4 Covers and Interlacing families

Two recent and relevant developments on spectral graph theory are going to be discussed in this section. Common to both is the use of signs in the adjacency matrix of a graph.

The first is a warm up, and also a review of our introductory section. It is a creative yet easy application, that solved a decades old problem originating in computer science.

The second represents an important development, that lead to the solution of several open problems.

4.1 Sensitivity

Let $Q_n = \{0,1\}^n$, and assume f is a boolean function, that is, $f : Q_n \to \{0,1\}$. For several applications in computer science, it is important to know or measure how f behaves if a coordinate of a point is changed. We say that $x, y \in Q_n$ are neighbours if they differ in one coordinate. In fact, we shall identify Q_n with a graph, henceforth called the n-dimensional hypercube.

The sensitivity of f at $x \in Q_n$, denoted by $s(f,x)$, is the number of neighbours y of x for which $f(x) \neq f(y)$. The sensitivity of the function f is the maximum sensitivity of f at all points of Q_n. In a sense, the sensitivity of a function is related to its complexity — for instance, a function which is determined by the first coordinate of an entry has sensitivity equal to 1.

There are several ways of measuring the complexity of a function, and an active field of research consists in deciding how two different ways compare. We introduce here below one other.

Every boolean function f can be uniquely represented a real multilinear polynomial. To see that, first perform the substitution $0 \to 1$ and $1 \to -1$. Let $S_n = \{1,-1\}^n$. Thus we see boolean functions as going from S_n to S_1.

Let χ_I be the function

$$\chi_I(x) = \prod_{i \in I} x_i,$$

Then, define the inner product that takes two functions and maps to a real number, given by

$$\langle f, g \rangle = 2^{-n} \sum_{x \in Q_n} f(x) g(x).$$

It is immediate to verify that $\langle \chi_I, \chi_I \rangle = 1$, and if $I \neq J$, then

$$\langle \chi_I, \chi_J \rangle = 2^{-n} \sum_{x \in Q_n} \chi_{I \Delta J}(x) = 0.$$

Because the space of all functions has dimension 2^n and we have just created 2^n orthogonal functions, they form a basis, and therefore there are coefficients α_I for each $I \subseteq [n]$ so that

$$f(x) = \sum_{I \subseteq [n]} \alpha_I \chi_I(x).$$

Moreover, each α_I is given by

$$\alpha_I = \langle f, \chi_I \rangle.$$

These are also called the Fourier transform of f at I. Let $d(f)$ denote the degree of the polynomial given above, that is,

$$d(f) = \max_{I \subseteq [n]} \{|I| : \alpha_I \neq 0\}.$$

In a sense this is also a measurement of complexity of the function f, and one wonders if there is any connection between $d(f)$ and $s(f)$. By standard methods, Nisam and Szegedy showed that $d(f) \geq \sqrt{s(f)/2}$. Our goal is to show the reverse inequality, which is an open problem that remained open for almost 30 years.

Gotsman and Linial showed the connection below.

Theorem 4.1. *For any $h : \mathbb{N} \to \mathbb{R}$, the following are equivalent.*

(a) Let $H \subseteq V(Q_n) = V$. Assume $|H| \neq 2^{n-1}$. Then we have that either $\Delta(Q_n[H]) \geq h(n)$ or $\Delta(Q_n[V - H]) \geq h(n)$.

(b) For any Boolean function $f : Q_n \to \{0, 1\}$, we have $s(f) \geq h(d(f))$.

Thus, in order to show that $s(f) \geq \sqrt{d(f)}$, it is enough to show that for any induced subgraph of the hypercube Q_n with more than 2^{n-1} vertices, the maximum degree is at least $\sqrt{n}$.

This problem is remarkably tricky. Note that there is an induced subgraph on 2^{n-1} vertices with maximum degree 0, because the hypercube is bipartite (the parity of the strings?). By adding one vertex to this graph you suddenly get something of degree n, which is as bad as it gets.

A better attempt is to split $[n]$ into (roughly) $\sqrt{n}$ sets of size $\sqrt{n}$. Say n is perfect square. Call these F_i, with $i = 1, ..., \sqrt{n}$. Consider the subset of $2^[n]$ formed by the sets of even cardinality that do not contain any F_i and those of odd cardinality that contain some F_i. It is possible to check that this family has size $2^{n-1} + 1$, and that the subgraph it induces has maximum degree $\sqrt{n}$. In the next section we show that this is best possible.

4.2 Cauchy Interlacing — a first contact

In the beginning of this course, we saw in Lemma 1.33 that the largest eigenvalue of a matrix is the maximum over a Rayleigh quotient, that is

$$\lambda_1(\mathbf{M}) = \max_{\mathbf{u} \neq 0} R_{\mathbf{M}}(\mathbf{u}).$$

Now assume $\mathbf{u}$ has been found, and we search for the maximum over all vectors orthogonal to $\mathbf{u}$. Then

$$\begin{aligned} R_{\mathbf{M}}(\mathbf{v}) &= \theta_0(\mathbf{v}^{\mathcal{T}}\mathbf{E}_0\mathbf{v}) + \theta_1(\mathbf{v}^{\mathcal{T}}\mathbf{E}_1\mathbf{v}) + \cdots + \theta_d(\mathbf{v}^{\mathcal{T}}\mathbf{E}_d\mathbf{v}) \\ &\leq \theta_1\big((\mathbf{v}^{\mathcal{T}}\mathbf{E}_0\mathbf{v}) + (\mathbf{v}^{\mathcal{T}}\mathbf{E}_1\mathbf{v}) + \cdots + (\mathbf{v}^{\mathcal{T}}\mathbf{E}_d\mathbf{v})\big) \\ &\leq \theta_1. \end{aligned}$$

Equality will be met if and only if $\mathbf{v}$ is eigenvector for θ_1. This argument generalizes for other eigenvalues and eigenvectors, and also for min instead of max if one looks for the smallest eigenvalues.

Theorem 4.2 (Cauchy's Interlacing — light version)**.** *Let $\mathbf{A}$ be a symmetric $n \times n$ matrix, and let $\mathbf{B}$ be a $m \times m$ principal square submatrix of $\mathbf{A}$. Let $\theta_1 \geq ... \geq \theta_n$ be the eigenvalues of $\mathbf{A}$ and $\lambda_1 \geq ... \geq \lambda_m$ be those of $\mathbf{B}$. Then, for all k with $1 \leq k \leq m$,*

$$\theta_{n-(m-k)} \leq \lambda_k \leq \theta_k.$$

Proof. Note there is $\mathbf{S}$ an $n \times m$ matrix satisfying $\mathbf{S}^T\mathbf{S} = \mathbf{I}$ and so that $\mathbf{B} = \mathbf{S}^T\mathbf{A}\mathbf{S}$. Let $\mathbf{u}_1, ..., \mathbf{u}_n$ be the eigenvectors of $\mathbf{A}$ corresponding to the θ_ks. The key thing now is to observe that, for all k, the subspace

$$\langle \mathbf{v}_1, ..., \mathbf{v}_k \rangle \cap \langle \mathbf{S}^T\mathbf{u}_1, ..., \mathbf{S}^T\mathbf{u}_{k-1} \rangle^\perp$$

contains at least one vector. Let $\mathbf{w}$ be such vector, which, in particular, implies $\mathbf{S}\mathbf{w} \in \langle \mathbf{u}_1, ..., \mathbf{u}_{k-1} \rangle^\perp$. Then, by the observation above, we have

$$\theta_k \geq \frac{(\mathbf{S}\mathbf{w})^T\mathbf{A}(\mathbf{S}\mathbf{w})}{(\mathbf{S}\mathbf{w})^T(\mathbf{S}\mathbf{w})} \geq \frac{\mathbf{w}^T\mathbf{B}\mathbf{w}}{\mathbf{w}^T\mathbf{w}} \geq \lambda_k.$$

The other inequalities follow from a symmetric argument. □

4.3 The Sensitivity Theorem

In 2019, Hao Huang shocked the graph theoretic community by providing a one-page proof of the Sensitivity Conjecture. Upon a clever yet elementary application of the interlacing principle, he showed that if H is a subgraph of the hypercube Q_n with more than 2^{n-1} vertices, then $\Delta(H) \geq \sqrt{n}$. To achieve this, he came up with a clever *signing* of the adjacency matrix of the hypercube.

Lemma 4.3. *Define the sequence $\mathbf{A}_i$ of matrices as follows:*

$$\mathbf{A}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \textit{and} \quad \mathbf{A}_n = \begin{pmatrix} \mathbf{A}_{n-1} & \mathbf{I} \\ \mathbf{I} & -\mathbf{A}_{n-1} \end{pmatrix}.$$

Then $\mathbf{A}_n$ is a $2^n \times 2^n$ matrix whose eigenvalues are $\pm\sqrt{n}$, each with multiplicity 2^{n-1}.

Proof. It follows immediately by induction that $\mathbf{A}_n^2 = n\mathbf{I}$. The result then follows by noticing that $\operatorname{tr} \mathbf{A}_n = 0$. □

The matrix $\mathbf{A}_n$ above is almost the adjacency matrix of the hypercube Q_n. In fact, changing all signs to $+1$, we get exactly $\mathbf{A}(Q_n)$.

Lemma 4.4. *Suppose H is a graph on m vertices, and $\mathbf{A}$ is an $m \times m$ symmetric matrix with entries 0, 1 and -1, whose nonzero entries correspond precisely to the edges of H. Then*

$$\Delta(H) \geq \lambda_1(\mathbf{A}).$$

Proof. Say $\mathbf{A}\mathbf{v} = \lambda_1 \mathbf{v}$, and assume $\mathbf{v}_a$ is the largest entry in absolute value of $\mathbf{v}$. Then

$$|\lambda_1 \mathbf{v}_a| = \left|(\mathbf{A}\mathbf{v})_a\right| \leq \sum_{b \sim a} |\mathbf{A}_{a,b}||\mathbf{v}_a| \leq \Delta(H)|\mathbf{v}_a|.$$

□

Theorem 4.5. *If H is a subgraph of the hypercube Q_n with more than 2^{n-1} vertices, then $\Delta(H) \geq \sqrt{n}$.*

Proof. Consider the "adjacency matrix" $\mathbf{A}_n$ of the hypercube, given in Lemma 4.3. Take a submatrix of $\mathbf{A}_n$ corresponding to the vertices in H, call it $\mathbf{B}$. Note that this is a ± 1 adjacency matrix of H, and therefore by the lemma above

$$\Delta(H) \geq \lambda_1(\mathbf{B}).$$

From interlacing, it follows that

$$\lambda_1(\mathbf{B}) \geq \lambda_{2^n - (2^{n-1}+1-1)}(\mathbf{A}_n) = \sqrt{n},$$

and the result follows. □

4.4 Covers

Start with a graph G, n vertices, with (standard) adjacency matrix $\mathbf{A}$, and assume you have also been given a signed adjacency matrix of G, say $\mathbf{B}$, as we did above for the hypercube.

Then replace every vertex of G by a pair of vertices. For each pair, decide which is the blue and which is the red, and form a new graph putting a matching between the two pairs of vertices that correspond to originally neighbouring vertices. Make it so that this matching connects vertices of the same colour if the corresponding entry in $\mathbf{B}$ was $+1$, and vertices of opposing colour if the corresponding entry in $\mathbf{B}$ was -1. This new graph is called a 2-lift or 2-cover of G.

How surprising is it to learn that the eigenvalues of the (standard) adjacency matrix of this new graph are precisely the union of the eigenvalues of $\mathbf{A}$ and those of $\mathbf{B}$?

The reason is not too sophisticated. Let $\overline{\mathbf{A}}$ be the adjacency matrix of this 2-cover. Picture $\overline{\mathbf{A}}$ as an $n \times n$ matrix, whose entries are equal to either

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Multiply $\overline{\mathbf{A}}$ from left and right by a $2n \times 2n$ matrix, block diagonal, with n blocks, all of them equal to

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Each block in $\overline{\mathbf{A}}$ gets transformed (diagonalized!), according to the rule

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Now, rows and columns of $\overline{\mathbf{A}}$ are organized so that red and blue vertices appear intercalated. Reorganize them so that red vertices are the first n vertices, and blue vertices are the last. It follows that $\overline{\mathbf{A}}$ becomes a block matrix: one block equal to $\mathbf{A}$ and the other equal to $\mathbf{B}$.

4.5 Expander graphs

Later in this course we will discuss random walks in graphs in more details. Our goal now is to motivate the discussion in the next sections. A random walk typically converges to a limit distribution, and when this occurs, a random sample from the vertices of the graph can be well approximated by a local procedure of walking randomly with few steps. Thus, one usually wishes that this convergence happens fast. Several fields of research rely on methods to study the speed of this convergence, and on the possibility to generate graphs for which fast mixing takes place.

It is reasonable to believe that any subset of vertices in graph for which convergency occurs fast cannot be too big compared to the number of edges in its boundary. Thus it is convenient to define the edge expansion ratio of graph as

$$h(G) = \min_{S \subseteq V(G) : |S| \leq n/2} \frac{|E(S, \overline{S})|}{|S|},$$

where $E(S, \overline{S})$ denotes the set of edges between S and its complement. One therefore wishes that $h(G)$ is large for a given graph that is a candidate to have the fast mixing process.

A sequence of d-regular graphs G_i of increasing size is a *family of expander graphs* if there $\varepsilon > 0$ so that $h(G_i) \geq \varepsilon$ for all i.

Here is an example: define graphs G_m so that $V(G_m) = \mathbb{Z}_m \times \mathbb{Z}_m$. The neighbours of (x, y) are $(x \pm y, y), (x, y \pm x), (x \pm y + 1, y), (x, y \pm x + 1)$, with all operations $\mod m$. This family is due to Margulis, and is the first explicitly constructed family of expander. The expansion ratio is not easily determined.

The following result associated expansion ratio and the second eigenvalue:

Theorem 4.6 (Dodziuk, Alon-Milman). *Let G be d-regular, with eigenvalues $d = \lambda_1 \geq \cdots \geq \lambda_n$. Then*

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

Thus, it is desirable that $d - \lambda_2$ is large, that is, λ_2 is small, in order for a graph to be an expander. How small can it be?

For a d-regular G with eigenvalues $lambda_1 \geq \cdots \geq \lambda_n$, let

$$\lambda = \max(|\lambda_2|, |\lambda_n|).$$

Theorem 4.7 (Alon-Boppana). *Let G be a d-regular graph on n vertices. Then*

$$\lambda_2 \geq 2\sqrt{d-1} - o_n(1),$$

with the understanding that $o_n(1)$ tends to 0 as $n \to \infty$, for any fixed d.

For a d-regular G with eigenvalues $lambda_1 \geq \cdots \geq \lambda_n$, let

$$\lambda = \max(|\lambda_2|, |\lambda_n|).$$

We will show the weaker version of the Theorem above for λ instead of λ_2.

Before showing the proof of this result, attempt the following easy exercise.

Exercise 4.8. Show that
$$\lambda \geq \sqrt{d}(1 - o_n(1)).$$

Hint: consider $\operatorname{tr} \mathbf{A}^2$.

Sketch of the proof.

- We have $\mathbf{A} = \mathbf{A}(G)$. Note that $\lambda(\mathbf{A}^{2\ell}) = \lambda(\mathbf{A})^{2\ell}$.
- Take two vertices a and b at distance D, and consider the vector $\mathbf{v} = \mathbf{e}_a - \mathbf{e}_b$. This vector is orthogonal to the all 1s vector, thus its Rayleigh quotient in $\mathbf{A}^{2\ell}$ lower bounds $\lambda^{2\ell}$. Assume $\ell < D/2$, thus $(\mathbf{A}^{2\ell})_{a,b} = 0$, and we have
$$\lambda^{2\ell} \geq \frac{\mathbf{v}^\mathsf{T} \mathbf{A}^{2\ell} \mathbf{v}}{2} = \frac{(\mathbf{A}^{2\ell})_{a,a} + (\mathbf{A}^{2\ell})_{a,b}}{2}.$$
- Let us lower bound $(\mathbf{A}^{2\ell})_{a,a}$. Picture a rooted tree at a, with d neighbours, each corresponding to the neighbours of a in G. To each, give them $d-1$ disjoint neighbours, each corresponding to their neighbours in G. Repeat. Note vertices of G will appear multiples times in this tree, and that is fine. Make this tree deep enough, say of depth larger than ℓ. Recall from Exercise 3.40 that this tree has largest eigenvalue smaller than $2\sqrt{d-1}$. Call this tree T.
- Any closed walk at a of length 2ℓ in T represents a closed walk at a of length 2ℓ in G. Thus
$$(\mathbf{A}(G)^{2\ell})_{a,a} \geq (\mathbf{A}(T)^{2\ell})_{a,a}.$$
Note that the tree is the same if you start at b instead.
- Any closed walk at a of length $2k$ in T can be mapped to a sequence of decisions of going either forward or backwards down the tree. It is easy to see that the number of such sequences are counted by the Catalan numbers (check wikipedia), which are equal to $C_\ell = \binom{2\ell}{\ell}/(\ell+1)$. Each step forward had at least $d-1$ options, therefore
$$(\mathbf{A}(T)^{2\ell})_{a,a} \geq (d-1)^\ell \binom{2\ell}{\ell} \frac{1}{\ell+1}.$$
- Gathering up:
$$\lambda^{2\ell} \geq (d-1)^\ell \binom{2\ell}{\ell} \frac{1}{\ell+1}.$$
Taking the 2ℓ-th root on both sides, and assuming that D and ℓ are largest possible, the result (eventually) follows.

□

4.6 Ramanujan graphs

A d-regular graph is a *Ramanujan graph* if $\lambda_2(G) \leq 2\sqrt{d-1}$. These are the best possible graphs you could find which guarantee a good expansion ratio. Infinite families of Ramanujan graphs therefore well sought after.

Lubotzky-Phillips-Sarnak (and independently Margulis) showed that arbitrarily large d-regular Ramanujan graphs exist when $d-1$ is prime (result later extender to prime powers).

For bipartite graphs, recall that the eigenvalues are symmetric about the origin. Thus an adjustment takes place in the definition of Ramanujan bipartite graphs: we say that a d-regular graph is a *bipartite Ramanujan graph* if all eigenvalue different than d and $-d$ lie within the interval $[-2\sqrt{d-1}, 2\sqrt{d-1}]$.

The goal of the rest of this section is to show the impressive result by Marcus, Spielman and Srivastava that infinite families of bipartite regular Ramanujan graphs exist for every fixed d.

So let us briefly review the strategy:

- Consider the matching polynomial of a graph G. It has real roots, and, moreover, if G has maximum degree d, then all roots have absolute value at most $2\sqrt{d-1}$.
- The average of the characteristic polynomials of the signed adjacency matrices of G is equal to $\mu_G(x)$.
- Start with G the complete bipartite graph, d-regular, and consider all possible signings. To each, a 2-cover is associated to the graph. The eigenvalues of the 2-covers are $d^{(1)}$,$0^{(2d-2)}$,$-d^{(1)}$, and the eigenvalues of the signed matrix.
- The average of these characteristic polynomials is μ, whose largest root is $2\sqrt{d-1}$.
- If we can show that at least one of them has largest root at most $2\sqrt{d-1}$, we choose this graph and repeat the procedure.
- The infinite family obtained this way is going to be an infinite family of bipartite Ramanujan graphs of degree d.

4.7 Interlacing families

Let $g(x) = \prod_{i=1}^{n-1}(x-\alpha_i)$, and $f(x) = \prod_{i=1}^{n}(x-\beta_i)$, with $\alpha_i, \beta_i \in \mathbb{R}$. We say that g *interlaces* f if

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

Real-rooted polynomials $f_1, \cdots, f_k$ of the same degree have a *common interlacer* if there is g so that g interlaces each f_i.

If $\beta_{i;j}$ denotes the jth smallest root of f_i, then it is easy to see that polynomials $f_1, \cdots, f_k$ have a common interlacer if and only if there are numbers $\alpha_0, \cdots, \alpha_n$ with $\beta_{i;j} \in [\alpha_{j-1}, \alpha_j]$ for all j. These numbers (except for α_0 and α_n will be the roots of a feasible interlacer g).

Lemma 4.9. *Let $f_1, \cdots, f_k$ be real-rooted polynomials of the same degree, all with positive leading coefficients. Define their sum*

$$f_\emptyset = \sum_{i=1}^{k} f_i.$$

If the polynomials f_i have a common interlacer, then there is an index j so that the largest root of f_j is upper bounded by the largest root of $f_\emptyset$.

Proof. Assume the common interlacer exists, let α be its largest root. It follows that for all i, $f_i(\alpha) \leq 0$, because the largest root of them all is at least α, and they are positive for larger values. Thus $f_\emptyset(\alpha) \leq 0$, and thus its largest root, say β, is at least α. As

$$0 = f_\emptyset(\beta) = \sum_{i=1}^{k} f_i(\beta),$$

it must be that for some j, $f_j(\beta) \geq 0$. Thus the largest root of f_j is between α and β. $\square$

Exercise 4.10. With the same assumptions (well, you may assume the polynomials have simple roots only), prove that $f_\emptyset$ is real-rooted. Also, show that there is an index j so that the mth largest root of f_j is upper bounded by the mth largest root of $f_\emptyset$.

We now want to extend the idea of a set of polynomials containing a common interlacer to the idea of a family of polynomials so that various subfamilies also have a common interlacer.

Let $S_1, \cdots, S_m$ be finite sets, and assume $\mathbf{s}_k$ denotes an element in $S = S_1 \times \cdots \times S_k$ (that is, $\mathbf{s}_k$ denotes a k-tuple of elements, with the jth element belonging to S_j, for $j = 1, ..., k$). For each such $\mathbf{s}_m$, assume polynomial $f_{\mathbf{s}_m}$ is defined, and it is real-rooted of degree n, with positive leading coefficient. For $k < m$, we define $f_{\mathbf{s}_k}$ as follows:

$$f_{\mathbf{s}_k} = \sum_{\mathbf{t} \,\in\, S_{k+1} \times \cdots \times S_m} f_{\mathbf{s}_k, \mathbf{t}}.$$

(Here we understand $\mathbf{s}_k, \mathbf{t}$ as the m-tuple obtained from concatenating the k elements from $\mathbf{s}_k$ to the $m-k$ elements from $\mathbf{t}$.)

We also interpret $f_{\mathbf{s}_0}$ as the sum of all polynomials in the family, or simply $f_\emptyset$. We say that the family $\{f_{\mathbf{s}_m}\}_{\mathbf{s}_m \in S}$ form an *interlacing family* if, for all $k \leq m-1$, and all $\mathbf{s}_k$, we have that

$$\{f_{\mathbf{s}_k, t}\}_{t \in S_{k+1}} \quad \text{have a common interlacer.}$$

Theorem 4.11. *Let $S_1, ..., S_m$ be finite sets, and $\{f_{\mathbf{s}_m}\}$ an interlacing family. Then there is one $\mathbf{s}_m \in S_1 \times \cdots \times S_m$ so that the largest root of $f_{\mathbf{s}_m}$ is at most the largest root of $f_\emptyset$.*

Proof. This is an immediate iterative application of Lemma 4.9. The family $\{f_t\}_{t \in S_1}$ has a common interlacer and sum to $f_\emptyset$. So one applies the lemma. Fix that $s_1 \in S_1$ so that the largest root is at most that of $f_\emptyset$. Now consider the family $\{f_{s_1,t}\}_{t \in S_2}$, which sum to f_{s_1}. Apply the lemma. Repeat. In the end, there will be a sequence $((s_j))$, with $s_j \in S_j$, so that the largest root of $f_{s_1,...,s_j}$ is at most the largest root of $f_{s_1,...,s_{j-1}}$, for all j. $\square$

Assume the graph G contains m edges, and define sets $S_j = \{+1, -1\}$, for $j = 1, \cdots, m$. For any $\mathbf{s}_m \in S_1 \times \cdots \times S_m$, we will define $f_{\mathbf{s}_m}$ as the characteristic polynomial of the signed adjacency matrix of G whose kth edge has received sign $\mathbf{s}_m(k)$. Our goal is to prove that this family is an interlacing family, and therefore that there is an $f_{\mathbf{s}_m}$ whose largest root is at most that of $f_\emptyset$, which in this case is precisely a multiple of $\mu_G(x)$, whose largest root is at most $2\sqrt{d-1}$.

In order to achieve this, we need to show that several common interlacers exist. The existence of these interlacers will be a consequence of the following lemma.

Lemma 4.12. *Let $f_1, ..., f_k$ be polynomials of the same degree with positive leading coefficient. They admit a common interlacer if and only if, for all coefficients $\lambda_i \geq 0$ that sum to 1, we have that $\sum_{i=1}^{k} \lambda_i f_i$ is real rooted.*

Proof. Assume for simplicity first that all roots of the f_i are distinct. Then one direction is easy and follows straight from Exercise 4.10. In general, consider small perturbations of the polynomials to have only polynomials of simple roots, and apply the theorem to these.

For the other direction, assume first that $k = 2$ and also that the polynomials have distinct roots. So, for all t, $0 \leq t \leq 1$,

$$g_t(x) = tf_1(x) + (1-t)f_2(x)$$

is real rooted. Consider the intervals i_k, the vary from the kth largest root of $g_0(x)$ to the kth largest of $g_1(x)$. These are disjoint (why?). The result then follows.

If they have shared roots, then factor them out, apply the argument above, and add them back in. Be careful if the shared roots fall into one of the i_k intervals.

It is not too difficult to extend the result for k polynomials, but we leave it out of these notes. □

As a consequence of this result above, real-rootedness is somehow equivalent to the existence of common interlacers.

Assume for now the following result.

Theorem 4.13. *Let $p_1, \cdots, p_m$ be numbers from 0 to 1. Then the following polynomial is real rooted:*

$$\sum_{\mathbf{s}_m \in \{1,-1\}^m} \left(\prod_{i:s_i=1} p_i \right) \left(\prod_{i:s_i=-1} (1-p_i) \right) f_{\mathbf{s}_m}(x).$$

Theorem 4.14. *The polynomials $f_{\mathbf{s}_m}$ form an interlacing family.*

Proof. We fix $\mathbf{s}_k \in \{1,-1\}^k$, and we need to show that $f_{\mathbf{s}_k,1}$ and $f_{\mathbf{s}_k,-1}$ have a common interlacer. To see that, it is enough to verify that

$$tf_{\mathbf{s}_k,1} + (1-t)f_{\mathbf{s}_k,-1}$$

is real-rooted for all $t \in [0,1]$.

This follows immediately from Theorem 4.13 by

- $p_i = \frac{1+s_i}{2}$ for i between 1 and k,
- $p_{k+1} = t$,
- $p_i = 1/2$, for $i \geq k+2$.

Thus, in the sum

$$\sum_{\mathbf{w}_m \in \{1,-1\}^m} \left(\prod_{i:w_i=1} p_i \right) \left(\prod_{i:w_i=-1} (1-p_i) \right) f_{\mathbf{w}_m}(x),$$

all terms whose prefix is different than $\mathbf{s}_k$ will be zero, and the remaining terms for $i \geq k+2$ only contribute with a constant. So this sum is a multiple of

$$t f_{\mathbf{s}_k,1} + (1-t) f_{\mathbf{s}_k,-1},$$

as we wanted. □

The result that we wanted to show now follows immediately.

- Start with $K_{d,d}$, the complete bipartite d-regular graph. Its eigenvalues are $d^{(1)}, 0^{(2d-2)}, -d^{(1)}$.
- Amongst all its signed adjacency matrices, pick that whose largest eigenvalue does not exceed that of $\mu_{K_{d,d}}$, which is $2\sqrt{d-1}$. It exists from a combination of Theorem 4.14, Theorem 4.11, Theorem 3.45, Exercise 3.42.
- Consider the covering graph determined by this signed adjacency matrix. All of its eigenvalues different from $\pm d$ lie within $\pm 2\sqrt{d-1}$, from Subsection 4.4.
- Repeat indefinitely.
- These will be a family of expanders, as $h(G) \geq (d - 2\sqrt{d-1})/2$ for all of them.
- Note that, implicitly here, we are using the fact that all these graphs are bipartite.

4.8 Real roots

Our goal in this section is to prove Theorem 4.13.

Theorem. *Let $p_1, \cdots, p_m$ be numbers from 0 to 1. Then the following polynomial is real rooted:*

$$\sum_{\mathbf{s}_m \in \{1,-1\}^m} \left(\prod_{i : s_i = 1} p_i \right) \left(\prod_{i : s_i = -1} (1-p_i) \right) f_{\mathbf{s}_m}(x).$$

A multivariate polynomial $f \in \mathbb{R}[z_1, ..., z_n]$ is called *real stable* if either $f \equiv 0$ or it is never zero whenever the imaginary parts of all z_i are positive. Note that a univariate real stable polynomial contains only real roots. We will expose some that the polynomial in the Theorem statement is real stable.

Theory of stable polynomials is rich and vast, and I do not intend to provide here an informative survey. I will however list some useful results, and guide you through an understanding of how they can be applied to our task.

Hurwitz's theorem is an important result about what happens on non-vanishing functions converge to another function. It basically implies in our context that a polynomial obtained as the limit of a convergent sequence of stable polynomials is itself stable.

Exercise 4.15. Using this fact, argue why, if $f(x_1, ..., x_k)$ is stable, then for all $c \in \mathbb{R}$, we also have $f(x_1, ..., x_{k-1}, c)$ stable.

A very useful fact is the following lemma.

Lemma 4.16. *Let $\mathbf{A}_1, ..., \mathbf{A}_m$ be positive semidefinite matrices, and $\mathbf{B}$ Hermitian. Then*

$$f(x_1, ..., x_k) = \det(x_1\mathbf{A}_1 + ... + x_k\mathbf{A}_k + \mathbf{B})$$

is a real stable polynomial.

Sketch of proof. It is easy to see that f is a real polynomial. It follows from Hurwitz theorem that it is enough to show this result for when $\mathbf{A}_i$ is positive definite for all i, meaning, these are Hermitian matrices with positive eigenvalues.

Pick $\mathbf{z} \in \mathbb{C}^k$, write it as $\mathbf{z} = \mathbf{a} + \mathrm{i}\mathbf{b}$, and assume $\mathbf{b} > \mathbf{0}$. Let $\mathbf{Q} = \sum b_i\mathbf{A}_i$, and $\mathbf{H} = \sum a_i\mathbf{A}_i + \mathbf{B}$. It follows that $\mathbf{Q}$ is positive definite, and $\mathbf{H}$ is Hermitian, and that

$$f(\mathbf{z}) = \det(\mathbf{Q})\det(\mathrm{i}\mathbf{I} + \mathbf{Q}^{-1/2}\mathbf{H}\mathbf{Q}^{-1/2}.$$

As $\det(\mathbf{Q}) \neq 0$, it follows that $f(\mathbf{z}) = 0$ if and only if $-\mathrm{i}$ is an eigenvalue of the Hermitian matrix $\mathbf{Q}^{-1/2}\mathbf{H}\mathbf{Q}^{-1/2}$, which is not true. $\square$

It is possible to show that differentiating with respect to one variable is is a linear map that preserves stability, and, with some more work, one obtains the following result

Lemma 4.17. *Let $p, q \in \mathbb{R}_+$, and consider variables x and y. Then if $f(x, y)$ is real stable, then so is*

$$(1 + p\partial_x + q\partial_y)f = f + p\partial_x f + q\partial_y f.$$

Exercise 4.18. Let $\mathbf{A}$ be invertible, and $\mathbf{u}$ and $\mathbf{v}$ vectors. The goal of this exercise is to show that

$$\begin{aligned}(x, y \to 0)(1 + p\partial_x + (1-p)\partial_y)\det(\mathbf{A} + x\mathbf{u}\mathbf{u}^\mathsf{T} + y\mathbf{v}\mathbf{v}^\mathsf{T}) = \\ = p\det(\mathbf{A} + \mathbf{u}\mathbf{u}^\mathsf{T}) + (1-p)\det(\mathbf{A} + \mathbf{v}\mathbf{v}^\mathsf{T}).\end{aligned}$$

($(x, y \to 0)$ means that after applying $(1 + p\partial_x + (1-p)\partial_y)$, set variables to 0).

(a) Recall Exercise 3.11. Show that

$$\partial_t \det(\mathbf{A} + t\mathbf{u}\mathbf{u}^\mathsf{T}) = \det(\mathbf{A})(\mathbf{u}^\mathsf{T}\mathbf{A}^{-1}\mathbf{u}).$$

(b) Show that the expression on the left hand side of the equality is equal to

$$\det(\mathbf{A})(1 + p(\mathbf{u}^\mathsf{T}\mathbf{A}^{-1}\mathbf{u}) + (1-p)(\mathbf{v}^\mathsf{T}\mathbf{A}^{-1}\mathbf{v})).$$

(c) Use Exercise 3.11 again.

Theorem 4.19. *Let $\mathbf{u}_i$, $\mathbf{v}_j$, $i, j = 1, ..., m$ be vectors. Let $p_1, ..., p_m$ be real numbers between 0 and 1, and let $\mathbf{D}$ be positive semidefinite. Then*

$$\begin{aligned}P(x) = \sum_{\mathbf{s}_m \in \{1,-1\}^m} \left(\prod_{i:s_i=1} p_i\right)\left(\prod_{i:s_i=-1}(1-p_i)\right) \cdot \\ \cdot \det\left(x\mathbf{I} + \mathbf{D} + \sum_{i:s_i=1}\mathbf{u}_i\mathbf{u}_i^\mathsf{T} + \sum_{i:s_i=-1}\mathbf{v}_i\mathbf{v}_i^\mathsf{T}\right)\end{aligned}$$

is real rooted.

Sketch of proof. First define the multivariate polynomial

$$q(x, s_1, ..., s_m, t_1, ..., t_m) = \det\left(x\mathbf{I} + \mathbf{D} + \sum_{i:s_i=1} s_i \mathbf{u}_i \mathbf{u}_i^\mathsf{T} + \sum_{i:s_i=-1} t_i \mathbf{v}_i \mathbf{v}_i^\mathsf{T} \right).$$

It is real stable from Lemma 4.16. Then, successively apply the operator from Exercise 4.18 taking each pair of variables (s_i, t_i). The resulting polynomial, upon setting s_i and t_i to 0, will be exactly the polynomial of the statement. $\square$

The only thing left now is to show that the polynomial in the statement of Theorem 4.13 is of the form of $P(x)$ as above.

Proof of Theorem 4.13. We will show that

$$\sum_{\mathbf{s}_m \in \{1,-1\}^m} \left(\prod_{i:s_i=1} p_i \right) \left(\prod_{i:s_i=-1} (1 - p_i) \right) \det(x\mathbf{I} + \Delta(G)\mathbf{I} - \mathbf{A}_{\mathbf{s}_m})$$

is real rooted. Let $\mathbf{D}$ be a diagonal matrix of degrees. Note that

$$\mathbf{D} - \mathbf{A}_\mathbf{s} = \sum_{ab \in E(G)} (\mathbf{e}_a \pm \mathbf{e}_b)(\mathbf{e}_a \pm \mathbf{e}_b)^\mathsf{T},$$

where the sign is $+$ if $\mathbf{s}(ab) = 1$, and $-$ if $\mathbf{s}(ab) = -1$. Thus $\Delta\mathbf{I} - \mathbf{A}_\mathbf{s}$ is of the form $\mathbf{E} + \sum_{i:s_i=1} \mathbf{u}_i \mathbf{u}_i^\mathsf{T} + \sum_{i:s_i=-1} \mathbf{v}_i \mathbf{v}_i^\mathsf{T}$ with $\mathbf{E}$ positive semidefinite, and the result follows. $\square$