MA431 Spectral Graph Theory: Lecture 9

Ahmad Abdi   Neil Olver*

Abstract

We introduce matching polynomials, and prove a striking theorem of Hielman and Lieb about the roots of such polynomials. We then show a result of Marcus, Spielman and Srivastava that shows that shows the existence of large bipartite Ramanujan graphs of arbitrary degree.

21 The matching polynomial

Definition 21.1. Let $G = (V, E)$ be a graph on n vertices, and let m_j denote the number of matchings in G with exactly j edges (with m_0 define to be 1). Then the *matching polynomial* of G, denoted μ_G, is given by

$$\mu_G(x) = \sum_{j \geq 0} m_j (-1)^j x^{n-2j}.$$

So we have just constructed a polynomial whose coefficients count matchings of different sizes. The reason we don't define the matching polynomial to be simply $\hat{\mu}_G(x) = \sum_{i \geq 0} m_i x^i$ will become apparent a bit later, but notice that $\mu_G(x) = x^n \hat{\mu}_G(-x^{-2})$, so properties of μ_G can be mapped to the more "obvious" $\hat{\mu}_G$.

The following basic identities are easily verified:

Lemma 21.2.

For any graph $G = (V, E)$, and any $u \in V$,

$$\mu_G(x) = x\mu_{G \setminus u}(x) - \sum_{w: vw \in E} \mu_{G \setminus \{v,w\}}(x). \tag{1}$$

For any graphs $G_1, \ldots, G_k$, if we take G to be their disjoint union, then

$$\mu_G(x) = \mu_{G_1}(x)\mu_{G_2}(x) \cdots \mu_{G_k}(x). \tag{2}$$

Proof. Exercise. □

The main fact about matching polynomials we will prove is the following striking result.

Theorem 21.3 (Heilman and Lieb)**.** *For any graph G, μ_G is real-rooted. Further, if G has maximum degree d, then all roots of μ_G lie in $[-2\sqrt{d-1}, 2\sqrt{d-1}]$.*

It seems somewhat miraculous that a polynomial defined in this "combinatorial" way should magically have real roots. The proof here we will give here is due to Godsil, which at least partially demystifies things. It relates the matching polynomial of G to the matching polynomial of a kind of cover of G called the path tree. Let's begin by understanding the matching polynomial of a tree. Here and in what follows, we will use $\chi_G(x)$ to denote the characteristic polynomial of G, i.e., $\chi_G(x) = \det(xI - A(G))$.

Theorem 21.4. *If G is a tree, then $\mu_G(x) = \chi_G(x)$.*

Proof. Let A be the adjacency matrix of $G = (V, E)$, and identify V with $\{1, 2, \ldots, n\}$.

Given a permutation $\pi \in \mathsf{S}_n$, let $\operatorname{sgn}(\pi)$ denote the sign of π (that is, 1 if the number of involutions of π is even, -1 if it is odd). Also let $\operatorname{Fix}(\pi)$ denote the set of fixed points of π: $\operatorname{Fix}(\pi) = \{i : \pi(i) = i\}$; and let $\operatorname{fix}(\pi) = |\operatorname{Fix}(\pi)|$.

We have

$$\begin{aligned}
\chi_G(x) &= \det(xI - A) \\
&= \sum_{\pi \in \mathsf{S}_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} (xI - A)_{i,\pi(i)} \\
&= \sum_{\pi \in \mathsf{S}_n} \operatorname{sgn}(\pi)(-1)^{\operatorname{fix}(\pi)} \prod_{i \notin \operatorname{Fix}(\pi)} \left(-\mathbf{1}_{\{i,\pi(i)\} \in E}\right) \\
&= \sum_{\substack{\pi \in \mathsf{S}_n : \\ \{i,\pi(i)\} \in E \ \forall i \notin \operatorname{Fix}(\pi)}} \operatorname{sgn}(\pi)(-1)^{n-\operatorname{fix}(\pi)} x^{\operatorname{fix}(\pi)} \\
&\overset{(*)}{=} \sum_{\substack{\pi \in \mathsf{S}_n \\ \pi(\pi(i)) = i \ \forall i}} \operatorname{sgn}(\pi)(-1)^{n-\operatorname{fix}(\pi)} x^{\operatorname{fix}(\pi)} \\
&= \sum_{k \geq 0} m_k (-1)^k \cdot (-1)^{2k} x^{n-2k} \\
&= \mu_G(x).
\end{aligned}$$

Regarding step $(*)$: the observation is that π defines a collection of directed cycles (including loops, which form $\operatorname{Fix}(\pi)$; but a tree has no cycles, except for cycles of length 2 obtained by traversing an edge in opposite directions. So the cycle decomposition of any π contributing to the sum consists only of loops and cycles of length 2, involving an edge of G. Such permutations can be identified with matchings of G in the obvious way, and the size of the matching will be the number of cycles of length 2. $\square$

This already shows the Heilman-Lieb theorem if G is a tree: we know that the characteristic polynomial is real-rooted (its roots are the eigenvalues of $A(G)$), and that all roots lie in $[-2\sqrt{d-1}, 2\sqrt{d-1}]$ if the tree has maximum degree d.

To extend to arbitrary graphs, we will define a tree associated with the graph.

Definition 21.5. Let $G = (V, E)$ be a graph, and $u \in V$. The *path tree* of G rooted at u, denoted $T(G, u)$, is defined as follows. The vertex set of $T(G, u)$ is the set of paths in G starting from u. Given two paths P_1, P_2

starting from u, P_1 and P_2 are connected in $T(G,u)$ if P_2 is obtained by adding a single edge to P_1 (or vice versa).

If G is a tree, then $T(G,u)$ is isomorphic to G, for any $u \in V$. If G is not connected, $T(G,u)$ will be fully determined by the component of G containing u.

Theorem 21.6.

1. *For any G, and any path tree $T := T(G,u)$,*

$$\frac{\mu_G(x)}{\mu_{G\setminus u}(x)} = \frac{\mu_T(x)}{\mu_{T\setminus u}(x)}. \tag{3}$$

2. *If G is connected, then μ_G divides μ_T.*

Before we prove this, let us observe that this implies the Hielman-Lieb theorem! If G is connected, then μ_G divides μ_T, where T is any path tree of G. But T is a tree of maximum degree equal to d, the maximum degree of G. Since $\mu_T = \chi_T$, and χ_T is real-rooted with all roots in $[-2\sqrt{d-1}, 2\sqrt{d-1}]$, the same is true for μ_G. If G is not connected, we can simply apply the above argument to each component of G separately, and then invoke (2).

Proof. **(1) ⇒ (2).** Assume the second claim holds for smaller graphs by induction (it certainly holds for trees and hence for all graphs on at most 2 nodes). Let $H_1, \ldots, H_\ell$ be the components of $G \setminus u$, and let v_i be any node in H_i that is adjacent to u in G. Then μ_{H_i} divides $\mu_{T(H_i,v_i)}$ by induction. Notice that $T(H_i, v_i)$ is a component of $T \setminus u$ for each i (any path in H_i starting from v_i can be extended with the edge uv_i to a path from u in T). It follows then by (2) that $\prod_{i=1}^{\ell} \mu_{T(H_i,v_i)}(x)$ divides $\mu_{T\setminus u}(x)$, whereas $\mu_{G\setminus u}(x) = \prod_{i=1}^{\ell} \mu_H(x)$. Thus $\mu_{G\setminus u}$ divides $\mu_{T\setminus u}$. Finally, we apply **(1)**.

(1). We will write $\mu(G,x)$ rather than $\mu_G(x)$ in what follows, for notational convenience.

Assume by induction that the claim holds for all smaller graphs (it certainly holds for all graphs on at most two nodes, since the claim is trivial for trees). Let $H = G \setminus u$, and let N be the set of neighbours of u in G. We will abuse notation slightly and identify any $v \in N$ with the corresponding node uv in $V(T)$. Let $T_v = T(H,v)$ for all $v \in N$.

We have

$$
\begin{aligned}
\frac{\mu(G,x)}{\mu(H,x)} &= \frac{x\mu(H,x) - \sum_{v\in N}\mu(H\setminus v,x)}{\mu(H,x)} && \text{by (1)}\\
&= x - \sum_{v\in N}\frac{\mu(H\setminus v,x)}{\mu(H,x)}\\
&= x - \sum_{v\in N}\frac{\mu(T_v\setminus v,x)}{\mu(T_v,x)} && \text{by induction}\\
&\overset{(*)}{=} x - \sum_{v\in N}\frac{\mu(T\setminus\{u,v\},x)}{\mu(T\setminus u,x)}\\
&= \frac{x\mu(T\setminus u,x) - \sum_{v\in N}\mu(T\setminus\{u,v\},x)}{\mu(T\setminus u,x)}\\
&= \frac{\mu(T,x)}{\mu(T\setminus u,x)} && \text{by (1)}
\end{aligned}
$$

Regarding equality $(*)$, the observation is that T_v is a component of $T\setminus u$, and in fact $T\setminus u$ is the disjoint union of the graphs T_w, for $w\in N$. Similarly $T\setminus\{u,v\}$ is the disjoint union of $T_v\setminus v$ and $\{T_w : w\in N\setminus\{v\}\}$. The equality then follows by recalling (2). □

22 The existence of bipartite Ramanujan graphs

We now present the proof by Marcus, Spielman and Srivastava of the following result.

Theorem 22.1. *For any $d\geq 3$, there are d-regular bipartite Ramanujan graphs of arbitrarily large size.*

Their proof is based on (partially) resolving a conjecture of Bilu and Linial. Their conjecture is centered around the following "lifting" construction.

Definition 22.2. Let $G=(V,E)$ be a graph, and $s\in\{-1,1\}^E$ a *signing* of the edges of G. Then the resulting *2-lift* of G, denoted by G^s, is the following.

- The vertex set of G^s consists of two disjoint copies of V, which we will denote V_1 and V_2. For any $v\in V$, we will use v_1 and v_2 to denote the corresponding copies of v in V_1 and V_2 respectively.
- For any edge $e=\{v,w\}\in E$, there will be two corresponding edges in G^s. If $s(e)=1$, these edges are $\{v_1,w_1\}$ and $\{v_2,w_2\}$; if $s(e)=-1$, they are $\{v_1,w_2\}$ and $\{v_2,w_1\}$.

The *signed adjacency matrix* associated with G and s is the matrix A^s with rows and columns indexed by the nodes of V defined by

$$
A^s_{ij} = \begin{cases} s(ij) & \text{if } ij\in E\\ 0 & \text{otherwise}\end{cases}.
$$

If $s(e)=1$ for all $e\in E(G)$, then G^s is simply two disjoint copies of G.

Lemma 22.3. *For any G with signing s, adjacency matrix A, and signed adjacency matrix A^s, the spectrum of G^s is the union (with multiplicities) of the spectrum of A and the spectrum of A^s.*

Proof. Let $G^+ = (V, E^+)$, where $E^+ = \{e \in E : s(e) = 1\}$, and let A^+ be the adjacency matrix of G^+. Similarly, let $G^- = (V, E^-)$, where $E^- = \{e \in E : s(e) = -1\}$, and let A^- be the adjacency matrix of G^-. Then $A = A^+ + A^-$ and $A^s = A^+ - A^-$. Further, it is easy to see that

$$A(G^s) = \begin{pmatrix} A^+ & A^- \\ A^- & A^+ \end{pmatrix}.$$

Straightforward calculations show that if v is an eigenvector of A with eigenvalue λ, $\begin{pmatrix} v \\ v \end{pmatrix}$ is an eigenvector of $A(G^s)$ with the same eigenvalue; and if w is an eigenvector of A^s with eigenvalue θ, then $\begin{pmatrix} w \\ -w \end{pmatrix}$ is an eigenvector of $A(G^s)$ with the same eigenvalue. It follows that given an orthonormal basis of eigenvalues $\{v_i : i = 1, \ldots, n\}$ and $\{w_i : i = 1, \ldots, n\}$ for A and A^s respectively,

$$\left\{ \begin{pmatrix} v_i \\ v_i \end{pmatrix} : i = 1, \ldots, n \right\} \cup \left\{ \begin{pmatrix} w_i \\ -w_i \end{pmatrix} : i = 1, \ldots, n \right\}$$

is an orthonormal basis of eigenvectors for $A(G^s)$. □

So the spectrum of a 2-lift of G consists of the "old" eigenvalues of G, plus "new" eigenvalues that come from the signed adjacency matrix. This is very convenient.

Conjecture (Bilu-Linial)**.** *For any d-regular graph G, (for any $d \geq 3$), there exists a signing of G such that the signed adjacency matrix A^s has spectral radius at most $2\sqrt{d-1}$.*

Thus, if G is Ramanujan, this conjecture says that there is a signing s of G so that G^s is Ramanujan. Since K^{d+1} is a d-regular Ramanujan graph for any $d \geq 3$, this implies arbitrarily large Ramanujan graphs of any degree. Similarly, if G is bipartite Ramanujan, there is a signing s so that G^s is as well (bipartiteness is preserved by the 2-lifting procedure). Since $K_{d,d}$ is bipartite Ramanujan, we get arbitrarily large bipartite Ramanujan graphs as well.

Here is what Marcus, Spielman and Srivastava proved.

Theorem 22.4. *For any d-regular graph G ($d \geq 3$), there exists a signing s so that the signed adjacency matrix A^s has maximum eigenvalue at most $2\sqrt{d-1}$, as well as a signing s' so that the minimum eigenvalue of $A^{s'}$ is at least $-2\sqrt{d-1}$.*

This is why their result only yields bipartite Ramanujan graphs (or "one-sided" Ramanujan graphs in general, though that doesn't seem to be a standard definition). For bipartite graphs, the symmetry of the spectrum implies that if the largest eigenvalue of A^s is small, the smallest eigenvalue is small in absolute value as well.

We will now see the connection to the matching polynomial; this is a result by Godsil and Gutman (which precedes the Bilu-Linial conjecture by quite a while).

Theorem 22.5. *Let G be any graph, and let s be a uniformly random signing of G, amongst all possible signings. Let A^s denote the signed adjacency matrix of G and s. Then*

$$\mathbb{E}[\chi_{A^s}(x)] = \mu_{G^s}(x).$$

Proof. Exercise. (This follows similar lines to the proof that $\mu_G = \chi_G$ for G a tree. It is useful to understand how the uniform averaging of signs "kills off" all cycles of length more than 2.) □

In what follows, we'll use the notation $\mathbb{E}_s[\cdots]$ to indicate that s is being chosen randomly.

Now: by the Heilman-Lieb theorem, we know that μ_G is real-rooted and that its roots are contained in $[-2\sqrt{d-1}, 2\sqrt{d+1}]$. So in some sense, we are good "on average": the roots of the averaged characteristic polynomial of signed adjacency matrices is good. But we want a single signing, and unfortunately it's just not true that the largest root of a polynomial behaves well under averaging. Let $\lambda_{\min}(f)$ and $\lambda_{\max}(f)$ denote the smallest and largest root of a real-rooted polynomial f. Then given two real-rooted polynomials f and g, even of the same degree and both with leading coefficient 1, it need not be true that $\frac{1}{2}(f+g)$ is even real-rooted. And if it is real-rooted, it need not be the case that $\min\{\lambda_{\max}(f), \lambda_{\max}(g)\} \leq \lambda_{\max}(\frac{1}{2}(f+g))$.

Fortunately, there are conditions where this does hold, based on interlacing. Before we go into that, we'll first massage our goal a bit.

First, it will be slightly more convenient to work with the "signed Laplacian". Given a signing σ and signed adjacency matrix A^σ, we define the signed Laplacian to be $L^\sigma = dI - A^\sigma$. It's not hard to see (check it) that $\lambda_{\max}(\chi_{L^\sigma}) = d - \lambda_{\min}(\chi_{A^\sigma})$ for any signing σ, and that $\lambda_{\max}(\mathbb{E}_s[\chi_{L^s}]) = d - \lambda_{\min}(\mathbb{E}_s[\chi_{A^s}])$. We make the definition

$$p_\sigma(x) := \chi_{L^s}(x);$$

so it suffices to find a signing σ for which $\lambda_{\max}(p_\sigma) \leq \lambda_{\max}(\mathbb{E}_s[p_s])$.

Next, instead of trying to go directly from $\mathbb{E}_s[p_s]$ to a fixed signing, we will go in a gradual way. Identify E with $\{1, 2, \ldots, m\}$, i.e., give a numbering of the edges. So we can view a signing as a vector in $\{\pm 1\}^m$. We now consider fixing the signs of *some* edges, and taking a uniformly random choice over the remaining edges. In particular, for $\sigma \in \{\pm 1\}^r$, for $0 \leq r \leq m$, define

$$p_\sigma(x) := \mathbb{E}_\rho[p_{\sigma,\rho}(x)];$$

here, ρ is uniform over $\{\pm 1\}^{m-r}$, and the subscript $p_{\sigma,\rho}$ refers to the complete signing obtained by concatenating σ and ρ. To show that there exists a complete signing $\sigma \in \{\pm 1\}^m$ with $\lambda_{\max}(p_\sigma) \leq \lambda_{\max}(p_\emptyset)$, it suffices to show that for any partial signing $\sigma \in \{\pm 1\}^r$, $\min\{\lambda_{\max}(p_{\sigma,1}), \lambda_{\max}(p_{\sigma,-1})\} \leq \lambda_{\max}(p_\sigma)$. Observe further that $p_\sigma = \frac{1}{2}(p_{\sigma,-1} + p_{\sigma,1})$. We will show (in an inductive fashion) that all these polynomials p_σ are real-rooted (at the moment, we only know this for $\sigma \in \{\pm 1\}^m$, and $\sigma = \emptyset$), and that the roots of $p_{\sigma,-1}$ and $p_{\sigma,1}$ have certain interlacing properties that allow us to draw the desired conclusion. So we first (re)-introduce some definitions and results for interlacing polynomials.

Definition 22.6. The sequence $\theta_{n-1} \leq \theta_{n-2} \leq \cdots \leq \theta_1$ *interlaces* the sequence $\mu_n \leq \mu_{n-1} \leq \cdots \leq mu_1$ if

$$\mu_n \leq \theta_{n-1} \leq \mu_{n-1} \leq \cdots \leq \mu_2 \leq \theta_1 \leq \mu_1.$$

The sequence $\theta_n \leq \theta_{n-1} \leq \cdots \leq \theta_1$ *left-interlaces* the sequence $\mu_n \leq \mu_{n-1} \leq \cdots \leq mu_1$ if

$$\theta_n \leq \mu_n \leq \theta_{n-1} \leq \mu_{n-1} \leq \cdots \leq \mu_2 \leq \theta_1 \leq \mu_1.$$

A real-rooted polynomial g interlaces (respectively left-interlaces) a real-rooted polynomial if the ordered roots of g interlace (respectively left-interlace) the ordered roots of f.

Definition 22.7. Let f, g be real-rooted polynomials of the same degree, both with positive leading coefficient. We say that f and g have a *common interlacing* if there is a real-rooted polynomial r which either interlaces both f and g, or left-interlaces both f and g.

Lemma 22.8. *If f and g are real-rooted polynomials with positive leading coefficient that have a common interlacing, then for all $t \geq 0$, $f + tg$ is real-rooted, and further,*

$$\lambda_{\max}(f + tg) \geq \min\{\lambda_{\max}(f), \lambda_{\max}(g)\}.$$

Proof. Exercise (see also Section 3). □

Lemma 22.9. *Let f, g be real-rooted polynomials with positive leading coefficient, with f having degree n and g having degree $n - 1$. Then g interlaces f if and only if $f - tg$ is real-rooted for all $t \in \mathbb{R}$. Further, if g does interlace f, then f left-interlaces $f - tg$ for all $t > 0$.*

Proof. Exercise. □

Given Lemma 22.8, it suffices to show that for any $r < m$ and partial signing $\sigma \in \{\pm 1\}^r$, $p_{\sigma,1}$ and $p_{\sigma,-1}$ are real-rooted and have a common interlacing.

In order to prove this, we will prove something slightly more general for inductive purposes. We will need only the following information about L^σ for $\sigma \in \{\pm 1\}^m$: there exist vectors $v_{i,s}$, for $i \in [m]$ and $s \in \{\pm 1\}$, such that

$$L^\sigma = \sum_{i=1}^{m} v_{i,\sigma(i)} v_{i,\sigma(i)}^\top.$$

This is obtained by setting $v_{i,s} = \mathrm{e}_w - s\mathrm{e}_z$, if i represents the edge $\{w, z\}$. It is easy to check that this agrees with the definition $L^\sigma = dI - A^\sigma$. In what follows, we will use $\chi(M, x)$ interchangeably with $\chi_M(x)$.

Theorem 22.10. *Let M be any symmetric $n \times n$ matrix, and let $w_{i,s}$ be any vector in $\mathbb{R}^n$, for $i \in [k]$, $s \in \{\pm 1\}$. Then*

$$\sum_{\rho \in \{\pm 1\}^k} \chi\left(M + \sum_{i=1}^{k} w_{i,\rho(i)} w_{i,\rho(i)}^\top, x\right)$$

is real-rooted, and for any $s \in \{\pm 1\}$,

$$\sum_{\rho \in \{\pm 1\}^{k-1}} \chi\left(M + \sum_{i=1}^{k-1} w_{i,\rho(i)} w_{i,\rho(i)}^{\top}, x\right)$$

left-interlaces

$$\sum_{\rho \in \{\pm 1\}^{k-1}} \chi\left(M + \sum_{i=1}^{k-1} w_{i,\rho(i)} w_{i,\rho(i)}^{\top} + w_{k,s} w_{k,s}^{\top}, x\right).$$

Observe that this indeed shows that p_σ is real-rooted for any partial signing, and that $p_{\sigma,1}$ and $p_{\sigma,-1}$ have a common interlacing, completing the proof of Theorem 22.4.

Proof. We begin with a claim about the impact of a rank-1 update on the characteristic polynomial of a matrix.

Claim 1. Let N be any symmetric $n \times n$ matrix, and $u \in \mathbb{R}^n$. Let

$$f_t(x) = \chi(N + tuu^{\top}, x).$$

for any t. Then there exists a degree $n-1$ polynomial g with positive leading coefficient such that

$$f_t(x) = \chi(N, x) - tg(x).$$

Proof. Since χ is invariant under orthogonal transformations, it suffices to show the claim for the case $u = \mathrm{e}_1$. Let $N^{(1)}$ denote the principal submatrix of N obtained by discarding the first row and column. Then

$$\begin{aligned}\chi(N + t\mathrm{e}_1\mathrm{e}_1^{\top}, x) &= \det(xI - N - t\mathrm{e}_1\mathrm{e}_1^{\top}) \\ &= \det(xI - N) - t\det(xI - N^{(1)}) \\ &= \chi(N, x) - t\chi(N^{(1)}, x).\end{aligned}$$

We used here that $\det X$ is linear as a function of the rows of X, and also Laplace expansion. □

Define, for any $s \in \{\pm 1\}$, $t \in \mathbb{R}$,

$$f_t^s(x) = \sum_{\rho \in \{\pm 1\}^{k-1}} \chi\left(M + \sum_{i=1}^{k-1} w_{i,\rho(i)} w_{i,\rho(i)}^{\top} + w_{k,s} w_{k,s}^{\top}, x\right).$$

Note that $f_0^{+1} = f_0^{-1}$, so we will write this as simply f_0. So we wish to show that $f_1^{+1} + f_1^{-1}$ is real-rooted, and that f_0 left-interlaces both f_1^{+1} and f_1^{-1}.

We proceed by induction on k. Observe that for any fixed t and sign s, f_t^s is real-rooted by induction, applied with the choice $M' = M + tw_{k,s}w_{k,s}^{\top}$ (or in the base case $k = 1$, because f_t^s is simply a characteristic polynomial). Now apply the above Claim to each term of f_t^s and sum, to deduce that

$$f_t^s(x) = f_0(x) - tg^s(x),$$

for some polynomial g^s which, being a sum of polynomials with positive leading coefficients, has positive leading coefficient as well. Since f_t^s is real-rooted for all t, we can deduce by Lemma 22.9 that f_0 left-interlaces f_t^s, for all $t > 0$. In particular, f_0 left-interlaces f_1^{+1} and f_1^{-1}. This shows the second part of the theorem, and the first part (that $f_t^{+1} + f_1^{-1}$ is real-rooted) follows immediately by Lemma 22.8. □