Spectral Graph Theory Lecture 24

Ramanujan Graphs of Every Size

Daniel A. Spielman December 2, 2015

Disclaimer

These notes are not necessarily an accurate representation of what happened in class. The notes written before class say what I think I should say. I sometimes edit the notes after class to make them way what I wish I had said.

There may be small mistakes, so I recommend that you check any mathematically precise statement before using it in your own work.

These notes were last revised on December 2, 2015.

24.1 Overview

We will *mostly* prove that there are Ramanujan graphs of every number of vertices and degree. The material in today's lecture comes from [MSS15b] and [MSS15a]. In those papers, we prove that for every even n and degree $d < n$ there is a **bipartite** Ramanujan graph of degree d on n vertices. A bipartite Ramanujan graph of degree d is an approximation of a complete bipartite graph. It's adjacency matrix thus has eigenvalues d and $-d$, and all other eigenvalues bounded in absolute value by $2\sqrt{d-1}$.

The difference between this result and that which we prove today is that we will show that for every $d < n$ there is a d-regular (multi) graph in whose second adjacency matrix eigenvalue is at most $2\sqrt{d-1}$. This bound is sufficient for many applications of expanders, but not all. We will not control the magnitude of the negative eigenvalues. The reason will simply be for simplicity: the proofs to bound the negative eigenvalues would take more lectures.

Next week we will see a different technique that won't produce a multigraph and that will produce a bipartite Ramanujan graph.

24.2 The Approach

We will consider the sum of d random perfect matchings on n vertices. This produces a d-regular graph that might be a multigraph. Friedman [Fri08] proves that such a graph is probably very close to being Ramanujan if n is big enough relative to d. In particular, he proves that for all d and $\epsilon > 0$ there is an n_0 so that for all $n > n_0$, such a graph will probably have all eigenvalues other than μ_1 bounded in absolute value by $2\sqrt{d-1}+\epsilon$. We remove the asymptotics and the ϵ, but merely prove the existence of one such graph. We do not estimate the probability with which such a graph is Ramanujan. But, it is predicted to be a constant [?].

The fundamental difference between our technique and that of Friedman is that Friedman bounds the moments of the distribution of the eigenvalues of such a random graph. I suspect that there is no true bound on these moments that would allow one to conclude that a random graph is probably Ramanujan. We consider the expected characteristic polynomial.

Let $\boldsymbol{M}$ be the adjacency matrix of a perfect matching, and let $\Pi_1, \dots, \Pi_d$ be independent uniform random permutation matrices. We will consider the expected characteristic polynomial

$$\mathbb{E}_{\Pi_1,\dots,\Pi_d} \chi_x(\Pi_1 \boldsymbol{M} \Pi_1^T + \cdots + \Pi_d \boldsymbol{M} \Pi_d^T).$$

In Lecture 22, we learned that this polynomial is real rooted. In Lecture 23, we learned a technique that allows us to compute this polynomial. Today we will prove that the second largest root of this polynomial is at most $2\sqrt{d-1}$. First, we show why this matters: it implies that there is some choice of the matrices $\Pi_1, \dots, \Pi_d$ so that resulting polynomial has second largest root at most $2\sqrt{d-1}$. These matrices provide the desired graph.

24.3 Interlacing Families of Polynomials

The general problem we face is the following. We have a large family of polynomials, say $p_1(x), \dots, p_m(x)$, for which we know each p_i is real-rooted and such that their sum is real rooted. We would like to show that there is some polynomial p_i whose largest root is at most the largest root of the sum, or rather we want to do this for the second-largest root. This is not true in general. But, it is true in our case. We will show that it is true whenever the polynomials form what we call an *interlacing family*.

Recall from Lecture 22 that we say that for monic degree n polynomials $p(x)$ and $r(x)$, $p(x) \to r(x)$ if the roots of p and r interlace, with the roots of r being larger. We proved that if $p_1(x) \to r(x)$ and $p_2(x) \to r(x)$, then every convex combination of p_1 and p_2 is real rooted. If we go through the proof, we will also see that for all $0 \leq s \leq 1$,

$$sp_2(x) + (1-s)p_1(x) \to r(x).$$

Proceeding by induction, we can show that if $p_i(x) \to r(x)$ for each i, then every convex combination of these polynomials interlaces $r(x)$, and is thus real rooted. That is, for every $s_1, \dots, s_m$ so that $s_i \geq 0$ (but not all are zero),

$$\sum_i s_i p_i(x) \to r(x).$$

Polynomials that satisfy this condition are said to have a *common interlacing*. By a technique analogous to the one we used to prove Lemma 22.3.2, one can prove that the polynomials $p_1, \dots, p_m$ have a common interlacing if and only if every convex combination of these polynomials is real rooted.

Lemma 24.3.1. *Let $p_1, \ldots, p_m$ be polynomials so that $p_i(x) \to r(x)$, and let $s_1, \ldots, s_m \geq 0$ be not identically zero. Define*

$$p_\emptyset(x) = \sum_{i=1}^{m} s_i p_i(x).$$

Then, there is an i so that the largest root of $p_i(x)$ is at most the largest root of $p_\emptyset(x)$. In general, for every j there is an i so that the jth largest root of $p_i(x)$ is at most the jth largest root of $p_\emptyset(x)$.

Proof. We prove this for the largest root. The proof for the others is similar. Let λ_1 and λ_2 be the largest and second-largest roots of $r(x)$. Each polynomial $p_i(x)$ has exactly one root between λ_1 and λ_2, and is positive at all $x > \lambda_1$. Now, let μ be the largest root of $p_\emptyset(x)$. We can see that μ must lie between λ_1 and λ_2. We also know that

$$\sum_i p_i(\mu) = 0.$$

If $p_i(\mu) = 0$ for some i, then we are done. If not, there is an i for which $p_i(\mu) > 0$. As p_i only has one root larger than λ_2, and it is eventually positive, the largest root of p_i must be less than μ. □

Our polynomials do not all have a common interlacing. However, they satisfy a property that is just as useful: they form an interlacing family. We say that a set of polynomials $p_1, \ldots, p_m$ forms an interlacing family if there is a rooted tree T in which

a. every leaf is labeled by some polynomial p_i,

b. every internal vertex is labeled by a nonzero, nonnegative combination of its children, and

c. all siblings have a common interlacing.

The last condition guarantees that every internal vertex is labeled by a real rooted polynomial. Note that the same label is allowed to appear at many leaves.

Lemma 24.3.2. *Let $p_1, \ldots, p_m$ be an interlacing family, let T be the tree witnessing this, and let $p_\emptyset$ be the polynomial labeling the root of the tree. Then, for every j there exists an i for which the jth largest root of p_i is at most the jth largest root of $p_\emptyset$.*

Proof. By Lemma 24.3.1, there is a child of the root whose label has a jth largest root that is smaller than the jth largest root of $p_\emptyset$. If that child is not a leaf, then we can proceed down the tree until we reach a leaf, at each step finding a node labeled by a polynomial whose jth largest root is at most the jth largest root of the previous polynomial. □

Our construction of permutations by sequences of random swaps provides the required interlacing family.

Theorem 24.3.3. *For permutation matrices $\Pi_1, \ldots, \Pi_d$, let*

$$p_{\Pi_1,\ldots,\Pi_d}(x) = \chi_x(\Pi_1 \boldsymbol{M} \Pi_1^T + \cdots + \Pi_d \boldsymbol{M} \Pi_d^T).$$

These polynomials form an interlacing family.

We will finish this lecture by proving that the second-largest root of

$$\mathbb{E} p_{\Pi_1,\ldots,\Pi_d}(x)$$

is at most $2\sqrt{d-1}$. This implies that there is a d-regular multigraph on n vertices in our family with second-largest adjacency eigenvalue at most $2\sqrt{d-1}$.

24.4 Root Bounds for Finite Free Convolutions

Recall from the last lecture that for n-dimensional symmetric matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ with uniform row sums a and b and characteristic polynomials $(x-a)p(x)$ and $(x-b)q(x)$,

$$\mathbb{E}_\Pi \chi_x(\boldsymbol{A} + \Pi \boldsymbol{B} \Pi^T) = (x - (a+b))p(x) \boxplus_{n-1} q(x).$$

This formula extends to sums of many such matrices. It is easy to show that

$$\chi_x(\boldsymbol{M}) = (x-1)^{n/2}(x+1)^{n/2} = (x-1)p(x), \quad \text{where } p(x) \stackrel{\text{def}}{=} (x-1)^{n/2-1}(x+1)^{n/2}.$$

So,

$$p_\emptyset(x) \stackrel{\text{def}}{=} \mathbb{E} p_{\Pi_1,\ldots,\Pi_d}(x) = (x-d)\left(p(x) \boxplus_{n-1} p(x) \boxplus_{n-1} p(x) \boxplus_{n-1} \cdots \boxplus_{n-1} p(x)\right),$$

where $p(x)$ appears d times above.

We would like to prove a bound on the largest root of this polynomial in terms of the largest roots of $p(x)$. This effort turns out not to be productive. To see why, consider matrices $\boldsymbol{A} = a\boldsymbol{I}$ and $\boldsymbol{B} = b\boldsymbol{I}$. It is clear that $\boldsymbol{A} + \Pi \boldsymbol{B} \Pi^T = (a+b)\boldsymbol{I}$ for every Π. This tells us that

$$(x-a)^n \boxplus (x-b)^n = (x-(a+b))^n.$$

So, the largest roots can add. This means that if we are going to obtain useful bounds on the roots of the sum, we are going to need to exploit facts about the distribution of the roots of $p(x)$. As in Lecture ??, we will use the barrier functions, just scaled a little differently.

For,

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

define the Cauchy transform of p at x to be

$$\mathcal{G}_p(x) = \frac{1}{d} \sum_{i=1}^{d} \frac{1}{x - \lambda_i} = \frac{1}{d} \frac{p'(x)}{p(x)}.$$

For those who are used to Cauchy transforms, I remark that this is the Cauchy transform of the uniform distribution on the roots of $p(x)$. As we will be interested in upper bounds on the Cauchy transform, we will want a number u so that for all $x > u$, $\mathcal{G}_p(x)$ is less than some specified value. That is, we want the *inverse Cauchy transform*, which we define to be

$$\mathcal{K}_p(w) = \max\{x : \mathcal{G}_p(x) = w\}.$$

For a real rooted polynomial p, and thus for real $\lambda_1, \ldots, \lambda_d$, it is the value of x that is larger than all the λ_i for which $\mathcal{G}_p(x) = w$. For $w = \infty$, it is the largest root of p. But, it is larger for finite w.

We will prove the following bound on the Cauchy transforms.

Theorem 24.4.1. *For degree n polynomials p and q and for $w > 0$,*

$$\mathcal{K}_{p \boxplus_n q}(w) \leq \mathcal{K}_p(w) + \mathcal{K}_q(w) - 1/w.$$

For $w = \infty$, this says that the largest root of $p \boxplus_n q$ is at most the sum of the largest roots of p and q. But, this is obvious.

To explain the $1/w$ term in the above expression, consider $q(x) = x^n$. As this is the characteristic polynomial of the all-zero matrix, $p \boxplus_n q = p(x)$. We have

$$\mathcal{G}_q(x) = \frac{1}{n} \frac{n x^{n-1}}{x^n} = \frac{1}{x}.$$

So,

$$\mathcal{K}_q(w) = \max\{x : 1/x = w\} = 1/w.$$

Thus,

$$\mathcal{K}_q(w) - 1/w = 0.$$

I will defer the proof of this theorme to next lecture (or maybe the paper [MSS15a]), and now just show how we use it.

24.5 The Calculation

For $p(x) = (x-1)^{n/2-1}(x+1)^{n/2}$,

$$\mathcal{G}_p(x) = \frac{1}{n-1}\left(\frac{n/2-1}{x-1} + \frac{n/2}{x+1}\right) \leq \frac{1}{n}\left(\frac{n/2}{x-1} + \frac{n/2}{x+1}\right),$$

for $x \geq 1$. This latter expression is simple to evaluate. It is

$$\frac{x}{x^2-1} = \mathcal{G}_{\chi(\boldsymbol{M})}(x).$$

We also see that

$$\mathcal{K}_p(w) \leq \mathcal{K}_{\chi(\boldsymbol{M})}(w),$$

for all $w \geq 0$.

Theorem 24.4.1 tells us that

$$\mathcal{K}_{p \boxplus_{n-1} \cdots \boxplus p}(w) \leq d\mathcal{K}_p(w) - \frac{d-1}{w}.$$

Using the above inequality, we see that this is at most

$$d\mathcal{K}_{\chi(\boldsymbol{M})}(w) - \frac{d-1}{w}.$$

As this is an upper bound on the largest root of $p \boxplus_{n-1} \cdots \boxplus_{n-1} p$, we wish to set w to minimize this expression. As,

$$\mathcal{G}_{\chi(\boldsymbol{M})}(x) = \frac{x}{x^2 - 1},$$

we have

$$\mathcal{K}_{\chi(\boldsymbol{M})}(w) = x \quad \text{if and only if} \quad w = \frac{x}{x^2 - 1}.$$

So,

$$d\mathcal{K}_{\chi(\boldsymbol{M})}(w) - \frac{d-1}{w}. \leq dx - d - 1\frac{x^2 - 1}{x}.$$

The choice of x that minimizes this is $\sqrt{d-1}$, at which point it becomes

$$d\sqrt{d-1} - \frac{(d-1)(d-2)}{\sqrt{d-1}} = d\sqrt{d-1} - (d-2)\sqrt{d-1} = 2\sqrt{d-1}.$$

24.6 Some explanation of Theorem 24.4.1

I will now have time to go through the proof of Theorem 24.4.1. So, I'll just tell you a little about it. We begin by transforming statements about the inverse Cauchy transform into statements about the roots of polynomials.

As $\mathcal{G}_p(x) = \frac{1}{d}\frac{p'(x)}{p(x)}$,

$$\mathcal{G}_p(x) = w \qquad \Longleftrightarrow \qquad p(x) - \frac{1}{wd}p'(x) = 0.$$

This tells us that

$$\mathcal{K}_p(w) = \text{maxroot}\left(p(x) - p'(x)/wd\right) = \text{maxroot}\left((1 - (1/wd)\partial_x)p\right).$$

As this sort of operator appears a lot in the proof, we give it a name:

$$U_\alpha = 1 - \alpha\partial_x.$$

In this notation, Theorem 24.4.1 becomes

$$\text{maxroot}\left(U_\alpha(p \boxplus_n q)\right) \leq \text{maxroot}\left(U_\alpha p\right) + \text{maxroot}\left(U_\alpha p\right) - n\alpha. \tag{24.1}$$

We, of course, also need to exploit an expression for the finite free convolution. Last lecture, we proved that if

$$p(x) = \sum_{i=0}^{n} x^{n-i}(-1)^i a_i \quad \text{and} \quad q(x) = \sum_{i=0}^{n} x^{n-i}(-1)^i b_i.$$

Then,

$$p(x) \boxplus_n q(x) = \sum_{k=0}^{n} x^{n-k}(-1)^k \sum_{i+j=k} \frac{(n-i)!(n-j)!}{n!(n-i-j)!} a_i b_j. \tag{24.2}$$

From this, one can derive a formula that plays better with derivatives:

$$p(x) \boxplus_n q(x) = \frac{1}{n!} \sum_{i=0}^{n} (n-i)! b_i p^{(i)}(x).$$

This equation allows us to understand what happens when p and q have different degrees.

Lemma 24.6.1. *If $p(x)$ has degree n and $q(x) = x^{n-1}$, then*

$$p(x) \boxplus_n q(x) = \partial_x p(x).$$

For the special case of $q(x) = x^{n-1}$, we have

$$U_\alpha q(x) = x^{n-1} - \alpha(n-1)x^{n-2},$$

so

$$\operatorname{maxroot}(U_\alpha q(x)) = \alpha(n-1).$$

So, in this case (24.1) says

$$\operatorname{maxroot}(U_\alpha \partial_x p) \leq \operatorname{maxroot}(U_\alpha p) + \operatorname{maxroot}(U_\alpha q) - n\alpha = \operatorname{maxroot}(U_\alpha p) - \alpha.$$

The proof of Theorem 24.4.1 has two major ingredients. We begin by proving the above inequality. We then show that the extreme case for the inequality is when $q(x) = (x-b)^n$ for some b. To do this, we consider an arbitrary real rooted polynomial q, and then modify it to make two of its roots the same. This leads to an induction on degree, which is essentially handled by the following result.

Lemma 24.6.2. *If $p(x)$ has degree n and the degree of $q(x)$ is less than n, then*

$$p \boxplus_n q = \frac{1}{n} (\partial_x p) \boxplus_{n-1} q.$$

The whose proof is fairly straightforward, and only requires 2 pages.

24.7 Some thoughts

I would like to reflect on the fundamental difference between considering expected characteristic polynomials and the distributions of the roots of random polynomials. Let $\boldsymbol{A}$ be a symmetric matrix of dimension $3k$ with k eigenvalues that are 1, 0, and -1. If you consider $\boldsymbol{A} + \Pi \boldsymbol{A} \Pi^T$ for a random Π, the resulting matrix will almost definitely have a root at 2 and a root at -2. In fact, the chance that it does not is exponentially small in k. However, all the roots of the expected characteristic polynomial of this matrix are strictly bounded away from 2. You could verify this by computing the Cauchy transform of this polynomial.

In our case, we considered a matrix $\boldsymbol{A}$ with k eigenvalues of 1 and k eigenvalues of -1. If we consider $\boldsymbol{A} + \Pi \boldsymbol{A} \Pi^T$, it will almost definitely have roots at 2 and -2, and in fact the expected characteristic polynomial has roots that are very close to this. But, if we consider

$$\boldsymbol{A} + \Pi_1 \boldsymbol{A} \Pi_1^T + \Pi_2 \boldsymbol{A} \Pi_2^T,$$

even though it almost definitely has roots at 3 and -3, the largest root of the expected characteristic polynomial is at most $2\sqrt{2} < 3$.

I should finish by saying that Theorem 24.4.1 is inspired by a theorem of Voiculescu that holds in the infinite dimensional case. In this limit, the inequality becomes an equality.

References

[Fri08] Joel Friedman. *A Proof of Alon's Second Eigenvalue Conjecture and Related Problems.* Number 910 in Memoirs of the American Mathematical Society. American Mathematical Society, 2008.

[MSS15a] A. W. Marcus, D. A. Spielman, and N. Srivastava. Finite free convolutions of polynomials. *arXiv preprint arXiv:1504.00350*, April 2015.

[MSS15b] Adam W Marcus, Nikhil Srivastava, and Daniel A Spielman. Interlacing families IV: Bipartite Ramanujan graphs of all sizes. *arXiv preprint arXiv:1505.08010*, 2015. appeared in Proceedings of the 56th IEEE Symposium on Foundations of Computer Science.