CS 798 - Algorithmic Spectral Graph Theory, Fall 2015, Waterloo

Lecture 20: Bipartite Ramanujan graphs

We prove the real-rootedness of the mixed characteristic polynomials using real stability, and then use the method of interlacing family to prove the existence of bipartite Ramanujan graphs.

The expected characteristic polynomials in this setting are equivalent to matching polynomials of graphs, which were well studied and we will discuss some basic results about them.

Ramanujan graphs

Given a d-regular undirected graph G, let $d = \alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_n$ be the eigenvalues of its adjacency matrix.

We say G is Ramanujan if $\max\{\alpha_2, |\alpha_n|\} \leq 2\sqrt{d-1}$.

We are interested in constructing an infinite family of d-regular graphs that are all Ramanujan.

This is best possible, as Alon and Boppana proved that for any $\varepsilon > 0$, every large enough d-regular graph has $\max\{\alpha_2, |\alpha_n|\} \geq 2\sqrt{d-1} - \varepsilon$. See the reference for a proof.

(We have seen a simple argument that $\max\{\alpha_2, |\alpha_n|\} \geq \Omega(\sqrt{d})$ in L10.)

There is a meaning of the value $2\sqrt{d-1}$. It is a bound on the absolute value of the eigenvalues of the infinite d-regular tree, intuitively the best possible d-regular expander graph.

There are known constructions of Ramanujan graphs of constant degree from Cayley graphs.

All known graphs are $(q+1)$-regular where q is a prime power.

The proofs use deep mathematical results and in particular some by Ramanujan (and hence the name).

They are explicit in that the neighbors of a vertex can be computed in $O(\log n)$ time.

See the survey by Hoory-Linial-Wigderson for more details.

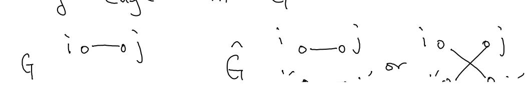
2-lift

It is of interest to find combinatorial constructions of Ramanujan graphs.

Bilu and Linial proposed a method to construct Ramanujan graphs using 2-lifts.

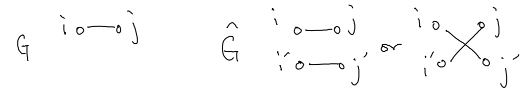
Given a graph $G = (V, E)$, a 2-lift of G is a graph $\hat{G} = (\hat{V}, \hat{E})$ where $\hat{V}$ is two copies of V, i.e. if $V = \{1, 2, \ldots, n\}$, then $V' = \{1, 2, \ldots, n, 1', 2', \ldots, n'\}$.

For each edge $ij \in E(G)$, there are two options to put corresponding edges in $\hat{G}$: either we put ij and $i'j'$, or put ij' and $i'j$.



For each edge $ij \in E(G)$, there are two options to put corresponding edges in $\hat{G}$:

either we put ij and $i'j'$, or put ij' and $i'j$.



So, given G, there are 2^m possible 2-lifts of G.

Bilu and Linial conjected that if G is Ramanujan, then there is a 2-lift of G that is also Ramanujan.

Note that if G is d-regular, any 2-lift of G is also d-regular with double number of vertices.

So, if the conjecture is true, it implies the existence of an infinite family of d-regular Ramanujan graphs for any degree d. Just start with the complete graph on $d+1$ vertices, which is Ramanujan, and keep doing a good 2-lift to double the graph size.

Bilu and Linial used probabilistic method (Lovász local lemma) to prove that there is a 2-lift with $\max\{\alpha_2, |\alpha_n|\} \leq O(\sqrt{d \log^3 d})$.

Bipartite Ramanujan graphs

Marcus, Spielman, Srivastava [MSS1] used the method of interlacing polynomials to prove a variant of Bilu-Linial conjecture.

Recall that the (adjacency matrix) spectrum of a bipartite graph is symmetric, so $\alpha_1 = d$ and $\alpha_n = -d$.

We say a bipartite graph is Ramanujan if $\max\{\alpha_2, |\alpha_{n-1}|\} \leq 2\sqrt{d-1}$.

The main theorem that we study today is:

Theorem Given a bipartite Ramanujan graph G, there is a 2-lift of G that is Ramanujan.

Note that a 2-lift of a bipartite graph is bipartite. So, starting from a complete bipartite graph with $2d$ vertices, which is Ramanujan, it implies an infinite family of d-regular bipartite Ramanujan graph for any degree d.

Spectrum of signed matrix

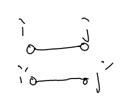
There is a nice formulation to analyze the spectrum of a 2-lift of a graph.

Let A be the adjacency matrix of G.

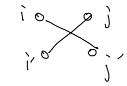
Let $\hat{G}$ be a 2-lift of G.

We encode the 2-lift $\hat{G}$ in a signed adjacency matrix A_s.

For $ij \in E(G)$, we set $(A_s)_{ij} = (A_s)_{ji} = 1$ if $ij \in E(\hat{G})$ and $i'j' \in E(\hat{G})$, i.e.



in $\hat{G}$;

otherwise, we set $(A_s)_{ij} = (A_s)_{ji} = -1$ if $ij' \in E(\hat{G})$ and $i'j \in E(\hat{G})$, i.e.  in G'

For $ij \notin E(G)$, we set $(A_s)_{ij} = (A_s)_{ji} = 0$.

<u>Lemma</u> The (adjacency matrix) spectrum of $\hat{G}$ is equal to the (disjoint) union of the spectrum of A (old eigenvalues) and the spectrum of A_s (new eigenvalues).

You are asked to prove this lemma in the homework.

With the lemma, to prove that there is a Ramanujan 2-lift of a Ramanujan graph, it is equivalent to proving that there is a signing of a Ramanujan graph (an assignment of ± 1 to each edge) so that the maximum absolute eigenvalue of A_s is at most $2\sqrt{d-1}$.

Bilu and Linial conjectured the stronger statement that any graph (not necessarily Ramanujan) has a signing such that all eigenvalues of A_s have absolute value at most $2\sqrt{d-1}$.

Marcus, Spielman, Srivastava proved this conjecture for bipartite graphs.

<u>Theorem</u> Any bipartite graph has a signing such that the maximum eigenvalue of A_s is $\leq 2\sqrt{d-1}$.

Note that for a bipartite graph, bounding the maximum eigenvalue is enough because the spectrum is symmetric.

This is the main reason that the result only holds for bipartite graphs, because the new probabilistic method using interlacing polynomials developed by MSS can only bound the maximum eigenvalue, or one eigenvalue, but not the maximum eigenvalue and the minimum eigenvalue at the same time.

We will prove this theorem using interlacing polynomials.

<u>Outline</u>

To use the method of interlacing polynomials, we need to establish the following two main steps.

1. Prove that there exists a signing such that $\text{maxroot}(\det(xI - A_s)) \leq \text{maxroot}\left(\mathop{\mathbb{E}}_{s \in \{\pm 1\}^n} \det(xI - A_s)\right)$.

   This involves the machinery that we have developed last time in L19.

2. Prove that $\text{maxroot}\left(\mathop{\mathbb{E}}_{s \in \{\pm 1\}^n} \det(xI - A_s)\right) \leq 2\sqrt{d-1}$.

   It turns out that the expected characteristic polynomial is the "matching polynomial" of the graph, a well-studied object in the literature, and existing results imply this bound.

We will first derive the second step using the results about matching polynomials.

Then, we will prove a general result about mixed characteristic polynomials, which will allow us to show that the new probabilistic method about maxroot works in this setting as well as in the setting of Weaver's conjecture. This will be the main focus of today.

Finally, we discuss the proofs of the known results about matching polynomials.

Expected characteristic polynomials and matching polynomials

Given a graph G, let m_i be the number of matchings in G with i edges (with $m_0 = 1$).

The matching polynomial is defined as $\mu_G(x) := \sum_{i \geq 0} x^{n-2i} (-1)^i m_i$.

Godsil and Gutman proved the following proposition.

Proposition $\quad \mathop{E}_{S \in \{\pm 1\}^n} \det(xI - A_S) = \mu_G(x)$.

proof We expand the determinant into permutations. Let $B_S = xI - A_S = \begin{pmatrix} x & & -1 & \\ & x & & \\ -1 & & x & 1 \\ & & 1 & x \end{pmatrix}$

Then $\mathop{E}_{S \in \{\pm 1\}^n} \det(xI - A_S) = \mathop{E}_{S \in \{\pm 1\}^n} \sum_{\sigma = \text{permutations}} \text{sgn}(\sigma) \prod_{i=1}^{n} (B_S)_{i,\sigma(i)} = \sum_{\sigma} \text{sgn}(\sigma) \prod_{i=1}^{n} E[(B_S)_{i,\sigma(i)}]$

where $\text{sgn}(\sigma) = (-1)^{\text{inv}(\sigma)}$ and $\text{inv}(\sigma) := |\{(i,j) \mid i<j \text{ and } \sigma(i) > \sigma(j)\}|$ is the number of inversions of σ.

Since $E[(B_S)_{ij}] = 0$ as it is ± 1 equally likely and each edge is independent, all the terms with at least one variable with degree one vanished.

Therefore, the terms remained can only be of the form $x^{n-2k} \prod_{l=1}^{k} (B_S)^2_{i_l, j_l} = x^{n-2k}$, where each edge appears as a degree two term $i_1 j_1 \quad i_2 j_2 \quad \ldots \quad i_l j_l$.

So, each matching of size k will contribute $\text{sgn}(\sigma)$ to the coefficient of x^{n-2k}.

We claim that every matching of size k has the same parity of the number of inversions, so that $\text{sgn}(\sigma) = -1$ if k is odd and $\text{sgn}(\sigma) = +1$ if k is even.

If the claim is true, then it follows that $\mathop{E}_{S \in \{\pm 1\}^n} \det(xI - A_S) = \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k m_k x^{n-2k} = \mu_G(x)$, proving the proposition.

It remains to check the claim. Given a permutation σ that corresponds to a matching of size k, we argue that its parity of the number of inversions is different from that of the identity permutation by one if k is odd and zero if k is even, and this would imply the claim.

To see this, let ij be a matched edge, so that $i<j$ and $\sigma(i) = j$ and $\sigma(j) = i$.

Consider the permutation that we just swap ij so that $\sigma(i)=i$ and $\sigma(j)=j$ while keeping other positions unchanged.

Then, observe that for each $l \neq i,j$, the number of inversions involving l is either unchanged or decrease by two, and hence the same parity for those inversions involving l.

However, the parity changes by one because of the pair ij (from ji to ij).

After k swaps, we get the identity permutation and the parity changes k times.

This implies that matchings of odd size have odd parity, and of even size have even parity.

This proves the claim and hence the proposition. $\square$

Heilmann and Lieb (1972) first studied the matching polynomials and prove some interesting results:

- the matching polynomials are real-rooted.
- for graphs with maximum degree d, all roots have absolute value at most $2\sqrt{d-1}$.

Combining the results by Godsil-Gutman and Heilmann-Lieb gives $\operatorname{maxroot}\left(\mathbb{E}_{s\in\{\pm1\}^n} \det(xI-A_s)\right) \leq 2\sqrt{d-1}$, and this completes the proof of the second step.

We will discuss the proofs of these existing results about matching polynomials later.

Interlacing family

Now we want to establish the first step: $\exists s \in \{\pm1\}^n$ with $\operatorname{maxroot}(\det(xI-A_s)) \leq \operatorname{maxroot}\left(\mathbb{E}_{s\in\{\pm1\}^n} \det(xI-A_s)\right)$.

We know from L19 that if $f_1,\ldots,f_m$ have a common interlacing, then $\exists i$ with $\operatorname{maxroot}(f_i) \leq \operatorname{maxroot}(\mathbb{E} f_i)$.

In the current setting, there are 2^n different signings, and we don't expect that they have a common interlacing. e.g. $A_{+,+,+,+,+}$ and $A_{-,-,-,-,-}$ would generally not having a common interlacing.

This does not mean that the conclusion is not true, just that we need to be more careful.

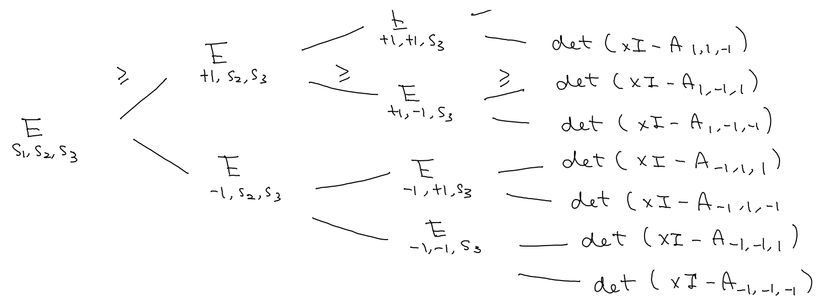
We will consider the "conditional" expectation polynomials and use a tree structure to establish it.

A conditional expectation polynomial is of the form $\mathbb{E}_{\substack{s\in\{\pm1\}^n \\ \text{given } s_1=a_1,\ldots,s_k=a_k}} \det(xI-A_s)$, i.e. the conditional expectation given the first k bits are fixed.

Informally, say $m=3$, we will prove the following interlacing tree structure:

$$\mathbb{E}_{+1,+1,s_3} \;\text{—}\; \det(xI-A_{1,1,1}),\ \det(xI-A_{1,1,-1})$$

$$\geq \;\mathbb{E}_{+1,s_2,s_3} \;\text{—}\; \mathbb{E}_{+1,+1,s_3},\ \geq\ \mathbb{E} \quad \geq\ \det(xI-A_{1,-1,1})$$



We will show that for each node in the tree, all the polynomials in its children have a common interlacing.

This is enough to establish the first step, since starting from the root, we know that there is a child with maxroot at most its parent, and we can repeat this until we reach a leaf, and this is the polynomial that corresponds to a concrete signing, and we are done.

We will execute the proof plan in a more general setting, so that the result can also be readily applied to the Weaver's setting.

Mixed characteristic polynomials

Let A_i be a random symmetric rank-one matrix (e.g. $A_i = \begin{cases} aa^T & \text{with prob } \frac{1}{4} \\ bb^T & \text{with prob } \frac{1}{3} \\ cc^T & \text{with prob } \frac{5}{12} \end{cases}$)

Let $A = \sum_{i=1}^m A_i$ be a sum of random rank-one matrices.

We are interested in showing that $\det(xI - \sum_{i=1}^m A_i)$ form an interlacing family.

The following identity is at the heart of this approach.

Theorem $\mathbb{E}_{A_1,\dots,A_m} \det(zI - \sum_{i=1}^m A_i) = \left(\prod_{i=1}^m \left(1 - \frac{\partial}{\partial z_i}\right)\right) \det\left(zI + \sum_{i=1}^m z_i \mathbb{E}[A_i]\right)\Big|_{z_1=z_2=\dots=z_m=0}$

We will prove this theorem in the next section.

We will first see how this implies the new probabilistic method about max-root would work.

Corollary The expected characteristic polynomial $\mathbb{E}_{A_1,\dots,A_m} \det(zI - \sum_{i=1}^m A_i)$ is real-rooted for any random symmetric rank-one matrices $A_1,\dots,A_m$.

proof We start from the RHS of the theorem.

Since A_i is a random symmetric rank-one matrix, $\mathbb{E}[A_i] = \sum p_i v_i v_i^T \succcurlyeq 0$ is PSD.

By the result in L19, $\det(zI + \sum_{i=1}^m z_i \mathbb{E}[A_i])$ is a real stable polynomial.

Then, from the last section in L19, differentiation and substitution of real numbers are real-stability preserving operations.

So, the LHS of the theorem is a real-stable univariate polynomial, and hence real-rooted. □

<u>Corollary</u> The polynomials $\det(zI - \sum_{i=1}^{m} A_i)$ form an interlacing family.

<u>proof</u> Following the picture in the previous section, we just need to prove that the children of a node have a common interlacing.

More precisely, suppose we fix the first k variables to be $A_1 = v_1 v_1^T, \ldots, A_k = v_k v_k^T$, and let A_{k+1} has l random choices $u_1 u_1^T, \ldots, u_l u_l^T$.

Then, we need to prove that the l conditional polynomials $\mathop{E}_{A_{k+2},\ldots,A_m} \det(zI - \sum_{i=1}^{k} v_i v_i^T - u_j u_j^T - \sum_{i=k+2}^{m} A_i)$ for $1 \le j \le l$ have a common interlacing.

By the results in L19, it is equivalent to proving that for any convex combination μ, the polynomial $\mu_j \mathop{E}_{A_{k+2},\ldots,A_m} \det(zI - \sum_{i=1}^{k} v_i v_i^T - u_j u_j^T - \sum_{i=k+2}^{m} A_i)$.

Note that this is just the expected characteristic polynomial $\mathop{E}_{B_1,\ldots,B_m} \det(zI - \sum_{i=1}^{m} B_i)$ for a related set of random symmetric rank-one matrices, where B_1 to B_k are just the (deterministic) random variables with $B_i = v_i v_i^T$ with probability one, B_{k+1} is the random variable with $B_{k+1} = u_j u_j^T$ with probability μ_j, and B_{k+2} to B_m are just the same as the random variables A_{k+2} to A_m.

So, by the previous corollary, this convex combination is real-rooted, and hence the children have a common interlacing, and hence the tree forms an interlacing family. □

So, the new probabilistic method about maxroot works for these polynomials, i.e.

$\exists A_1, \ldots, A_m$ such that $\operatorname{maxroot}\left(\det(xI - \sum_{i=1}^{m} A_i)\right) \le \operatorname{maxroot}\left(\mathop{E}_{A_1 \ldots A_m} \det(xI - \sum_{i=1}^{m} A_i)\right)$.

It remains to prove the theorem and show that the signing family is a special case, which we will do in the following two sections.

<u>Multilinear formula</u>

We will prove that $\mathop{E}_{A_1,\ldots,A_m} \det(zI - \sum_{i=1}^{m} A_i) = \left(\prod_{i=1}^{m} \left(1 - \frac{\partial}{\partial z_i}\right)\right) \det(zI + \sum_{i=1}^{m} z_i E[A_i]) \Big|_{z_1 = z_2 = \ldots = z_m = 0}$.

We will do this in two steps, first deterministic, then take expectation.

One variable

First, in a simple form, we observe $\det(B + tvv^T)$ is a linear function in terms of t.

To see this, we recall that $\det(B + tvv^T) = \det(B)(1 + tv^T B^{-1} v) = \det(B)(1 + t\,\mathrm{Tr}(B^{-1} vv^T))$,

and hence it is a linear function in terms of t and vv^T.

When a function $f(t)$ is linear in t, we can write it as $f(t) = f(0) + tf'(0) = \left(1 + t\frac{\partial}{\partial z}\right) f(z)\Big|_{z=0}$

So, $\det(B + tvv^T) = \left(1 + t\frac{\partial}{\partial z}\right) \det(B + zvv^T)\Big|_{z=0}$.

This formula is true when B is non-singular, and by a continuity argument (similar to the one in L19 about substituting real numbers), it is true for all B.

This is where the rank-one condition is used, to show that the function is linear.

Many variables

Now, consider $\det(B + t_1 v_1 v_1^T + \dots + t_m v_m v_m^T)$ where t_i are variables.

Fixing all but one t_i, the above argument shows that it is linear in t_i.

Therefore, this is a multilinear polynomial in terms of $t_1, \dots, t_m$.

If $f(t_1, \dots, t_m)$ is a multilinear polynomial, we can write it as

$f(t_1, \dots, t_m) = \sum_{S \subseteq [m]} \left(\prod_{i \in S} t_i\right)\left(\prod_{i \in S} \frac{\partial}{\partial z_i}\right) f(z_1, \dots, z_m)\Big|_{z_1 = z_2 = \dots = z_m = 0}$, where each term corresponds to taking off a subset of variables, and putting in zeros to get the coefficient of that subset, e.g.

$f(t_1, t_2, t_3) = 3t_1t_2 + 4t_2t_3 + 2t_1t_3$, then to get the coefficient of t_1t_3, we can compute $\frac{\partial}{\partial z_1}\frac{\partial}{\partial z_3} f(z_1, z_2, z_3) = 3z_2 + 4z_2 + 2$, and then substitute $z_1 = z_2 = z_3 = 0$ to get 2.

Note that $f(t_1, \dots, t_m) = \sum_{S \subseteq [m]} \left(\prod_{i \in S} t_i\right)\left(\prod_{i \in S} \frac{\partial}{\partial z_i}\right) f(z_1, \dots, z_m)\Big|_{z_1 = z_2 = \dots = z_m = 0} = \prod_{i=1}^{m}\left(1 + t_i \frac{\partial}{\partial z_i}\right) f(z_1, \dots, z_m)\Big|_{z_1 = z_2 = \dots = z_m = 0}$

Therefore, we can write $\det(B + t_1 A_1 + \dots + t_m A_m) = \prod_{i=1}^{m}\left(1 + t_i \frac{\partial}{\partial z_i}\right) \det(B + z_1 A_1 + \dots + z_m A_m)\Big|_{z_1 = \dots = z_m = 0}$

when A_i are rank-one symmetric matrices.

Setting $B = xI$ and $t_i = -1$ for all i, we get $\det\left(xI - \sum_{i=1}^{m} A_i\right) = \prod_{i=1}^{m}\left(1 - \frac{\partial}{\partial z_i}\right) \det\left(xI - \sum_{i=1}^{m} z_i A_i\right)\Big|_{z_1 = z_2 = \dots = z_m = 0}$.

Expectation

To get the theorem, we take the expectation of both sides of the above equation.

The LHS is then the LHS of the theorem.

For the RHS, before we do the substitution, it is a multilinear polynomial in z_i and A_i.

Since the A_i are independent random variables, the expectation can be pushed inside and get the RHS of the theorem.

For example, suppose we have a term $\frac{\partial}{\partial z_1 \partial z_2 \partial z_3} z_1 \mathrm{Tr}(B_1 A_1) z_2 \mathrm{Tr}(B_2 A_2) z_3 \mathrm{Tr}(B_3 A_3)$,

then $E\left[\frac{\partial}{\partial z_1 \partial z_2 \partial z_3} z_1 \mathrm{Tr}(B_1 A_1) z_2 \mathrm{Tr}(B_2 A_2) z_3 \mathrm{Tr}(B_3 A_3) \right]$

$= \frac{\partial}{\partial z_1 \partial z_2 \partial z_3} E\left[z_1 \mathrm{Tr}(B_1 A_1) z_2 \mathrm{Tr}(B_2 A_2) z_3 \mathrm{Tr}(B_3 A_3) \right]$ since differentiation is lineer

$= \frac{\partial}{\partial z_1 \partial z_2 \partial z_3} z_1 z_2 z_3 E[\mathrm{Tr}(B_1 A_1)] E[\mathrm{Tr}(B_2 A_2)] E[\mathrm{Tr}(B_3 A_3)]$ since A_i are independent

$= \frac{\partial}{\partial z_1 \partial z_2 \partial z_3} z_1 z_2 z_3 \mathrm{Tr}(B_1 E[A_1]) \mathrm{Tr}(B_2 E[A_2]) \mathrm{Tr}(B_3 E[A_3])$ since trace is linear.

Therefore, after taking expectation on RHS, we can push it to the innermost, and the RHS is just the original RHS with A_i replaced by $E[A_i]$.

This proves the theorem.

There are other proofs [MSS2] which are probably more precise and shorter, but this proof (presented by Tao) is more insightful about why this is true and where is it coming from.

Finishing step one

Now, we have proved that the family of polynomials $\det(B - \sum_{i=1}^{m} A_i)$ where A_i are random rank-one matrices form an interlacing family, which establishes the foundation of the probabilistic method.

We can use it as is in the Weaver's setting.

Here we see that it also applies in the Ramanujan setting (the signed matrices).

We consider the family $\det(xI - A_s)$ for $s \in \{\pm 1\}^n$.

We can write $A_s = \sum_{e \in G} A_e$ where $A_e = \begin{array}{c|cc} & i & j \\ \hline i & & 1 \\ j & 1 & \end{array}$ or $A_e = \begin{array}{c|cc} & i & j \\ \hline i & & -1 \\ j & -1 & \end{array}$, but the A_e are not rank-one.

In short, we do a transformation to turn adjacency matrices into Laplacian matrices and then each edge will be a rank-one PSD matrix and the result will apply.

Let d be the maximum degree.

Note that $E \det(xI - A_s)$ is real-rooted iff $E \det(xI + dI - A_s)$ is real-rooted, as we just shift the roots by d.

Now the matrix $dI - A_s = D + \sum_{e \in G} L_e'$, where $L_e' = \begin{cases} (\chi_i - \chi_j)(\chi_i - \chi_j)^T = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} & \text{if } (A_s)_{i,j} = 1 \\ (\chi_i + \chi_j)(\chi_i + \chi_j)^T = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} & \text{if } (A_s)_{i,j} = -1. \end{cases}$

and D is a diagonal matrix where $D_{i,i} = d - \deg(i) \geq 0$ and so $D \succcurlyeq 0$.

Therefore, $\det(xI + dI - A_s) = \det(\underbrace{xI + D} + \sum_{e \in G} L_e')$

and D is a diagonal matrix where $D_{i,i} = d - \deg(i) \geq 0$ and so $D \succcurlyeq 0$.

Therefore, $\det(xI + dI - A_s) = \det(\underbrace{xI + D}_{\succcurlyeq 0} + \sum_{e \in G} L_e')$ ← sum of rank one PSD.

By the results (two corollaries) about mixed characteristic polynomials, we conclude that they form an interlacing family, and so is the family $\det(xI - A_s)$.

We have completed the proofs about bipartite Ramanujan graphs.

Matching polynomials

We discuss Heilmann and Lieb's results about matching polynomials that they are real-rooted and have maxroot $\leq 2\sqrt{d-1}$ when the maximum degree of the graph is d.

The original proof uses recursion and induction.

We present an approach by Godsil, which is more systematic and consists of three steps:

1. The matching polynomial of a graph of maximum degree d divides the matching polynomial of an associated tree of maximum degree d.
2. The matching polynomial of a tree is equal to its characteristic polynomial.
3. The maximum eigenvalue of a tree of maximum degree d is at most $2\sqrt{d-1}$.

Since the characteristic polynomial is real-rooted, 1+2 implies that the matching polynomials are real-rooted, and 3 implies that the maxroot of the characteristic polynomial of the tree is at most $2\sqrt{d-1}$, and hence implies Heilmann and Lieb's results.

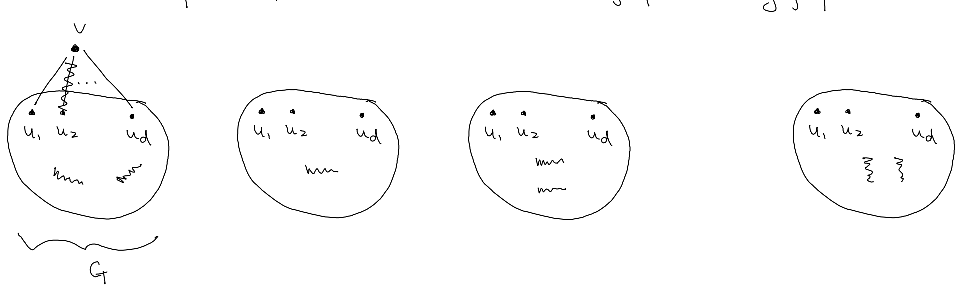
The second step is an exercise.

The third step is left as a homework problem.

We sketch the proof of the first step.

Given a graph G, we look at an arbitrary vertex v.

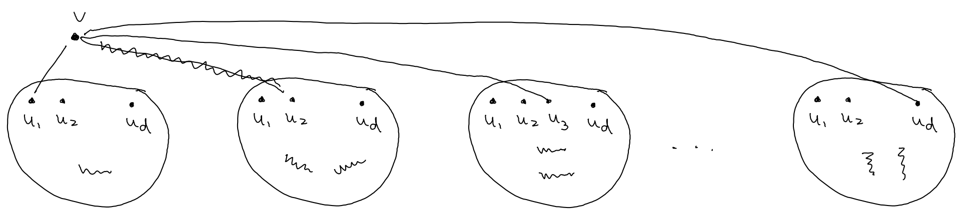
We make $d-1$ copies of $G-v$ and call the graph resulting graph H.



H is the disjoint union of all these.

It should be clear that the matching polynomial of G divides the polynomial of H, as it is easy to verify that $\mu_H(x) = \mu_G(x) \cdot \left(\mu_{G-v}(x)\right)^{d-1}$, i.e. product of the matching polynomials.

Now, consider the following graph H', where vu_i in the first copy is replaced by vu_i in the i-th copy.



The claim is that the matching polynomials of H and H' are the same.

The reason is that there is a one-to-one correspondence between matchings in H and matchings in H', as v can only be matched to one vertex (see the matchings in the pictures).

Now, in H', there is no (simple) cycles involving v.

Applying the same operations (duplicate and "branch") on all the copies of u_1 and so on (on all the copies of another vertex), the resulting (huge) graph will have no cycles and is a tree.

All these operations preserve the property that the matching polynomial of the small graph divides that of the bigger graph, and so eventually the matching polynomial of G divides the matching polynomial of the final graph, which is a tree. This proves the first step.

The proof can be made more formal but I think it is easier to understand without symbols.

References

- [MSS1] Interlacing families I: Bipartite Ramanujan graphs of all degrees. by Marcus, Spielman, Srivastava
- [MSS2] Interlacing families II: Mixed characteristic polynomials and the Kadison-Singer problem.
- Real stable polynomials and the Kadison-Singer problem, blogpost by Terence Tao.
- Algebraic combinatorics (chapter 5 and 6 for matching polynomials), by Godsil.