Interlacing Families and the Kadison–Singer Problem

Contents

1 **Introduction** 1

2 **Stable Polynomials** 3
- 2.1 Real Rootedness and Real Stability . . . 4
- 2.2 Barrier Argument . . . 7

3 **The Kadison–Singer Problem** 11
- 3.1 C*-algebras and States . . . 11
- 3.2 Kadison–Singer Problem . . . 12
- 3.3 Paving Conjecture . . . 13
- 3.4 Proof of the Paving Conjecture . . . 14

4 **Ramanujan Graphs** 17
- 4.1 Expander Graphs . . . 18
- 4.2 2-lifts and the Construction of Ramanujan Graphs . . . 20
- 4.3 Matching Polynomials . . . 21
- 4.4 Existence of bipartite Ramanujan graphs of arbitrary degree . . . 23
- 4.5 Generic bound . . . 23

Bibliography 27

Chapter 1

Introduction

Given m balls $[m]$ and any collection of m subsets $S_1, \dots S_m \subset [m]$ of these balls, can we colour the m balls in red and blue in such a way that the number of red and blue balls approximately agree in all of the subsets S_k? Ideally, we would colour the balls all half-red, half-blue. But how close can we come to this situation with a discrete colouring?

This would be a typical question from the field of *discrepancy theory*. An answer to the above question due to Spencer [16], is that we can find a colouring such that in all S_k the difference between the numbers of red and blue balls is at most $6\sqrt{m}$.

A result of a similar kind lies at the core of a sequence of the two papers [10, 11] by Marcus, Spielman and Srivastava (see also the review article [12] by the same authors and the blog article [17] by Srivastava) this essay is about. Instead of partitioning a set of m balls, we want to partition m vectors $v_1, \dots, v_m \in \mathbb{R}^n$ which are small ($\|v_k\|^2 \le \epsilon$ for all k) and decompose the identity matrix $\mathbb{1}$ in the sense that

$$v_1 v_1^* + \cdots + v_m v_m^* = \mathbb{1}$$

into two subsets $\{v_k\}_{k \in S_1} \cup \{v_k\}_{k \in S_2} = \{v_1, \dots, v_m\}$ in such a way that

$$\sum_{k \in S_j} v_k v_k^* \approx \frac{1}{2} \mathbb{1}$$

for $j = 1, 2$ in some approximative sense. Intuitively, it should be clear that the smallness condition $\|v_k\|^2 \le \epsilon$ is necessary for such a result to hold since any single vector v_k of relatively big norm makes any partition rather unbalanced. As a special case of the discrepancy result Corollary 2.1 which lies at the core of chapter two, we can prove

$$\left\| \sum_{k \in S_j} v_k v_k^* \right\| \le \left(1/\sqrt{2} + \sqrt{\epsilon} \right)^2 \tag{1.1}$$

and therefore

$$-(\sqrt{2\epsilon} + \epsilon)\mathbb{1} \le \sum_{k \in S_j} v_k v_k^* - \frac{1}{2}\mathbb{1} \le (\sqrt{2\epsilon} + \epsilon)\mathbb{1}$$

for $j = 1, 2$ where $A \le B$ means that the matrix $B - A$ is positive semidefinite.

An often successful approach for proving the bound from eq. (1.1) could be to pick the partition randomly in the sense that we consider independent Bernoulli distributed random variables $\epsilon_1, \ldots, \epsilon_m$ and to compute the probability of $\left\| \sum_{k=1}^m \epsilon_k v_k v_k^* \right\|$ being small. In the present case, however, often successful methods like Matrix Chernoff bounds or more sophisticated concentration results all only give highly probable bounds on eq. (1.1) growing as $O(\sqrt{\epsilon \log n})$ with the dimension n (see, for example [20, 21]). At the expense of replacing *high probability* with *positive probability*, Marcus, Spielman and Srivastava managed to improve this bound to be constant in the dimension n. Phrased probabilistically the main result the sequence of papers is the following Theorem:

Theorem 1.1. *Let* $\epsilon > 0$, $m, n \in \mathbb{N}$, $(\Omega, \mathcal{F}, \mathbf{P})$ *be a finite probability space and* $v_1, \ldots, v_m$ *be independent random* $\mathbb{C}^n$*-vectors such that* $\mathbf{E} \left\| v_i \right\|^2 \leq \epsilon$ *and*

$$\sum_{k=1}^m \mathbf{E}\, v_k v_k^* = \mathbb{1}.$$

Then

$$\mathbf{P}\left(\left\| \sum_{k=1}^m v_k v_k^* \right\| \leq (1+\sqrt{\epsilon})^2 \right) > 0.$$

Note that this can not be proved via the first moment method since $\mathbf{E} \left\| \sum_{k=1}^m v_k v_k^* \right\|$ grows with the dimension n. Instead, Marcus, Spielman and Srivastava used the characterization of $\left\| \sum_{k=1}^m v_k v_k^* \right\|$ as the largest root of the characteristic polynomial

$$x \mapsto P_{\sum_{k=1}^m v_k v_k^*}(x) := \det \left(x\mathbb{1} - \sum_{k=1}^m v_k v_k^* \right)$$

and showed that the largest root of the expected characteristic polynomial $\mathbf{E}\, P_{\sum_{k=1}^m v_k v_k^*}$ is at most given by $(1+\sqrt{\epsilon})^2$. To conclude the proof they used interlacing properties of characteristic polynomials to show that – while the determinant generally behaves very non-linearly – the largest root of $\mathbf{E}\, P_{\sum_{k=1}^m v_k v_k^*}$ still is an average of the largest roots of the different realizations of $P_{\sum_{k=1}^m v_k v_k^*}$.

This innovative method enabled Marcus, Spielman and Srivastava to resolve two long-open problems from rather different areas of Mathematics – they gave a positive answer to the Kadison–Singer problem and managed to prove the existence of bipartite Ramanujan graphs of arbitrary degree.

The purpose if this essay is to understand the method of interlacing polynomials, as well as how it can be applied to solve the two famous and unrelated problems. The author does, of course, not claim any originality and follows the original papers [10, 11] as well as subsequently published simplifications or alternative accounts [19, 15, 18, 22, 3] closely.

The second chapter of this essay establishes the method of interlacing polynomials, while in the third and fourth chapter the two problems are (briefly) introduced and resolved using results from chapter two.

Chapter 2

Stable Polynomials

The goal of this chapter is to give a self contained proof of Theorem 1.1. Before starting with the proof, however, we want to prove the following deterministic Corollary, already advertised in the introduction on discrepancy.

Corollary 2.1. *Let $n, m \in \mathbb{N}$ and $u_1, \dots, u_m \in \mathbb{C}^n$ be vectors with $u_1u_1^* + \cdots + u_m u_m^* = \mathbb{1}$ and $\|u_k\|^2 \leq \epsilon$ for all $k \in [m]$. Then for any $r \in \mathbb{N}$ there exists a partition $[m] = S_1 \cup \cdots \cup S_r$ such that*

$$\left\| \sum_{k \in S_j} u_k u_k^* \right\| \leq \left(\frac{1}{\sqrt{r}} + \sqrt{\epsilon} \right)^2$$

for all $j \in [r]$.

Proof. Take $\Omega = [r]^m$ equipped with the uniform distribution and set $v_k : \Omega \to \mathbb{C}^n \otimes \mathbb{C}^r$ to map $v_k(\omega) := \sqrt{r} u_k \otimes e_{\omega_k}$ for $k \in [m]$ with $e_1, \dots e_r$ being the standard basis of $\mathbb{C}^r$. Then $\mathbf{E} \|v_k\|^2 = \mathbf{E}\, r \|u_k\|^2 \leq r\epsilon$ and

$$\sum_{k=1}^m \mathbf{E}\, v_k v_k^* = r \sum_{k=1}^m (u_k u_k^*) \otimes \left(\mathbf{E}_\omega\, e_{\omega_k} e_{\omega_k}^* \right) = r \sum_{k=1}^m (u_k u_k^*) \otimes \left(\frac{1}{r} \mathbb{1} \right) = \mathbb{1}$$

and thus by Theorem 1.1 for some $\omega \in [r]^m$ we have

$$\begin{aligned} \sum_{j=1}^r \left\| \sum_{k \in [m], \omega_k = j} r u_k u_k^* \right\|^2 &= \left\| \left(\sum_{k \in [m], \omega_k = 1} r u_k u_k^* \right) \otimes e_1 e_1^* + \cdots + \left(\sum_{k \in [m], \omega_k = r} r u_k u_k^* \right) \otimes e_r e_r^* \right\|^2 \\ &= \left\| \sum_{k=1}^m v_k(\omega) v_k(\omega)^* \right\|^2 \leq (1 + \sqrt{r\epsilon})^4 \end{aligned}$$

implying

$$\left\| \sum_{k \in [m], \omega_k = j} u_k u_k^* \right\| \leq \left(\frac{1}{\sqrt{r}} + \sqrt{\epsilon} \right)^2$$

for all $j \in [r]$. □

Before going into the details, we will give a rough sketch of the proof strategy for Theorem 1.1. The proof of the probabilistic bound on the operator norm of a matrix utilizes a rather unusual approach. It uses the interpretation of the operator norm of the hermitian matrix $\sum_{k=1}^m v_k v_k^*$ as the largest root of its characteristic polynomial

$$x \mapsto P_{\sum_{k=1}^m v_k v_k^*}(x) := \det\left(x - \sum_{k=1}^m v_k v_k^*\right). \tag{2.1}$$

The proof will proceed in two stages. Firstly, we prove the claimed bound for the expected polynomial:

Proposition 2.2. *Under the assumptions of Theorem 1.1 we have that the largest root of the expected characteristic polynomial*

$$x \mapsto \mathbf{E}\, P_{\sum_{k=1}^m v_k v_k^*}(x) = \mathbf{E} \det\left(x - \sum_{j=1}^m v_k v_k^*\right)$$

satisfies

$$\operatorname{maxroot} \mathbf{E}\, P_{\sum_{k=1}^m v_k v_k^*} \le (1+\sqrt{\epsilon})^2.$$

Secondly we will relate the maximal root of $\mathbf{E}\, P_{\sum_{k=1}^m v_k v_k^*}$ to the maximal roots of the realizations of $P_{\sum_{k=1}^m v_k v_k^*}$:

Proposition 2.3. *Let $m, n \in \mathbb{N}$, $(\Omega, \mathcal{F}, \mathbf{P})$ be a finite probability space and $v_1, \dots, v_m$ be independent random $\mathbb{C}^n$-vectors. Then the polynomial $\mathbf{E}\, P_{\sum_{k=1}^m v_k v_k^*}$ has only real roots and it holds that*

$$\min \operatorname{maxroot} P_{\sum_{k=1}^m v_k v_k^*} \le \operatorname{maxroot} \mathbf{E}\, P_{\sum_{k=1}^m v_k v_k^*} \le \max \operatorname{maxroot} P_{\sum_{k=1}^m v_k v_k^*}$$

where the minimum and maximum are taken over all values the random vectors v_k take with positive probability. In particular,

$$\left\| \sum_{k=1}^m v_k v_k^* \right\| \le \operatorname{maxroot} \mathbf{E}\, P_{\sum_{k=1}^m v_k v_k^*}$$

with positive probability.

Remark 2.4. *Note that this result is rather non-trivial and relies on specific properties of characteristic polynomials. In general, roots of polynomials need not to behave well under convex combinations. For example, the average of polynomials with only real roots can have complex roots. Also note that the expected characteristic polynomial is, in general, not a characteristic polynomial.*

Clearly both Propositions together imply the Theorem. We shall begin with the proof of Proposition 2.3 by showing that characteristic polynomials exhibit certain interlacing properties.

2.1 Real Rootedness and Real Stability

We begin by better understanding the characteristic polynomial of sums of rank one matrices.

Lemma 2.5. *The characteristic polynomial of the sum of rank-one matrices $\sum_{k=1}^m v_k v_k^*$ can be written as*

$$\det\left(z - \sum_{k=1}^m v_k v_k^*\right) = \mu[v_1 v_1^*, \dots, v_m v_m^*](z)$$

where

$$\mu[A_1, \dots, A_m](z) := \left[\prod_{k=1}^{m} \left(1 - \frac{\partial}{\partial z_k}\right)\right] \det\left(z\mathbb{1} + \sum_{k=1}^{m} z_k A_k\right)\Big|_{z_1=\dots=z_m=0} \tag{2.2}$$

is the so called mixed characteristic polynomial *of positive semidefinite matrices* $A_1, \dots, A_m$.

Proof. Since for any invertible matrix A and vector v we have

$$\begin{aligned}\det(A - vv^*) &= \det(A)\det(\mathbb{1} - A^{-1}vv^*) = (1 - v^*A^{-1}v)\det A\\ &= (1 - \partial/\partial z)(1 + zv^*A^{-1}v)\big|_{z=0}\det A = (1 - \partial/\partial z)\det(A + zvv^*)\big|_{z=0}\end{aligned}$$

the same formula holds true for any matrix A by density of invertible matrices. Applying this formula iteratively then proves the Lemma. □

While the determinant of some matrix generally depends in a very non-linear fashion on the matrix, the mixed characteristic polynomial depends affine-linearly on rank-one matrices:

Lemma 2.6 (Affine-linearity of mixed characteristic polynomials). *Let* $\lambda_1, \dots, \lambda_l > 0$ *with* $\lambda_1 + \dots + \lambda_l = 1$ *and* $A_1, \dots, A_{m-1}$ *be positive semidefinite matrices and* $w_1, \dots, w_l$ *be vectors of the corresponding dimension. Then*

$$\sum_{j=1}^{l} \lambda_j \mu[A_1, \dots, A_{m-1}, w_j w_j^*] = \mu[A_1, \dots, A_{m-1}, \lambda_1 w_1 w_1^* + \dots + \lambda_l w_l w_l^*]$$

and the same holds true in the other entries since μ *is invariant under permutations of the matrices. In particular, if the vectors are chosen independently randomly and assume only finitely many values we have that*

$$\mathbf{E}\det\left(z\mathbb{1} - \sum_{k=1}^{m} v_k v_k^*\right) = \mu[\mathbf{E}\, v_1 v_1^*, \dots, \mathbf{E}\, v_m v_m^*](z). \tag{2.3}$$

Proof. For invertible matrices A we find

$$\begin{aligned}\sum_{j=1}^{l}\lambda_j \det(A + zw_jw_j^*) &= \sum_{j=1}^{l} \lambda_j \det(A)(1 + z\operatorname{Tr}(A^{-1}w_jw_j^*))\\ &= \det(A)\left[1 + z\operatorname{Tr}\left(A^{-1}\sum_{j=1}^{l}\lambda_j w_j w_j^*\right)\right] = \det\left(A + z\sum_{j=1}^{l}\lambda_k w_j w_j^*\right) + \mathcal{O}\left(z^2\right)\end{aligned}$$

from the straight-forward expansion

$$\det(\mathbb{1} + zB) = 1 + z\operatorname{Tr} B + \mathcal{O}\left(z^2\right).$$

The same result for general matrices and thereby the claim follows, again, by density of invertible matrices. Applying this formula iteratively using independence also yields eq. (2.3). □

Clearly all realizations of the random characteristic polynomial $\det(z\mathbb{1} - \sum_{k=1}^{m} v_k v_k^*)$ have only real roots. The expected characteristic polynomial is a convex combination of these, so called *real rooted*, polynomials and can be shown to also only having real roots. It should be emphasized that this is not the case for general polynomials.

Lemma 2.7. *Let $A_1, \dots, A_m$ be positive semidefinite matrices of the same dimension. Then*

(i) the polynomial

$$(z, z_1, \dots, z_m) \mapsto \det(z\mathbb{1} + z_1 A_1 + \cdots + z_m A_m)$$

has no zeros with $\Im z, \Im z_1, \dots, \Im z_m$ all positive (such a polynomial is called real stable*),*

(ii) the polynomial $\mu[A_1, \dots, A_m]$ is real rooted.

Proof.

(i) The coefficients of the polynomial are real since the polynomial takes real values for all choices of real variables. Also, for all vectors $0 \neq x \in \mathbb{C}^k$ we have

$$\begin{aligned}\Im \langle x, (z\mathbb{1} + z_1 A_1 + \cdots + z_m A_m)x \rangle &= \langle x, (\Im z\mathbb{1} + \Im z_1 A_1 + \cdots + \Im z_m A_m)x \rangle \\ &\geq \Im z \, \|x\|^2 > 0\end{aligned}$$

in case that $\Im z, \Im z_1, \dots, \Im z_m > 0$ implying that the determinant is non-zero.

(ii) Assume that p is an univariate polynomial with no roots in the upper half plane. Thus we can write $p(z) = c\prod_{j=1}^{l}(z - x_j)$ for some $x_1, \dots, x_l$ with $\Im x_j \leq 0$ and some $c \in \mathbb{C}$. Then

$$(1 - \partial/\partial z)p(z) = c\left(1 - \sum_{j=1}^{l} \frac{1}{z - x_j}\right)\prod_{j=1}^{l}(z - x_j)$$

also has no roots in the upper half plane. Applying this to the m variables of the polynomial from part (i) while fixing the other variables, we find that

$$\left(1 - \frac{\partial}{\partial z_1}\right) \dots \left(1 - \frac{\partial}{\partial z_m}\right) \det\left(z + \sum_{k=1}^{m} z_k v_k v_k^*\right) \tag{2.4}$$

is a real stable polynomial.

If $(z_1, \dots, z_m) \mapsto p(z_1, \dots, z_m)$ is a real stable polynomial then $p(\cdot, \dots, \cdot, 0)$ can be locally uniformly approximated by the sequence $p(\cdot, \dots, \cdot, i/n)$, $n \in \mathbb{N}$ and since the latter has no zeros in $\{\, z \in \mathbb{Z} \mid \Im z > 0 \,\}^{m-1}$ by Hurwitz's Theorem from complex analysis (see, for example, [6, Theorem 2.5]) neither does the former unless it's identically zero. Specialising this to eq. (2.4) shows that the restriction to $z_1 = \cdots = z_m = 0$ produces a real stable (and thus real rooted since complex roots of univariate polynomials come in pairs of complex conjugates) univariate polynomial since characteristic polynomials cannot vanish identically. □

The last ingredient we need for the proof of Proposition 2.3 is a control on the maximal root of convex combinations of our polynomials.

Lemma 2.8. *Let p, q be polynomials of the same degree with leading coefficient 1 such that $\lambda p + (1 - \lambda)q$ is real rooted for all $0 \leq \lambda \leq 1$. If $\operatorname{maxroot} p \leq \operatorname{maxroot} q$, then*

$$\operatorname{maxroot} p \leq \operatorname{maxroot}(\lambda p + (1 - \lambda)q) \leq \operatorname{maxroot} q$$

for all $0 \leq \lambda \leq 1$.

Proof. Since $p, q > 0$ in $(\operatorname{maxroot} q, \infty)$ the second inequality is immediate. Assume for contradiction that the first inequality fails, i.e., that $p_\lambda := \lambda p + (1-\lambda)q$ has no roots in $[\operatorname{maxroot} p, \operatorname{maxroot} q]$ for some $0 < \lambda < 1$. Then $p_\lambda > 0$ in $[\operatorname{maxroot} p, \operatorname{maxroot} q]$ which in turn implies that $q(\operatorname{maxroot} p) > 0$ and consequently that $q = p_0$ has at least two zeros in $(\operatorname{maxroot} p, \operatorname{maxroot} q]$. Note that since $p_\mu(\operatorname{maxroot} p) > 0$ and p_μ has only real roots for all $\mu < 1$, there has to be some point $\mu < \lambda$ where at least two roots of p_μ from $(-\infty, \operatorname{maxroot} p)$ jump to $(\operatorname{maxroot} p, \operatorname{maxroot} q]$. This contradict, for example, Hurwitz' Theorem. □

Proof of Proposition 2.3. It follows from Lemmata 2.5, 2.6 and 2.7 that all convex combinations of the random realizations of the characteristic polynomial are real rooted. Thus (and since characteristic polynomials have leading coefficient 1) we can apply Lemma 2.8 iteratively to conclude the proof. □

2.2 Barrier Argument

The goal of this section is to give a proof of Proposition 2.2. According to the preceding section and after a straightforward translation we can rewrite the expected characteristic polynomial as

$$\mathbf{E}\, P_{\sum_{k=1}^r v_k v_k^*}(z) = \left[\prod_{k=1}^m \left(1 - \frac{\partial}{\partial z_k}\right)\right] \det\left(\sum_{k=1}^m z_k \,\mathbf{E}\, v_k v_k^*\right)\Bigg|_{z_1=\cdots=z_m=z}$$

After defining $p(z_1, \ldots, z_m) := \det(z_1 \,\mathbf{E}\, v_1 v_1^* + \cdots + z_m \,\mathbf{E}\, v_m v_m^*)$ and introducing the shorthand notation $\partial_k := \partial/\partial z_k$ we have to show that $(1-\partial_1)\ldots(1-\partial_m)p$ is non-zero on the diagonal

$$\left\{\, (t, \ldots, t) \in \mathbb{R}^m \;\middle|\; t > (1+\sqrt{\epsilon})^2 \,\right\}.$$

In fact we will prove the stronger statement that $(1-\partial_1)\ldots(1-\partial_m)p$ is nonzero in the orthant $\left\{\, (z_1, \ldots, z_m) \in \mathbb{R}^m \;\middle|\; z_k > (1+\sqrt{\epsilon})^2 \text{ for all } k \in [m] \,\right\}$, in other words, that $((1+\sqrt{\epsilon})^2, \ldots, (1+\sqrt{\epsilon})^2)$ *lies above the zeros* of $(1-\partial_1)\ldots(1-\partial_m)p$.

Proof of Proposition 2.2. We want to bound the zeros of $(1-\partial_1)\ldots(1-\partial_m)p$ using an iterative argument. Firstly, p has no zeros above $(0, \ldots, 0)$ since $z_1 \,\mathbf{E}\, v_1 v_1^* + \cdots + z_m \,\mathbf{E}\, v_m v_m^* \geq \mathbb{1} \min_k z_k$ is non-singular in this orthant. For the next iteration note that in this orthant $(1-\partial_k)p = 0$ if and only if $\frac{\partial_k p}{p} = \partial_k \log \circ p = 1$. Clearly

$$\begin{aligned}
\partial_k \log(p(z_1, \ldots, z_m))\big|_{z_1,\ldots,z_m=t} &= \frac{\partial_k \det(z_1 \,\mathbf{E}\, v_1 v_1^* + \cdots + z_m \,\mathbf{E}\, v_m v_m^*)\big|_{z_1=\cdots=z_m=t}}{\det(t \,\mathbf{E}\, v_1 v_1^* + \cdots + t \,\mathbf{E}\, v_m v_m^*)} \\
&= \partial_k \det\left(1 + \frac{z_k - t}{t} \,\mathbf{E}\, v_k v_k^*\right)\Bigg|_{z_k=t} = \frac{\operatorname{Tr} \mathbf{E}\, v_k v_k^*}{t} \leq \frac{\epsilon}{t}
\end{aligned}$$

and we can choose $t > 0$ and $\delta > 0$ such that $\epsilon/t < 1 - 1/\delta$. This is the basis for the iteration step, as summarized in the following Proposition.

Proposition 2.9. *Suppose that $x \in \mathbb{R}^m$ lies above the roots of some real stable polynomial $(z_1, \ldots, z_m) \mapsto q(z_1, \ldots, z_m)$ and*

$$\partial_k \log(q(z_1, \ldots, z_m))\big|_{z=x} < 1 - 1/\delta$$

for all k and some $\delta > 0$. Then x also lies above the zeros of $(1-\partial_k)q$ and

$$\partial_j \log\left[(1-\partial_k)q(z_1, \ldots, z_m)\right]\big|_{z=(x_1,\ldots,x_{k-1},x_k+\delta,x_{k+1},\ldots,x_m)} \leq \partial_j \log q(z_1, \ldots, z_m)\big|_{z=x}$$

for all k, j.

Applying this result to our case yields that $(t, \dots, t)$ lies above the roots of $(1 - \partial_1)p$ and that

$$\partial_i \log\left[(1 - \partial_1)p(z_1, \dots, z_m)\right]\big|_{z=(t+\delta,t,\dots,t)} \le 1 - 1/\delta.$$

Thus we are able to iterate and apply Proposition 2.9 again to the still real stable (see the proof of Lemma 2.7) polynomial $(1 - \partial_1)p$ showing that $(t + \delta, t + \delta, t, \dots, t)$ lies above the zeros of $(1 - \partial_1)(1 - \partial_2)p$. Iterating this we find that $(t + \delta, \dots, t + \delta)$ lies above the zeros of $(1-\partial_1)\dots(1-\partial_r)p$. In particular, the maximal root of the expected characteristic polynomial $\mathbf{E}\, P_{\sum_{k=1}^m v_k v_k^*}$ is at most $t+\delta$. We can now choose $t = \epsilon + \sqrt{\epsilon}$ and $\delta = 1 + \sqrt{\epsilon}$ to complete the proof. ☐

It remains to prove the non-trivial Proposition 2.9. The proof will follow from the following monotonicity and convexity properties of the so called barrier functions $\partial_k q$.

Lemma 2.10. *Suppose $x \in \mathbb{R}^m$ lies above the zeros of some real stable polynomial $(z_1, \dots, z_m) \mapsto q(z_1, \dots, z_m)$. Then the functions*

$$t \mapsto \partial_1 \log(q(z_1, \dots, z_m))\big|_{z=(x_1+t,x_2,\dots,x_m)} \tag{2.5}$$

and

$$t \mapsto \partial_1 \log(q(z_1, \dots, z_m))\big|_{z=(x_1,x_2+t,x_3,\dots,x_m)} \tag{2.6}$$

are positive, decreasing and convex for $t \ge 0$.

Proof. The univariate polynomial $z_1 \mapsto q(z_1, \dots, z_m)$ can be factorized as

$$q(z_1, \dots, z_m) = c(z_2, \dots, z_m) \prod_{j=1}^{l} (z_1 - \rho_j(z_2, \dots, z_m))$$

in terms of its real roots $\rho_1(z_2, \dots, z_m) \ge \dots \ge \rho_l(z_2, \dots, z_m)$. Then

$$\partial_1 \log(q(z_1, \dots, z_m)) = \sum_{j=1}^{l} \frac{1}{z_1 - \rho_j(z_2, \dots, z_m)} > 0$$

for $z \ge x$. Also, obviously $\partial_1^2 \log(q(z_1, \dots, z_m)) < 0$ and $\partial_1^3 \log(q(z_1, \dots, z_m)) > 0$ for $z \ge x$ proving the monotonicity and convexity of the function from eq. (2.5).

While this first statement was effectively (after freezing the remaining variables) a statement about univariate real rooted polynomials, the second statement is effectively a statement about real stable polynomials in two variables and considerably harder to prove. Let $(z_1, z_2) \mapsto q'(z_1, z_2)$ denote the bivariate restriction of q which can be factorized as $q'(z_1, z_2) = c(z_1) \prod_{j=1}^{l} (z_2 - \rho_j(z_1))$ in terms of its real roots $\rho_1(z_1) \ge \dots \ge \rho_n(z_1)$. Then

$$\partial_2 \partial_1 \log(q'(z_1, z_2)) = \partial_1 \sum_{j=1}^{l} (z_2 - \rho_j(z_1))^{-1}$$

and

$$\partial_2^2 \partial_1 \log(q'(z_1, z_2)) = -\partial_1 \sum_{j=1}^{l} (z_2 - \rho_j(z_1))^{-2}$$

and the proof would be complete if we knew that $z_1 \mapsto \rho_j(z_1)$ was decreasing for all j.

Assume for contradiction that this is not the case, i.e., that there exists $a_1 \in \mathbb{R}$ such that some root satisfies $c := \partial_1 \rho_i(z_1)\big|_{z_1=a} > 0$ then (a_1, a_2) with $a_2 := \rho_i(a_1)$ is a root of q'. Then we find

$$\partial_2 q'(z_1, z_2)\big|_{z=a} = c(a_1) \prod_{j \neq i} (a_2 - \rho_j(a_1)) =: c'$$

and

$$\partial_1 q'(z_1, z_2)\big|_{z=a} = -c(a_1) c \prod_{j \neq i} (a_2 - \rho_j(a_1)) = -cc'$$

and consequently we can expand q' around a as

$$q'(z_1, z_2) = c' \left[-c(z_1 - a_1) + (z_2 - a_2)\right] + \mathcal{O}\left((|z_1 - a_1| + |z_2 - a_2|)^2\right).$$

Now define the univariate polynomial

$$q_\epsilon(x) := \epsilon^{-1} q'(i\epsilon + a_1, x\epsilon + a_2) = \epsilon^{-1} c'(-ic\epsilon + x\epsilon) + \mathcal{O}(\epsilon) = -icc' + c'x + \mathcal{O}(\epsilon)$$

with a root in the upper half plane for small ϵ. But this would imply that q' also had a root (z_1, z_2) with z_1 in the upper half plane contradicting the real stability. □

Now we have all ingredients to prove the remaining Proposition 2.9 and thereby concluding the proof of Proposition 2.2 and consequently also Theorem 1.1.

Proof of Proposition 2.9. From the monotonicity in all variables (using Lemma 2.10 after an appropriate relabelling) it follows that $\partial_k \log \circ q < 1 - 1/\delta$ in the orthant above x and therefore $(1 - \partial_k) q$ cannot have any roots in this area. For the claimed bound note that

$$\partial_j \log \circ \left[(1 - \partial_k) q\right] = \partial_j \log \circ \left[q \cdot (1 - \partial_k \log \circ q)\right] = \partial_j \log \circ q - \frac{\partial_j \partial_k \log \circ q}{1 - \partial_k \log \circ q}$$

where the denominator is, by assumption, in the orthant above x at least $1/\delta$. Thus

$$\begin{aligned} &\partial_j \log\left[(1 - \partial_k) q(z_1, \ldots, z_m)\right] \big|_{z=(x_1, \ldots, x_{k-1}, x_k + \delta, x_{k+1}, \ldots, x_m)} \\ &\quad \leq (\partial_j - \delta \partial_k \partial_j) \log(q(z_1, \ldots, z_m))\big|_{z=(x_1, \ldots, x_{k-1}, x_k + \delta, x_{j+1}, \ldots, x_m)} \leq \partial_j \log(q(z_1, \ldots, z_m))\big|_{z=x} \end{aligned}$$

from convexity in the form $f(t + \delta) - \delta f'(t + \delta) \leq f(t)$. □

Chapter 3

The Kadison–Singer Problem

3.1 C*-algebras and States

The Kadison–Singer Problem originated as a conjecture about pure states on the C*-algebra $\mathcal{B}(\ell_2)$ of bounded operators on the Hilbert space ℓ_2. We start by giving a short introduction to the notion of abstract C*-algebras and their states.

A *Banach algebra* $\mathcal{A}$ is complex Banach space together with a distributive and associative multiplication satisfying $\lambda(AB) = (\lambda A)B = A(\lambda B)$ and $\|AB\| \leq \|A\| \|B\|$ for all $A, B \in \mathcal{A}$ and $\lambda \in \mathbb{C}$. A C*-algebra is a Banach algebra $\mathcal{A}$ together with a map $^* : \mathcal{A} \to \mathcal{A}$ satisfying

- * is an involution, i.e., $(A^*)^* = A$ for all $A \in \mathcal{A}$;
- * is conjugate linear, i.e., $(\lambda A)^* = \overline{\lambda} A^*$ for all $A \in \mathcal{A}$, $\lambda \in \mathbb{C}$;
- $(AB)^* = B^* A^*$ for all $A, B \in \mathcal{A}$;
- $\|A^* A\| = \|A\|^2$.

If $\mathcal{A}$ is unital, i.e., there exists a multiplicative identity $\mathbb{1} \in \mathcal{A}$ satisfying $\mathbb{1}A = A\mathbb{1}$ for all $A \in \mathcal{A}$, then this identity automatically satisfies $\mathbb{1} = \mathbb{1}^*$ and $\|\mathbb{1}\| = 1$. To avoid introducing the concept of *approximate identities*, we shall only consider unital C*-algebras. For the discussion of the Kadison–Singer this is sufficient since all discussed C*-algebras will be unital.

Example 3.1. *The bounded linear operators $\mathcal{B}(\mathcal{H})$ acting on a Hilbert space $\mathcal{H}$ with composition as the multiplication form a Banach algebra. Recall that to every operator $A \in \mathcal{B}(\mathcal{H})$ corresponds its adjoint operator $A^* \in \mathcal{B}(\mathcal{H})$ uniquely defined by Riesz Representation Theorem such that $\langle x, A^* y\rangle = \langle Ax, y\rangle$ for all $x, y \in \mathcal{H}$. It can be easily checked that the map $A \mapsto A^*$ satisfies all the properties requested above, giving $\mathcal{B}(\mathcal{H})$ also the structure of a C*-algebra. Clearly any closed *-subalgebra $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ (meaning that the subalgebra is norm-closed and stable under the *-operation) of $\mathcal{B}(\mathcal{H})$ is also a C*–algebra.*

According to the Gelfand–Naimark Theorem (see, for example, [8]) the above examples are indeed the only examples of C*-algebras. The theorem states that any C*-algebra is isometrically *-isomorphic (meaning that the isomorphism respects the *-operation) to a closed *-subalgebra of $\mathcal{B}(\mathcal{H})$, the bounded operators on some Hilbert space $\mathcal{H}$. Therefore, from now on, we shall assume that every C*-algebra is realized as a *-closed subalgebra of $\mathcal{B}(\mathcal{H})$, for some Hilbert space $\mathcal{H}$.

Remark 3.2. *The commutative Gelfand–Naimark Theorem (see, for example, [4, Theorem 2.1.11B]) gives an alternative characterization of commutative C*-algebras. It states that any unital commutative C*-algebra is isometrically *-isomorphic to $C(X)$ for some compact Hausdorff space X.*

A linear functional $\phi\colon \mathcal{A} \to \mathbb{C}$ on a C*-algebra $\mathcal{A}$ is called *positive* if it satisfies $\phi(A^*A) \geq 0$ for all $A \in \mathcal{A}$. Alternatively positive functionals are functionals mapping positive elements of $\mathcal{A}$ to non-negative scalars, where $A \in \mathcal{A}$ is called positive if one of the following equivalent (see [4, Proposition 2.2.10]) conditions is satisfied:

- $A = B^*B$ for some $B \in \mathcal{A}$;
- $A = B^2$ for some $B \in \mathcal{A}$ with $B^* = B$;
- $\sigma(A) \subset [0, \infty)$, i.e., $A - \lambda\mathbb{1}$ is not invertible only for $\lambda \in [0, \infty)$;
- (in case $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$) A is positive semidefinite, i.e., $\langle x, Ax\rangle \in [0, \infty)$ for all $x \in \mathcal{H}$.

Due to the rich structure of C*-algebras this purely algebraic property already implies (see [4, Proposition 2.3.11]) the topological property of ϕ being continuous and that ϕ attains its norm

$$\|\phi\| = \sup\{\, |\phi(A)| \mid A \in \mathcal{A}, \|A\| \leq 1 \,\} = \phi(\mathbb{1})$$

on the multiplicative identity $\mathbb{1}$. If a positive functional ϕ is normalized in the sense $\|\phi\| = \phi(\mathbb{1}) = 1$ it is called a *state*. Clearly any convex combination $t\phi_1 + (1-t)\phi_2$ of two states ϕ_1, ϕ_2 is again a positive functional and is also a state since

$$\|t\phi_1 + (1-t)\phi_2\| = (t\phi_1 + (1-t)\phi_2)(\mathbb{1}) = t\phi_1(\mathbb{1}) + (1-t)\phi_2(\mathbb{1}) = 1.$$

A state is called *pure* if it cannot be written as the convex combination of any two other states. The set of states $E_{\mathcal{A}}$ on $\mathcal{A}$ is a weak*-closed subset of the dual $\mathcal{A}^*$. Clearly the set of pure states $P_{\mathcal{A}}$ on $\mathcal{A}$ is equal to the set of extreme points of $E_{\mathcal{A}}$ and it follows from Krein–Milman that $E_{\mathcal{A}} = \overline{\operatorname{conv} P_{\mathcal{A}}}$ (for details, see [4, Theorem 2.3.15]).

Example 3.3.

(i) Consider the Hilbert space $\ell_2 n = \mathbb{C}^n$ equipped the Euclidean inner product and the C-algebra $M_n := \mathcal{B}(\ell_2^n)$ of $n \times n$ matrices equipped with the operator norm. Then for any positive semidefinite matrix ρ with unit trace $\operatorname{Tr}\rho = 1$, the map $\phi(A) = \operatorname{Tr}\rho A$ defines a state on $\mathcal{A}$. Moreover, ϕ is pure if and only if $\operatorname{Tr}\rho^2 = 1$ if and only if ρ is a rank 1 matrix, i.e., there exists some unit vector $x \in \ell_2^n$ such that $\operatorname{Tr}\rho A = \langle x, Ax\rangle$.*

(ii) By the Riesz–Markov Representation Theorem the states on $C(X)$ for X compact and Hausdorff are precisely given by maps of the form $f \mapsto \int_X f \,\mathrm{d}\mu$ for Borel probability measures on X. It can be seen that the pure states are those corresponding to Dirac measures δ_x, $x \in X$. Thus in particular all pure states on $C(X)$ are homeomorphisms (in fact, this is an equivalence). Remark 3.2 allows to generalize this result to all unital commutative C-algebras.*

3.2 Kadison–Singer Problem

According to Dirac's fundamental vision of quantum mechanics [7]:

> *To introduce a representation in practice*
>
> - *We look for observables which we would like to have diagonal either because we are interested in their probabilities or for reasons of mathematical simplicity;*
> - *We must see that they all commute – a necessary condition since diagonal matrices always commute;*

- *We then see that they form a complete commuting set, and if not we add some more commuting observables to make them into a complete commuting set;*
- *We set up an orthogonal representation with this commuting set diagonal.*

> *The representation is then completely determined [...] by the observables that are diagonal [...]*

In the language of C*-algebras Dirac thought of the diagonal operators $\mathcal{D}(\ell_2)$ on ℓ_2 as a complete commuting set of observables and of a very special class of pure states on $\mathcal{D}(\ell_2)$, those of the form $A \mapsto \langle e_k, Ae_k \rangle$ corresponding to the *orthogonal representation*, i.e., the canonical basis $(e_k)_{k\in\mathbb{N}}$ of ℓ_2. These special, so called *vector*, pure states on $\mathcal{D}(\ell_2)$ can be uniquely extended to $\mathcal{B}(\ell_2)$. But, what about the others? That is the Kadison-Singer Problem:

Conjecture 3.4 (Kadison–Singer (KS)). *Every pure state on the algebra of diagonal operators $\mathcal{D}(\ell_2)$ on ℓ_2 can be uniquely extended to a state on the algebra of bounded operators $\mathcal{B}(\ell_2)$.*

More generally, Dirac is suggested that pure states on any maximally commuting subalgebra of $\mathcal{B}(\ell_2)$ can be uniquely extended to states on $\mathcal{B}(\ell_2)$. Kadison and Singer managed to prove that this is false for any subalgebra not equal to $\mathcal{D}(\ell_2)$, see [9, 5]. Therefore it should not be too surprising that Kadison and Singer actually believed KS to be wrong:

> *[...] The results that we have obtained leave the question of uniqueness of extension of the singular pure states of $\mathcal{D}(\ell_2)$ open. We incline to the view that such extensions are non-unique [...]* (page 397 of [9])

Remark 3.5. *Note that the important statement in KS is the uniqueness since any state ϕ on $\mathcal{D}(\ell_2)$ can be trivially extended to $\phi \circ \operatorname{diag}$ on $\mathcal{B}(\ell_2)$, where* diag *maps any matrix to its diagonal part. Therefore the Kadison–Singer conjecture can be easily rephrased as:*

Conjecture 3.6 (KS'). *Every extension of a pure state ϕ on $\mathcal{D}(\ell_2)$ vanishes on operators with zero diagonal.*

Also note that the extension is automatically pure. Indeed, if we could write the unique extension ψ as a non-trivial convex combinations of states $t\psi_1 + (1-t)\psi_2$, the pureness of ϕ would imply that $\psi_1\big|_{\mathcal{D}(\ell_2)} = \psi_2\big|_{\mathcal{D}(\ell_2)} = \phi$, contradicting the uniqueness of the extension.

3.3 Paving Conjecture

Instead of proving the Kadison–Singer conjecture directly we will prove the following Paving conjecture, which turns out to be equivalent to KS.

Conjecture 3.7 (Paving Conjecture (PC)). *For all $\epsilon > 0$ there exists $r \in \mathbb{N}$ such that for all $n \in \mathbb{N}$ and $T \in M_n = \mathcal{B}(\ell_2^n)$ with $\operatorname{diag} T = 0$ there exist diagonal projections $Q_1, \ldots, Q_r \in := \mathcal{D}(\ell_2^n)$ with $Q_1 + \cdots + Q_r = \mathbb{1}$ satisfying*

$$\|Q_j T Q_j\| \leq \epsilon \|T\|$$

for all $1 \leq j \leq r$.

Since n is assumed to be independent from r PC can be shown to be equivalent to the following Paving Conjecture on $\mathcal{B}(\ell_2)$. With $(e_n)_{n\in\mathbb{N}}$ being the standard basis of ℓ_2, we can clearly identify ℓ_2^n as $\operatorname{span}\{e_1, \ldots, e_n\} \subset \ell_2$ and have the projections $P_n = e_1e_1^* + \ldots e_ne_n^*$ mapping $\ell_2 \to \ell_2^n \subset \ell_2$.

Conjecture 3.8 (PC'). *For all $\epsilon > 0$ there exists $r \in \mathbb{N}$ such that for all $T \in \mathcal{B}(\ell_2)$ with* $\operatorname{diag} T = 0$ *there exist diagonal projections $Q_1, \ldots, Q_r \in \mathcal{D}(\ell_2)$ with $Q_1 + \cdots + Q_r = \mathbb{1}$ satisfying*

$$\|Q_j T Q_j\| \leq \epsilon \|T\|$$

for all $1 \leq j \leq r$.

Proof of Equivalence of PC and PC'. Assume that PC holds and let $T \in \mathcal{B}(\ell_2)$ be some operator with $\operatorname{diag} T = 0$. Then for $T^{(n)} = P_n T P_n$ there exist projections $Q_1^{(n)}, \ldots, Q_r^{(n)} \in \mathcal{D}(\ell_2^n) \subset \mathcal{D}(\ell_2)$ satisfying $Q_1^{(n)} + \cdots + Q_r^{(n)} = P_n$ and $\left\|Q_j^{(n)} T^{(n)} Q_j^{(n)}\right\| \leq \epsilon \left\|T^{(n)}\right\|$ for all j. The diagonal projections in the norm topology can be canonically identified with $\{0,1\}^{\mathbb{N}}$ (with $\{0,1\}$ being equipped with the discrete topology) which is compact by Tychonoff and we can therefore find a subsequence such that $Q_j^{(n_k)}$ converges to some projection $Q_j \in \mathcal{D}(\ell_2)$ for all j as $k \to \infty$ which obviously still satisfy $Q_1 + \cdots + Q_r = 1$. Thus

$$\|Q_j T Q_j\| = \lim_{k\to\infty} \left\|Q_j^{(n_k)} T^{(n_k)} Q_j^{(n_k)}\right\| \leq \lim_{k\to\infty} \epsilon \left\|T^{(n_k)}\right\| = \epsilon \|T\|.$$

The converse follows easily from the identification $\ell_2^n \subset \ell_2$. □

We now are ready to study the relationship of PC and KS:

Equivalence of PC' and KS'. Assume that PC' holds and fix any $\epsilon > 0$. Let ψ be an extension of a pure state ϕ on $\mathcal{D}(\ell_2)$ to $\mathcal{B}(\ell_2)$ and let $T \in \mathcal{B}(\ell_2)$ be some operator with $\operatorname{diag} T = 0$. Choose $Q_1, \ldots, Q_r$ to be the projections from PC' and observe that $\psi(Q_j) = \phi(Q_j) \in \{0,1\}$ since ϕ is a pure state on the commutative C*-algebra $\mathcal{D}(\ell_2)$ and thus a homeomorphism (see Example 3.3(ii)) implying $\phi(Q_j) = \phi(Q_j^2) = \phi(Q_j)^2$. Since $1 = \sum_{1\leq j\leq r} \phi(Q_j)$ it follows that $\phi(Q_j) = \delta_{ij}$ for some $1 \leq i \leq r$. The Cauchy-Schwarz inequality now implies

$$|\psi(SQ_j)| \leq \psi(SS^*)\psi(Q_j^* Q_j) = \psi(SS^*)\psi(Q_j) = 0$$

and similarly $\psi(Q_j S) = 0$ for all $j \neq i$ and $S \in \mathcal{B}(\ell_2)$ implying

$$|\psi(T)| = |\psi(Q_i T Q_i)| \leq \epsilon \|T\|.$$

We therefore proved KS' since $\psi(T) = 0$ by the arbitrariness of ϵ.

Since the converse is not needed for the scope of this essay, details of the implication $KS' \Rightarrow PS'$ shall be omitted. The interested reader is referred to [9, Lemma 5] □

3.4 Proof of the Paving Conjecture

The study of interlacing families of polynomials allows us to give a positive solution to PC' and thereby also KS.

Proposition 3.9 (PC for orthogonal projections (PC_P)). *For any orthogonal projection $P \in M_n$, $n \in \mathbb{N}$ and $r \in \mathbb{N}$ there exist diagonal projections $Q_1, \ldots, Q_r \in M_n$ with $Q_1 + \cdots + Q_r = 1$ and*

$$\|Q_j P Q_j\| \leq \left(1/\sqrt{r} + \sqrt{\|\operatorname{diag} P\|}\right)^2$$

for all $j \in [r]$.

Proof. For $u_k := Pe_k$, $k \in [n]$ we have $\|u_k\|^2 = \langle e_k, Pe_k \rangle = P_{k,k} \leq \|\operatorname{diag} P\|$ and

$$\sum_{k\in[n]} u_k u_k^* \Big|_{\operatorname{Ran} P} = \sum_{k\in[n]} Pe_k e_k^* P\big|_{\operatorname{Ran} P} = \mathbb{1}_{\operatorname{Ran} P}.$$

From Corollary 2.1 we then find a partition $[n] = S_1 \cup \cdots \cup S_r$ and corresponding diagonal projections $Q_j := \sum_{k\in S_j} e_k e_k^*$ which satisfy

$$\begin{aligned} \|Q_j P Q_j\| &= \|(PQ_j)^*(PQ_j)\| = \|(PQ_j)(PQ_j)^*\| = \|PQ_jP\| \\ &= \left\| \sum_{k\in S_j} Pe_k e_k^* P \right\| = \left\| \sum_{k\in S_j} u_k u_k^* \right\| \leq \left(1/\sqrt{r} + \sqrt{\|\operatorname{diag} P\|}\right)^2. \qquad \square \end{aligned}$$

This special case of the paving conjecture can be amplified to prove the general case:

Theorem 3.10 (PC holds true). *For all $r \in \mathbb{N}$ and $T \in M_n$ with $\operatorname{diag} T = 0$ there exist diagonal projections $Q_1, \ldots, Q_{r^4}$ with $Q_1 + \cdots + Q_{r^4} = 1$ satisfying*

$$\|Q_j T Q_j\| \leq 4\left(1/r + \sqrt{2/r}\right)\|T\|$$

for all $1 \leq j \leq r^4$.

Proof. First assume that T is hermitian with $\|T\| \leq 1$. Then

$$P := \frac{1}{2}\begin{pmatrix} \mathbb{1} + T & \sqrt{\mathbb{1} - T^2} \\ \sqrt{\mathbb{1} - T^2} & \mathbb{1} - T \end{pmatrix}$$

is a Hermitian projection with $\|\operatorname{diag} P\| = 1/2$ and for any $r \in \mathbb{N}$ we can apply Proposition 3.9 to find diagonal projections $Q_1, \ldots, Q_r \in M_{2n}$ with $\sum_{j\in[r]} Q_j = 1$ and $\|Q_j P Q_j\| \leq \left(1/\sqrt{r} + 1/\sqrt{2}\right)^2$ for all $j \in [r]$. If we decompose

$$Q_j = \begin{pmatrix} Q_j^{(1)} & 0 \\ 0 & Q_j^{(2)} \end{pmatrix}$$

this implies

$$\max\{\left\|Q_j^{(1)}(\mathbb{1} + T)Q_j^{(1)}\right\|, \left\|Q_j^{(2)}(\mathbb{1} - T)Q_j^{(2)}\right\|\} \leq \epsilon_r := 2\left(1/\sqrt{r} + 1/\sqrt{2}\right)^2.$$

It follows from the first inequality that $0 \leq Q_j^{(1)}(\mathbb{1} + T)Q_j^{(1)} \leq \epsilon_r Q_j^{(1)}$, i.e., $-Q_j^{(1)} \leq Q_j^{(1)} T Q_j^{(1)} \leq (\epsilon_r - 1) Q_j^{(1)}$. The second inequality yields analogously that $Q_j^{(1)} \geq Q_j^{(1)} T Q_j^{(1)} \geq -(\epsilon_r - 1) Q_j^{(1)}$ and therefore with $Q_{i,j} := Q_i^{(1)} Q_j^{(2)}$ we can conclude

$$-(\epsilon_r - 1)Q_{i,j} \leq Q_{i,j} T Q_{i,j} \leq (\epsilon_r - 1)Q_{i,j}$$

and consequently also $\|Q_{i,j} T Q_{i,j}\| \leq \epsilon_r - 1$ for all $i, j \in [r]$. By scaling we can immediately generalize this to $\|Q_{i,j} T Q_{i,j}\| \leq (\epsilon_r - 1)\|T\|$ for Hermitian T of arbitrary norm. Since $\sum_{i,j\in[r]} Q_{i,j} = 1$ and $\epsilon_r - 1$ becomes arbitrarily small for large r this already proves the conjecture for hermitian matrices.

Now consider an arbitrary matrix T which can be decomposed as $T = U + iV$ with U, V being hermitian and $\|U\|, \|V\| \leq \|T\|$. From the above argument we find $2r^2$

diagonal projections $Q_1', Q_1'', \ldots, Q_{r^2}', Q_{r^2}''$ paving U and V, respectively. After defining $\tilde{Q}_{i,j} \coloneqq Q_i' Q_j''$, $i, j \in [r^2]$ we finally find

$$\left\| \tilde{Q}_{i,j} T \tilde{Q}_{i,j} \right\| \leq \left\| \tilde{Q}_{i,j} U \tilde{Q}_{i,j} \right\| + \left\| \tilde{Q}_{i,j} V \tilde{Q}_{i,j} \right\| \leq 2(\epsilon_r - 1) \|T\|$$

for the diagonal projections $\tilde{Q}_{i,j}$ still decomposing the identity $\sum_{i,j \in [r^2]} \tilde{Q}_{i,j} = 1$. □

Chapter 4

Ramanujan Graphs

To a (undirected) graph $G = (V, E)$ with vertex set V and edge set E, there corresponds the so called *adjacency matrix* $A(G)$ with $A(G)_{i,j} = 1$ if $\{i, j\} \in E$ and $A(G)_{i,j} = 0$ otherwise. The characteristic polynomial of the graph is defined to be the characteristic polynomial of its adjacency matrix, i.e.,

$$P_G(z) \coloneqq P_{A(G)}(z) \coloneqq \det(z - A(G)).$$

The spectrum $\operatorname{Spec} G$ of G is defined to be the spectrum of its adjacency matrix, i.e., the set of zeros of $P_{A(G)}$ and similarly, by eigenvalues and eigenvectors of G we just mean eigenvalues and eigenvectors of $A(G)$. The spectrum of a graph is real since adjacency matrices are obviously symmetric. As a convention, we order the eigenvalues by size in such a way that

$$\lambda_1(G) \geq \cdots \geq \lambda_{|V|}(G).$$

Example 4.1 (Spectrum of complete (bipartite) graphs).

(i) *The complete graph K_{d+1} on $d + 1$ vertices (which is d-regular) only has eigenvalues -1 and d. Indeed, let α_k denote the column vector of length k with all α's and $\mathbb{1}$ denote the identity. Then*

$$A(K_{d+1}) = 1_{d+1} \cdot 1_{d+1}^T - \mathbb{1}$$

where the rank one matrix $1_{d+1} \cdot 1_{d+1}^T$ has eigenvalues $d + 1$ and 0.

(ii) *The spectrum of the complete d-regular bipartite graph $K_{d,d}$ is $\{-d, 0, d\}$. Indeed,*

$$A(K_{d,d}) = \begin{pmatrix} 1_d \\ 0_d \end{pmatrix} \cdot \begin{pmatrix} 0_d^T & 1_d^T \end{pmatrix} + \begin{pmatrix} 0_d \\ 1_d \end{pmatrix} \cdot \begin{pmatrix} 1_d^T & 0_d^T \end{pmatrix}$$

is a rank two matrix with eigenvectors $\left(\begin{smallmatrix} 1_d \\ 1_d \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1_d \\ -1_d \end{smallmatrix}\right)$ corresponding to the eigenvalues $d, -d$.

Proposition 4.2 (Spectral properties of d-regular graphs). *Let $G = (V, E)$ be a d-regular graph, i.e., all vertices have degree d. Then*

(i) $\operatorname{Spec} G \subset [-d, d]$.

(ii) $d = \lambda_1(G) \in \operatorname{Spec} G$. *Also, $\lambda_2(G) < d$ if and only if G is connected.*

(iii) $\operatorname{Spec} G$ *is symmetric about* 0 *if G is bipartite.*

(iv) *A connected graph G is bipartite if and only if $-d \in \operatorname{Spec} G$.*

Proof.

(i) For $|\lambda|> d$ the matrix $\lambda - A(G)$ is strictly diagonal dominant and thereby non-singular.

(ii) The vector of all 1's is an eigenvector corresponding to the eigenvalue d since all rows sums up to d. If $f = (f_v)_{v\in V}$ is an eigenvector corresponding to eigenvalue d with $f_u = \max_{v\in V}|f_u|$ (otherwise $-f$ is also an eigenvector), then $f_u = (A(G)f)_u/d = \sum_{\{u,v\}\in E} f_v/d \leq \sum_{\{u,v\}\in E} f(u)/d = f_u$ implies $f_v = f_u$ for all v adjacent to v. Thus we find that for connected graphs the eigenspace corresponding to eigenvalue d is one-dimensional. Conversely in disconnected d-regular graphs one can trivially find independent eigenvectors corresponding to eigenvalue λ.

(iii) The adjacency matrix of a bipartite graph takes the form $A(G) = \left(\begin{smallmatrix} 0 & A \\ A^T & 0 \end{smallmatrix}\right)$. If $\left(\begin{smallmatrix} v \\ w \end{smallmatrix}\right)$ is an eigenvector with eigenvalue λ, then

$$A(G)\begin{pmatrix} -v \\ w \end{pmatrix} = \begin{pmatrix} Aw \\ -A^T v \end{pmatrix} = \begin{pmatrix} \lambda v \\ -\lambda w \end{pmatrix} = -\lambda \begin{pmatrix} -v \\ w \end{pmatrix}$$

and therefore $-\lambda$ is also an eigenvalue.

(iv) One implication follows trivially from the previous ones. Conversely any eigenvector f corresponding to eigenvalue $-d$ satisfies $f_u = (-1)^{\mathrm{dist}(u,v)} f_v$ for u, v in V (recall that G is assumed to be connected). This clearly induces a bipartition of the graph. □

4.1 Expander Graphs

In various areas of mathematics and computer science the question arises to find sparse d-regular graphs $G = (V, E)$ on a large number of vertices with good connectivity properties. Explicitly, one often asks for graphs with a large expansion coefficient

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

where $e(S, \overline{S})$ is the number of edges between S and its complement. For d-regular graphs, clearly $e(S, \overline{S}) \leq d|S|$ and therefore d is a natural upper limit on the expansion coefficient. We can, however, also find a lower limit in terms of an spectral gap of the graph:

Proposition 4.3. *For a d-regular graph $G = (V, E)$ we have $h(G) \geq (d - \lambda_2(G))/2$.*

Proof. Let S be a minimizer of $h(G)$. The graph Laplacian $\Delta := d\mathbb{1} - A(G)$ has $d - \lambda_1(G) = 0$ as its smallest eigenvalue corresponding to the constant vectors as eigenvectors. Define a vector f by $f_v = -|S|$ for $v \in \overline{S}$ and $f_v = |\overline{S}|$ for $v \in S$ which is obviously orthogonal to any constant vector and thereby it follows, for example from the min-max Theorem, that

$$d - \lambda_2(G) \leq \frac{\langle f, \Delta f\rangle}{\|f\|^2} = \frac{\sum_{\{u,v\}\in E}(f_v - f_u)^2}{\sum_{v\in V} f_v^2} = \frac{|V|^2 e(S, \overline{S})}{|V||S||\overline{S}|} \leq 2\frac{e(S, \overline{S})}{|S|} = 2h(G)$$

since $|S|\leq |V|/2$. □

Remark 4.4. *We can also improve the upper limit in terms of the spectral gap. Indeed, it holds that $h(G) \leq \sqrt{d(d - \lambda_2(G))}$. For details the reader is referred to [2].*

This reduces the problem of finding graphs with good expander properties to finding graphs with small non-trivial spectrum $\{\lambda_2(G), \ldots, \lambda_{|V|}(G)\}$ or $\{\lambda_2(G), \ldots, \lambda_{|V|-1}(G)\}$ for bipartite graphs, respectively. The following Theorem due to Alon and Boppana [1] shows that there is a theoretical limit on how small the non-trivial spectrum can be.

Theorem 4.5. *Let $G = (V, E)$ be a d-regular graph such there exist two edges $\{u_1, u_2\}, \{v_1, v_2\} \in E$ with distance $\operatorname{dist}(\{u_1, u_2\}, \{v_1, v_2\}) \geq 2k + 2$. Then*

$$\lambda_2(G) \geq 2\sqrt{d-1} + \frac{2\sqrt{d-1}-1}{k+1}.$$

Proof (after Nilli, [14]). The result is trivial for disconnected graphs. Therefore from now on assume G to be connected and note that in this case the eigenspace corresponding to eigenvalue d is one dimensional. Define $V_i = \{ v \in V \mid \min_{j=1,2} \operatorname{dist}(v, v_j) = i \}$ and analogously U_i for $0 \leq i \leq k$. Clearly $\bigcup U_i$ and $\bigcup V_i$ are disjoint, in fact, these vertex sets are not even adjacent. Now define a vector $f = (f_v)_{v \in V}$ such that $f_v = (d-1)^{-i/2}$ if $v \in V_i$ and $f_v = \alpha(d-1)^{-i/2}$ if $v \in U_i$ and $f_v = 0$ otherwise with $\alpha < 0$ chosen such that $\langle f, 1_{|V|} \rangle = \sum_{v \in V} f_v = 0$, i.e., f being orthogonal to the eigenvector corresponding to eigenvalue d. Thus it follows, for example from the min-max Theorem (using that the eigenspace corresponding to eigenvalue λ is one-dimensional), that $\lambda_2(G) \geq \langle Af, f \rangle / \langle f, f \rangle$. Clearly

$$\langle f, f \rangle = \sum_{v \in V} f_v^2 = \sum_{i=0}^{k} \frac{|V_i|}{(d-1)^i} + \sum_{i=0}^{k} \frac{\alpha^2 |U_i|}{(d-1)^i} =: C_{V,1} + C_{U,1}.$$

For a bound on $\langle Af, f \rangle$ it is most convenient to write $\Delta = d\mathbb{1} - A$ and then compute

$$\langle \Delta f, f \rangle = \sum_{\{u,v\} \in E} (f_u - f_v)^2 = \sum_{\substack{\{u,v\} \in E, \\ u \text{ or } v \in \bigcup V_i}} (f_u - f_v)^2 + \sum_{\substack{\{u,v\} \in E, \\ u \text{ or } v \in \bigcup U_i}} (f_u - f_v)^2 =: C_{V,2} + C_{U,2}.$$

We can reorder the sum $C_{V,2}$ in in terms with $u \in V_i, v \in V_{i+1}$ and $u \in V_k, v \notin \bigcup V_i$ to obtain

$$\begin{aligned}
C_{V,2} &= \sum_{i=0}^{k-1} \sum_{u \in V_i} \sum_{\substack{\{u,v\} \in E, \\ v \in V_{i+1}}} \left((d-1)^{-i/2} - (d-1)^{-i/2-1/2} \right)^2 + \sum_{u \in V_k} \sum_{\substack{\{u,v\} \in E, \\ v \notin V_{k-1}}} (d-1)^{-k} \\
&\leq \sum_{i=0}^{k-1} |V_i|(d-1) \left((d-1)^{-i/2} - (d-1)^{-i/2-1/2} \right)^2 + |V_k|(d-1)^{-k+1} \\
&= \sum_{i=0}^{k-1} \frac{|V_i|}{(d-1)^i} \left(d - 2\sqrt{d-1} \right) + \frac{|V_k|}{(d-1)^k} \left((d - 2\sqrt{d-1}) + (2\sqrt{d-1} - 1) \right) \\
&= (d - 2\sqrt{d-1}) C_{V,1} + (2\sqrt{d-1} - 1) \frac{|V_k|}{(d-1)^k}.
\end{aligned}$$

Since $|V_i| \leq (d-1)|V_{i-1}|$ for all i, we find $|V_k|/(d-1)^k \leq C_{V,1}/(k+1)$ and consequently $C_{V,2}/C_{V,1} \leq d - 2\sqrt{d-1} + \frac{2\sqrt{d-1}-1}{k+1}$. The same inequality holds for $C_{U,2}/C_{U,1}$ and we can conclude

$$\frac{\langle Af, f \rangle}{\langle f, f \rangle} = d - \frac{\langle \Delta f, f \rangle}{\langle f, f \rangle} \geq 2\sqrt{d-1} + \frac{2\sqrt{d-1}-1}{k+1}. \qquad \square$$

Remark 4.6. *If $\{G_n\}_{n\in\mathbb{N}}$ is an infinite family of d-regular graphs, the vertex number is unbounded and so is the maximal edge distance. Therefore* $\liminf_{n\to\infty} \lambda_2(G) \geq 2\sqrt{d-1}$.

Graphs with non-trivial spectrum, that is optimal in the sense of this Theorem are called Ramanujan:

Definition 4.7. *A d-regular graph $G = (V, E)$ is called* Ramanujan *if* $\max\{\lambda_2(G), |\lambda_{|V|}(G)|\} \leq 2\sqrt{d-1}$. *$G$ is called* bipartite Ramanujan *if it is bipartite and* $\lambda_2(G) \leq 2\sqrt{d-1}$.

4.2 2-lifts and the Construction of Ramanujan Graphs

It is an open question whether for all $d \in \mathbb{N}$ there exist infinite families of Ramanujan graphs. The question is known to have an affirmative answer in the case of $d = p^r + 1$, with $r \in \mathbb{N}$ and p prime (see [13]). The theory developed in chapter two allowed Marcus, Spielman, Srivastava to prove the existence of infinite families of bipartite Ramanujan graphs for all $d \in \mathbb{N}$:

Theorem 4.8. *For all $d \geq 2$ and $k \in \mathbb{N}$ there exists a d-regular bipartite graph $G = (V, E)$ with $|V| = d2^{k+1}$ and $\lambda_2(G) \leq 2\sqrt{d-1}$.*

The construction method was first suggested by Bilu and Linial. We start with some Ramanujan graph $G = (V, E)$ and want to construct another Ramanujan graph on vertex set $V \sqcup V = \{\, (v, i) \mid v \in V, i \in \{1, 2\} \,\}$. Giving an edge signing $s\colon E \to \{-1, 1\}$ we add for some given edge $\{u, v\}$ the edges $\{(u, 1), (v, 1)\}$ and $\{(u, 2), (v, 2)\}$ to the new edge set E_s if $s(\{u, v\}) = 1$ and $\{(u, 1), (v, 2)\}$ and $\{(u, 2), (v, 1)\}$ otherwise to obtain the new graph $G_s = (V \sqcup V, E_s)$. The graph G_s with twice as many vertices is called a 2-lift of G. It is evident from the construction that the 2-lift of a d-regular graph remains d-regular and that 2-lifts on bipartite graphs remain bipartite. The spectrum of a 2-lift G_s also can be easily related to the spectrum of G:

Lemma 4.9. *Define $A_s(G)_{u,v} = s(\{u, v\})A(G)_{u,v}$ to be the signed adjacency matrix of G. Then*

$$A(G_s) = \frac{1}{2}\begin{pmatrix} A(G) + A_s(G) & A(G) - A_s(G) \\ A(G) - A_s(G) & A(G) + A_s(G) \end{pmatrix}$$

and $\operatorname{Spec} G_s = \operatorname{Spec} A(G) \cup \operatorname{Spec} A_s(G)$.

Proof. Clearly

$$\frac{1}{2}(A(G) + A_s(G)) = A((V, s^{-1}\{1\}))$$

and

$$\frac{1}{2}(A(G) - A_s(G)) = A((V, s^{-1}\{-1\})).$$

If f is an eigenvector of $A(G)$ corresponding to eigenvalue λ, then $A(G_s)\begin{pmatrix} f \\ f \end{pmatrix} = \lambda \begin{pmatrix} f \\ f \end{pmatrix}$ and also, if f is an eigenvector of $A_s(G)$ corresponding to eigenvalue λ, then $\begin{pmatrix} f \\ -f \end{pmatrix}$ is an eigenvector of $A(G_s)$ corresponding to the same eigenvalue. This produces $2|V|$ linearly independent eigenvectors proving the assertion. $\square$

From this Lemma we can already outline the strategy for proving Theorem 4.8. We know that $K_{d,d}$ is bipartite Ramanujan since $\lambda_2(K_{d,d}) = 0$. Thus if we could find a signing s with $\lambda_1(A_s(K_{d,d}))\begin{pmatrix} f \\ f \end{pmatrix} = \operatorname{maxroot} P_{A_s(K_{d,d})} \leq 2\sqrt{d-1}$, we would have a bipartite Ramanujan graph G_s on $4d$ vertices and could hope that we can continue in an inductive manner. It has been conjectures by Bilu and Linial that this procedure can be used to produce infinite families of d-regular Ramanujan graphs for arbitrary d.

Conjecture 4.10 (BL). *Every d-regular graph $G = (V, E)$ has a signing $s\colon E \to \{-1, 1\}$ with $\|A_s(G)\| \le 2\sqrt{d-1}$.*

As a consequence of the theory developed in chapter two Marcus, Spielman and Srivastava could give a partial answer to this conjecture and thereby showed that the construction suggested in the previous paragraph allows yields arbitrary big bipartite Ramanujan graphs, proving Theorem 4.8.

Theorem 4.11. *Every d-regular graph $G = (V, E)$ has a signing $s\colon E \to \{-1, 1\}$ with $\lambda_1(A_s(G)) = \operatorname{maxroot} P_{A_s(G)} \le 2\sqrt{d-1}$.*

The idea behind the proof of this Theorem is choosing the signing uniformly randomly.

Proposition 4.12. *Let $G = (V, E)$ be a d-regular Graph with $d \ge 2$. If we choose the $2^{|E|}$ signings $s\colon E \to \{-1, 1\}$ randomly all with probability $2^{-|E|}$, then we have that $\mathbf{E}_s\, P_{A_s(G)}$ is a real rooted polynomial with $\operatorname{maxroot} \mathbf{E}_s\, P_{A_s(G)} \le 2\sqrt{d-1}$.*

Proposition 4.13. *For some $s \in \{-1, 1\}^E$, we have that $\operatorname{maxroot} P_{A_s(G)} \le \operatorname{maxroot} \mathbf{E}_s\, P_{A_s(G)}$ and therefore also $\lambda_1(A_s(G)) \le \operatorname{maxroot} \mathbf{E}_s\, P_{A_s(G)}$.*

4.3 Matching Polynomials

Before giving proofs of these two Propositions, we want to begin by better understanding the expected characteristic polynomial:

Lemma 4.14. *Und the assumptions of Lemma 4.12 we have that*

$$\mathbf{E}_s\, P_{A_s(G)}(z) = \mu_G(z) := \sum_{k=0}^{\lfloor |V|/2 \rfloor} z^{|V|-2k}(-1)^k m_k(G)$$

where $m_k(G)$ is the number of matchings in G with k edges.

Proof. By definition we have,

$$\begin{aligned}\mathbf{E}_s \det(z\mathbb{1} - A_s(G)) &= \sum_{\sigma \in \mathrm{Sym}(V)} \operatorname{sgn}(\sigma)\, \mathbf{E}_s \prod_{v \in V} (z - A_s(G)_{v,\sigma(v)}) \\ &= \sum_{U \subset V} z^{|V \setminus U|} \sum_{\sigma \in \mathrm{Sym}^*(U)} \operatorname{sgn}(\sigma)\, \mathbf{E}_s \prod_{v \in U} [-\chi_E(\{v, \sigma(v)\}) s(\{v, \sigma(v)\})]\end{aligned}$$

where Sym and Sym* are the permutation and derangement groups. The expected product is non-zero if and only if $\{\ \{v, \sigma(v)\} \mid v \in U\ \} \subset E$ and $\sigma^2(v) = v$ for all $v \in U$ (using independence), i.e., for derangements σ consisting only of cycles of length 2. Clearly in this case $|U|$ has to be even, $\operatorname{sgn}(\sigma) = (-1)^{|U|/2}$ and the expected products are all 1 since $\mathbf{E}\, s(\{v, \sigma(v)\})^2 = 1$. These 2-cyclic derangements are in one-to-one correspondence to $|U|/2$-matchings on $G(U)$ and thus

$$\mathbf{E}_s \det(z\mathbb{1} - A_s(G)) = \sum_{\substack{U \subset V, \\ |U| \text{ even}}} (-1)^{|U|/2} z^{|V \setminus U|} m_{|U|/2}(G(U)) = \mu_G(z). \qquad \square$$

Thus the expected characteristic polynomial is precisely given by the so called *matching polynomial* of the graph. The claimed bound on the maximal root of $2\sqrt{d-1}$ is a special case of the following property of general matching polynomials:

Proposition 4.15. *Let* $G = (V, E)$ *be a graph with maximal degree* $\Delta(G)$*. Then* $\mu_G(z) > 0$ *for* $z > \max\{2\sqrt{\Delta(G) - 1}, 1\}$.

Proof. It is evident from the definition of the matching polynomial that

$$\mu_G(z) = \mu_{G_1}(z) \dots \mu_{G_k}(z)$$

where $G_1 \sqcup \dots \sqcup G_k = G$ are the connected components of G. It therefore suffices to prove the statement for all connected components separately. In the case that some component G_l only has one vertex, we clearly have $\mu_{G_l}(z) = z$ for which the claim holds. In case some connected component G_l consists of two vertices, we have $\mu_{G_l}(z) = z^2 - 1$ for which the claim holds as well. From now on we concentrate on connected components G_l with at least three vertices and therefore $\Delta(G_l) \geq 2$ and aim to prove $\mu_{G_l}(z) > 0$ for $z > 2\sqrt{\Delta(G_l) - 1}$.

We are aiming to achieve this via induction and therefore claim that we have the recursive relation

$$\mu_G(z) = z\mu_{G(V\setminus\{v\})}(z) - \sum_{\{u,v\}\in E} \mu_{G(V\setminus\{u,v\})}(z)$$

for all $v \in V$. Indeed, there are $m_k(G(V \setminus \{v\}))$ matchings of size k on G not containing v and $\sum_{\{u,v\}\in E} m_{k-1}(G \setminus \{u, v\})$ matchings of size k containing v.

The induction step basically follows from the following Lemma:

Lemma 4.16. *Let* $H = (U, F)$ *be a graph,* $v \in U$ *a vertex and* $\delta \in \mathbb{N}$ *some integer such that* $\max\{2, \Delta(H), \deg v + 1\} \leq \delta$*. Then*

$$\frac{\mu_H(z)}{\mu_{H(U\setminus\{v\})}(z)} > \sqrt{\delta - 1}$$

for all $z > 2\sqrt{\delta - 1}$.

Proof of Lemma. The proof goes via induction on $|U|$. If $|U|= 1$, then the only vertex has degree zero and the quotient is simply given by z. For the induction step note that

$$\frac{\mu_H(z)}{\mu_{H(U\setminus\{v\})}(z)} = z - \sum_{\{u,v\}\in F} \frac{\mu_{H(U\setminus\{u,v\})}(z)}{\mu_{H(U\setminus\{v\})}(z)} > 2\sqrt{\delta - 1} - \frac{\deg v}{\sqrt{\delta - 1}} \geq \sqrt{\delta - 1}$$

which follows from the induction hypothesis since $\Delta(H(U \setminus \{v\})) \leq \Delta(H) \leq \delta$ and $\deg_{H(U\setminus\{v\})} u + 1 \leq \deg_H u \leq \Delta(H) \leq \delta$ for all u with $\{u, v\} \in F$. ☐

To conclude the proof pick any $v \in V$, and then for any $u \in V$ with $\{u, v\} \in E$ the Lemma can be applied to $H = G(V \setminus \{v\})$, u and $\delta = \Delta(G)$ since $\Delta(H) \leq \Delta(G)$ and $\deg_H u + 1 \leq \deg_G u \leq \Delta(G)$. Thus

$$\begin{aligned}\frac{\mu_G(z)}{\mu_{G(V\setminus\{v\})}(z)} &= z - \sum_{\{u,v\}\in E} \frac{\mu_{G(V\setminus\{u,v\})}(z)}{\mu_{G(V\setminus\{v\})}(z)} > 2\sqrt{\Delta(G) - 1} - \frac{\deg v}{\sqrt{\Delta(G) - 1}} \\ &\geq 2\sqrt{\Delta(G) - 1} - \frac{\Delta(G)}{\sqrt{\Delta(G) - 1}} > 0\end{aligned}$$

for $z > 2\sqrt{\Delta(G) - 1}$ and the claim follows from induction over $|V|$. ☐

This obviously proves Proposition 4.12:

Proof of Proposition 4.12. Observe that the matching polynomial is either even or odd and therefore $|\mu_G(z)|> 0$ for all $|z|> 2\sqrt{d - 1}$ from Proposition 4.15. ☐

4.4 Existence of bipartite Ramanujan graphs of arbitrary degree

We now turn to the proof of Proposition 4.13. Clearly

$$A_s(G) = -d\mathbb{1} + \sum_{\{u,v\}\in E} (e_u + s(\{u,v\})e_v)(e_u + s(\{u,v\})e_v)^T$$

can be written as a sum of a multiple of the identity and independent rank one matrices in terms of the standard basis $(e_v)_{v\in V}$. We now apply Proposition 2.3 to this sum of independent rank one matrices to conclude the proof:

Proof of Proposition 4.13. From Proposition 2.3 it follows that

$$\begin{aligned}\lambda_1(A_s(G)) = \lambda_1(A_s(G) + d\mathbb{1}) - d &\le \operatorname{maxroot} \mathbf{E}\, P_{A_s(G)+d\mathbb{1}} - d \\ &= \operatorname{maxroot} \mathbf{E}\, P_{A_s(G)}(\cdot - d) - d = \operatorname{maxroot} \mathbf{E}\, P_{A_s(G)}\end{aligned}$$

with positive probability. □

Now the existence of bipartite Ramanujan graphs of arbitrary degrees is almost an immediate consequence:

Proof of Theorems 4.8 and 4.11. Start with $K_{d,d}$ which bipartite Ramanujan after Example 4.1. Then after Propositions 4.12 and 4.13 there exists some signing with $\lambda_1(A_s(K_{d,d})) \le 2\sqrt{d-1}$ and thus the corresponding 2-lift $(K_{d,d})_s$ is still bipartite and satisfies

$$\lambda_s((K_{d,d})_s) = \max\{\lambda_1(A_s(K_{d,d})), \lambda_2(K_{d,d})\} \le 2\sqrt{d-1}.$$

We therefore have found a bipartite Ramanujan graph $(K_{d,d})_s$ on $4d$ vertices. We can continue inductively to obtain bipartite Ramanujan graphs of order $2^k d$ for all $k \in \mathbb{N}$. □

Finally, it should be remarked why the introduced technique does not prove the existence of non-bipartite Ramanujan graphs of arbitrary degree. Theorem 4.11 also applies to non-bipartite graphs and can be easily generalized to prove the existence of signings with $\lambda_1(A_s(G)) \le 2\sqrt{d-1}$ or $\lambda_{|V|}(A_s(G)) \ge -2\sqrt{d-1}$, respectively. The problem lies in the fact that the interlacing argument does not give the existence of a signing which satisfies these two conditions simultaneously.

4.5 Generic bound

It is also interesting to note that while the bound on the maximal root of the matching polynomial was necessary to prove the precise eigenvalue bound, a more generic bound using Theorem 1.1 gives the same asymptotic behaviour. Indeed,

$$\mathbf{E}_s \sum_{\{u,v\}\in E} \frac{(e_u + s(\{u,v\})e_v)(e_u + s(\{u,v\})e_v)^T}{d} = \mathbb{1}$$

and

$$\mathbf{E}_s \left\| \frac{e_u + s(\{u,v\})e_v}{\sqrt{d}} \right\|^2 = \frac{2}{d} =: \epsilon.$$

Thus Theorem 1.1 gives that

$$\sum_{\{u,v\}\in E} \frac{(e_u + s(\{u,v\})e_v)(e_u + s(\{u,v\})e_v)^T}{d} \le (1 + \sqrt{2/d})^2 \mathbb{1}$$

with positive probability and therefore also $A_s(G) \leq \sqrt{8d} + 2$ with positive probability, which is of the same order as the optimal bound $A_s(G) \leq 2\sqrt{d-1}$. An interesting consequence of the generic bound applied to the signed adjacency matrices is that the ϵ-dependence in Theorem 1.1 is necessary since we would otherwise have a contradiction to Theorem 4.5.

Bibliography

[1] Noga Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986.

[2] Noga Alon. On the edge-expansion of graphs. *Combin. Probab. Comput.*, 6(2):145–152, 1997.

[3] Petter Brändén. Lecture Notes for Interlacing Families. https://people.kth.se/ pbranden/notes.pdf.

[4] O. Bratteli and Robinson. *Operator Algebras and Quantum Statistical Mechanics 1. C*-and W*-Algebras. Symmetry Groups. Decomposition of States*. Springer, 1987.

[5] P. G. Casazza and J. Crandell Tremain. The Kadison-Singer Problem in mathematics and engineering. *Proceedings of the National Academy of Science*, 103:2032–2039, February 2006.

[6] J.B. Conway. *Functions of One Complex Variable I*. Graduate Texts in Mathematics. Springer New York, 2012.

[7] P. A. M. Dirac. *The Principles of Quantum Mechanics*. International series of monographs on physics (Oxford, England). 1930.

[8] I. Gelfand and M. Neumark. On the imbedding of normed rings into the ring of operators in Hilbert space. *Mat. Sb., Nov. Ser.*, 12:197–213, 1943.

[9] Richard V. Kadison and I. M. Singer. Extensions of pure states. *American Journal of Mathematics*, 81(2):pp. 383–400, 1959.

[10] A. Marcus, D. A. Spielman, and N. Srivastava. Interlacing Families I: Bipartite Ramanujan Graphs of All Degrees. *ArXiv e-prints*, April 2013.

[11] A. Marcus, D. A Spielman, and N. Srivastava. Interlacing Families II: Mixed Characteristic Polynomials and the Kadison-Singer Problem. *ArXiv e-prints*, June 2013.

[12] A. W. Marcus, D. A. Spielman, and N. Srivastava. Ramanujan Graphs and the Solution of the Kadison-Singer Problem. *ArXiv e-prints*, August 2014.

[13] M. Morgenstern. Existence and Explicit Constructions of q + 1 Regular Ramanujan Graphs for Every Prime Power q. *Journal of Combinatorial Theory, Series B*, 62(1):44 – 62, 1994.

[14] Alon Nilli. On the second eigenvalue of a graph. *Discrete Mathematics*, 91(2):207–210, 1991.

[15] Stochastic Analyis Seminar Princeton University. Proof of Kadison-Singer. https://blogs.princeton.edu/sas/2014/10/05/lectures-2-and-3-proof-of-kadison-singer-1/.

[16] Joel Spencer. Six standard deviations suffice. *Transactions of the American Mathematical Society*, 289(2):679–679, feb 1985.

[17] Nikhil Srivastava. Discrepancy, Graphs, and the Kadison-Singer Problem. http://windowsontheory.org/2013/07/11/discrepancy-graphs-and-the-kadison-singer-conjecture-2/, July 2013.

[18] Terence Tao. Real stable polynomials and the Kadison-Singer problem. https://terrytao.wordpress.com/2013/11/04/real-stable-polynomials-and-the-kadison-singer-problem/, November 2013.

[19] D. Timotin. The solution to the Kadison–Singer Problem: yet another presentation. *ArXiv e-prints*, January 2015.

[20] J. A. Tropp. The random paving property for uniformly bounded matrices. *ArXiv Mathematics e-prints*, December 2006.

[21] J. A. Tropp. User-friendly tail bounds for sums of random matrices. *ArXiv e-prints*, April 2010.

[22] A. Valette. Le problème de Kadison-Singer (The Kadison-Singer problem). *ArXiv e-prints*, September 2014.