conjecture. Proceedings of the thirty-fifth annual ACM symposium on Theory of computing, June 09-11, 2003, San Diego, CA, USA.
- [7] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bull. AMS*, 43(2006) 439–561.
- [8] N. Linial and D. Puder. Word maps and spectra of random graph lifts. *Random Structures and Algorithms*, accepted for publication.
- [9] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261-277, 1988.
- [10] B. McKay. The expected eigenvalue distribution of a large regular graph. *Lin. Alg. Appl.*, 40:203-216, 1981.
- [11] A. Nilli. On the second eigenvalue of a graph. *Discrete Math.*, 91(2):207-210, 1991.
- [12] O. Reingold. Undirected st-connectivity in log-space. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pages 376-385, 2005.