Part III Essay: Ramanujan Graphs

September 6, 2016

1 Introduction

In this essay we present background, history and importance of Ramanujan graphs focusing on the recent developments made by Marcus, Spielman and Srivastava in [1]. We present a proof of their result about existence of infinite families of Ramanujan graphs of every degree.

In Section 2 we present preliminary results and motivation for seeking Ramanujan graphs, by the means of showing which properties they satisfy. Here, we also present a novel idea for constructing Ramanujan graphs, which provides several classes of small examples, we use this idea to illustrate the issues we need to deal with when seeking Ramanujan graphs.

In Subsection 2.1 we start by introducing basic concepts and prove a series of basic results we will use throughout the essay. We also comment on the relation of this results to the largest eigenvalues of a graph.

In Subsection 2.2 we introduce the concept of the second largest eigenvalue and link it to various important properties of the graph. We present one such, in particular we improve on a result by Chung [3] linking the second eigenvalue to the diameter of the graph.

In Subsection 2.3 we present the result of Alon [7] showing that in some sense Ramanujan graphs are best possible in terms of second eigenvalue size for large values of n.

In subsection 2.4 we present our own work on construction of small Ramanujan graphs, having the goal of presenting the problem and pointing out the issues with constructions.

In Section 3 we present the result of Marcus, Spielman and Srivastava from [1] and develop several concepts instrumental in the proof. We start this section with an overview of the proof, as we believe it would be easier to follow after considerations of Section 2.

2 Eigenvalues of a Graph

2.1 Preliminaries

Throughout this essay $G \equiv G(V, E)$ will denote a simple graph with vertex set $V(G)$ and edge set $E(G)$. We note that by a simple graph we denote a graph which has no loops or multiple edges joining the same pair of vertices.

The *adjacency matrix* of a graph G on n vertices indexed by the set $[n] = \{1, 2, \cdots, n\}$ is an $n \times n$ matrix A with entries in $\{0, 1\}$ where $A_{ij} = 1$ if and only if there is an edge between vertices i and j.

Adjacency matrix contains information about various important properties of the graph. For example, its k-th power counts the number of walks of length k in the graph. Furthermore, we notice that when we defined the adjacency matrix we needed to index the vertices, which means that adjacency matrix of a graph is only specified up to permutations of rows and columns. This motivates the importance of matrix invariants under these permutations, some of the most important of which are its eigenvalues.

We note that $\det(\lambda I - A)$ is invariant under a transposition of indices of a pair of vertices as this corresponds to one swap of rows and one corresponding swap of columns, each only changing the sign of the determinant. This shows the eigenvalues of the adjacency matrix are invariant under the choice of indexing and as such are a property of the graph itself. The set of eigenvalues of a graph are called its *spectrum*. The relation between eigenvalues and eigenvectors with various properties of a graph is the subject of study of spectral graph theory.

The adjacency matrix of a graph is a real symmetric matrix, as such all its eigenvalues are real. Furthermore, we can find an orthonormal basis of real eigenvectors, as proved in [2]. This implies that the geometric multiplicities match the algebraic and each eigenspace is of full rank.

A graph is *regular* if every vertex has the same number of neighbours, we say it is *d-regular* if this number equals d. We notice that a graph is regular if and only if $(11 \cdots 1)$ is an eigenvector of the adjacency matrix, where the graph is d-regular if the corresponding eigenvalue equals d, this is a simple example of a spectral result.

In the following lemma we turn to the significance of the largest, in absolute value, eigenvalue. We give elementary, so slightly longer, proofs using several different ideas in an attempt to provide basic motivation and intuition for the following sections.

Lemma 2.1. Given a d-regular graph G, with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$.

a) $d = \lambda_1$

b) Let k be the largest integer such that $\lambda_k = d$. Then k equals the number of connected components of G.

c) $\lambda_n \geq -d$ with equality if and only if G is bipartite.

d) If G is bipartite λ and $-\lambda$ have matching multiplicities for any real number λ.

Proof.

a) Let $\overline{v} = (v_1 \cdots v_n)$ be a non-zero eigenvector of an eigenvalue λ, let us assume without loss of generality that v_1 maximizes $|v_i|$, as v is non-zero this implies $|v_1| > 0$. Let $N(i)$ denote the set of neighbours of i, note that the graph is d-regular so $|N(i)| = d$. Then

$$|\lambda_1 v_1| = |(A\overline{v})_1| = \left| \sum_{i \in N(1)} v_i \right| \leq \sum_{i \in N(1)} |v_i| \leq |N(1)||v_1| = d|v_1| \tag{1}$$

where we used the triangle inequality. This implies $|\lambda_1| \leq d$. Noticing that $v_i = 1$ for all i is an eigenvector with eigenvalue d so $\lambda_1 = d$.

b) Let us first assume that the graph is connected.

If $\lambda = d$ we have equality in the above inequality, taking further, as $\overline{v}$ was an arbitrary eigenvector, that $\overline{v}$ is a real eigenvector, we get $v_1 = v_i$ for all $i \in N(1)$.

Let S be the set of vertices i with $v_i = v_1$. If $S \neq [n]$ as G is connected there is an edge ij with $i \in S$ and $j \notin S$ but then repeating the above argument for i we reach a contradiction, hence $S = [n]$. This implies that any eigenvalue of $\overline{v}$ is a multiple of $(11 \cdots 1)$ so the eigenspace of λ is 1-dimensional which combined with the fact there is a basis of eigenvectors implies multiplicity of λ is 1.

Finally, we notice that if we reorder the indices, so that the vertices in the connected components are consecutive in the indexing, the adjacency matrix becomes a block diagonal matrix with each block corresponding to the adjacency matrix of a connected component. This implies that the set of eigenvalues of G is a union, taken with multiplicity, of the eigenvalues of its connected components. This, combined with the fact a connected graph has eigenvalue d with multiplicity 1, proves the result.

c) The argument in part **a)** shows $|\lambda_n| \leq d$ so $\lambda_n \geq -d$.

If $-d$ is an eigenvalue we have equality in the inequality of part **a)**, as before we may assume $\overline{v}$ is real. This in turn implies $|v_1| = |v_i|$ for all $i \in N(1)$ and repeating the argument of part **b)** we conclude $|v_i|$ is constant for all i so the same inequality can be applied to any vertex.

Furthermore in order to have equality we need v_i to be constant on $N(j)$ for all j, in particular any 2 vertices at distance 2 have the same value of v_i. Repeated application of this implies any 2 vertices joined by a walk of even length have same v_i.

We note that v_i can not be constant for all i as then the eigenvalue would be d. Hence there are 2 values that v_i can take, namely $\pm v_1$, we claim the vertices taking the same value define the bipartitions.

Let us assume that there is an edge between i, j with $v_i = v_j$. As the graph is connected, there needs to exist a k and a path from i to k, such that $v_k \neq v_i$. If this path is even we get a contradiction to $v_i \neq v_k$, else by appending edge ij to the path, we find an even walk from j to k implying $v_j = v_k$ again giving us a contradiction. This shows that G is indeed bipartite.

If G is bipartite, taking $v_i = 1$ if i is in the first bipartition and $v_i = -1$ if it is in the second. The resulting vector is an eigenvector with eigenvalue $-d$.

d) If $\lambda = 0$ the claim is trivially true, we now assume $\lambda \neq 0$.

Indexing the vertices in a way such that members of the same bipartition are consecutive in the ordering, the adjacency matrix A will be a block matrix of the following form

$$A = \begin{pmatrix} 0 & X \\ X^t & 0 \end{pmatrix}$$

Where X is an $a \times b$ matrix where a, b are the sizes of the bipartitions. If $\overline{v} = \begin{pmatrix} \overline{u} \\ \overline{w} \end{pmatrix}$ is an eigenvector, where $u \in \mathbb{R}^a$ and $w \in \mathbb{R}^b$.

$$A\overline{v} = \begin{pmatrix} 0 & X \\ X^t & 0 \end{pmatrix} \begin{pmatrix} \overline{u} \\ \overline{w} \end{pmatrix} = \begin{pmatrix} X\overline{w} \\ X^t\overline{u} \end{pmatrix}$$

In particular $X\overline{w} = \lambda\overline{u}$ and $X^t\overline{u} = \lambda\overline{w}$.

Let $\overline{v}_1, \cdots, \overline{v_l}$ be a basis for the eigenspace of λ. Let $\overline{u}_i, \overline{w_i}$ be decompositions as above. Let us assume $\sum_{i=1}^{l} k_i\overline{u}_i = 0$ which gives $0 = X^t \sum_{i=1}^{l} k_i\overline{u}_i = \sum_{i=1}^{l} k_i X^t\overline{u}_i = \lambda \sum_{i=1}^{l} k_i\overline{w_i}$. As $\lambda \neq 0$ this implies $\sum_{i=1}^{l} k_i\overline{w_i} = 0$, so $\sum_{i=1}^{l} k_i\overline{v}_i = 0$ implying $k_i = 0 \, \forall \, i$. This shows $\overline{u}_i$ are linearly independent.

But $\overline{v}'_i = \begin{pmatrix} \overline{u}_i \\ -\overline{w}_i \end{pmatrix}$ is an eigenvector of $-\lambda$. As $\overline{u}_i$ are independent, so are $\overline{v'_i}$ implying that multiplicity of $-\lambda$ is larger or equal then that of λ. Repeating the argument for $-\lambda$ in place of λ we conclude they are in fact equal, completing the proof.

□

Eigenvalues $\pm d$ are often called *trivial* as the above lemma answers completely the question of how they relate to graph properties.

We notice that using the above lemma we are able to deduce certain properties of the graph simply by knowing something about its (largest) eigenvalue. For example we can read out the number of connected components or know if it is bipartite.

We also notice that by using Gaussian elimination we can find the eigenvalues in worst case complexity of $O(n^3)$. In this case both bipartiteness and number of connected components can be found faster using direct algorithms, but there are properties in which the method using eigenvalues is very useful computationally.

We also add that, in some sense, by imposing a condition of regularity we are controlling the largest eigenvalue, the above lemma simply showcases how well this control works. This motivates a natural question about how the second largest eigenvalue varies with the graph. We are specifically interested by how much it can vary and which properties of the graph can be controlled using it.

2.2 The Second Largest Eigenvalue

We start by specifying what exactly we mean by the *second largest eigenvalue*. In particular we mean the largest, in absolute value, non-trivial eigenvalue. Throughout this essay μ stands for this eigenvalue, taken to be positive if we have a choice. We will see later that this corresponds to the smallest non-zero eigenvalue of the Laplacian matrix.

Throughout this essay we will mostly work with connected graphs, but we add that most results generalise trivially to arbitrary graphs by applying the connected versions to each connected component of the graph.

A first property that we try to control using the second eigenvalue is the *diameter* of a graph. For 2 vertices $a, b \in V(G)$ we define $d(a, b)$ to be the length of a shortest path between them, where the length of a path is the number of edges it contains. We define the diameter $D(G)$ to be the maximal $d(a, b)$ over all pairs of vertices $a, b \in V(G)$.

The following theorem is a refinement of the work of Chung [3]. In particular our improvement applies when a graph is bipartite, in that case their result gives a vacuous $D(G) \leq \infty$ while our is barely weaker then the general bound.

We draw attention to the fact Chung uses a different definition of what the second largest eigenvalue is, if we index the eigenvalues so that $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$ their result applies for $|\lambda_2|$ which, for example if G is bipartite equals d and is still trivial. Our improvement lies in showing the result is still true with our meaning of the second largest eigenvalue, in particular the largest non-trivial one.

Theorem 2.2. For a d-regular connected graph on n vertices, with the second largest eigenvalue $\mu \neq 0$, we have:

$$D(G) \leq \left\lceil \frac{\log (n-1)}{\log \frac{d}{|\mu|}} \right\rceil + \mathbb{1}(G \text{ is bipartite}) \tag{1}$$

Where the indicator function $\mathbb{1}(G \text{ is bipartite}) = 1$ if G is bipartite and 0 otherwise.

Proof. Let A be the adjacency matrix, $\lambda_1, \lambda_2, \cdots \lambda_n$ its eigenvalues and $\overline{v}_1, \cdots \overline{v}_n$ the corresponding orthonormal basis of real eigenvectors.

As we mentioned before $(A^m)_{ij}$ equals the number of walks of length m between vertices i, j in the graph. So, we conclude that if $(A^m)_{ij} > 0$, for all pairs i, j, then $D(G) \leq m$.

Using this observation the idea of the proof is to diagonalise the adjacency matrix, so we can take its power and split it into the part dominated by the trivial eigenvalues and the remainder and finally bound the remainder in terms of μ, as $|\mu| < d$ we notice that powers of eigenvalue d will dominate and hence, ensure positivity of entries for moderately large m.

We can write $A = \sum_{i=1}^{n} \lambda_i \overline{v}_i \overline{v}_i^t$ and $A^m = \sum_{i=1}^{n} \lambda_i^m \overline{v}_i \overline{v}_i^t$

- Let us assume the graph is not bipartite.

  Hence, there is a unique trivial eigenvalue $\lambda_1 = d$ with corresponding eigenvector $\overline{v}_1 = \frac{1}{\sqrt{n}}(11\cdots 1)^t$. Also $|\lambda_i| \leq |\mu|$ for all $i > 1$. This implies:

$$\begin{aligned}
(A^m)_{rs} &= \sum_{i=1}^{n} \lambda_i^m \left(\overline{v}_i \overline{v}_i^t\right)_{rs} \\
&= \frac{d^m}{n} + \sum_{i>1} \lambda_i^m (\overline{v}_i)_r (\overline{v}_i)_s \\
&\geq \frac{d^m}{n} - \left| \sum_{i>1} \lambda_i^m (\overline{v}_i)_r (\overline{v}_i)_s \right| \\
&\geq \frac{d^m}{n} - |\mu|^m \sum_{i>1} |(\overline{v}_i)_r| |(\overline{v}_i)_s| \\
&\geq \frac{d^m}{n} - |\mu|^m \left(\sum_{i>1} |(\overline{v}_i)_r|^2 \right)^{1/2} \left(\sum_{i>1} |(\overline{v}_i)_s|^2 \right)^{1/2} \\
&= \frac{d^m}{n} - |\mu|^m \left(1 - |(\overline{v}_1)_r|^2\right)^{1/2} \left(1 - |(\overline{v}_1)_s|^2\right)^{1/2} \\
&= \frac{d^m}{n} - |\mu|^m \left(1 - \frac{1}{n}\right)
\end{aligned}$$

  Here the second inequality follows by the triangle inequality and the third by an application of Cauchy-Schwartz inequality.

  By careful consideration of when the equality is possible in the above sequence of inequalities we can deduce that it is possible only if $r = s$ or $\mu = 0$. The former is irrelevant in our argument as we are only interested in paths between different vertices and the latter is excluded in the statement as it would require by one higher bound in the left hand side of (1).

  Using this, given $\left(\frac{d}{|\mu|}\right)^m \geq n - 1$, we have $(A^m)_{rs} > 0$ for all $r \neq s$ implying $D(G) \leq m$ giving us inequality (1).

- We are left with the case when G is bipartite.

  Firstly we notice that the above argument needs a slight modification. It is possible that A^m has zero entries off the diagonal for all m, in fact this happens for any non trivial bipartite graph with $n \geq 3$, as any two members of the same partition can only be connected by walks of even length while members of different bipartitions can only have walks of odd length connecting them.

  In order to tackle this, we note that if for some $m > 0$ we show that at least one of the $(A^m)_{rs} > 0$ and $(A^{m+1})_{rs} > 0$ is true, for all pairs of vertices r, s, then $D(G) \leq m + 1$.

As G is connected and bipartite from **Lemma** 1.1 we know that there are exactly 2 trivial eigenvalues, namely $\lambda_1 = d$ with eigenvector $\overline{v}_1 = \frac{1}{n}(11\cdots 1)^t$ and $\lambda_2 = -d$ with eigenvector $\overline{v_2} = \frac{1}{n}(11\cdots 1-1-1\cdots -1)^t$.

Noticing that $\lambda_2^m(v_2)_r(v_2)_s = \pm\frac{d^m}{n}$ and is always of different sign than the corresponding value for $m+1$ using the same inequalities as before we conclude that either

$$(A^m)_{rs} \geq \frac{2d^m}{n} - |\mu|^m\left(1-\frac{2}{n}\right)$$

or

$$(A^{m+1})_{rs} \geq \frac{2d^{m+1}}{n} - |\mu|^{m+1}\left(1-\frac{2}{n}\right).$$

So given $\left(\frac{d}{|\mu|}\right)^m \geq n-1$ we know that $\left(\frac{d}{|\mu|}\right)^{m+1} \geq \left(\frac{d}{|\mu|}\right)^m \geq n-1 > \frac{n-2}{2}$ so the inequality 1 holds in this case as well. We add that the indicator function in the bipartite case comes in as we deduce that $D(G) \leq m+1$ rather than just m.

□

Remark For completeness, the condition $\mu = 0$ is possible only if the graph is complete or complete bipartite (with same size bipartitions) in which cases the diameter is trivially 1 and 2 respectively.

This theorem shows that by having a small second eigenvalue we can bound the diameter of the graph. We mention several other properties that can be controlled by the second eigenvalue. For example how good an expander a graph is, in the sense of the following theorem[1]. The following theorem is due to Tanner [5], its proof is very similar to the one given above so we omit it here.

Theorem 2.3. Given a d-regular graph with second largest eigenvalue μ. Given a subset of vertices $X \subset V(G)$, let $N(X)$ denote the neighbourhood of X then:

$$N(X) \geq \frac{d^2|X|}{(d^2-\mu^2)|X|/n + \mu^2}$$

Proof. We point the reader to [5], or for more details to the work of Alon and Millman [4]. □

There are several other important properties which can be controlled by $|\mu|$ in a similar fashion. For example the smaller $|\mu|$ is, the graph can be made a better expander, magnifier, to have large girth (in some specific cases), concentrator, for detailed expositions of these results we point the reader to [4], [5], [6].

We add that each of the above mentioned properties are desirable in certain constructions of various types of networks, for example having a short diameter would ensure that information can travel between any 2 clients via a short path, expansion property can be

[1] We add that there are several different definitions of expander graphs depending very much on the circumstances, most are very much positively correlated or even equivalent.

used to ensure redundancy or increased bandwidth. So it makes a lot of sense to attempt to construct graphs with small second eigenvalue.

2.3 Lower Bound

We start by presenting a result of Alon [7] which shows that we can not get $|\mu|$ to be too small as n increases. Following [7] we actually show a slightly stronger statement, in particular the second largest positive eigenvalue can not be too small.

Theorem 2.4. Given a d-regular connected graph G with diameter $D(G) \geq 2k+2$. If $d = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ are the eigenvalues of G, then we have:

$$\lambda_2 \geq 2\sqrt{d-1} - \frac{2\sqrt{d-1}-1}{k}.$$

Proof. For simplicity and to follow [7] we introduce the Laplacian matrix $L = dI - A$, where I is the identity matrix and A is the adjacency matrix.

We note that eigenvalues of L are given by $\mu_i = d - \lambda_i$. So in particular μ_2 is the smallest strictly positive eigenvalue of L. Let $\{\overline{v}_i\}_{i=1}^n$ be an orthonormal basis of real eigenvectors for L, which is the same as for A.

For any $\overline{x} \neq 0 \in \mathbb{R}^n$ orthogonal to $\overline{v}_1$, in other words such that $(\overline{x}, \overline{v}_1) = \overline{x}^t\overline{v}_1 = \sum_{i=1}^n x_i = 0$, let $\overline{x} = \sum_{i=1}^n a_i\overline{v}_i$ then the orthogonality condition is equivalent to $a_1 = 0$. We have

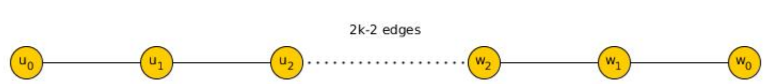
$$\frac{(L\overline{x}, \overline{x})}{(\overline{x}, \overline{x})} = \frac{\overline{x}^t L\overline{x}}{\overline{x}^t\overline{x}} = \frac{\sum_{i=1}^n \mu_i a_i^2}{\sum_{i=1}^n a_i^2} = \frac{\sum_{i=2}^n \mu_i a_i^2}{\sum_{i=2}^n a_i^2} \geq \frac{\sum_{i=2}^n \mu_2 a_i^2}{\sum_{i=1}^n a_i^2} = \mu_2 \tag{1}$$

Where the third equality follows from $a_1 = 0$ and the inequality from $\mu_2 \leq \mu_i$ for all $i \geq 2$.[2]

Based on this, our strategy is to find an $\overline{x}$ subject to above conditions for which $(L\overline{x}, \overline{x})/(\overline{x}, \overline{x})$ is small.

Let u_0, w_0 be vertices of the graph such that $d(u_0, w_0) = D(G) \geq 2k+2$. Let u_1, w_1 be vertices of the shortest path joining u_0, w_0 adjacent to u_0, w_0 respectively.

Figure 1: Definition of u_i, w_i.



We now define $U_0 = \{u_0, u_1\}$ and $W_0 = \{w_0, w_1\}$ we define U_i to be the set of vertices having shortest distance i from U_0. We define W_i analogously.

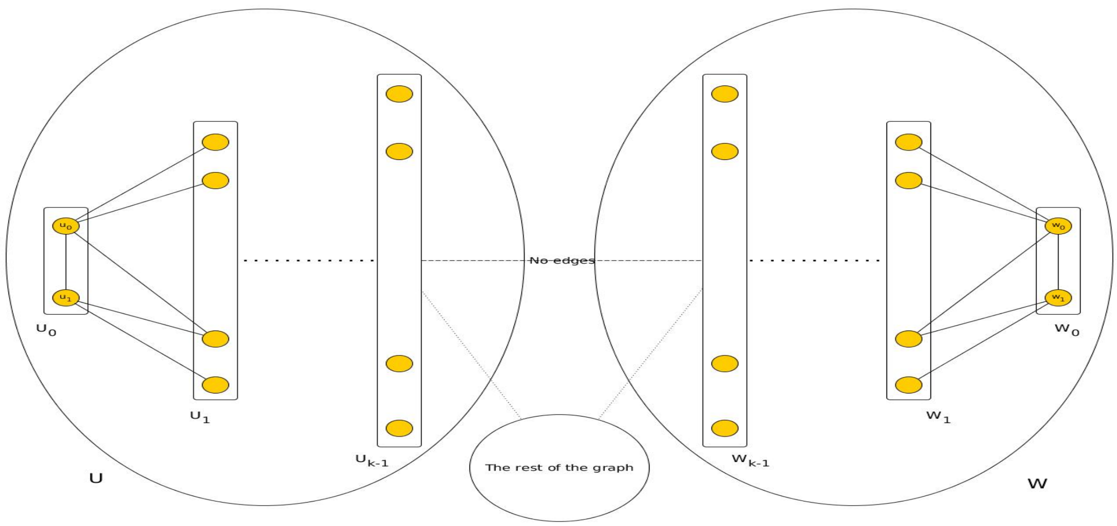
The important property we require is $|U_i| \leq (d-1)|U_{i-1}|$ for $i \geq 1$, this follows as each element of U_i needs to have at least one edge towards U_{i-1} and the graph being d regular

[2] We add that taking $\overline{x} = \overline{v_2}$ we can achieve the equality. This argument is in fact a simple case of the Courant-Fischer theorem giving an alternative, variational, definition of eigenvalues.

implies every member can add at most $d-1$ other elements to U_i. Clearly, the same inequality is true for $|W_i|$.

We also notice that $U = \bigcup_{i=0}^{k-1} U_i$ is disjoint from $W = \bigcup_{i=0}^{k-1} W_i$ and there are no edges between members of U and W, as otherwise we would have a path joining U_0 and W_0 of length at most $k-1+k-1+1 = 2k-1$ which is not possible as $d(u_0, w_0) \geq 2k+2$, so consequently $d(u_1, w_1) \geq 2k$, the following figure illustrates the setting.

Figure 2: Illustration of the problem setting.



We define $x_v = \frac{1}{(d-1)^{i/2}}$ for all $v \in U_i$ and $x_v = \frac{c}{(d-1)^{i/2}}$ for all $v \in W_i$ and all $i \leq k$, we set $x_v = 0$ for all other vertices, c here is a real parameter chosen such that $\sum_{v\in V(G)} x_v = 0$. Our plan is to use inequality (1) with this $\overline{x}$.[3] In order to do this we need to estimate $(L\overline{x}, \overline{x})$ and $(\overline{x}, \overline{x})$.

$$(\overline{x}, \overline{x}) = \sum_{v=1} n x_v^2 = \sum_{v\in U} x_v^2 + \sum_{v\in W} x_v^2 = \sum_{i=0}^{k-1} \sum_{v\in U_i} x_v^2 + \sum_{i=0}^{k-1} \sum_{v\in W_i} x_v^2 = \underbrace{\sum_{i=0}^{k-1} \frac{|U_i|}{(d-1)^i}}_{A} + c^2 \underbrace{\sum_{i=0}^{k-1} \frac{|W_i|}{(d-1)^i}}_{B}$$

Where we used that $x_v = 0$ outside $U \cup W$ and define A, B to be the summands of the right hand side.

$$(L\overline{x}, \overline{x}) = \sum_{v=1}^{n} \sum_{u=1}^{n} x_v L_{vu} x_u = \sum_{v=1}^{n} d x_v^2 - \sum_{vu\in E(G)} 2 x_v x_u = \sum_{vu\in E(G)} (x_v - x_u)^2 \tag{2}$$

[3]We attempt to give some motivation for this choice of $\overline{x}$. If we consider an infinite d-ary tree (an infinite tree in which every vertex has exactly d neighbours), which is in some sense, ignoring the fact it is infinite, an example showing the above result is tight, so in order to find finite borderline examples it makes a lot of sense to work with graphs which at least locally look like a tree. With this in mind it makes a lot of sense making $\overline{x}$ constant on each of U_i, W_j as in some sense we can not distinguish these vertices, making it 0 elsewhere should not have a large impact as by our choice of u_0, w_0 we should cover almost all the graph in these cases. Graphs looking locally like a tree should also in turn have equality in the $|U_i| \leq (d-1)|U_{i-1}|$, which is the reason we need $\sqrt{(d-1)^{-i}}$ to counter the increase in maximal number of vertices in $|U_i|$, note that we will be working with squares in the scalar products so we need to take the root.

Where we used the fact each vertex is an endpoint of exactly d edges and the factor of 2 comes in, as the summation $vu \in E(G)$ goes over the edges of G, so we need to count both uv and vu.

Now noticing that all the edges with one vertex in U_i have the second vertex in either U_{i-1}, U_i or U_{i+1} for $i \le k-2$. and that as we argued before there are no edges between U_{k-1} and W_{k-1} we get that contribution to 2 from edges incident with U is:

$$\sum_{i=0}^{k-2}\sum_{v\in U_i}\left(\underbrace{\sum_{u\in U_i:vu\in E(G)}(x_v-x_u)^2}_{=0 \text{ as } x \text{ is constant on } U_i}+\sum_{u\in U_{i+1}:vu\in E(G)}(x_v-x_u)^2\right)+\sum_{v\in U_{k-1}}\sum_{u\in V(G)-U:vu\in E(G)}(x_v-x_u)^2 =$$

$$\sum_{i=0}^{k-2}\sum_{v\in U_i}\sum_{u\in U_{i+1}}(x_v-x_u)^2+\sum_{v\in U_{k-1}}\sum_{u\in V(G)-U:vu\in E(G)}x_v^2 \le$$

$$\sum_{i=0}^{k-2}|U_i|(d-1)\left(\frac{1}{(d-1)^{i/2}}-\frac{1}{(d-1)^(i+1)/2}\right)^2+|U_{k-1}|(d-1)\left(\frac{1}{(d-1)^{(k-1)/2}}\right)^2 =$$

$$\sum_{i=0}^{k-2}\frac{|U_i|(\sqrt{d-1}-1)^2}{(d-1)^i}+\frac{|U_{k-1}|}{(d-1)^{k-2}} =$$

$$\sum_{i=0}^{k-2}\frac{|U_i|}{(d-1)^i}(d-2\sqrt{d-1})+\frac{|U_{k-1}|}{(d-1)^{k-1}}(d-2\sqrt{d-1}+2\sqrt{d-1}-1) =$$

$$\sum_{i=0}^{k-1}\frac{|U_i|}{(d-1)^i}(d-2\sqrt{d-1})+\frac{|U_{k-1}|}{(d-1)^{k-1}}(2\sqrt{d-1}-1) =$$

$$(d-2\sqrt{d-1})A+\frac{|U_{k-1}|}{(d-1)^{k-1}}(2\sqrt{d-1}-1) \le$$

$$(d-2\sqrt{d-1})A+k\frac{|U_{k-1}|}{(d-1)^{k-1}}\frac{2\sqrt{d-1}-1}{k} \le$$

$$(d-2\sqrt{d-1})A+\sum_{i=0}^{k-1}\frac{|U_i|}{(d-1)^i}\frac{2\sqrt{d-1}-1}{k} =$$

$$\left(d-2\sqrt{d-1}+\frac{2\sqrt{d-1}-1}{k}\right)A$$

Where the last inequality comes from $|U_i| \le (d-1)|U_{i-1}|$ which implies that $\frac{|U_i|}{(d-1)^i}$ is non-increasing.

Analogously, as the c^2 factors out, we conclude that the contribution of edges incident with W satisfies an equivalent inequality:

$$\sum_{i=0}^{k-2}\sum_{v\in W_i}\left(\underbrace{\sum_{u\in W_i:vu\in E(G)}(x_v-x_u)^2}_{=0 \text{ as } x \text{ is constant on } W_i}+\sum_{u\in W_{i+1}:vu\in E(G)}(x_v-x_u)^2\right)+\sum_{v\in W_{k-1}}\sum_{u\in V(G)-W:vu\in E(G)}(x_v-x_u)^2$$

$$\leq\left(d-2\sqrt{d-1}+\frac{2\sqrt{d-1}-1}{k}\right)B$$

Combining these inequalities we get:

$$(L\overline{x},\overline{x})\leq\left(d-2\sqrt{d-1}+\frac{2\sqrt{d-1}-1}{k}\right)(A+B)\leq\left(d-2\sqrt{d-1}+\frac{2\sqrt{d-1}-1}{k}\right)(\overline{x},\overline{x})$$

This using the above bound on μ_2 shows $\lambda_2=d-\mu_2\geq\left(d-2\sqrt{d-1}+\frac{2\sqrt{d-1}-1}{k}\right)$ as claimed. $\square$

Let u_0,w_0 be vertices having distance equal to the diameter, then U_i,W_i for $i\leq k+1$ cover the whole graph. Bounding $|U_i|,|W_i|$, using above inequalities, we can deduce $D(G)\geq\frac{\log(n)}{\log(d)}$. This shows $D(G)\to\infty$ as $n\to\infty$. This combined with the above result implies that for any infinite sequence of d-regular graphs G we have $\limsup\lambda_2(G)\geq 2\sqrt{d-1}$.

2.4 Small Ramanujan Graphs

We now turn to actually finding graphs having small second eigenvalue. Lubotzky, Phillips and Sarnak in [8] first defined *Ramanujan graph* to be a connected d-regular graph whose second eigenvalue satisfies $|\mu|\leq 2\sqrt{d-1}$.

In this section we offer several small examples in order to get some feeling on how elusive these graphs really are. We note that on their own such graphs are not particularly useful as they are not necessarily optimal, or even close to optimal, as the result of the previous subsection only shows that $2\sqrt{d-1}$ is optimal in the limit $n\to\infty$. This is the reason we are actually looking for infinite sequences of such graphs.

On the other hand, in the following section we require such a small example as a base to build infinite examples, so the examples we offer here give rise to existence of several new infinite families.

The following lemma will help us find the eigenvalues of certain graphs.

A matrix is called *circular* if $A_{i,j}=A_{i+1,j+1}$ for all $i,j\leq n$ where the indices are taken modulo n.

Lemma 2.5. Let A be a circular matrix then $\overline{v}_i=(1,\omega^i,\omega^{2i},\cdots,\omega^{(n-1)i})^t$ for $1\leq i\leq n$ make a basis of eigenvectors, where $\omega=e^{\frac{2\pi i}{n}}$.

Proof. We have

$$(A\overline{v}_i)_j = \sum_{k=1}^{n} A_{j,k}(\overline{v}_i)_k = \sum_{k=1}^{n} A_{j,k}\omega^{ki} = \sum_{k=1}^{n} A_{1,k-j+1}\omega^{(k-j+1)i}\omega^{(j-1)i} = \left(\sum_{k=1}^{n} A_{1,k}\omega^{ki}\right)(\overline{v}_i)_j$$

showing that $\overline{v}_i$ is an eigenvector with eigenvalue $\sum_{k=1}^{n} A_{1,k}\omega^{ki}$.

It remains to show v_i are linearly independent, which we notice by listing them in a matrix which is then a Vandermonde matrix having determinant $\prod_{i<j}(\omega^i - \omega^j) \neq 0$[4] implying the columns are independent as desired. $\square$

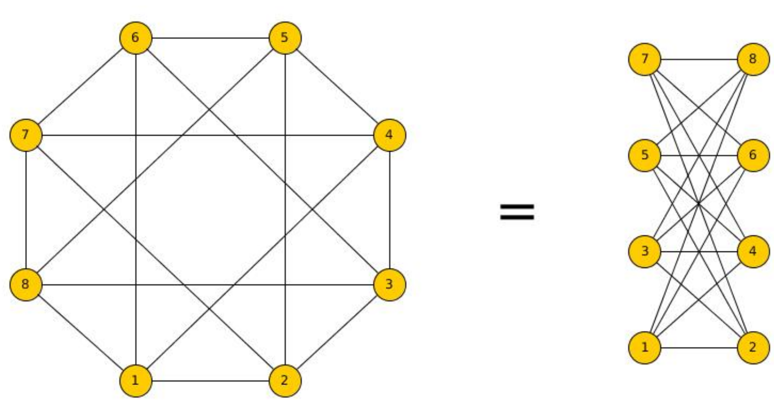
The only 2-regular connected graph on n vertices is an n-cycle C_n. We notice that the adjacency matrix of the cycle is circular so by **Lemma** 2.5 we deduce the eigenvalues are given by $\lambda_i = \omega^i + \omega^{n-i} = 2\cos\left(\frac{2\pi i}{n}\right)$ for $i \leq n$ and hence $|\lambda_i| \leq 2 = 2\sqrt{2-1}$ showing the cycles are in fact Ramanujan graphs.

With this in mind we consider a generalisation of a cycle graph, which we name *circular graphs*. A circular graph on n vertices is given by a sequence of k numbers $1 \leq a_1 < a_2 < \cdots < a_k < \frac{n}{2}$ such that if we take $V(G) = [n]$ the edges of G are given by $i \sim i + a_j$ for all $1 \leq i \leq n, 1 \leq j \leq k$, vertices taken modulo n. An alternative definition is a graph having a circular adjacency matrix with first row $A_{1,1+a_i}, A_{1,n+1-a_i} = 1$ and $A_{1,j} = 0$ else, note that the remaining entries are fixed by this and the condition that A is circular. We denote such a circular graph by $(a_1, a_2, \cdots, a_k)$.

To illustrate the idea we describe several standard classes of graph in terms of this framework:

- C_N can be written as (1) or in fact (i) for $i < \frac{n}{2}$.
- K_{2n-1} - the complete graph on $2n-1$ vertices can be written as $(1, 2, \cdots, n-1)$ on $2n-1$ vertices
- $K_{2n,2n}$ - the complete bipartite graph with $2n$ vertices in each bipartition can be written as $(1, 3, \cdots, 2n-1)$ on $4n$ vertices. **Figure** 3 illustrates this example for $n = 4$.

Figure 3: $K_{4,4} \equiv (1, 3)$.



[4]Shown, for example, in [2] Section 1.2, page 13.

We note that circular graphs are $2k$-regular and are given by k interleaved Hamilton cycles, furthermore what makes them interesting to us is that we can easily find their eigenvalues using **Lemma** 2.5.

We further allow $a_k = \frac{n}{2}$, which enables us to work with $2k-1$ regular graphs as well, so for example allows us to specify K_{2n} and $K_{2n-1,2n-1}$ in this framework as well. We note that in this case the graph consists of $k-1$ Hamilton cycles and a perfect matching.

Considering a complete graph on n vertices, its spectrum is $d = n-1$ of multiplicity 1 and -1 of multiplicity $n-1$, by **Lemma** 2.5. So complete graphs are always Ramanujan. Noticing the second largest eigenvalue is very far from $2\sqrt{d-1}$ it makes sense to look at graphs similar to a complete graph, this motivates the following lemma.

Lemma 2.6. Any circular graph $G = (a_1 \cdots a_k)$ with $k \geq \frac{n}{2} + 2 - \sqrt{n+1}$ is Ramanujan.

Proof. G is a d-regular graph, where $d = 2k$ or $d = 2k-1$ if $a_k = \frac{n}{2}$.

Let $B = [n-1] - \bigcup_{j=1}^{k} \{a_j, n-a_j\}$, note that $|B| = n-d$.

Let $\omega_i = e^{\frac{2\pi i}{n}}$. The eigenvalues of G are, using **Lemma** 2.5, for $1 \leq i \leq n-1$

$$\lambda_i = \sum_{j=1}^{k} \left(\omega_i^{a_j} + \omega_i^{n-a_j}\right) = \sum_{j=1}^{n-1} \omega_i^j - \sum_{j\in B} \omega_i^j = -1 - \sum_{j\in B} \omega_i^j$$

and $|\lambda_n| = d$. Where we used $\sum_{j=0}^{n-1} \omega_i^j = 0$ for $1 \leq i \leq n-1$. We can conclude from this that, for $i < n$, we have $|\lambda_i| \leq 1 + n - d \leq 2\sqrt{d-1} + n + 1 - (\sqrt{d-1}+1)^2 \leq 2\sqrt{d-1} + n + 1 - (\sqrt{2k-2}+1)^2 \leq 2\sqrt{d-1} + n + 1 - (\sqrt{n+1} - 1 + 1)^2 = 2\sqrt{d-1}$. Where we used $k \geq \frac{n}{2} + 2 - \sqrt{n+1}$ which is equivalent to $2k - 2 \geq (\sqrt{n+1} - 1)^2$.

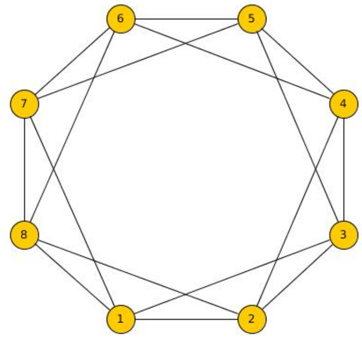
This shows that all the non-trivial eigenvalues satisfy $|\lambda_i| \leq 2\sqrt{d-1}$ so the graph is Ramanujan. $\square$

The above lemma provides Ramanujan examples for each d, but we note that while this lemma gives an infinite number of Ramanujan graphs, we seek infinite families while keeping d fixed and this lemma only gives a finite number of examples for each fixed d.

We offer several examples showing it is possible to find circular Ramanujan graphs with k significantly smaller than required for **Lemma** 2.6.:

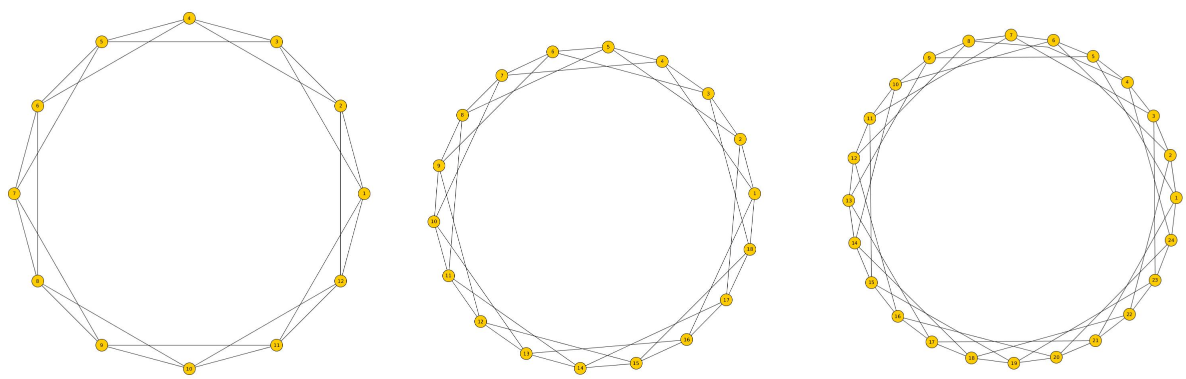
- $(1,2)$ circular graph on $n = 8$ vertices is Ramanujan - this follows from Lemma 1.5 showing its eigenvalues are $\lambda_i = 2\cos\frac{2\pi i}{8} + 2\cos\frac{4\pi i}{8}$ for $0 \leq i \leq 7$ finally noticing $\frac{\sqrt{2}}{2} + 1 \leq \sqrt{3} = \sqrt{d-1}$ we only need to check i for which both cos arguments are multiples of π so in particular only $i = 0, 4$ first one corresponding to the trivial eigenvalue and the second one to $\lambda_4 = 0$.
- $(1, m)$ circular graph on $n = 8m$ for $2 \leq m \leq 8$ is Ramanujan, by similar considerations as in the previous example.

Figure 4: $(1, 2)$ circular graph on 8 vertices.



- $(1, m)$ circular graph on $n = 6m$ for $2 \le m \le 6$ is Ramanujan again following similar considerations.

Figure 5: $(1, m)$ circular Ramanujan graphs for $m = 2, 3, 4$ and $n = 6m$.



We notice that none of the examples above are bipartite, in case of **Lemma** 2.6 they have too many edges as a bipartite graph can have at most half of the possible edges. Being bipartite is important for the following section as there we require a bipartite Ramanujan graph for a base for the argument.

Similarly as above we start by noticing, using **Lemma** 2.5, that $K_{n,n}$ has eigenvalues $d, -d$ with multiplicity 1 and 0 with multiplicity $2n-2$, this shows $K_{n,n}$ is Ramanujan. The following lemma provides us with another class of Ramanujan graphs, but this time they are bipartite.

Lemma 2.7. Let $G = (a_1 \cdots a_k)$ be a circular graph on $2n$ vertices. If each a_i is odd and $k \ge \frac{n+1}{2} + 1 - \sqrt{n}$ then G is a bipartite Ramanujan graph.

Proof. We note that the graph is bipartite with bipartitions given by even/odd vertices, as a_i being odd implies each edge joins only vertices of different parity.

The proof now carries on in a similar fashion as that of **Lemma** 2.6. The main difference is the use of identity $\omega_i + \omega_i^3 + \cdots + \omega_i^{2n-1} = 0$ unless $\omega_i = \pm 1$. (Note that this follows as

$(1+\omega_i)(\omega_i+\omega_i^3+\cdots+\omega_i^{2n-1})=\omega_i+\cdots+\omega_i^{2n}=0.)$

The remainder of the proof goes along exactly the same lines as in the **Lemma** 2.6, the eignvalues consist of a negative sum of $n-d$ powers of ω_i not-attained by eigenvalues so the triangle inequality gives us $|\lambda_i| \leq n-d$ for $1<i<n$. Which now using the given inequality gives the graph is Ramanujan as claimed. □

We finally note that there is another condition we require for a graph to be Ramanujan, namely the condition for the graph to be connected, we did not specifically mention this before as it was obviously true in the examples above, for example by looking at the eigenvalues and invoking **Lemma** 2.1. We bring this up here as it is easy to construct an infinite family of circular graphs of fixed degree which have all non-trivial eigenvalues in the desired interval, for example $(k, 2k)$ on $8k$ vertices is a 4 regular graph with all non-trivial eigenvalues smaller than $\sqrt{d-1}$ in absolute value, but this graph is not connected.

We also note that circular graphs offer some sort of a stability result, in the sense that changing one value of a_i can only change the eigenvalues by at most 2, this is yet another consequence of **Lemma** 2.5. This combined with the scaling argument of the previous paragraph also unfortunately fails as there we have k eigenvalues which are equal to d so if we connect the graph by changing some value of a_i, we would have non-trivial eigenvalues larger than $d-2$.

The idea behind circular graphs is that we can control the eigenvalues really well but still have considerable freedom in choosing the graph, which might give us enough choice to find a Ramanujan graph. Unfortunately as the Ramanujan graphs become closer to being tight examples, by **Theorem** 2.4, it gets harder to find them and this method does not seem to offer enough freedom to achieve it. For example a similar idea rests behind the explicit constructions of Lubotzky, Phillips and Sarnak in [8] who consider Cayley graphs of certain groups, so they provides more freedom than our circular graphs.

3 The Main Result

3.1 History

As the previous section illustrated, construction of infinite families of Ramanujan graphs of fixed degree is, to the best of our knowledge, a hard problem. We do hope that the previous section illustrated many of the useful properties that make them important, not only in mathematics but also in theoretical computer science. It is then not surprising that that they have been a subject of substantial study in these fields.

The term Ramanujan graphs was coined by Lubotzky, Phillips and Sarnak in 1988 [8] but there was substantial study of the second eigenvalue already under way even before that, for example the paper by Alon and Milman [4] in 1985.

Lubotzky, Phillips and Sarnak in [8] first show existence of infinite families of d-regular

Ramanujan graphs. The main idea of their work is to look at Cayley graphs of suitable number theoretic groups, their argument requires $d = p + 1$ for p a prime.

We note that Margulis in 1988 [10] has independently shown the same result using what, in his own words, can be considered equivalent to the work of Lubotzky, Phillips and Sarnak. Margulis also notices that similar constructions also work for $d = q + 1$ for q a power of a prime number. This has been made explicit by building on the work of Lubotzky, Phillips and Sarnak by Morgenstern in 1994 [9].

Pizer in 1990 [11] constructed a different family of Ramanujan graphs for $d = p + 1$ with lower girth than the ones found by Lubotzky, Phillips and Sarnak but he finds a significantly larger number of families. Also their construction provides "almost Ramanujan" graphs for $d = m + 1$ when m is not a prime.

There have been very few further constructions of infinite families of Ramanujan graphs, the only ones we are aware of are by Chiu [13] in 1992 and Jordan and Livné [12] in 1997. Both producing examples only with $d = p + 1$ for p prime.

We note once again that none of the constructions above work for $d = q + 1$ when q is not a power of a prime.

Bilu and Linial [14] in 2006 suggested a method of constructing Ramanujan graphs based on 2-lifts of a graph, which we introduce in the following subsection. They managed to find infinite families of almost Ramanujan d-regular graphs, in the sense that the second largest eigenvalue is at most $c\sqrt{d \log^3 d}$.

In a remarkable paper of Marcus, Spielman and Srivastava [1] of 2015 they show the existence of infinite families of Ramanujan graphs of any degree. This is the result we present throughout the following subsections.

They build on the idea of 2-lifts given by Bilu and Linial but they introduce a novel existence argument, namely the method of interlacing polynomials which shows much promise, for example Marcus, Spielman and Srivastava settled the famous Kadison-Singer conjecture [15] using it.

Marcus, Spielman and Srivastava have recently, in a 2015 preprint [16], further improved their result to show that there exist bipartite Ramanujan graphs of every number of vertices and every degree.

We do add that their construction only gives bipartite Ramanujan graphs and unlike most previous examples their construction is not explicit and produces only bipartite graphs.

We finally remark that bipartite Ramanujan graphs are in a sense weaker examples than non-bipartite ones as they actually have 2 large, in absolute value, eigenvalues, namely d and $-d$. In most applications similar results hold for both bipartite and non-bipartite Ramanujan graphs. An example of this is the argument used in **Theorem** 2.2 which demonstrates this well and uncovers one reason why the results carry over, namely the contribution of the $-d$ eigenvalue is countered by that of the d for specific parities of the path lengths.

3.2 Overview

In this subsection we attempt to give a short overview of the proof of Marcus, Spielman and Srivastava in an attempt to make it easier to follow the actual proof. We split it into parts which will correspond to our subsections as the proof uses several very different techniques.

1. **2-lifts**

   In this section we introduce a transformation of a graph, called a 2-lift, with several properties which interest us. A 2-lift results in a new graph with double the number of vertices and edges, where for each original edge we have 2 options for the new pair of edges. The important result we show in this section is that the new graph has n eigenvalues matching those of the original graph and n new eigenvalues corresponding to a signing matrix which depends on the choices we made for the new edges.

2. **Matching Polynomial**

   In this subsection we introduce the matching polynomial and prove that it equals the expected characteristic polynomial of the signing matrix of 2-lifts, where we make the choices uniformly at random. We then prove 2 important results about the matching polynomial, namely that its roots are all real and bounded by $2\sqrt{d-1}$. Our goal is to show that this bound is satisfied by at least one of the characteristic polynomials which sum up to the matching polynomial as part of the expected characteristic polynomial identity.

3. **Interlacing**

   Here we define a property of interlacing. It interests us because, if satisfied by a sequence of polynomials, we can conclude that maximal root of their sum bounds the roots of at least one of them. We prove a version of this statement and give a way of showing the interlacing property is satisfied by linking it to real-rootedness.

4. **Main Result**

   We complete the argument by using a technical result which we prove in the following subsection. Starting from a graph which is bipartite and Ramanujan, we can using the consideration of the previous subsections find a 2-lift which also gives a bipartite Ramanujan graph, by repeating this argument we find an infinite sequence of d-regular Ramanujan graphs.

5. **Real Stability**

   The goal of this section is to present the technical result we needed in the previous subsection. We introduce a generalisation of real-rootedness to polynomials of multiple variables called real stability. We show that certain general determinantal polynomials are always real stable and present several operators which preserve this property. We then use these operators on the real stable polynomials to obtain that an univariate polynomial we needed for showing interlacing is real stable, we then show that for univariate polynomials real stability is equivalent to real-rootedness which completes the proof.

3.3 2-lifts.

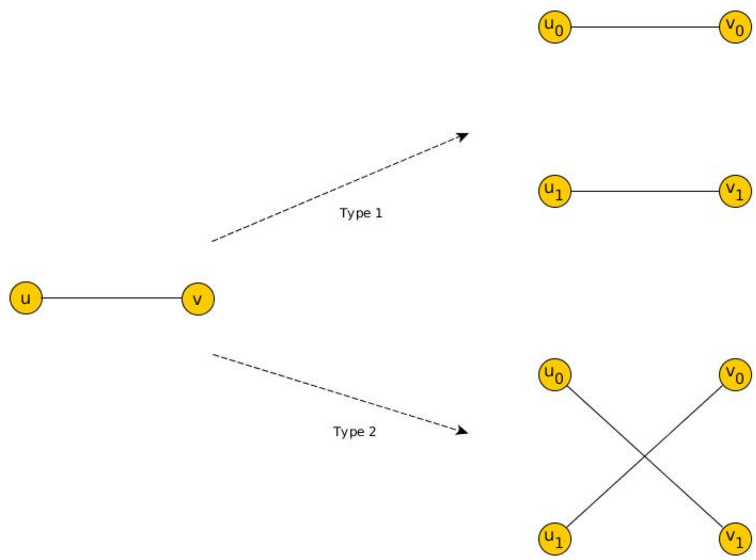
The ideas of this section are due to Bilu and Linial [14] who build on the work by Friedman [17].

The basic idea is to look for a transformation of the graph which would enable us to control the second eigenvalue while growing the number of vertices and keeping the degrees fixed. It turns out that 2-lifts are such a transformation.

2-lift of a graph G is a graph H which has two vertices v_0, v_1 for each vertex $v \in V(G)$. The pair $\{v_0, v_1\}$ is called *the fibre* of the original vertex v and the fibres of separate vertices are disjoint. Every edge in $E(G)$ gives rise to two edges in H. If $u \sim v$ is an edge of G, $\{u_0, u_1\}$, $\{v_0, v_1\}$ are the fibres of u, v respectively, then H can either contain the pair of edges

of type 1: $\{u_0 \sim v_0, u_1 \sim v_1\}$, or of type 2: $\{u_0 \sim v_1, u_1 \sim v_0\}$.

Figure 6: Possible types of edges in a 2-lift:



For example, if we consider 2 extreme examples, if only edges of type 1 appear in the 2-lift H consists of 2 disjoint copies of G, while if only edges of type 2 appear H is a double cover of G.

The double cover of a graph is a bipartite graph in which each vertex is represented in both bipartitions and each original edge joins the same vertices in different bipartitions. Double cover is an often useful construction in combinatorics when we want to relate general graphs to bipartite ones. An alternative way to define a double cover is using the above notion of 2-lifts so this might offer some motivation for considering the 2-lifts in our problem.

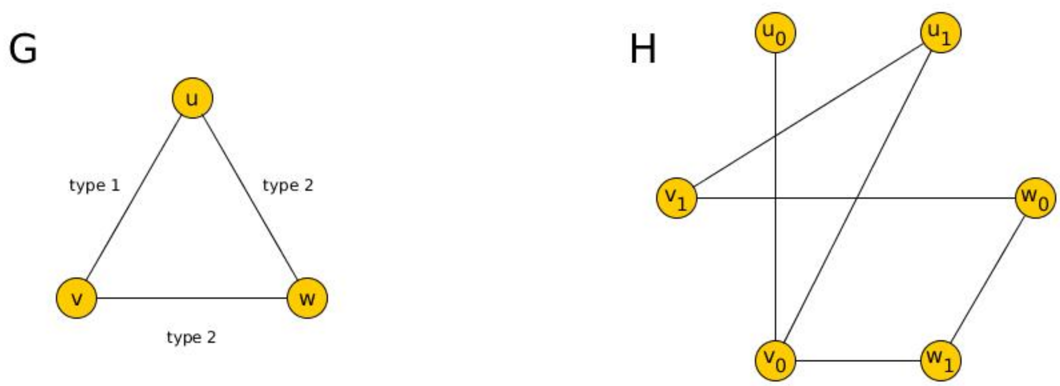
The reason 2 lifts are important is because of the control over eigenvalues they offer. In order to make this precise, we encode the different lifts as signings $s : E(G) \to \{\pm 1\}$ of the edges of G, such that a 2-lift H corresponds to a signing:

$$s(u,v) = 1 \text{ if edges of type 1 corresponding to } u \sim v \text{ appear in } H.$$
$$s(u,v) = -1 \text{ if edges of type 2 corresponding to } u \sim v \text{ appear in } H.$$

We now define the signed adjacency matrix A_s to equal the adjacency matrix except that entries corresponding to an edge $u \sim v$ are replaced by $s(u,v)$.

The following figure illustrates these definitions on an example of G a triangle, and a 2-lift H corresponding to the signing $s(u,v) = 1, s(u,w) = -1$ and $s(v,w) = -1$.

Figure 7: A triangle and one of its 2-lifts.



The adjacency matrix of G, the signed adjacency matrix and the adjacency matrix of H are:

$$A(G) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \; A_s = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix} \text{ and } A(H) = \left(\begin{array}{ccc|ccc} 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \end{array}\right)$$

Where we the rows and columns of $A(G), A_s$ correspond to the ordering u, v, w while rows and columns of $A(H)$ correspond to the ordering $u_0, v_0, w_0, u_1, v_1, w_1$. We have written $A(H)$ as a block matrix to hint at the proof of the following theorem.

Theorem 3.1. *Given a graph G with adjacency matrix A and a 2-lift H of G corresponding to the signed adjacency matrix A_s the eigenvalues of H are given as a union of eigenvalues of A and A_s taken with multiplicities.*

Proof. Let A_1, A_2 be defined as $A_1 = \frac{A+A_s}{2}$ and $A_2 = \frac{A-A_s}{2}$ so that A_1 equals A_s except for the entries of -1 which are replaced by zeros, while A_2 equals $-A_s$ under the same transformation. Note that A_i is the adjacency matrix of G restricted only to edges which give rise to edges of type i in the lift.

If row i of A corresponds to vertex v_i and we define $\{v_i', v_i''\}$ to be the fiber of v_i if we index the rows of $A(H)$ by the ordering $v_1', v_2', \cdots, v_n', v_1'', v_2'', \cdots, v_n'$ we get

$$A(H) = \begin{pmatrix} A_1 & A_2 \\ A_2 & A_1 \end{pmatrix}$$

As the entries of the upper left and lower right $n \times n$ submatrix equal 1 only if the corresponding edge was of type 1 while upper right and lower left submatrix only occur when the edge was of type 2.

We now notice that if $\overline{u}$ is a real eigenvector of A then $(A_1 + A_2)\overline{u} = \lambda\overline{u}$. So

$$A(H)\begin{pmatrix} \overline{u} \\ \overline{u} \end{pmatrix} = \begin{pmatrix} A_1 & A_2 \\ A_2 & A_1 \end{pmatrix}\begin{pmatrix} \overline{u} \\ \overline{u} \end{pmatrix} = \begin{pmatrix} (A_1 + A_2)\overline{u} \\ (A_2 + A_1)\overline{u} \end{pmatrix} = \lambda\begin{pmatrix} \overline{u} \\ \overline{u} \end{pmatrix}$$

Similarly if $\overline{v}$ is a real eigenvector of A_s then $(A_1 - A_2)\overline{v} = \mu\overline{v}$. So

$$A(H)\begin{pmatrix} \overline{v} \\ -\overline{v} \end{pmatrix} = \begin{pmatrix} A_1 & A_2 \\ A_2 & A_1 \end{pmatrix}\begin{pmatrix} \overline{v} \\ -\overline{v} \end{pmatrix} = \begin{pmatrix} (A_1 - A_2)\overline{v} \\ (A_2 - A_1)\overline{v} \end{pmatrix} = \mu\begin{pmatrix} \overline{v} \\ -\overline{v} \end{pmatrix}$$

Further noticing $\begin{pmatrix} \overline{u} \\ \overline{u} \end{pmatrix}$ and $\begin{pmatrix} \overline{v} \\ -\overline{v} \end{pmatrix}$ are orthogonal we notice that if $u_1, u_2, \cdots, u_n$ is a real eigenvector basis for A and $v_1, v_2, \cdots, v_n$ is a real eigenvector basis for A_s then $\begin{pmatrix} \overline{u}_1 \\ \overline{u}_1 \end{pmatrix}, \cdots \begin{pmatrix} \overline{u}_n \\ \overline{u}_n \end{pmatrix}, \begin{pmatrix} \overline{v}_1 \\ -\overline{v}_1 \end{pmatrix}, \cdots, \begin{pmatrix} \overline{v}_n \\ -\overline{v}_n \end{pmatrix}$ are independent and there are $2n$ of them so they make an eigenvector basis for $A(H)$. This in turn implies their eigenvalues are all the eigenvalues of $A(H)$ and as $\begin{pmatrix} \overline{u}_i \\ \overline{u}_i \end{pmatrix}$ have all the eigenvalues of A and $\begin{pmatrix} \overline{v}_i \\ -\overline{v}_i \end{pmatrix}$ all the eigenvalues of A_s the result follows. □

Following Friedman [17], the eigenvalues of A are called *old eigenvalues* and eigenvalues of A_s are called *new eigenvalues.*

In the following subsections we prove that we can choose a 2-lift such that all the new eigenvalues satisfy $\mu \leq 2\sqrt{d-1}$.

We are only able to prove the upper bound, so in order to control the lower bound we start with a bipartite graph and notice that any 2-lift of a bipartite graph is also bipartite, now using **Lemma** 2.1 d), stating that λ is an eigenvalue if and only if $-\lambda$ is, we deduce that the lower bound will also be satisfied.

So starting with a bipartite Ramanujan graph, for example $K_{n,n}$, and choosing a suitable 2-lift in each step we construct a new Ramanujan graph with double the number of vertices. repeating the process gives an infinite family as desired.

Now it only remains to show that we can choose such a 2-lift and in order to do this we find a relation between 2-lifts and the matching polynomial.

3.4 Matching Polynomial

Given a graph G, let m_i denote the number of matchings of size i in the graph. Where a matching of size i is a set of i independent edges. We further define $m_0 = 1$.

We now define the *matching polynomial* as

$$\mu_G(x) \equiv \sum_{i \geq 0} x^{n-2i}(-1)^i m_i$$

following Heilman and Lieb [18] who first defined it in relation to a problem in chemistry.

We start by proving the result, of Godsil and Gutman [19], relating the matching polynomial to 2-lifts which is the real reason we are interested in the matching polynomial.

Theorem 3.2. Let G be a graph with $m = E(G)$. Let A_s denote the signed adjacency matrix corresponding to a signing s. Let $f_s = \det xI - A_s$ be the characteristic polynomial of A_s

$$\frac{1}{2^m} \sum_{s \in \pm 1^m} f_s(x) = \mu_G(x)$$

We notice that we can write the left hand side as an expected characteristic polynomial, where we choose the signing uniformly at random.

For an example we consider G as a triangle graph depicted in Figure 7. The matching polynomial is $\mu_G(x) = x^3 - 3x$, as the only matchings are 1-matchings of which there are exactly 3. The expected characteristic polynomial is given by

$$\mathbb{E} f_s(x) = \frac{4(x-2)(x+1)^2 + 4(x+2)(x-1)^2}{2^3} = \frac{x^3 - 3x - 2 + x^3 - 3x + 2}{2} = x^3 - 3x = \mu_G(x)$$

Proof. We denote by $\mathrm{sym}(S)$ the set of permutations of a set S. Let $\mathrm{sg}(\sigma)$ equal 0 if σ is an even permutation and 1 if it is odd. We expand the determinant as a sum over permutations $\sigma \in \mathrm{sym}([n])$ and get:

$$\begin{aligned}
\frac{1}{2^m} \sum_{s \in \pm 1^m} f_s(x) &= \mathbb{E} f_s(x) \\
&= \mathbb{E}(\det(xI - A_s)) \\
&= \mathbb{E}\left(\sum_{\sigma \in \mathrm{sym}([n])} (-1)^{\mathrm{sg}(\sigma)} \prod_{i=1}^{n} (xI - A_s)_{i,\sigma(i)} \right) \\
&= \mathbb{E}\left(\sum_{k=0}^{n} x^{n-k} \sum_{S \subset [n], |S| = k} \sum_{\pi \in \mathrm{sym}(S)} (-1)^{\mathrm{sg}(\pi)} \prod_{i \in S} (A_s)_{i,\pi(i)} \right) \\
&= \sum_{k=0}^{n} x^{n-k} \sum_{S \subset [n], |S| = k} \sum_{\pi \in \mathrm{sym}(S)} \mathbb{E}\left((-1)^{\mathrm{sg}(\pi)} \prod_{i \in S} s_{i,\pi(i)} \right)
\end{aligned}$$

Where π corresponds to the part of σ with $\sigma(i) \neq i$ and we extend the definition of $s_{u,v}$ to equal 0 if (u, v) is not an edge.

The first three equalities come from the definitions of the characteristic polynomial and the determinant of a matrix while the fourth one comes from the linearity of expectation and expanding the product.

We now notice that $\mathbb{E}(s_{i,j}) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot (-1) = 0$ if (i, j) is an edge of G while $s_{i,j} = 0$ if (i, j) is not an edge, in particular $\mathbb{E}(s_{i,j}) = 0$.

We further notice that $s_{i,j}$ is independent from any collection of $s_{k,l}$ if it does not contain $s_{j,i}$. If π is such that there exists an i such that $\pi(\pi(i)) \neq i$ then $s_{i,\pi(i)}$ is independent from $\{s_{j,\pi(j)} | j \in S - \{i\}\}$ implying that

$$\mathbb{E}\left((-1)^{\mathrm{sg}(\pi)} \prod_{j \in S} s_{j,\pi(j)}\right) = \underbrace{\mathbb{E}(s_{i,\pi(i)})}_{=0} \mathbb{E}\left((-1)^{\mathrm{sg}(\pi)} \prod_{j \in S - \{i\}} s_{i,\pi(i)}\right) = 0$$

So the only products which make a contribution to the sum are those for which $\pi \circ \pi = e$, where e denotes the trivial permutation.

We further notice that $s_{i,\pi(i)}$ is non zero only if it corresponds to an edge of G, while $\pi \circ \pi = e$ gives that only contribution to the sum comes from π being a matching function, of size $\frac{|S|}{2}$. This explains the first following equality.

$$\begin{aligned}
\frac{1}{2^m} \sum_{s \in \pm 1^m} f_s(x) &= \sum_{k=0}^{n} x^{n-k} \sum_{|S|=k, \pi \text{ a matching of size } \frac{|S|}{2}} (-1)^{|S|/2} \mathbb{E}\left(\prod s_{i,\pi(i)}^2\right) \\
&= \sum_{l=0}^{\lfloor \frac{n}{2} \rfloor} x^{n-2l} \sum_{|S|=2l,\, \pi \text{ a matching of size } l} (-1)^l \prod \mathbb{E}\left(s_{i,\pi(i)}^2\right) \\
&= \sum_{l=0}^{\lfloor \frac{n}{2} \rfloor} x^{n-2l} \cdot m_l \cdot (-1)^l \prod 1 = \mu_G(x)
\end{aligned}$$

The second equality follows from noticing that a matching can only have even number of vertices so k odd contribution is zero while we use independence, once again, to exchange the product and expectation.

The third equality follows from noticing that $\mathbb{E}\left(s_{i,\pi(i)}^2\right) = \frac{1}{2}1^2 + \frac{1}{2}(-1)^2 = 1$ where we assume $i \sim \pi(i)$ when we say π corresponds to a matching. $\square$

We now present 2 results which we will need in the following subsection. They were first proved by Heilmann and Lieb [18] in a slightly more analytical fashion so our proofs will follow ideas of Godsil and Gutman [19].

Theorem 3.3. All the roots of a matching polynomial are real.

Theorem 3.4. All the roots of the matching polynomial of a d-regular graph have absolute value at most $2\sqrt{d-1}$.

We state these theorems here as the ideas of their proofs are not instrumental in our main proof so the reader will not lose much in terms of continuity if they avoid the rest of the section. We need to build up a few results before we can prove them.

The next lemma contains recurrence relations instrumental in showing the remaining lemmas of this subsection.

Given a graph $G(V,E)$ and a vertex $v \in V$ we denote by $G-v$ the induced subgraph on $V-\{v\}$, while if e is an edge $G-e$ denotes the graph defined by $(V, E-\{e\})$.

Lemma 3.5. Given a graph $G(V,E)$, $u, v \in V$ such that $u \sim v = e \in E$ then

a) $\mu_G(x) = \mu_{G-e}(x) - \mu_{G-u-v}(x)$

b) If $u_1, u_2, \cdots u_d$ denote the neighbours of v then $\mu_G(x) = \mu_{G-v}(x) - \sum_{i=1}^{d} \mu_{G-v-u_i}(x)$

c) If $G_1, G_2, \cdots, G_k$ are disjoint graphs and $G = \bigcup_{i=1}^{k} G_i$ then

$$\mu_G(x) = \prod_{i=1}^{k} \mu_{G_i}(x).$$

Proof. **a)** By considering whether a matching includes e or not we obtain the following relation $m_k(G) = m_{k-1}(G-u-v) + m_k(G-e)$, substituting in the definition $\mu_G(x) = \sum_{k\geq 0} x^{n-2k}(-1)^k m_k(G) = x^n + \sum_{k\geq 1} x^{n-2k}(-1)^k (m_k(G-e) + m_{k-1}(G-u-v)) = \mu_{G-e}(x) + \sum_{l\geq 0} x^{n-2-2l}(-1)^{l+1} m_l(G-u-v) = \mu_{G-e}(x) - \mu_{G-u-v}(x)$

b) It follows by applying part **a)** to edges $e_i \equiv v \sim u_i$ for $1 \leq i \leq d$ and noticing that an isolated vertex can not contribute to any matching. $\mu_G(x) = \mu_{G-e_1}(x) - \mu_{G-v-u_1}(x) = \mu_{G-e_1-e_2}(x) - \mu_{G-v-u_2}(x) - \mu_{G-v-u_1}(x) = \cdots = \mu_{G-e_1-\cdots-e_d}(x) - \sum_{i=1}^{d} \mu_{G-v-u_i}(x) = \mu_{G-v}(x) - \sum_{i=1}^{d} \mu_{G-v-u_i}(x)$ as desired.

c) We first prove the result for $k = 2$.

$$\mu_{G_1}(x)\mu_{G_2}(x) = \left(\sum_{i\geq 0} x^{n_1-2i}(-1)^i m_i(G_1)\right)\left(\sum_{i\geq 0} x^{n_2-2j}(-1)^j m_j(G_2)\right) =$$

$$\sum_{k\geq 0} x^{n_1+n_2-2k} \sum_{i=0}^{k} (-1)^i m_i(G_1)(-1)^{k-i} m_{k-i}(G_2) = \sum_{k\geq 0} x^{n-2k}(-1)^k m_k(G_1 \cup G_2) = \mu_{G_1\cup G_2}(x)$$

Where n_1, n_2, n denote the number of vertices of G_1, G_2 and G respectively and hence satisfy $n_1 + n_2 = n$. We first expand the product given by the definition and then used the identity $m_k(G_1 \cup G_2) = \sum_{i=0}^{k} m_i(G_1) m_{k-i}(G_2)$ which follows upon conditioning that i edges of the k matching on G are among G_1.

To get the general result we use the $k = 2$ version repeatedly:

$$\mu_G(x) = \mu_{G_1\cup\cdots\cup G_{k-1}}(x)\mu_{G_k}(x) = \mu_{G_1\cup\cdots\cup G_{k-2}}(x)\mu_{G_{k-1}}(x)\mu_{G_k}(x) = \cdots = \prod_{i=1}^{k}\mu_{G_i}(x)$$

as desired.

□

We note that we rarely revisit the actual definition of the matching polynomial after this point, the remaining results simply make use of the above properties and, as it is often case with useful graph related polynomials, the arguments consist of induction which exploits the recurrence relation.

In order to prove **Theorem** 3.3 we will show that the matching polynomial divides the characteristic polynomial of a certain tree, hence as characteristic polynomial has only real roots we will be done. The following lemma is a special case of this idea, which we will use to prove it in general.

The proof we offer is our own work and is simpler than the one offered by Godsil and Gutman, whose proof looks more like the one we used to show **Lemma** 3.2.

Lemma 3.6. Let G be a tree and let A be its adjacency matrix. Then G's matching polynomial matches its characteristic polynomial

$$\mu_G(x) = \chi_G(x) = \det(xI - A)$$

Proof. We prove the result by induction on n, the number of vertices. As a basis we notice that if $n = 1, 2$ both polynomials equal $x, x^2 - 1$. For the step we assume the result is true for all trees with less than n vertices.

Let v be a leaf of the tree and u the only vertex adjacent to v. Using **Lemma** 3.5 a) for $e = u \sim v$ we obtain $\mu_G(x) = \mu_{G-e}(x) - \mu_{G-u-v}(x) = x\mu_{G-v}(x) - \mu_{G-u-v}(x)$.

If we let the first row/column of A correspond to v and second to u we have:

$$\det(xI_n - A) = \det \left(\begin{array}{cc|cccccc} x & -1 & 0 & 0 & 0 & \cdots & 0 \\ -1 & x & . & . & . & ... & . \\ \hline 0 & . & & & & & \\ 0 & . & & & & & \\ 0 & . & \multicolumn{5}{c}{xI_{n-2} - A(G-v-u)} \\ \vdots & \vdots & & & & & \\ 0 & . & & & & & \end{array} \right)$$

$$= x\det(xI_{n-1} - A(G-v)) + (-1)^2 \det \left(\begin{array}{c|ccccc} -1 & . & . & . & \cdots & . \\ \hline 0 & & & & & \\ 0 & & & & & \\ 0 & \multicolumn{5}{c}{xI_{n-2} - A(G-v-u)} \\ \vdots & & & & & \\ 0 & & & & & \end{array} \right)$$

$$= x\det(xI_{n-1} - A(G-v)) - \det(xI_{n-2} - A(G-v-u))$$

Where dots represent entries we did not determine but which cancelled out during the Laplace expansion. We used Laplace's formula to expand along the first row for the second equality and first column for the third.

This shows $\chi_G(x) = x\chi_{G-v}(x) - \chi_{G-u-v}(x) = x\mu_{G-v}(x) - \mu_{G-u-v}(x) = \mu_G(x)$ where the second equality comes from applying inductive assumption to $G - v$ and $G - u - v$ having $n - 1, n - 2 < n$ vertices respectively. $\square$

We start by defining the *path tree* $T_v(G)$ of a graph G starting at vertex v as follows, vertices of $T_v(G)$ are paths of G which do not pass through any vertex twice, while two paths are connected by an edge if one extends the other by one vertex. As an example we present Figure 8: Our goal is to show that the matching polynomial divides the characteristic polynomial of any of its path trees. We need an intermediary result first. It shows in some sense that the matching polynomial changes the same for a graph as it does for its path tree upon a removal of a vertex.

Lemma 3.7. For a graph G and a vertex v we have

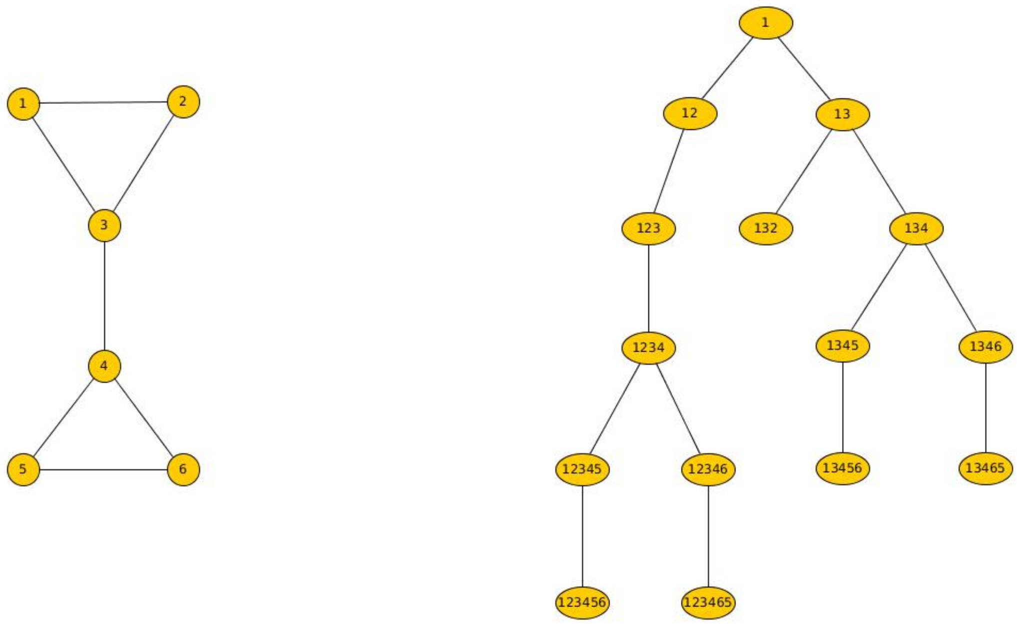
$$\frac{\mu_{G-v}(x)}{\mu_G(x)} = \frac{\mu_{T_v(G)-v}(x)}{\mu_{T_v(G)}(x)}.$$

Where the nominator of the right hand side denotes a forest created by removing v from $T_v(G)$, so we can write it as a disjoint union of trees

$$\bigcup_{u \sim v} T_u(G - v).$$

The proof of this lemma might seem technical and complicated, but there is very little more than following the standard recurrence and induction method which leads us to which terms to analyse.

Figure 8: An example graph with its path tree.



Proof. From the definition of a path tree it follows that a path tree of a tree graph equals the graph itself. So the claim of this lemma is true for trees. As all graphs on at most 2 vertices are trees the claim is true for any graph on at most 2 vertices.

We now proceed by induction using the above consideration as a basis, and assume that the result is true for any graph on less than n vertices.

Throughout this proof all the polynomials will have parameter x so we suppress it in the notation, so $\mu(G)$ will stand for $\mu_G(x)$.

Using **Lemma** 3.5 b) we obtain:

$$\frac{\mu(G)}{\mu(G-v)} = \frac{x\mu(G-v) - \sum_{u\sim v}\mu(G-v-u)}{\mu(G-v)} =$$

$$x - \sum_{u\sim v}\frac{\mu(G-v-u)}{\mu(G-v)} = x - \sum_{u\sim v}\frac{\mu(T_u(G-v)-u)}{\mu(T_u(G-v))} \tag{1}$$

Where the last equality follows from the inductive assumption applied to $G-v$. Applying the observation before this proof we get:

$$T_u(G-v) - u = \bigcup_{w\sim u, w\neq v} T_w(G-u-v) \tag{2}$$

$$T_v(G) - v = \bigcup_{w\sim v} T_w(G-v)$$

This equation combined with **Lemma** 3.3 c), as the union on the right hand side is disjoint, implies the following:

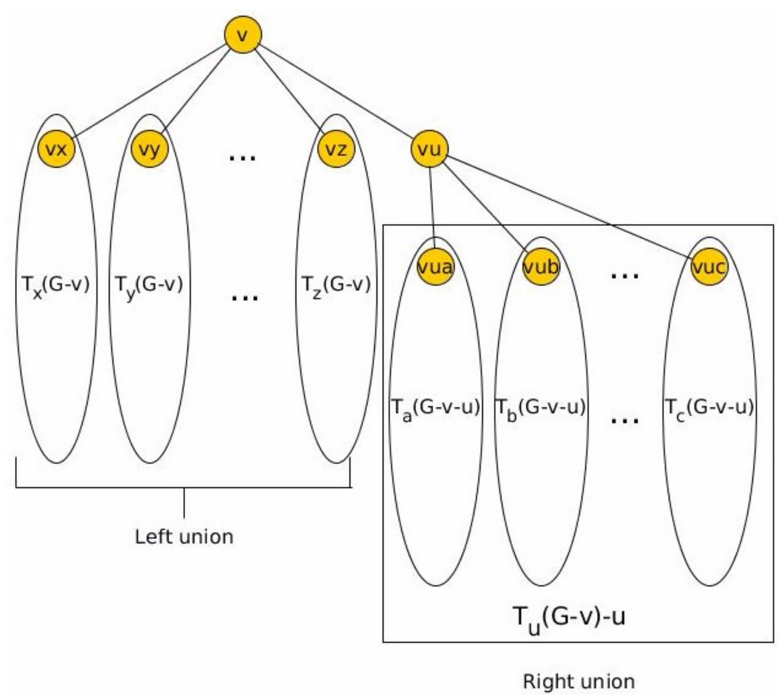
$$\mu(T_v(G) - v) = \prod_{w \sim v} \mu(T_w(G - v)) \tag{3}$$

For $u \sim v$ we denote by vu the vertex in $T_v(G)$ corresponding to the path from v to u.

$$\begin{aligned} T_v(G) - v - vu &= \left(\bigcup_{w \sim v, w \neq u} T_w(G - v) \right) \cup \left(\bigcup_{w \sim u, w \neq v} T_w(G - v - u) \right) \\ &= \left(\bigcup_{w \sim v, w \neq u} T_w(G - v) \right) \cup (T_u(G - v) - u) \end{aligned} \tag{4}$$

Where the first equality follows by considering components of the tree $T_v(G)$ after removal of v and vu, the left union consisting of components consisting of descendants of children of v, except u, and the right by components consisting of descendants of children of vu. While the last equality follows from (2). **Figure** 9 might help with following this argument.

Figure 9: Illustration of the component structure of $T_v(G) - v - vu$.



Now taking the matching polynomials of trees on both sides of (4) and invoking **Lemma** 3.3 c) we obtain

$$\mu(T_v(G) - v - vu) = \left(\prod_{w \sim v, w \neq u} \mu(T_w(G - v)) \right) \cdot \mu(T_u(G - v) - u)$$

Dividing this by (3) we obtain:

$$\frac{\mu(T_v(G)-v-vu)}{\mu(T_v(G)-v)} = \frac{\left(\prod_{w\sim v, w\neq u} \mu(T_w(G-v))\right)\cdot \mu(T_u(G-v)-u)}{\prod_{w\sim v}\mu(T_w(G-v))} = \frac{\mu(T_u(G-v)-u)}{\mu(T_u(G-v))}$$

Now plugging this into (1) gives us:

$$\begin{aligned}\frac{\mu(G)}{\mu(G-v)} &= x - \sum_{u\sim v}\frac{\mu(T_u(G-v)-u)}{\mu(T_u(G-v))}\\ &= x - \sum_{u\sim v}\frac{\mu(T_v(G)-v-vu)}{\mu(T_v(G)-v)}\\ &= \frac{x\mu(T_v(G)-v)-\sum_{u\sim v}\mu(T_v(G)-v-vu)}{\mu(T_v(G)-v)}\\ &= \frac{\mu(T_v(G))}{\mu(T_v(G)-v)}\end{aligned} \tag{5}$$

as desired. □

We now state and prove the result we hinted about before,

Theorem 3.8. Let G be a graph and $v \in V(G)$ then $\mu_G(x) \mid \mu_{T_v(G)}(x)$.

Proof. We again proceed by induction, as the basis, the result is trivial for G a tree so holds on graph with at most 2 vertices. We now assume inductively that it holds for $G-v$ so for $u \sim v$

$$\mu_{G-v}(x) \mid \mu_{T_u(G-v)}(x) \mid \prod_{w\sim v}\mu_{T_w(G-v)} = \mu_{T_v(G)-v}$$

Where we used **Lemma** 3.3 a) and c) as in the previous proof for the last equality. In particular we have shown that $\frac{\mu_{T_v(G)-v}}{\mu_{G-v}(x)}$ is a polynomial in x. So in particular using **Theorem** 3.7:

$$\mu_{T_v(G)}(x) = \mu_{T_v(G)-v}(x)\frac{\mu_G(x)}{\mu_{G-v}(x)} = \mu_G(x)\frac{\mu_{T_v(G)-v}(x)}{\mu_{G-v}(x)}$$

which shows the claim. □

We now summarise the proof of **Theorem** 3.3.

Proof of Theorem 3.3. For v a By **Theorem** 3.8. and **Lemma** 3.6 we have

$$\mu_G(x) \mid \mu_{T_v(G)}(x) = \chi_{T_v(G)}(x).$$

As all the roots of the characteristic polynomials are real so are the roots of $\mu_G(x)$ as claimed. □

To prove **Theorem** 3.4. we need another small ingredient.

Lemma 3.9. If T is a tree with maximal degree at most d then its eigenvalues are at most $2\sqrt{d-1}$ in absolute value.

Proof. We root the tree, which means we pick a vertex v which we call root and assign *depth* $d(u)$ to each other vertex u, defined as distance to v.

Let A be the adjacency matrix. We define a diagonal matrix D as follows

$$D(u,u) \equiv \left(\sqrt{d-1}\right)^{d(u)}.$$

We notice that A and DAD^{-1} have the same eigenvalues as:

$$\det(xI - DAD^{-1}) = \det(D(xI)D^{-1} - DAD^{-1}) = \det(D(xI-A)D^{-1})$$
$$= \det(D)(\det(xI-A))\det\left(D^{-1}\right) = \det(xI-A).$$

We take a look at the entries of DAD^{-1}

$$DAD^{-1}_{i,j} = \sum_{k,l} D_{i,k}A_{kl}D^{-1}_{l,j} = D(i,i)A_{i,j}D(j,j)^{-1} = A_{i,j}\left(\sqrt{d-1}\right)^{d(i)-d(j)}$$

We now look at the row sums of DAD^{-1}, our goal is to prove that each row sum is at most $2\sqrt{d-1}$. We have three cases to consider:

- Row corresponding to the root v. We have $D(v,v) = 1$ and $D(u,u) = \sqrt{d-1}$ for each neighbour of v. $A_{v,u}$ is non-zero only when $v \sim u$ so the row sum equals $\sum_{u=1}^{n} A_{v,u}\left(\sqrt{d-1}\right)^{d(v)-d(u)} = \sum_{u\sim v} 1\cdot\left(\sqrt{d-1}\right)^{0-1} = \frac{d}{\sqrt{d-1}} \le 2\sqrt{d-1}$
- Row corresponding to a leaf v. v only has one neighbour w which has $d(w) = d(v)-1$. Hence
$$\sum_{u=1}^{n} A_{v,u}\left(\sqrt{d-1}\right)^{d(v)-d(u)} = 1\cdot\left(\sqrt{d-1}\right)^{d(v)-d(w)} = \sqrt{d-1} \le 2\sqrt{d-1}$$
- Row corresponding to an intermediary vertex v. If w denotes the parent of v and $u_1, \cdots u_k$ denote its children, which are defined by the rooting of the tree. We note that $d(w) = d(v)-1$ and $d(u_i) = d(v)+1$ and $k+1 \le d$.
$$\sum_{u=1}^{n} A_{v,u}\left(\sqrt{d-1}\right)^{d(v)-d(u)} = \left(\sqrt{d-1}\right)^{d(v)-d(w)} + \sum_{i=1}^{k}\left(\sqrt{d-1}\right)^{d(v)-d(u_i)} =$$
$$\left(\sqrt{d-1}\right)^{d(v)-(d(v)-1)} + \sum_{i=1}^{k}\left(\sqrt{d-1}\right)^{d(v)-(d(v)+1)} = \sqrt{d-1} + k\left(\sqrt{d-1}\right)^{-1} \le$$
$$\sqrt{d-1} + (d-1)\left(\sqrt{d-1}\right)^{-1} = 2\sqrt{d-1}$$

  Finally, if λ is an eigenvalue and $\overline{x}$ a corresponding non-zero eigenvector. We reindex the vertices so that $|x_1|$ is maximal, then $x_1 \neq 0$

  Now $|\lambda x_1| = |(A\overline{x})_1| = |\sum_{i=1}^{n} A_{1,i}x_i| \le \sum_{i=1}^{n} A_{1,i}|x_i| \le 2\sqrt{d-1}|x_1|$ cancelling out the $|x_1|$ we obtain $|\lambda| \le 2\sqrt{d-1}$ as desired.

□

We are now ready to prove **Theorem** 3.4.

Proof of Theorem 3.4. If G is a d regular graph, given $v \in V(G)$. Let x be a vertex of $T_v(G)$, corresponding to a path ending in $u \in V(G)$ then x can have at most d neighbours in $T_v(G)$ as an edge from x corresponds to either appending x to its parent or appending a neighbour of u to x. Hence x has degree less than d in $T_v(G)$. As x was arbitrary, $T_v(G)$ has maximal degree d.

We now apply **Lemma** 3.9 to $T_v(G)$ which implies that all roots of $\chi(T_v(G))$ are smaller than $2\sqrt{d-1}$. Now **Theorem** 3.8 and **Theorem** 3.6 give us $\mu_G(x) \mid \mu_{T_v(G)}(x) = \chi(T_v(G))$ implying that all roots of μ_G also satisfy this inequality, completing the proof. □

3.5 Interlacing

This section contains the basis of the novel idea of Marcus, Spielman and Srivastava which enables them to prove the existence of infinite families of d-regular Ramanujan graph.

We note that using interlacing to control roots, as itself is not novel, for example the proof of Heilmann and Lieb of **Theorem** 3.3 is based on a similar idea. The novelty is in the way we use it, in particular through interlacing families.

The only missing ingredient in our main proof is relation between the roots of μ_G and the roots characteristic polynomials of signings $f_s(x)$. We know by **Theorem** 3.2 that $\mu_G(x) = \mathbb{E} f_s(x)$.

Unfortunately given polynomials $f + g = h$ it is possible that h has a larger root, in absolute value than f and g.

Example: Let

$$f(x) = (x-1)(x-8) = x^2 - 9x + 8 \text{ and}$$
$$g(x) = (x-9)(x-10) = x^2 - 19x + 90 \text{ then}$$
$$h(x) = f(x) + g(x) = 2x^2 - 28x + 98 = 2(x-7)^2$$

and in particular largest root of h is 7 which is smaller than both largest roots of f and g, being $8, 10$ respectively.

Luckily for us, the polynomials f_s in our sum are quite mutually dependent and as such we are able to find an explicit relation between them, allowing the above argument.

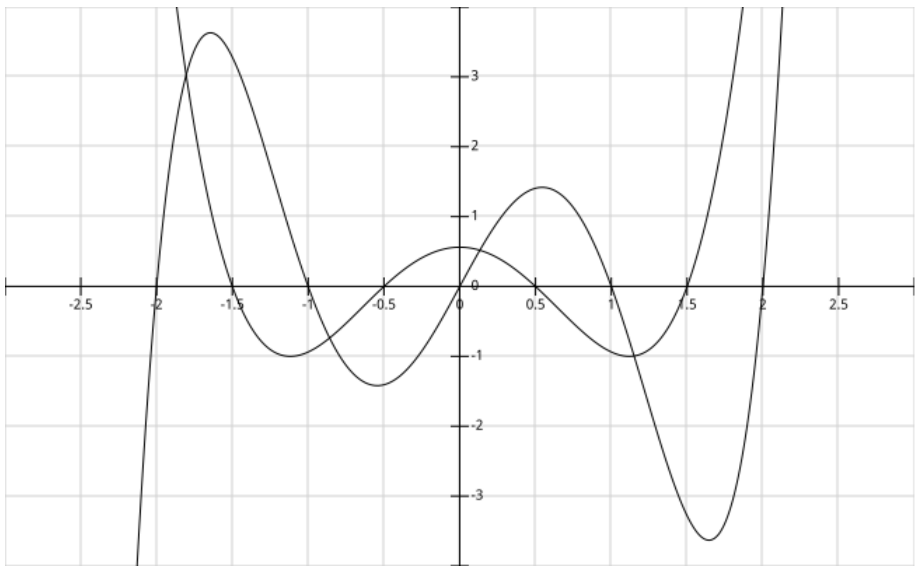
Definition We say that a polynomial $g(x) = \prod_{i=1}^{n-1}(x - \alpha_i)$ *interlaces* a polynomial $f(x) = \prod_{i=1}^{n}(x - \beta_i)$ if

$$\beta_1 \leq \alpha_1 \leq \beta_2 \leq \alpha_2 \leq \cdots \leq \alpha_{n-1} \leq \beta_n.$$

We say that polynomials $f_1, f_2, \cdots, f_k$ have a *common interlacing* if there is a polynomial f so that f interlaces f_i for all i.

We include the following plot to give some feeling on how interlacing polynomials "look like".

Figure 10: $(x-3/2)(x-1/2)(x+1/2)(x+3/2)$ interlaces $(x-2)(x-1)x(x+1)(x+2)$



Let $\beta_{i,1} \leq \cdots \leq \beta_{i,n}$ denote the roots of f_i, then f_i have a common interlacing if and only if there exist reals $\alpha_0 \leq \alpha_1 \leq \cdots \leq \alpha_n$ such that $\beta_{i,j} \in [\alpha_{j-1}, \alpha_j]$ for all j. Note that α_j, $1 \leq j \leq n-1$ will correspond to roots of the common interlace, while α_0, α_n can be taken arbitraily as long as they are smaller/larger than all $\beta_{i,j}$.

The following lemma shows that having a common interlace is a sufficient extra condition guaranteeing that one of the summands has a smaller root then the sum. The proof is very easy and is more of an exercise in using the definitions than anything else.

We define $r(f)$ to be the largest real root of a polynomial f, we set $r(f) = -\infty$ if f has no real roots.

Lemma 3.10. Let f, g be polynomials of the same degree, having only real roots and with positive leading coefficients. Let $h = f + g$. If f, g have a common interlacing then $\min(r(f), r(g)) \leq r(h)$.

Proof. Let n be the degree of the polynomials involved and let i denote a polynomial interlacing f, g. Let $\alpha_{n-1} = r(i)$, which is real by the definition of the interlacing. We also define $f_{n-1} \leq f_n$ to be 2 largest roots of f, similarly define g_{n-1}, g_n. We now have $g_{n-1}, f_{n-1} \leq \alpha_{n-1} \leq g_n, f_n$

As f, g have positive leading coefficient $f(x), g(x)$ are non-negative on $[f_n, \infty]$ and $[g_n, \infty]$ respectively so they are non-positive on $[f_{n-1}, f_n]$ and $[g_{n-1}, g_n]$. As α_{n-1} is contained in both these intervals we deduce $f(\alpha_{n-1}), g(\alpha_{n-1}) \leq 0$ so $h(\alpha_{n-1}) = f(\alpha_{n-1}) + g(\alpha_{n-1}) \leq 0$.

h also has a positive leading coefficient so is eventually positive so the intermediate value theorem implies there is a root of h larger than α_{n-1}. In particular $\beta_n \equiv r(h) \geq \alpha_{n-1}$.

Now $f(\beta_n) + g(\beta_n) = 0$ so either $f(\beta_n) \geq 0$ or $f(\beta_n) \geq 0$, without loss of generality let us assume this holds for f.

Now as $f(\alpha_{n-1}) \leq 0$ there is a root of f in the interval $[\alpha_{n-1}, \beta_n]$ and by the interlacing property there can be only one x such that $f(x) = 0$ and $x \geq \alpha_{n-1}$ so the largest root of f is in th interval $[\alpha_{n-1}, \beta_n]$ and in particular $r(f) \leq \beta_n = r(h)$ as desired. $\square$

Remarks:

- This proposition is true for n polynomials instead of just f, g and the proof goes in exactly the same way. We have proved the 2 polynomial version as we believe it is easier to visualise and we only need this version in our proof.
- A further reason we have chosen to prove the 2 polynomial version rather than n, is to emphasize that this is an intermediate result, we are not able to apply the n polynomial version to our characteristic polynomials f_s as it is possible there is no common interlace for all the summands.
- The consideration of this Lemma can be extended to further roots, showing f or g's k-th root is bounded above by h's. The proof is along the same lines as the one above.
- We also note that in our previous example the lemma does not apply as the polynomials do not admit a common interlace.

The idea for the proof is to apply the above lemma in layers. In each step we split the current sum of polynomials into 2 sums f_1, f_{-1} by fixing a sign of an edge and apply the above lemma to f_1, f_{-1} and repeat for the one guaranteed to have smaller maximal root than the sum.

We remind the reader that we defined a signing as an assignment of ± 1 to each of the m edges of the graph, if we number the edges as $e_1, \cdots, e_m$ we can represent a signing as an element of $s = (s_1, \cdots, s_n) \in \{\pm 1\}^m$, furthermore we defined $f_s = f_{s_1, \cdots, s_m}$ as the characteristic polynomial of the corresponding lift.

Definition. Given a partial signing $s' = (s_1, \cdots, s_k) \in \{\pm 1\}^k$ for $k < m$ we define the sums

$$f_{s_1, \cdots, s_k} = \sum_{s_{k+1}, \cdots, s_m \in \{\pm 1\}} f_{s_1, s_2, \cdots, s_m}$$

as well as

$$f_{\emptyset} = \sum_{s_1, \cdots, s_m \in \{\pm 1\}} f_{s_1, s_2, \cdots, s_m}$$

So in particular $f_{s_1, \cdots, s_k}$ corresponds to the sum of the characteristic polynomials of signings having fixed first k signs.

Theorem 3.11. If for all $0 \leq k \leq m-1$ and all partial signings $(s_1, \cdots, s_k) \in \{\pm 1\}^k$ polynomials $f_{s_1, \cdots, s_k, 1}$ and $f_{s_1, \cdots, s_k, -1}$ have a common interlacing then there is a signing $s = (s_1, \cdots, s_m) \in \{\pm 1\}^m$ such that $r(f_s) \leq r(f_{\emptyset})$.[5]

[5] *Remark.* If the conditions of the theorem are satisfied the polynomials $\{f_s\}_{s \in \{\pm 1\}^m}$ are said to form an

Proof. By the assumption we know that f_1 and f_{-1} have a common interlacing and $f_\emptyset = f_1 + f_{-1}$ so **Lemma** 3.10 implies there is an $s_1 \in \{\pm 1\}$ such that $r(f_{s_1}) \leq r(f_\emptyset)$.

We now proceed inductively using the above consideration as a basis. Our inductive assumption is that there exists a partial signing $s' = (s_1, \cdots, s_k)$ such that $r(f_{s'}) \leq r(f_\emptyset)$.

We have $f_{s'} = f_{s_1,\cdots,s_k,1} + f_{s_1,\cdots,s_k,-1}$ and the assumption of the theorem is that the summands have an interleaver. Hence we can apply **Lemma** 3.10. to conclude there is an s_{k+1} such that $r(f_{s_1,\cdots,s_k,s_{k+1}}) \leq r(f_{s'}) \leq r(f_\emptyset)$ completing the step, the induction and the proof as the step $k = m$ is exactly the claim of the theorem. ☐

We are now concerned with proving the assumption of the **Theorem** 3.11 actually holds. There is a systematic way of doing this that we follow.

We start with the following Lemma, it appears as a special case of **Theorem** 3.1 of Dedieu [20], Theorem 2 of Fell [21] and as consequence of Theorem 3.5 or a special case of **Theorem** 3.6 of Chudnovsky and Seymour [22]. Each author offers a proof based on significantly different underlying ideas. We present our version of the proof based on the one given by Chudnovsky and Seymour.

We say polynomials f, g are *compatible* if $tf + (1-t)g$ has only real roots for every $t \in [0, 1]$.

Lemma 3.12. Let f, g be polynomials of the same degree with positive leading coefficients. If f, g are compatible then f, g have a common interlacing.[6]

Before the proof we need a couple of easy propositions.

We define $z_f(x)$ to denote the number of zeros of a polynomial f larger or equal then x.

We say f, g *agree* at a point x if $f(x)g(x) > 0$ so in particular if f, g are both strictly positive or strictly negative at x.

Proposition 3.13. Let f, g be compatible polynomials which agree at $a, b \in \mathbb{R}$, $a < b$. Then the following holds:

$$z_f(b) - z_f(a) = z_g(b) - z_g(a)$$

Proof. We define $p_t(x) = tf(x) + (1-t)g(x)$ for each $0 \leq t \leq 1$. We note that $p_t(a), p_t(b) \neq 0$ as the definition of agreeing requires strict positivity/negativity.

As t varies from 0 to 1 the roots of p_t move continuously in the complex plane, but by the compatibility assumption all the roots of each p_t are real so the roots need to move continuously on the real line. This combined with the fact no zeros of any p_t are at the boundary implies the number of zeros in the interval $[a, b]$ is constant as p varies and in

---

interlacing family. Marcus, Spielman and Srivastava in [1] defined it in a more general setting, replacing $\{\pm 1\}^m$ with a product of not necessarily different finite sets and removing the relation to signings. The theorem remains true in this setting as well. We decided to prove it in this form as it covers all the ideas but is less cumbersome on the notational side.

[6] The only if statement is also true. The lemma is also true for n polynomials. But we decide to present the simplest version necessary for our main proof.

particular for $p = 0, 1$ and $p_0 = g$, $p_1 = f$. Hence

$$z_f(b) - z_f(a) = z_g(b) - z_g(a).$$

□

Proposition 3.14. Let f, g be compatible polynomials of equal degrees and positive leading coefficients. Then $|z_f(x) - z_g(x)| \leq 1$ for all real x.

Proof. We proceed by induction, the claim is trivial for $n = 1$. We assume it to be true for $n - 1$.

If f, g have a common root r then $f(x) = (x - r)f_1$ and $g(x) = (x - r)g_1$. We note that $|z_f(x) - z_g(x)|$ is never changed by r, also $tf(x) + (1 - t)g(x) = (x - r)(tf_1(x) + (1 - t)g_1(x))$ shows that compatibility of f, g implies compatibility of f_1, g_1. We now apply induction on f_1, g_1 and are done.

So we may assume f, g have no common roots.

We will need that for a real rooted polynomial f, if x is its root and there are k roots larger or equal than x, then there are $k - 1$ roots of its derivative f' larger or equal than x. This follows by observing that multiplicity of a root in f' is by one less than in f to handle repeated roots and applying mean value theorem several times to place zeros of the derivative between different roots.

Using this we see that f, g compatible implies f', g' are compatible.

Suppose, for the sake of contradiction, that there exists a y such that $z_f(y) - z_g(y) \geq 2$. By the fact $z_f(y) - z_g(y)$ only decreases at roots of f we may assume that y is a root of f.

Suppose that $z_f(y) - z_g(y) \geq 3$. We have $z_{f'}(y) = z_f(y) - 1$ and $z_{g'}(y) \leq z_g(y)$, hence $z_{f'}(y) - z_{g'}(y) \geq z_f(y) - z_g(y) - 1 \geq 2$ which is a contradiction to the inductive assumption applied to f', g' which are both of degree $n - 1$ and are compatible.

So the only option is $z_f(y) - z_g(y) = 2$. Let b be larger than all the roots of either of f, g, then as both f, g have positive leading coefficient f, g agree on b.

We notice that if x is not a zero of f,

$$\begin{aligned} 2 \mid z_f(x) &\Rightarrow f(x) > 0 \text{ and} \\ 2 \nmid z_f(x) &\Rightarrow f(x) < 0 \end{aligned}$$

As the sign of f changes at each zero, fictitiously for repeated roots, and f is ultimately positive as it has a positive leading coefficient. Hence if we pick $a < y$ but close enough to y so that there are no zeros of f or g in $[a, y)$, f, g agree at a.

Now applying **Proposition** 3.13 we get

$$0 = z_f(b) - z_g(b) = z_f(a) - z_g(a) = z_f(y) - z_g(y) = 2$$

which is a contradiction. □

Proof of Lemma 3.12. Let $a_1 \leq a_2, \leq \cdots a_n$ be the roots of f and $b_1 \leq b_2, \leq \cdots \leq b_n$ the roots of g.

Let us assume for the sake of contradiction that there exists an $i \leq n-1$ such that $[a_i, a_{i+1}] \cap [b_i, b_{i+1}] = \emptyset$ as otherwise we can easily find an interlacing.

Without loss of generality in this case we have $a_{i+1} < b_i$. If we choose a $c \in (a_{i+1}, b_i)$ then $z_f(c) = n-i-1$ and $z_g(c) = n-i+1$ giving $z_f(c) - z_g(c) = 2$ which is a contradiction to **Proposition** 3.14. and completes the proof. □

3.6 Main Result

We need the last ingredient before proving the main theorem. We state it here and provide a proof in the following subsection. This is the only result for which we do not provide a self-contained proof as certain parts of the proof, while being standard results in the context, rely on results with long and very analytical proofs.

Theorem 3.15. Given $p_1, \cdots p_m \in [0,1]$. Then the following polynomial is real rooted.

$$\sum_{s\in\{\pm 1\}^m} \left(\prod_{i\,:s_i=1} p_i\right)\left(\prod_{i\,:s_i=-1} (1-p_i)\right) f_s(x)$$

We notice that this polynomial is the expected polynomial of random signings, but this time not taken uniformly. Setting $p_i = \frac{1}{2}$ and combining with **Theorem** 3.2 we obtain an alternative proof of the **Theorem** 3.3.

Theorem 3.16. $\{f_s\}_{s\in\{\pm1\}^m}$ form an interlacing family, in particular the assumption of **Theorem** 3.11 is satisfied.

Proof. Applying the **Theorem** 3.15 with $p_i = \frac{s_i+1}{2}$ for $1 \leq i \leq k$, $p_{k+1} = t$ and $p_i = \frac{1}{2}$ for $i \geq k+2$ we get the following polynomial is real rooted:

$$\sum_{s'\in\{\pm 1\}^m} \left(\prod_{i\,:s'_i=1} p_i\right)\left(\prod_{i\,:s'_i=-1} (1-p_i)\right) f_s(x) = 2^{n-k-1}(tf_{s_1,s_2,\cdots,s_k,1} + (1-t)f_{s_1,s_2,\cdots,s_k,-1})$$

As $p_i = 0$ if $s_i = -1$ and $1 - p_i = 0$ if $s_i = 1$ for $i \leq k$ so the summand is 0 unles $s'_i = s_i$ for all $i \leq k$, the factor 2^{n-k-1} is due to the p_i for $i \geq k+2$ and the t combination comes from p_{k+1}.

So this shows $f_{s_1,s_2,\cdots,s_k,1}$ and $f_{s_1,s_2,\cdots,s_k,-1}$ are compatible. They are of the same degree and have positive leading coefficients so **Lemma** 3.12 applies showing f, g have a common interlacing as desired. □

We are finally ready to prove the main result

Theorem 3.17. For each $d \geq 2$ there is an infinite family of d-regular bipartite Ramanujan graphs.

Proof. For $d = 2$ taking G an even cycle graph C_{2n} is bipartite and Ramanujan by **Lemma** 2.6. We now assume $d \geq 3$.

We start with a Ramanujan bipartite graph G_1 such as $K_{d,d}$, or in fact any provided by **Lemma** 2.7.

We proceed by induction, given a bipartite Ramanujan graph G_k, using **Theorem** 3.11 we can find a signing s such that largest root of f_s is smaller than largest root of $f_\emptyset$.

Theorem 3.2 implies $f_\emptyset = 2^m \mu_G$ so they have the same roots and in particular from **Theorem** 3.4 we know the largest root of $f_\emptyset$ is at most $2\sqrt{d-1}$.

This implies eigenvalues of A_s are bounded above by $2\sqrt{d-1}$, while all non-trivial eigenvalues of the adjacency matrix A of G_k are bounded above by $2\sqrt{d-1}$ by the inductive assumption that G_k is Ramanujan.

We define G_{k+1} to be a 2-lift of G_k corresponding to s. We note that this implies G_{k+1} is also bipartite, it is also d-regular and has double the number of vertices.

By **Theorem** 3.1 eigenvalues of G_{k+1} consist of a union of eigenvalues of A and A_s implying that all the non-trivial eigenvalues of G_{k+1} are smaller than $2\sqrt{d-1}$. The final ingredient is to use **Lemma** 2.1 d) to conclude that, as G_{k+1} is bipartite, all non-trivial eigenvalues are larger than $-2\sqrt{d-1}$. Hence G_{k+1} is also Ramanujan.

Hence $G_1, G_2, \cdots$ make an infinite family of bipartite Ramanujan graphs. □

3.7 Real Stability

This section is concerned with proving **Theorem** 3.15. We need to develop a machinery of a certain multivariate generalization of real-rootedness called *real stability*. It was first studied by Borcea and Branden in [23] and [24]. We will present a small portion we require for more details about the method we point an interested reader to a survey paper by Wagner [25].

Definition. A multivariate polynomial $f \in \mathbb{R}[x_1, \cdots, x_n]$ is said to be *real stable* if it is the zero polynomial or if $f(x_1, \cdots, x_n) \neq 0$ whenever $\Im(x_i) > 0$ for every i.

Note that in the definition the polynomial is constrained to have real coefficients, but we may evaluate it at complex numbers.

We start with an easy proposition, on the way real-stability generalizes real-rootedness.

Proposition 3.18. If $n = 1$ real-stability is equivalent to real rootedness.

Proof. By the funadmental theorem of algebra f has n complex roots, if z is a root then so is its complex conjugate z^* but unless z is real one of z, z^* have positive imaginary part and f vanishes so a real stability implies real-rootedness.

For the other direction, if f is real rooted it only vanishes at points with imaginary part 0 so it is real stable. $\square$

We now present the Hurwitz's theorem, a standard theorem of complex analysis. We need it for several limiting arguments in what follows.

Theorem 3.19. Let $\sigma \subseteq \mathbb{C}^m$ be a connected open set, and let $f_n, n \in \mathbb{N}$ be a sequence of functions, each analytic and non-vanishing on σ, which converges to a limit f uniformly on compact subsets of σ. Then f is either non-vanishing on σ or identically zero.

Proof. This is a standard result in complex analysis, we point the reader to Theorem 1.3.8 of [26]. $\square$

There are several closure properties satisfied by real stable polynomials, for example a product of 2 real stable polynomials is trivially also real stable. We prove the following specification result we will make us of later, the proof demonstrates why Hurwitz is useful in this setting.

Proposition 3.20. If $f(x_1, x_2, \cdots, x_n)$ is a real stable polynomial in n variables then $f(x_1, x_2, \cdots, x_{n-1}, r)$ is a real stable polynomial in $n-1$ variables, for arbitrary, fixed real r.

Proof. $f(x_1, x_2, \cdots, x_{n-1}, r)$ is clearly real stable if we allow r to be complex with $\Im(r) > 0$ as any zero with $\Im(x_i) > 0$ for $i \leq n-1$ would give appended with r a zero of $f(x_1, x_2, \cdots, x_n)$ with this property contradicting its real stability.

Now we apply Hurwitz to a sequence $f_k = f(x_1, x_2, \cdots, x_{n-1}, r + i2^{-k})$ to conclude the result. $\square$

The following lemma will provide us with a certain family of real stable polynomials on which we will build later on.

Lemma 3.21. Let $A_1, A_2. \cdots, A_m$ be positive semi-definite matrices. Then the following polynomial is real stable:

$$f(x_1, x_2, \cdots, x_n) = \det(x_1 A_1 + \cdots + x_m A_m).$$

Proof. Notice that

$$\overline{f} = \det(x_1 \overline{A_1} + \cdots + x_m \overline{A_m} + \overline{H}) = \det(x_1 A_1^t + \cdots + x_m A_m^t) = \det(x_1 A_1 + \cdots + x_m A_m) = f$$

where we used $\overline{A_i} = A_i^t$ implied by the semi-definite assumption and $\det(A) = \det(A^t)$.

This implies $f \in \mathbb{R}[\overline{x}]$.

By Hurwitz's Theorem applied to f modified by taking $A_i + \frac{1}{k} I$ it suffices to prove that f is stable when each A_i is positive definite.

Let $x_j = a_j + ib_j$ such that $a_j, b_j \in \mathbb{R}$ for all j. Assuming $b_j > 0$ for all j we get that $B = \sum_{j=1}^m b_j A_j$ is still positive definite. This implies B has a positive definite square root matrix $C = B^{1/2}$. While $A = \sum_{j=1}^m a_j A_j$ is Hermitian. Now

$$f(\overline{x}) = \det(A + iB) = \det\left(CC^{-1}AC^{-1}C + CiIC\right) = \det(C)\det\left(C^{-1}AC^{-1} + iI\right)\det(C)$$
$$= \det(B)\det\left(C^{-1}AC^{-1} + iI\right)$$

Now if $f(\overline{x}) = 0$ then as B is positive definite $\det(B) \neq 0$ so $-i$ is an eigenvalue of $C^{-1}AC^{-1}$, which is a Hermitian matrix, as C Hermitian implies C^{-1} is Hermitian as well and taking a conjugate transpose of $C^{-1}AC^{-1}$ gives itself. This is impossible as Hermitian matrices have only real roots, in particular f is real stable. □

We now state a result of Borcea and Brändén which characterizes an entire class of differential operators that preserve real stability. To simplify notation we write ∂_{x_i} for partial differentiation with respect to x_i.

Theorem 3.22. Let $T : \mathbb{R}[x_1, \cdots, x_n] \to \mathbb{R}[x_1, \cdots, x_n]$ be an operator of the form

$$\sum_{\alpha,\beta\in\mathbb{N}^n} c_{\alpha,\beta} \prod_{i=1}^n x_i^{\alpha_i} \prod_{i=1}^n (\partial_{x_i})^{\beta_i}$$

where $c_{\alpha,\beta} \in \mathbb{R}$ is zero for all but finitely many terms. We define

$$F_T(\overline{x}, \overline{y}) = \sum_{\alpha,\beta} c_{\alpha,\beta} \prod_{i=1}^n x_i^{\alpha_i} \prod_{i=1}^n y_i^{\beta_i}$$

T preserves real stability if and only if $F_T(\overline{x}, -\overline{y})$ is real stable.

Proof. This is one of the key results of the theory and requires far more machinery then we present here. We point the reader to Theorem 1.3 in Borcea and Brändén [24]. □

We note that the above theorem for T a single partial derivative is not that hard to prove, it is in fact an easy consequence of the Gauss-Lucas theorem stating that roots of the derivative of a polynomial are contained in the convex hull of its roots. The hard part is combining the operators. We will only need the following special case of the above theorem.

Corollary 3.23. For non-negative real numbers a and b and variables x_i and x_j, the operator $T = 1 + a\partial_{x_i} + b\partial_{x_j}$ preserves real stability.

Proof. Applying the above theorem for T, so there are only three non-zero $c_{\alpha,\beta}$, first equal to one and corresponding to $\alpha, \beta = \overline{0}$ and two equal a, b and corresponding to $\alpha = \overline{0}$ and β's all zeros except 1 in i-th/j-th position. Hence by the above theorem we only need to show $f(x_i, x_j) = 1 - ax_i - bx_j$ is real stable. If $\Im(x_i), \Im(x_j) > 0$ the imaginary part of $f(x_i, x_j)$ must also be negative, or $a = b = 0$, where the former implies $f(x_i, x_j) \neq 0$ so f is real stable and in the latter case T is identity so trivially real stability preserving. □

This finishes setting up the machinery we need, we now proceed to generate the expected characterisitc polynomial of **Theorem** 3.15 using the above operators.

We will need the following easy result, called the matrix determinant lemma.

Lemma 3.24. Let A be an invertible $n \times n$ matrix, and $\overline{u}, \overline{v}$ column vectors. Then

$$\det(A + \overline{u}\overline{v}^t) = (1 + \overline{v}^t A^{-1}\overline{u}) \det A$$

Proof. We first prove it for the case when $A = I_n$ by observing:

$$\begin{pmatrix} I_n & 0 \\ \overline{v}^t & 1 \end{pmatrix} \begin{pmatrix} I_n + \overline{u}\overline{v}^t & \overline{u} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} I_n & 0 \\ -\overline{v}^t & 1 \end{pmatrix} = \begin{pmatrix} I_n + \overline{u}\overline{v}^t & \overline{u} \\ \overline{v}^t + \overline{v}^t\overline{u}\overline{v}^t & 1 + \overline{v}^t\overline{u} \end{pmatrix} \begin{pmatrix} I_n & 0 \\ -\overline{v}^t & 1 \end{pmatrix} = \begin{pmatrix} I_n & \overline{u} \\ 0 & 1 + \overline{v}^t\overline{u} \end{pmatrix}$$

And taking the determinant of respective sides we obtain

$$\det(I_n + \overline{u}\overline{v}^t) = 1 + \overline{v}^t\overline{u}$$

which completes the proof for $A = I_n$. Now using this for vectors $A^{-1}\overline{u}$ and $\overline{v}$. we get

$$\det(A + \overline{u}\overline{v}^t) = \det A \det(I_n + (A^{-1}\overline{u})\overline{v}^t) = (1 + \overline{v}^t A^{-1}\overline{u}) \det A$$

completing the proof. □

Let Z_x be the operator, acting on multivariate polynomials, corresponding to setting variable x to 0, for $i \in \mathbb{N}$ we abbreviate by Z_i the operator $Z_{x_i} : \mathbb{R}[x_1, \cdots, x_n] \to \mathbb{R}[x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n]$. We note that by the **Proposition** 3.20 operators Z_i preserve real stability. The following lemma might look arbitrary but it is in fact exactly what comes up as the inductive step in the proof of the following theorem.

Lemma 3.25. Let A be an invertible matrix, $\overline{u}, \overline{v}$ be vectors and $p \in [0, 1]$. Then

$$Z_i Z_j (1 + p\partial_{x_i} + (1-p)\partial_{x_j}) \det(A + x_i\overline{u}\overline{u}^t + x_j\overline{v}\overline{v}^t) = p\det(A + \overline{u}\overline{u}^t) + (1-p)\det(A + \overline{v}\overline{v}^t)$$

Proof. Using **Lemma** 3.24 for $t\overline{u}$ and $\overline{u}$ and taking a derivative with respect to t we obtain the Jacobi's formula for the derivative of the determinant:

$$\partial_t(A + t\overline{u}\overline{u}^t) = (\overline{u}^t A^{-1}\overline{u}) \det A$$

Expanding the left hand side of our desired equality and applying this several time we get:

$$Z_i Z_j (1 + p\partial_{x_i} + (1-p)\partial_{x_j}) \det(A + x_i\overline{u}\overline{u}^t + x_j\overline{v}\overline{v}^t) =$$
$$Z_i Z_j \det(A + x_i\overline{u}\overline{u}^t + x_j\overline{v}\overline{v}^t) + pZ_i\partial_{x_i} Z_j \det(A + x_i\overline{u}\overline{u}^t + x_j\overline{v}\overline{v}^t) + (1-p)Z_j\partial_{x_j} Z_i \det(A + x_i\overline{u}\overline{u}^t + x_j\overline{v}\overline{v}^t) =$$
$$\det A + pZ_i\partial_{x_i} \det(A + x_i\overline{u}\overline{u}^t) + (1-p)Z_j\partial_{x_j} \det(A + x_j\overline{v}\overline{v}^t) =$$
$$\det A + pZ_i(\overline{u}^t A^{-1}\overline{u}) \det A + (1-p)Z_j(\overline{v}^t A^{-1}\overline{v}^t) \det A =$$
$$p(1 + \overline{u}^t A^{-1}\overline{u}) \det A + (1-p)(1 + \overline{v}^t A^{-1}\overline{v}^t) \det A =$$
$$p\det(A + \overline{u}\overline{u}^t) + (1-p)\det(A + \overline{v}\overline{v}^t)$$

Where we used the fact operators Z, ∂ acting on different variables commute when acting on the polynomials for the first equality, the above Jacobi's formula for the third and **Lemma** 3.24 directly for the fifth. □

We now present our main result on real-rootedness which we were building up to and which will give the desired proof of **Theorem** 3.15.

Theorem 3.26. Let $\overline{u}_1, \cdots, \overline{u}_m$ and $\overline{v}_1, \cdots, \overline{v}_m$ be vectors in $\mathbb{R}^m$, let $p_1, \cdots, p_m \in [0,1]$ be real numbers. Then

$$P(z) \equiv \sum_{S \subseteq [m]} \left(\prod_{i \in S} p_i \right) \left(\prod_{i \notin S} (1 - p_i) \right) \det \left(zI_m + \sum_{i \in S} \overline{u}_i \overline{u}_i^t + \sum_{i \notin S} \overline{v}_i \overline{v}_i^t \right)$$

is real-rooted.

Proof. We define Q as

$$Q(z, x_1, \cdots, x_m, y_1, \cdots, y_m) \equiv \det \left(zI + \sum_{i \in S} x_i \overline{u}_i \overline{u}_i^t + \sum_{i \notin S} y_i \overline{v}_i \overline{v}_i^t \right)$$

Let $T_i = 1 + p_i \partial_{x_i} + (1 - p_i)\partial_{y_i}$, we show that we can write

$$P(z) = \left(\prod_{i=1}^{m} Z_{x_i} Z_{y_i} T_i \right) Q(z, x_1, \cdots, x_m, y_1, \cdots, y_m) \tag{1}$$

To do this we prove, by induction on k the following claim:

$$\left(\prod_{i=1}^{k} Z_{x_i} Z_{y_i} T_i \right) Q(z, x_1, \cdots, x_m, y_1, \cdots, y_m) =$$

$$\sum_{S \subseteq [k]} \left(\prod_{i \in S} p_i \right) \left(\prod_{i \in [k] \setminus S} (1 - p_i) \right) \det \left(zI_m + \sum_{i \in S} \overline{u}_i \overline{u}_i^t + \sum_{i \in [k] \setminus S} \overline{v}_i \overline{v}_i^t + \sum_{i > k} x_i \overline{u}_i \overline{u}_i^t + y_i \overline{v}_i \overline{v}_i^t \right)$$

The base case $k = 0$ is simply the definition of Q. For the inductive step we apply the operator $Z_{x_{k+1}} Z_{y_{k+1}} T_{k+1}$ to the right hand side, all the operators are additive so we can fix an $S \subseteq [k]$ and consider

$$zI_m + \sum_{i \in S} \overline{u}_i \overline{u}_i^t + \sum_{i \in [k] \setminus S} \overline{v}_i \overline{v}_i^t + \sum_{i > k} x_i \overline{u}_i \overline{u}_i^t + y_i \overline{v}_i \overline{v}_i^t =$$

$$\underbrace{zI_m + \sum_{i \in S} \overline{u}_i \overline{u}_i^t + \sum_{i \in [k] \setminus S} \overline{v}_i \overline{v}_i^t + \sum_{i > k+1} x_i \overline{u}_i \overline{u}_i^t + y_i \overline{v}_i \overline{v}_i^t}_{A} + x_{k+1} \overline{u}_{k+1} \overline{u}_{k+1}^t + y_{k+1} \overline{v}_{k+1} \overline{v}_{k+1}^t =$$

$$A + x_{k+1} \overline{u}_{k+1} \overline{u}_{k+1}^t + y_{k+1} \overline{v}_{k+1} \overline{v}_{k+1}^t$$

So applying **Lemma** 3.25[7] for A, x_{k+1}, y_{k+1} we obtain:

$$Z_{x_{k+1}} Z_{y_{k+1}} T_{k+1} \det \left(zI_m + \sum_{i \in S} \overline{u}_i \overline{u}_i^t + \sum_{i \in [k] \setminus S} \overline{v}_i \overline{v}_i^t + \sum_{i > k} x_i \overline{u}_i \overline{u}_i^t + y_i \overline{v}_i \overline{v}_i^t \right) =$$

[7]Note that we are not guaranteed that A here is invertible, but we do know that it can be singular for at most n values of z, but polynomials which are equal for all but n values are indeed equal, by noticing their difference has at most n roots, so the equality is true even for the values of z for which **Lemma** 3.25 does not apply.

$$Z_{x_{k+1}} Z_{y_{k+1}} T_{k+1} \det\left(A + x_{k+1}\overline{u}_{k+1}\overline{u}_{k+1}^t + y_{k+1}\overline{v}_{k+1}\overline{v}_{k+1}^t\right) =$$
$$p_{k+1}\det(A + \overline{u}_{k+1}\overline{u}_{k+1}) + (1 - p_{k+1})\det(A + \overline{v}_{k+1}\overline{v}_{k+1})$$

So taking into account $\prod_{i\in S} p_i \prod_{i\in[k]\setminus S}(1-p_i)$ the first term is the contribution of $S \cup \{k+1\}$ while the second of just S in the sum of $S \subseteq [k+1]$. Finally summing over S completes the step.

So taking the claim for $k = m$ we obtain (1). Each operator in right hand side of (1) preserves real stability, by **Proposition** 3.20 and **Corollary** 3.23 and Q is real stable so we conclude P is real-stable. But P is an univariate polynomial so **Proposition** 3.18 implies it is real rooted, as desired. □

We are now ready to prove **Theorem** 3.15.

Proof of Theorem 3.15. We want to prove that the polynomial

$$\sum_{s\in\{\pm1\}^m} \left(\prod_{i\,:s_i=1} p_i\right)\left(\prod_{i\,:s_i=-1} (1-p_i)\right)\det(xI - A_s)$$

is real rooted. This is equivalent to proving that the following polynomial is real rooted:

$$\sum_{s\in\{\pm1\}^m} \left(\prod_{i\,:s_i=1} p_i\right)\left(\prod_{i\,:s_i=-1} (1-p_i)\right)\det(xI + dI - A_s) \tag{2}$$

as their roots only differ by x, we remind the reader that d is the degree of each vertex as we are working with a d-regular graph.

For each edge (u, v) we define the rank 1-matrices

$$L_{u,v}^{1} = (\overline{e}_u - \overline{e}_v)(\overline{e}_u - \overline{e}_v)^t, \text{ and}$$
$$L_{u,v}^{-1} = (\overline{e}_u + \overline{e}_v)(\overline{e}_u + \overline{e}_v)^t.$$

Where e_v is the vector with all entries 0 except the one corresponding to v which is 1. For a signing s, let $s(u, v)$ denote the sign it assigns to the edge (u, v). We now have

$$dI - A_s = \sum_{u\sim v} L_{u,v}^{s(u,v)}.$$

If we now set $\overline{a}_{u,v} = \overline{e_u} - \overline{e}_v$ and $\overline{b}_{u,v} = \overline{e_u} + \overline{e}_v$ we can express the polynomial in (2) as

$$\sum_{s\in\{\pm1\}^m} \left(\prod_{i\,:s_i=1} p_i\right)\left(\prod_{i\,:s_i=-1} (1-p_i)\right)\det\left(xI + \sum_{s(u,v)=1} \overline{a}_{u,v}\overline{a}_{u,v}^t + \sum_{s(u,v)=-1} \overline{b}_{u,v}\overline{b}_{u,v}^t\right)$$

Which is of the form for which **Theorem** 3.26 applies showing this polynomial is real rooted as desired. □

3.8 Final Remarks

We add that Marcus, Spielman and Srivastava in [1] often show slightly stronger results than we do, as their goal is to present the method of using interlacing families. We on the other hand are interested in Ramanujan graphs themselves and wanted to present their proof in the shortest possible version while not losing on the substance of their ideas. We also include all but one proof of the intermediary results, which they do not do.

As an example, they extended the definition of Ramanujan graphs to arbitrary graphs and shown existence of infinite families of biregular Ramanujan graphs, where a graph is biregular if it is bipartite and vertices in each bipartition have equal degrees. The proof is essentially the same as the one we presented above.

They also draw an analogy between this proof technique and the probabilistic method. Probabilistic method, in its basic form, relies on the fact that for every random variable $X : \Sigma \to \mathbb{R}$, there is an $\omega \in \Sigma$ such that $X(\omega) \leq \mathbb{E}[X]$. While this method gives for certain special polynomial random variables $P : \Sigma \to \mathbb{R}[x]$ that there exists a $\omega \in \sigma$ such that $\lambda_{\max}(P(\omega)) \leq \lambda_{\max}(\mathbb{E}(P))$.

We note that in many applications the probabilistic method, by using a suitable concentration inequality, often gives very good random algorithms which simply repeatedly sample. Here though, because of the way we show existence of a good 2-lift, it is very hard to determine how many 2-lifts are actually good, it is very hard to determine how good or bad such an algorithm would be.

While this construction provides larger families of Ramanujan graph then the construction of Lubotzky, Phillips and Sarnak in [8] their Ramanujan graphs are explicit so are much less computationally expensive to construct for larger n. Their examples also have further, often useful, properties such as having large girth, which we can not determine from this construction, which controls only the eigenvalues.

Marcus, Spielman and Srivastava have recently [16] extended this method to show that there are infinite families of Ramanujan graphs of all sizes in addition to all degrees, we note that the above method only produces graphs of sizes $2^n N$ where N is the size of the initial graph. Our **Lemma** 2.7 allows us to construct Ramanujan graphs of more sizes than their original paper but it does not produce all sizes.

References

[1] A. Marcus, D. Spielman, and N. Srivastava. *Interlacing Families I: Bipartite Ramanujan Graphs of All Degrees.*, (2015), Annals of Mathematics, 182 (1), pp. 307-325.

[2] Prasolov, V. V., Ivanov, Simeon *Problems and theorems in linear algebra.*, (1994). American Mathematical Society, Providence, R.I

[3] F. R. K. Chung, *Diameters and Eigenvalues*, J. of the American Mathematical Society, Volume 2, Number 2 (1998),

[4] N Alon, V.D Milman, λ_1, *Isoperimetric inequalities for graphs, and superconcentrators*, Journal of Combinatorial Theory, Series B, Volume 38, Issue 1, 1985, Pages 73-88.

[5] R. M. Tanner, *Explicit Concentrators from Generalized N-Gons*, SIAM. J. on Algebraic and Discrete Methods, 5(3), 287293.

[6] F. Bien, *Constuction of Telephone Networks by Group Representations*, Notices of the Amer. Math. Society, 36 (1989), No. 1, 522.

[7] A. Nilli, *On the second eigenvalue of a graph*, Discrete Math, (1991), 207-210

[8] A. Lubotzky, R. Phillips and P. Sarnak. *Ramanujan graphs.* Combinatorica, 8(3):261277, 1988.

[9] M. Morgenstern. *Existance and explicit constructions of* $q+1$ *regular Ramanujan graphs for every prime power* q. Journal of Combinatorial Theory, Series B, 62:4462, 1994.

[10] G. A. Margulis. *Explicit group theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators.* Problems of Information Transmission, 24(1):3946, July 1988

[11] A. K. Pizer. *Ramanujan graphs and Hecke operators.* Bulletin of the AMS, 23(1), 1990.

[12] B. W. Jordan and R. Livné. *Ramanujan local systems on graphs.* Topology, 36(5): 10071024, 1997.

[13] P. Chiu. *Cubic Ramanujan graphs.* Combinatorica, 12(3):275285, 1992.

[14] Y. Bilu and N. Linial. *Lifts, discrepancy and nearly optimal spectral gap.* Combinatorica, 26(5):495519, 2006.

[15] A. Marcus, D. A. Spielman, and N. Srivastava. *Interlacing families II: Mixed characteristic polynomials and the Kadison-Singer problem.* arXiv preprint, arXiv:1306.3969, 2013.

[16] A. Marcus, D. A. Spielman, and N. Srivastava. *Interlacing Families IV: Bipartite Ramanujan Graphs of All Sizes* arXiv preprint, arXiv:1505.08010 2015

[17] J. Friedman. *Relative expanders or weakly relatively Ramanujan graphs.* Duke Mathematical Journal, 118(1):1935, 2003

[18] O. J. Heilmann and E. H. Lieb. *Theory of monomer-dimer systems.* Communications in Mathematical Physics, 25(3):190232, 1972.

[19] C. D. Godsil and I. Gutman. *On the matching polynomial of a graph.* In L. Lovász and V. T. Sós, editors, Algebraic Methods in graph theory, volume I of Colloquia Mathematica Societatis János Bolyai, 25, pages 241249. János Bolyai Mathematical Society, 1981.

[20] J. P. Dedieu. *Obreschkoffs theorem revisited: what convex sets are contained in the set of hyperbolic polynomials?* Journal of Pure and Applied Algebra, 81(3):269278, 1992.

[21] H. J. Fell. *Zeros of convex combinations of polynomials.* Pacific J. Math., 89(1):4350, 1980.

[22] M. Chudnovsky and P. Seymour. *The roots of the independence polynomial of a clawfree graph.* Journal of Combinatorial Theory, Series B, 97(3):350357, 2007.

[23] J. Borcea and P. Brändén. *Applications of stable polynomials to mixed determinants: Johnsons conjectures, unimodality, and symmetrized Fischer products.* Duke Mathematical Journal, 143(2):205223, 2008.

[24] P. Brändén and J. Borcea. *Multivariate Polya-Schur classification problems in the Weyl algebra.* Proceedings of the London Mathematical Society, (3) 101(1):73104, 2010.

[25] D. G. Wagner. *Multivariate stable polynomials: theory and applications. Bulletin of the* American Mathematical Society, 48(1):5384, 2011.

[26] Q.I. Rahman and G. Schmeisser. *Analytic Theory of Polynomials.* London Math. Soc. Monographs (N.S.) 26, Oxford U.P., New York NY, 2002.