CHENNAI MATHEMATICAL INSTITUTE
Department of Computer Science

MASTER OF COMPUTER SCIENCE

Geometry of Polynomials

Dissertation of:
Bhargav C S

Advisor:
Prof. Partha Mukhopadhyay

May 2019

Acknowledgements

I am immensely thankful to my supervisor Prof. Partha Mukhopadhyay for his guidance and support. He encouraged me to pursue my interests and helped me make good academic decisions whenever I felt lost.

I am also very thankful to Prof. Prajakta Nimbhorkar for an inspiring course on Complexity Theory that made me want to pursue it further.

I am grateful to CMI for providing a great environment to learn and grow. I learnt a lot, made great friends and enjoyed being part of the CMI culture. The faculty never treated us like students, but like future colleagues listening patiently to all our stupid ideas. Coming to CMI was one of the best decisions I ever made.

Finally, I am thankful to my parents for their constant love and support. Thanks for encouraging me to always do what I loved, no matter the local repercussions.

Abstract

What is the role of real rooted polynomials and their multivariate generalizations as related to problems of algorithms, combinatorics and probability? In this survey we aim to give an overview of the techniques involved and some recent results using these techniques.

In particular, we discuss the theory of interlacing families of polynomials and give their application in proving the existence of bipartite ramanujan graphs of all degrees.

We also discuss the concept of real stability and some properties and techniques related to that.

Contents

List of Figures **vi**

1 Introduction **1**
- 1.1 Real Rooted polynomials 1
  - 1.1.1 Newton's inequalities 3
  - 1.1.2 Connection with Independent Bernoulli variables 4
  - 1.1.3 Counting Matchings 4
  - 1.1.4 Spanning Tree Polynomial 7

2 Real Stable Polynomials **9**
- 2.1 Stable Polynomials 9
  - 2.1.1 Closure properties of stable polynomials . . 11

3 Bipartite Ramanujan Graphs of all degrees **13**
- 3.1 Expanders 13
- 3.2 Ramanujan Graphs 14
- 3.3 Bilu-Linial's approach 15
- 3.4 Proof Idea 16
  - 3.4.1 Idea 1: Random signings 16
  - 3.4.2 Idea 2: Look at the (expected) characteristic polynomial 16
- 3.5 The Proof 17
  - 3.5.1 The expected characteristic polynomial . . 17
  - 3.5.2 Largest root of the expected polynomial . . 18
  - 3.5.3 Interlacing families and Real Stability . . . 19
- 3.6 Some observations 26

Bibliography **27**

List of Figures

3.1 Expanders 13
3.2 Non expanders 13
3.3 Polynomials with common interlacing 20

Introduction 1

The geometry of polynomials in general studies the zero loci of polynomials with complex coefficients using complex and real analysis. We focus on polynomials with real roots and their multivariate generalization. We are interested in questions like : How are the properties of a graph or probability distribution or a matrix related to the zeros of its generating polynomials?

1.1 Real Rooted polynomials

Definition 1 *A polynomial* $p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots a_1 x + a_0$, $a_n \neq 0$ *is real rooted if* $p(x) = a_n \prod_{j=1}^{n} (x - \lambda_j)$, $\lambda_j \in \mathbb{R}$.

Some facts about polynomials:

- Roots cannot be expressed in closed form using radicals for $n \geq 5$.
- However roots can be computed very well using numerical methods and techniques.

- Useful relations between roots and coefficients exist. Suppose $p(x) = \prod_{j=1}^{n}(x - \lambda_j)$, then

$$a_{n-1} = -\sum_j \lambda_j$$
$$a_{n-2} = \sum_{j_1<j_2} \lambda_{j_1}\lambda_{j_2}$$
$$\dots$$
$$a_0 = (-1)^n \prod_{j=1}^{n} \lambda_j$$

There are also some more identities by Newton that are useful: Suppose $p(x) = \prod_{j=1}^{n}(x - \lambda_j) = x^n + a_{n-1}x^{n-1} + \dots a_1x + a_0$. Let $m_k = \sum_{j=1}^{n} \lambda_j^k$.

Clearly, $m_1 = \sum \lambda_j = -a_{n-1}$.

Then, $m_2 = \sum \lambda_j^2 = (\sum \lambda_j)^2 - 2\sum_{j_1<j_2} \lambda_{j_1}\lambda_{j_2} = a_{n-1}^2 - 2a_{n-2}$.

In general, $m_k = -ka_{n-k} - \sum_{i=1}^{k-1} a_{n-k-i}m_i$.

Checking Real rootedness

There is a nice criterion to check real rootedness of a polynomial. This is a theorem by Hermite and Sylvester.

Theorem 2 $p(x) = \prod_{j=1}^{n}(x - \lambda_j)$, $\lambda_j \in \mathbb{C}$ *is real rooted iff the following matrix is positive semi-definite.*

$$H = (m_{i+j-2})_{i,j=1}^{n}, \text{ where } m_k = \sum_{j=1}^{n} \lambda_j^k$$

Hence, this is a computationally tractable problem. We can check real rootedness by computing m_k using Newton's identities and then checking whether H is PSD.

1.1.1 Newton's inequalities

Newton noticed a few more interesting properties of real rooted polynomials:

If the coefficients of a real rooted polynomial are positive, then they are always unimodal, i.e.,

$$a_n \leq a_{n-1} \leq \ldots \leq a_{k-1} \leq a_k \geq a_{k+1} \geq \ldots \geq a_1 \geq a_0$$

Also, that they are log concave:

$$\frac{a_n}{a_{n-1}} \leq \frac{a_{n-1}}{a_{n-2}} \leq \ldots \leq \frac{a_1}{a_0}$$

Infact, something stronger called ultra log concavity (ULC) holds:

Theorem 3 *If* $p(x) = \sum_{j=0}^{n} a_j x_j$ *is real rooted,* $a_j \geq 0$ *then*

$$\left(\frac{a_k}{\binom{n}{k}}\right)^2 \geq \frac{a_{k-1}}{\binom{n}{k-1}} \frac{a_{k+1}}{\binom{n}{k+1}}$$

Proof: The inequalities are infact consequences of two simple closure properties of real rooted polynomials.

- **Differentiation**: If $p(x)$ is real rooted, so is $p'(x)$. The proof uses Rolle's theorem.
- **Inversion**: If $p(x)$ has degree n and is real rooted, then so is $q(x) = x^n p(1/x)$, which has the same coefficents, but in reverse order. This is because the roots of q are the reciprocals of the (nonzero) roots of p.

Now the inequality is got by differentiating the polynomial p $k-1$ times, reversing the coefficients, and differentiating another $n-k-1$ to get a quadratic polynomial with coefficients equal to a_{k-1}, a_k, a_{k+1} times some binomial coefficients. From above closure properties, this polynomial must be real rooted which means it's discriminant must be nonnegative which gives us the inequalities. ∎

1.1.2 Connection with Independent Bernoulli variables

There is a nice conection between real rooted polynomials and independent bernoulli random variables:

Lemma 4 *A polynomial $p(x) = \sum a_j x^j$ with $a_j \geq 0$ and $\sum a_j = 1$ is real rooted iff there are independent Bernoulli random variables $Y_1, \ldots, Y_k$ such that $a_k = Pr[\sum Y_j = k]$ $\forall k$.*

Proof: If $a_k = Pr[\sum Y_j = k]$, Y_j's independent, then $p(x) = \sum a_j x^j = \mathbb{E}[x^{\sum_j Y_j}] = \prod_j \mathbb{E}[x^{Y_j}] = \prod_j (1 - p_j + p_j x)$ where p_j is the probability that $Y_j = 1$. This is a real rooted polynomial.
On the other hand, if $p(x) = \sum a_j x^j$ is real rooted, $a_j \geq 0$ then we can write it as $p(x) = C \prod_j (x + b_j)$, $b_j \geq 0$ and $C = 1/\prod_j (1 + b_j)$ since $p(1) = 1$. So, $p(x) = \prod_j \frac{x+b_j}{1+b_j}$ which is nothing but the generating function of of Bernoulli variables with parameters $p_j = 1/(1 + b_j)$. ■

One consequence of this is that the sequence of sums of bernoulli variables forms a log concave sequence.

Corollary 5 *$a_k = Pr[\sum Y_j = k]$ form a log-concave sequence.*

1.1.3 Counting Matchings

We consider a more combinatorial situation. A matching is a subset of edges such that no two edges share a common vertex. Given a graph G on n vertices, let m_k be the number of matchings with k edges in G. For example, $m_{n/2}$ is the number of perfect matchings. Now, we could ask if the m_k's are unimodal. But, since edges in a random matching are not independently distributed, we can't use the results from above.

Curiously, this was actually studied in the context of statistical mechanics. The motivation comes from studying phase transitions. In the context of general graphs with nonnegative weights, we define

$$m_k := \sum_{\text{matching } M, |M|=k} \prod_{e \in M} w_e$$

and then the generating function is the **matching polynomial**:

$$\mu_G(x) := \sum_{k=0}^{n/2} x^{n-2k}(-1)^k m_k$$

Heilmann and Lieb (1972) showed that the matching polynomial is real rooted.

Theorem 6 *For every weighted graph G with nonnegative edge weights, $\mu_G(x)$ is real rooted.*

Proof: The proof is based on a recurrence satisfied by the matching polynomial (can also be thought of as an alternate definition of it). For any vertex $v \in G$,

$$\mu_G(x) = x\mu_{G\backslash v}(x) - \sum_{u\sim v} w_{uv}\mu_{G\backslash uv}(x)$$

where $G\backslash v$ and $G\backslash uv$ are vertex deleted graphs and $u \sim v$ denotes the vertices adjacent to v.

We need the concept of **interlacing polynomials** to proceed.

Definition 7 *Let $p(x) = \prod_{i=1}^{n}(x - \alpha_i)$ and $q(x) = \prod_{i=1}^{m}(x - \beta_i)$ be real rooted polynomials of degrees differing by atmost 1 (say $n \geq m$). We say q interlaces p if*

$$\alpha_1 \geq \beta_1 \geq \ldots \beta_m \geq \alpha_n, \text{ when } m = n - 1$$

and

$$\alpha_1 \geq \beta_1 \geq \ldots \alpha_n \geq \beta_m, \text{ when } m = n$$

Now, assume that G is a complete graph on n vertices with $w_{uv} > 0$ for all pairs $u, v \in V$ (we will remove this assumption later). Assume inductively that for every such graph H with atmost $n - 1$ vertices, the following holds:

1. $\mu_H(x)$ is real rooted with all roots distinct.

2. For every $w \in H$, $\mu_{H\backslash w}(x)$ strictly interlaces $\mu_H(x)$.

For the base case, there is a single edge uv, for which $\mu_G(x) = x^2 - w_{uv}$ and $\mu_{G\setminus v}(x) = x$. So, the claim is true since w_{uv} is positive.

Now, fix a vertex $v \in G$ and let $\lambda_{n-1} < \ldots < \lambda_1$ be the roots of $\mu_{G\setminus v}$. By induction, we know that each $\mu_{G\setminus uv}$ strictly interlaces $\mu_{G\setminus v}$. This implies that $\mu_{G\setminus uv}(\lambda_1) > 0$ (since each of these polynomials is monic) and since each interval $(\lambda_i, \lambda_{i+1})$ contains exactly one root of each $\mu_{G\setminus uv}$, we get that

$$sign(\mu_{G\setminus uv}(\lambda_i)) = (-1)^{i+1}, i = 1, \ldots, n-1$$

for all $u \sim v$. The weights w_{uv} are positive. So, the sum

$$r(x) = \sum_{u \sim v} w_{uv}\mu_{G\setminus uv}(x)$$

must also alternate sign at λ_i's. And, using our recurrence from before, we get

$$\mu_G(\lambda_i) = \lambda_i \mu_{G\setminus v}(\lambda_i) - r(\lambda_i) = -r(\lambda_i)$$

and hence we get $sign(\mu_G(\lambda_i)) = (-1)^i$. By the intermediate value theorem, we see that μ_G has atleast one root in each interval $(\lambda_i, \lambda_{i+1})$, which yields $n-2$ roots. Since $\mu_G(\lambda_1) < 0$ and $\mu_G(x)$ goes to infinity as x goes to infinity, it must also be that $\mu_G(\lambda_0) = 0$ for some $\lambda_0 > \lambda_1$. A similar argument yields another root $\lambda_n < \lambda_{n-1}$ giving us a total of n real roots that are distinct and are interlaced by the roots of $\mu_{G\setminus v}$.

Now, we need to remove the assumption of positive weights. We consider a sequence of graphs $G^{(k)}$ with weights $w_{uv}^{(k)}$ equal to w_{uv} if w_{uv} is positive and equal to $1/k$ if $w_{uv} = 0$. These weights converge to the weights of G. Now, the polynomials $\mu_{G^{(k)}}(x)$ converge to $\mu_G(x)$ coefficient-wise. Since a limit of real rooted polynomials is either zero or real rooted, we also get that $\mu_G(x)$ is real rooted. ■

We mention another combinatorial polynomial that has similar properties:

1.1.4 Spanning Tree Polynomial

For a graph G and any fixed set of edges F (think of a *cut*), we define the spanning tree polynomial as

$$\mathcal{T}_F(x) = \sum_{\text{spanning tree } T \text{ of } G} x^{F \cap T}.$$

We define it this way since without F, the polynomial is uninteresting.

It turns out that this polynomial is also real rooted.

Real Stable Polynomials 2

2.1 Stable Polynomials

We denote the upper half plane by $\mathcal{H} = \{z : Im(z) > 0\}$. Also, let $\mathbf{z}$ denote the vector $(z_1, \ldots, z_n)$.

Definition 8 *A non zero polynomial $f \in \mathbb{C}[z_1, \ldots, z_n]$ is called stable if it has no zeros in $\mathcal{H}^n$, i.e.,*

$$Im(z_i) > 0 \quad \forall i \Rightarrow f(\mathbf{z}) \neq 0$$

A stable polynomial with real coefficients is called *Real Stable.*

We state here, without proof, an equivalent definition of real stable polynomials.

Definition 9 *A polynomial is real stable (resp. stable) iff for every $\mathbf{e} \in \mathbb{R}^n_{>0}$ and $\mathbf{x} \in \mathbb{R}^n$, the univariate restriction*

$$t \mapsto f(te + \mathbf{x})$$

is real rooted (resp. stable).

Note that a univariate real stable polynomial must be real rooted, since the roots of real polynomials occur in conjugate pairs. And since there is no root in the upper half plane, there is no root in the lower half plane too.

A few examples of stable polynomials:

- The monomials $\prod_{i=1}^{n} z_i$.
- Linear polynomials $\sum_i a_i z_i$ with $a_i > 0$.
- The polynomial $1 - z_1 z_2$ (since a product of two numbers in $\mathcal{H}$ can't be positive).

There is a very important example that is useful in many appplications:

Theorem 10 *For positive semidefinite matrices $A_1, \ldots, A_n \succeq 0$, and Hermitian B, the polynomial*

$$\det\left(\sum_{i=1}^{n} z_i A_i + B\right)$$

is real stable.

Proof: Assume that the A_i are positive definite (we'll correct this assumption later). Now, consider a univariate restriction

$$t \mapsto \det\left(t \sum_{i=1}^{n} e_i A_i + \left(\sum_{i=1}^{n} x_i A_i + B\right)\right).$$

Since the e_i are positive, $M := \sum_{i=1}^{n} e_i A_i > 0$ has a negative square root $M^{-1/2}$ and so, we can write the above as

$$t \mapsto \det\left(M^{-1/2}\right) \det\left(tI + M^{1/2}\left(\sum_{i=1}^{n} x_i A_i + B\right) M^{1/2}\right) \det\left(M^{-1/2}\right).$$

And since this is a multiple of a characteristic polynomial of a Hermitian matrix, it must be real rooted.

The general case of positive semidefinite matrices is handled by taking limit of positive definite matrices and observing that the limit along each univariate restriction must be real rooted or zero. ■

2.1.1 Closure properties of stable polynomials

There are many operations under which stable polynomials are closed and this makes the set of transformations that can be performed on them very rich. This will be useful to us in the future.

Theorem 11 *The following linear transformations on $\mathbb{C}[\mathbf{z}]$ map every stable polynomial to another stable polynomial or zero.*

1. ***Permutation:*** $f(z_1, z_2, \ldots, z_n) \mapsto f\left(z_{\sigma(1)}, \ldots, z_{\sigma(n)}\right)$ *for some permutation* $\sigma : [n] \mapsto [n]$.
2. ***Scaling:*** $f(z_1, \ldots, z_n) \mapsto f(az_1, \ldots, z_n)$ *where* $a > 0$.
3. ***Inversion:*** $f(z_1, \ldots, z_n) \mapsto z_1^d f(-1/z_1, \ldots, z_n)$ *where* d *is the degree of* z_1 *in* f.
4. ***Instantiation:*** $f \mapsto f(a, z_2, \ldots, z_n) \in \mathbb{C}[z_2, \ldots, z_n]$ *where* $a \in \mathcal{H} \cup \mathbb{R}$.
5. ***Differentiation:*** $f \mapsto \frac{\partial}{\partial z_1} f$.

Proof: (1) and (2) hold directly from the definition itself. (4) holds since the map of z to $-1/z$ preserves the upper half plane. For (5), if $a \in \mathcal{H}$, then the result is definitely true. In the case when $a \in \mathbb{R}$, we take the limit of the polynomials $f(a + 1/k, z_2, \ldots, z_n)$ and then use the characterization of real stability from above. (6) is true because of a very useful result called George-Lucas theorem which we prove next. ■

Theorem 12 *If* $f \in \mathbf{C}[z]$, *then the roots of* $f'(z)$ *lie in the convex hull of the roots of* $f(z)$.

Proof: Let $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ be the roots of f. We can assume without loss of generality that f and f' have no common roots. Then, if $f'(z) = 0$, we get that

$$0 = \frac{f'(z)}{f(z)} = \sum_{i=1}^{n} \frac{1}{z - \lambda_i} = \sum_{i=1}^{n} \frac{\overline{z - \lambda_i}}{|z - \lambda_i|^2}$$

If we rearrange and simplify, we get

$$z = \sum_{i=1}^{n} \frac{|z - \lambda_i|^{-2}}{\sum_{i=1}^{n} |z - \lambda_i|^{-2}} \lambda_i$$

which is a convex combination. ■

We also state here, without proof, a very curious fact about random spanning tree distributions. The result uses the closure properties we described above.

Suppose $G = (V, E)$ is a graph, $F \subset E$ is a subset of its edges, and say the tree T is uniformly random and spanning. Then, what can we say about the distribution of the random variable $F \cap T$? Note that the edges that appear in a random spanning tree are not dependent.

It turns out that the distribution is a Poisson Binomial Distribution! That is, it is a summation of independent Bernoulli variables. This can also be seen as a corollary to the fact we saw in the last chapter about the spanning tree polynomial being real rooted.

3.1 Expanders

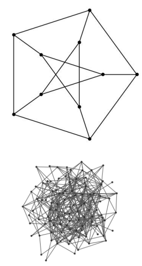

FIGURE 3.1: Expanders

Expanders are sparse regular graphs that are well connected. An expander sort of behaves like a random graph. There are many ways to define the notion of well connectedness. A combinatorial way to define it is to say that every set of vertices has many outgoing edges. A probabilistic way is to say that a graph is well connected if a random walk on a graph mixes quickly. And there are also algebraic ways of defining it, which we'll see shortly. Expanders are very useful objects in complexity theory. They are useful in creating pseudorandom generators, error correcting codes and in general in a lot of places in mathematics.

What we will be interested in are Spectral Expanders. To define this notion, we look at the adjacency matrix of G, a d-regular graph on n verices. The adjacency matrix is a symmetric $n \times n$ matrix and hence has n *real* eigenvalues. Let's call them $\lambda_1, \geq \lambda_2 \geq \ldots \geq \lambda_n$. There are some general facts about them. It's easy to check that for any d-regular graph, d is an eigenvalue and it's the largest. Also, if the graph is *bipartite*, then the eigenvalues are symmetric about zero. We call the eigenvalue d (also $-d$ if G is bipartite) "trivial". Now, a good expander has small non-trivial eigenvalues. More formally,

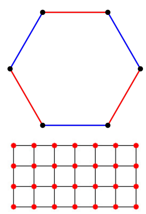

FIGURE 3.2: Non expanders

Definition 13 *A graph G is a* (n, d, λ) *expander if G is a d-regular graph on n vertices and has all it's non-trivial eigenvalues in the range* $[-\lambda, \lambda]$.

What are some good expanders we know of? The complete graphs K_d and $K_{d,d}$ have all non-trivial eigenvalues equal to zero and are hence very good expanders. But they're not sparse (their degree is a function of the number of vertices). And it is known that we cannot have an infinite sequence of d-regular graphs all of which have non-trivial eigenvalues equal to zero. Also, from the Alon-Bopanna bound [Nilli (1991)], we know that for every $\epsilon > 0$, every sufficiently large d-regular graph has a non-trivial eigenvalue of absolute value atleast $2\sqrt{d-1} - \epsilon$. So, absolute value atmost $2\sqrt{d-1}$ is the best we can hope for. We call such graphs *Ramanujan*.

3.2 Ramanujan Graphs

Definition 14 *A graph G is Ramanujan if all the non-trivial eigenvalues of G have absolute value atmost* $2\sqrt{d-1}$.

So, Ramanujan graphs are the best expanders and we would like to construct infinite families of them with constant degree. Lubotsky-Philips-Sarnak[Lubotzky et al. (1988)], Margulis[Margulis (1988)] were the first to give such a construction. They built these from Cayley graphs and the graphs were regular and had degree $p + 1$, where p is a prime. The analysis was based on number theory and was pretty sophisticated. This was a breakthrough result, but the question was, what about other degrees? Is this a strictly number theoretic phenomenon? And is there a simpler analysis? This result was extended to graphs of degree $q+1$ where q is a prime power, but there were very few other constructions. Friedman[Friedman (2004)] showed that a random d-regular graph is almost Ramanujan, i.e., for every $\epsilon > 0$ the absolute value of all the non-trivial eigenvalues of almost every sufficiently large d-regular graph is at most $2\sqrt{d-1} + \epsilon$. But nothing was known about existence of infinite families of Ramanujan Graphs of every degree greater than 2, until Marcus, Spielman and Shrivastava proved [Marcus et al. (2015)] that infinte families of *bipartite* Ramanujan graphs exist of every degree.

3.3 Bilu-Linial's approach

The approach Marcus, et.al take is by proving a variant of a conjecture by Bilu and Linial (2006). This consists of finding an operation which doubles the size of a graph without blowing up its eigenvalues. Bilu and Linial suggested that this can be done by using a sequence of 2-lifts of a base graph.

Given a graph $G = (V, E)$, a 2-lift of G is a graph that has two vertices for each vertex in V. The pair of vertices is called *fibre* of the original vertex. Every edge in E corresponds to two edges in the 2-lift. If (u, v) is an edge in E, $\{u_0, u_1\}$ is the fibre of u, and $\{v_0, v_1\}$ is the fibre of v, then the 2-lift can either contain the pair of edges

$$\{(u_0, v_0), (u_1, v_1)\} \tag{3.1}$$

or

$$\{(u_0, v_1), (u_1, v_0)\} \tag{3.2}$$

If only edges of the first type appear, then we get just two disjoint copies of the original graph and if only edges of the second type appear, then we obtain the double cover of G. If there are m edges in the graph, then we get 2^m 2-lifts.

To analyse the eigenvalues of a 2-lift, we study the signed adjacency matrix of G. We define *signings* $s : E \Rightarrow \{\pm 1\}$, functions on the edges of G. $s(u, v) = 1$ if edges of type 3.1 appear and $s(u, v) = -1$ if edges of type 3.2 appear. The signed adjacency matrix A_s is same as the adjacency matrix of A except the entries corresponding to edge (u, v) are $s(u, v)$. Now, we can show that the eigenvalues of the signed adjacency matrix are exactly the new eigenvalues when you take a 2-lift.

Lemma 15 *Let A be the adjacency matrix of graph G and let A_s be the signed adjacency matrix of the associated 2-lift $\hat{G}$. Then, every eigenvalue of A and every eigenvalue of A_s are eigenvlaues of $\hat{G}$ (with multiplicity).*

Proof: We can see that the adjacency matrix of $\hat{G}$ is:

$$\hat{A} = \begin{bmatrix} A_1 & A_2 \\ A_2 & A_1 \end{bmatrix}$$

where A_1 is adjacency matrix of $(V, s^{-1}(1))$ and similarly A_2 is adjacency matrix of $(V, s^{-1}(-1))$. So, $A = A_1 + A_2$ and $A_s = A_1 - A_2$. Now, if u is an eigenvector of A, then $\hat{u} = (u, u)$ is an eigenvector of $\hat{A}$, and similarly, if w is an eigenvector of A_s then $\hat{w} = (w, -w)$ is an eigenvector of $\hat{A}$ (all with same eigenvlaues). And since all these $\hat{u}$ and $\hat{w}$'s are perpendicular and $2n$ in number, they span all the eigenvectors of $\hat{A}$. ■

So now, the problem of finding a good lift is the same as the problem of finding a good signing. Bilu and Linial conjectured that every d-regular graph has a signing in which all the new eigenvalues have absolute value atmost $2\sqrt{d-1}$. Equivalently, there is a signing s such that the norm of the matrix A_s (which is the absolute value of the max eigenvalue) is bounded by $2\sqrt{d-1}$. Marcus et.al give this bound only for the positive eigenvalue. Hence the result is only for bipartite graphs (since the eigenvalues are symmetric about zero in this case). If we repeatedly keep taking two lifts of a base d-regular bipartite Ramanujan graph, then we get an infinite family of d-regular bipartite Ramanujan graphs.

3.4 Proof Idea

Let's now see how we could prove such a statement.

3.4.1 Idea 1: Random signings

One reasonable idea might be to try to use the probabilistic method. Choose a random signing $s \in \{\pm 1\}^m$ and then argue that the expected value of the norm of A_s is bounded by $2\sqrt{d-1}$. This then gives that there is some s for which the norm is bounded. But unfortunately, this does not work; Bilu and Linial (2006) showed that the expected value could be much larger: $\mathbb{E}[\|A_s\|] \gg 2\sqrt{d-1}$. So we need some other idea.

3.4.2 Idea 2: Look at the (expected) characteristic polynomial

We observe that

$$\lambda_1(A_s) = \lambda_{max}(\chi A_s), \text{ where } \chi A_s = det(xI - A_s)$$

Now, this is not a novel observation at all. Before the work of Marcus, Spielman and Srivastava, it was considered in Random Matrix theory that this representation of the norm is not very useful due to the highly non linear nature of the characteristic polynomial. As an example, a basic fact like triangle inequality is very hard to establish using the characteristic polynomial definition above. But it turns out that for certain special type of random matrices, there is a rich structure which is extremely useful and this is what Marcus et.al. exploited.

Consider the expected characteristic polynomial: $\mathbb{E}_{s\in\{\pm1\}^m}\chi A_s(x)$. Usually this is useless, but in this case, $\{\chi A_s\}_{s\in\{\pm1\}^m}$ is what they called an *intelacing family.* A consequence of this is that there exists an s such that $\lambda_{max}(\chi A_S) \leq \lambda_{max}(\mathbb{E}\chi A_S)$. So now, this gives a proof strategy to follow:

1. Show that $\{\chi A_s\}_{s\in\{\pm1\}^m}$ is an interlacing family.

2. Calculate the expected characteristic polynomial.

3. Bound the largest root of the expected polynomial.

If we can do this and get the required bounds, we'd be done.

3.5 The Proof

We will prove parts (2) and (3) first. They are easier and also give motivation for part (1).

3.5.1 The expected characteristic polynomial

It turns out that the expected characteristic polynomial is actually the matching polynomial that we saw before! This was proved long ago by Godsil and Gutman (1978) and holds for all graphs.

Theorem 16

$$\mathbb{E}\chi A_s(x) = \mu_G(x)$$

Proof: The proof is not very hard. We expand $[det(xI - A_s)]$ using permutations.

$$\begin{aligned}
&\mathbb{E}[det(xI - A_s)] \\
&= \mathbb{E}_s\left[\sum_{\sigma \in \mathrm{sym}([n])} (-1)^{|\sigma|} \prod_{i=1}^{n} (xI - A_s)_{i,\sigma(i)}\right] \\
&= \sum_{k=0}^{n} x^{n-k} \sum_{S \subset [n], |S|=k} \sum_{\pi \in \mathrm{sym}(S)} \mathbb{E}_s\left[(-1)^{|\pi|} \prod_{i \in S} (A_s)_{i,\pi(i)}\right] \\
&\text{where } \pi \text{ denotes the part of } \sigma \text{ with } \sigma(i) \neq i \\
&= \sum_{k=0}^{n} x^{n-k} \sum_{S \subset [n], |S|=k \pi \in \mathrm{sym}(S)} \mathbb{E}_s\left[(-1)^{|\pi|} \prod_{i \in S} s_{i,\pi(i)}\right]
\end{aligned}$$

Since the s_{ij} are independent with $\mathbb{E}[s_{ij}] = 0$, the only products which survive are those which contain an even power of s_{ij}. So, we only need to look at permutations with orbit size 2, which are just perfect matchings on S. There are no perfect matchings when $|S|$ is odd. And since expectation of s_{ij}^2 is 1, we get

$$\begin{aligned}
&\mathbb{E}[det(xI - A_s)] \\
&= \sum_{k=0}^{n} x^{n-k} \sum_{|S|=k \text{ matching } \pi \text{ on } S} (-1)^{|S|/2} \cdot 1 \\
&= \mu_G(x)
\end{aligned}$$

■

3.5.2 Largest root of the expected polynomial

This was also proved long ago by Heilmann and Lieb (1972) and it turns out that the bound is exactly what we need!

Theorem 17 $\lambda_{max}(\mu_G(x)) \leq 2\sqrt{d-1}$

The proof is based on similar recurrences we used in the introductory chapter to show that the matching polynomial is real rooted. We omit the proof here. But we note that the occurence of this number $2\sqrt{d-1}$ is not a coincidence. It is actually the spectral radius of an infinite d-ary tree and this is the same object that is used to prove the Alon-Bopanna bound.

Now, let's see how to prove the first part.

3.5.3 Interlacing families and Real Stability

We can ask a general question: given real rooted polynomials p_0 and p_1, when are the roots of $p_i(x)$ related to the roots of $\mathbb{E}_i p_i(x)$? Definitely not always. For example, the average of $(x-1)^2$ and $(x+1)^2$ is x^2+1, which doesn't have any real roots. But, sometimes it works. If the roots alternate, then the average polynomial will have real roots and the roots will lie in the convex hull of the pairs of alternating roots. We call this phenomenon "common interlacing". Recall the definition of interlacing polynomials.

The polynomial

$$q = \prod_{i=1}^{n-1} (x - \alpha_i)$$

interlaces

$$p = \prod_{i=1}^{n} (x - \beta_i)$$

if

$$\beta_n \leq \alpha_{n-1} \leq \beta_{n-1} \ldots \leq \alpha_1 \leq \beta_1$$

Now, we say that two polynomials have a *common interlacing* if there is a third polynomial that interlaces both the polynomials.

And it turns out this is the criterion we need for reasoning about the roots of the expected polynomial.

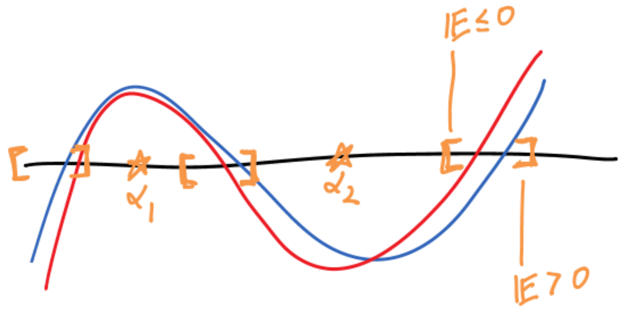

Figure 3.3: Polynomials with common interlacing

Theorem 18 *If p_0, p_1 are monic polynomials that have a common intelacing, then there exists an i such that $\lambda_{max}(p_i) \leq \lambda_{max}(\mathbb{E}_i p_i)$*

Proof:

The proof is basically an intermediate value theorem. Refer to the figure 3.3. The two polynomials are interlaced by the common polynomial whose roots are α_1 and α_2. In this case, the roots divide the line into three intervals each of which has exactly one root of each of the polynomial. Consider the right most interval. Since the polynomials are monic, at infinity, they are both positive. So, to the right of the right most interval, the average (expectation) is also positive. But in the interval, each polynomial changes sign exactly once. So, at the beginning, each polynomial is nonpositive. So the average (expectation) is also nonpositive. Since the average went from being nonpositive to positive, it has to be zero somewhere in between. Now, we take the interval right up to the roots of the polynomials, we see that the average has a zero in between the two roots! So, the maximum root of one of them is atmost the maximum root of the average, which is what we need. ∎

Now we define what it means for a family of polynomials to be interlacing.

Definition 19 $\{p_s\}_{s\in\{\pm1\}^m}$ *is an interlacing family if it can be* *placed on the leaves of a tree such that every node is the sum of the children below, and sets of siblings have common interlacings.*

The main property we use of such families is the following theorem:

Theorem 20 *If* $\{p_s\}_{s\in\{\pm1\}^m}$ *is an interlacing family, there is an* s *so that* $\lambda_{max}(p_s) \leq \lambda_{max}(p_\phi)$. *Here,* p_ϕ *is just the expectation of all* p_s*'s.*

Proof: The proof is by induction. Let's start at the root of the tree. The node is p_ϕ and it is the sum of p_0 and p_1. Since p_0 and p_1 have a common intelacing, we know that the maximum root of one of them (say p_0) is atmost the max root of p_ϕ. Now, we consider p_0 and reason similarly about its children. We will get that the max root of one of the children p_{00} or p_{01} is atmost the max root of p_0 and hence also of p_ϕ. And we continue walking down the tree till we reach the leaf which then gives us our required polynomial. ■

So now, if we can prove that our family $\{\chi A_s\}_{s\in\{\pm1\}^m}$ is an interlacing family, we will be done. Turns out, there is a very nice characterization of interlacing families. This is a folklore lemma which has been around quite some time though it appears in the book of (Fisk'08). It says that two polynomials have a common interlacing iff all their convex combinations are real rooted.

Lemma 21 *Two polynnomials $p_0(x)$ and $p_1(x)$ with roots all real have a common interlacing iff $\lambda p_0(x) + (1-\lambda)p_1(x)$ is real rooted for all $0 \leq \lambda \leq 1$.*

Proof: This proof needs a fact which we state without proof.

Lemma 22 *Let $\epsilon > 0$ and f a real rooted polynomial of degree n. Then the polynomial $f_\epsilon = (1-\epsilon\partial)^n f$ is real rooted and has simple roots (multiplicity 1).*

($\Rightarrow$): This is straightforward and the proof is the same as the proof of Theorem 18. The polynomials have a common interlacing which divides the real line into intervals each of which contains exactly one root of the polynnomials. And like before, the average polynomial $(\lambda p_0(x) + (1-\lambda)p_1(x))$ changes sign once in every interval. Hence, it has to vanish in the interval by the intermediate value theorem. Therefore it is real rooted.
($\Leftarrow$): We first assume that f and g have no common roots and that their roots are simple. Now, the roots of the average polynomial trace out n different intervals I_i on the real line as λ goes from 0 to 1. It starts with the roots of p_1 and ends at the roots of p_0. Now, each of these intervals must contain exactly one root of p_0 and one root of p_1 only, since otherwise, there would exist a λ such that for some $x \in \mathbb{R}$, the average and p_0 (say) both have a zero at x. But this forces p_1 to also have a zero at x which contradicts our assumption of no common roots. So now, we can choose subintervals of I_i such that they don't overlap and each contain exactly one root of p_0 and p_1. Hence, we get interlacing. The no roots assumption can be overcome by factoring out the common roots, getting the interlacing of the factored out polynomials, and then, add a root of the common interlacing polynomial on any one side of the common root. For root multiplicity, we use the fact from above. We take a sequence of polynomials p_{0_ϵ} and p_{1_ϵ}. These have simple roots for any $\epsilon > 0$ and converge to p_0 and p_1 respectively, uniformly on any bounded interval as ϵ goes to zero. We get the result by using the continuity of roots as a function of the coefficients. ■

Now let's see how we prove that our family of characteristic polynnomials of signed matrices is infact interlacing.

- Let
$$p_{s_1,\ldots,s_k}(x) = \mathbb{E}_{s_{k+1},\ldots,s_m}[p_{s_1,\ldots,s_m}(x)]$$
- The leaves of the tree are signings $s_1, \ldots, s_m$.
- The internal nodes are partial signings $s_1, \ldots, s_k$.

We want to prove that for all $s_1, \ldots, s_k$ and $\lambda \in [0, 1]$,

$$\lambda p_{s_1,\ldots,s_k,1}(x) + (1 - \lambda)p_{s_1,\ldots,s_k,-1}(x)$$

is real rooted, i.e., both extensions (children) have a common interlacing.

But we observe that any such polynomial is an expected characteristic polynomial over *all* the signings, but with a *bias* distribution:

- $s_1, \ldots, s_k$ are fixed.
- s_{k+1} is 1 with probability λ and -1 with probability $(1 - \lambda)$.
- $s_{k+2}, \ldots, s_m$ are uniformly ± 1.

So, it suffices to prove that $\mathbb{E}_{s\in\{\pm1\}^m}[p_s(x)]$ is real rooted for every *product* distribution on the entries of s. By a *product* distribution, we mean that for every sign, we choose it with some bias, but they're independent. So, we need to show that

$$\sum_{a\in\{\pm1\}^m} p_s(x) \prod_{i:s_i=1} \lambda_i \prod_{i:s_i=-1} (1 - \lambda_i)$$

for $\lambda_1, \ldots, \lambda_m \in [0, 1]$ is real rooted.

And this can be seen as a generalization of Heilman and Leib because when the distribution is uniform, this is just the matching polynomial. We have now generalized it for any product distribution. So, we can view this as the generalized matching polynomial.

Now, it suffices to show real rootedness of

$$\mathbb{E}_{s\in\{\pm1\}^m}[p_s(x - d)] = \mathbb{E}_{s\in\{\pm1\}^m} det(xI - (dI - A_s)),$$

since changing the variable from x to $x - d$ just shifts the roots by d and hence doesn't change the real rootedness.

Now, this is useful because $dI - A_s$ can be written as signed sum of rank one matrices instead of signed sum of rank two matrices, which is what A_s is:

$$A_s = \sum_{ij \in E} s_{ij}(\delta_i \delta_j^T + \delta_j \delta_i^T).$$

But,

$$dI - A_s = \sum_{s_{ij}=1} (\delta_i - \delta_j)(\delta_i - \delta_j)^T + \sum_{s_{ij}=-1} (\delta_i + \delta_j)(\delta_i + \delta_j)^T$$

This is similar to what we do when we turn an adjacency matrix into a laplacian matrix. In this case, if all signs were 1, then A_s is the adjacency matrix and $dI - A_s$ is the laplacian matrix. So we can think of this as the signed laplacian matrix.

Now, we can write our polynomial of interest as an expected characteristic polynomial of a sum of independent rank one matrices:

$$\mathbb{E}_{s \in \{\pm 1\}^m} det(xI - (dI - A_s)) = \mathbb{E} det(xI - \sum_{ij \in E} v_{ij} v_{ij}^T),$$

where v_{ij} is $(\delta_i - \delta_j)$ with probability λ_{ij} and is $(\delta_i + \delta_j)$ with probability $(1 - \lambda_{ij})$.

Now, we have a master real rootedness theorem that all such polynomials have real roots.

Theorem 23 *Given "any" independent random vectors $v_1, \dots, v_m \in \mathbb{R}^d$, their expected characteristic polynomial*

$$\mathbb{E}det(xI - \sum_i v_i v_i^T)$$

has real roots.

Proof: The proof is by the multivariate method. Whenever we want to prove real rootedness of univariate polynomials, it is usually useful to view the univariate polynomial as a restriction of a multivariate polynomial and prove that the multivariate polynomial has "real roots". A multivariate real rooted polynomial is called a "Real Stable Polynomial", which we saw previously. And we saw that for positive semidefinite matrices, the determinant of a linear combination (w.r.t. to some variables z_i) is real stable.

So, the plan is to apply closure properties of real stable polynomials to such an initial real stable polynomial to get our required polynomial.

For this, we use the following identity:
Suppose $v_1, \dots, v_m$ are **independent** random vectors with $A_i := \mathbb{E}v_i v_i^T$. Then,

$$\mathbb{E}det(xI - \sum_i v_i v_i^T) = \prod_{i=1}^{m} \left(1 - \frac{\partial}{\partial z_i}\right) det(xI + \sum_i z_i A_i)|_{z_1 = \dots z_m = 0}$$

The left hand side is a univariate polynomial that is a sum of possibly exponentially many polynomials. The right hand side is a multivariate polynomial with $m+1$ variables (the z_i's are m in number). But we set all the new variables to zero at the end. So we get a univariate polynomial. And rank one and independence of the vectors is critical here, since if we drop any of these, nothing like this is true.

We omit the proof of this identity (but note that it goes through the Cauchy-Binet formula). And, once we have this, the result is immediate since on the right hand side, the determinant is real stable as the matrices inside are all PSD. Then, from the closure properties, the differentiation and the setting of variables preserves real stability. Hence the whole polynomial is real stable. But since it is univariate, this means that it is real rooted.

■

3.6 Some observations

We note some key features of the proof techniques and some further results.

- The proof is non constructive. It's only existential. Cohen (2016) gave a constructive polynomial time algorithm to find such Ramanujan Graphs.
- In the proof, there are two distinct parts. We reduced the existence of a good matrix to:
  1. Proving real rootedness of an expected polynomial.
  2. Bounding roots of the expected polynomial.
- The first part of proving real rootedness was done by a very general means (Real Stability). It works for any sum of independent rank one matrices. But the second part was done using very specific techniques of combinatorics catered to the matching polynomial and graphs. It turns out we can actually even do the second part in a general way.
- And Marcus, Spielman and Srivastava used this to prove some other things in combinatorics and linear algebra. They also resolved an old foundational conjecture in Physics called the Kadison-Singer problem using the same techniques.

Bibliography

Yonatan Bilu and Nathan Linial. Lifts, discrepancy and nearly optimal spectral gap*. *Combinatorica*, 26:495–519, 2006.

M. B. Cohen. Ramanujan graphs in polynomial time. In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 276–281, Oct 2016. doi: 10.1109/FOCS.2016.37.

Joel Friedman. A proof of alon's second eigenvalue conjecture and related problems. *CoRR*, cs.DM/0405020, 2004. URL `http://arxiv.org/abs/cs.DM/0405020`.

Chris Godsil and Ivan Gutman. On the matching polynomial of a graph. *Algebraic Methods in Graph Theory*, 25, 01 1978.

Ole J. Heilmann and Elliott H. Lieb. Theory of monomer-dimer systems. *Comm. Math. Phys.*, 25(3):190–232, 1972. URL `https://projecteuclid.org:443/euclid.cmp/1103857921`.

A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988. ISSN 0209-9683. doi: 10.1007/BF02126799. URL `https://doi.org/10.1007/BF02126799`.

Adam W. Marcus, Daniel A. Spielman, and Nikhil Srivastava. Interlacing families I: Bipartite Ramanujan graphs of all degrees. *Ann. of Math. (2)*, 182(1):307–325, 2015. ISSN 0003-486X. doi: 10.4007/annals.2015.182.1.7. URL `https://doi.org/10.4007/annals.2015.182.1.7`.

G. A. Margulis. Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators. *Problemy Peredachi Informatsii*, 24(1):51–60, 1988. ISSN 0555-2923.

A. Nilli. On the second eigenvalue of a graph. *Discrete Math.*, 91 (2):207–210, 1991. ISSN 0012-365X. doi: 10.1016/0012-365X(91) 90112-F. URL `https://doi.org/10.1016/0012-365X(91)90112-F`.