The Solution of the Kadison-Singer Problem

Adam Marcus (Crisply, Yale)
Daniel Spielman (Yale)
Nikhil Srivastava (MSR India)

Outline

Disclaimer

The Kadison-Singer Problem, defined.

Restricted Invertibility, a simple proof.

Break

Kadison-Singer, outline of proof.

The Kadison-Singer Problem (‘59)

A positive solution is equivalent to:

- Anderson’s Paving Conjectures (‘79, ‘81)
- Bourgain-Tzafriri Conjecture (‘91)
- Feichtinger Conjecture (‘05)
- Many others

Implied by:

- Akemann and Anderson’s Paving Conjecture (‘91)
- Weaver’s KS_2 Conjecture

The Kadison-Singer Problem (‘59)

A positive solution is equivalent to:

Anderson’s Paving Conjectures (‘79, ‘81)

Bourgain-Tzafriri Conjecture (‘91)

Feichtinger Conjecture (‘05)

Many others

Implied by:

Akemann and Anderson’s Paving Conjecture (‘91)

Weaver’s KS_2 Conjecture

The Kadison-Singer Problem ('59)

A positive solution is equivalent to:

- Anderson's Paving Conjectures ('79, '81)
- Bourgain-Tzafriri Conjecture ('91)
- Feichtinger Conjecture ('05)
- Many others

Implied by:

- Akemann and Anderson's Paving Conjecture ('91)
- Weaver's KS_2 Conjecture

The Kadison-Singer Problem ('59)

Let $\mathcal{A}$ be a maximal Abelian subalgebra of $\mathcal{B}(\ell^2(\mathbb{N}))$, the algebra of bounded linear operators on $\ell^2(\mathbb{N})$

Let $\rho : \mathcal{A} \to \mathbb{C}$ be a pure state.
Is the extension of ρ to $\mathcal{B}(\ell^2(\mathbb{N}))$ unique?

See Nick Harvey's Survey or Terry Tao's Blog

Anderson's Paving Conjecture '79

For all $\epsilon > 0$ there is a k so that for every
n-by-n symmetric matrix A with zero diagonals,

there is a partition of $\{1, ..., n\}$ into $S_1, ..., S_k$

$$\|A(S_j, S_j)\| \leq \epsilon \|A\| \quad \text{for} \quad j = 1, \ldots, k$$

Recall $\|A\| = \max_{\|x\|=1} \|Ax\|$

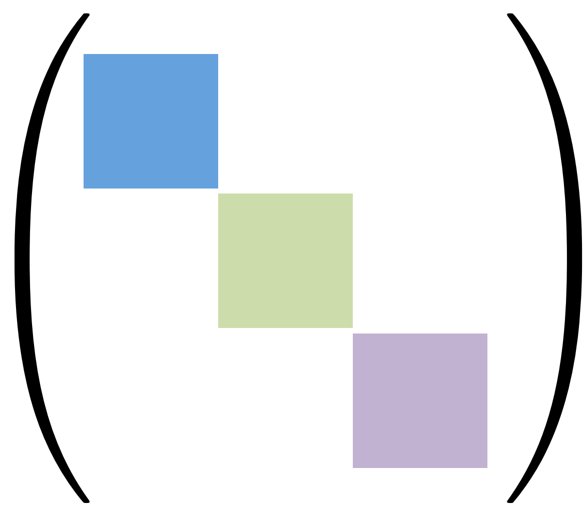
Anderson's Paving Conjecture '79

For all $\epsilon > 0$ there is a k so that for every
n-by-n symmetric matrix A with zero diagonals,

there is a partition of $\{1, ..., n\}$ into $S_1, ..., S_k$

$$\|A(S_j, S_j)\| \leq \epsilon \|A\| \quad \text{for} \quad j = 1, \ldots, k$$

Recall $\|A\| = \max_{\|x\|=1} \|Ax\|$



Anderson's Paving Conjecture '79

For all $\epsilon > 0$ there is a k so that for every self-adjoint bounded linear operator A on ℓ_2,

there is a partition of $\mathbb{N}$ into $S_1, ..., S_k$

$$\|A(S_j, S_j)\| \leq \epsilon \|A\| \quad \text{for} \quad j = 1, \ldots, k$$

$$\|A\| = \sup_{\|x\|=1} \|Ax\|$$

Anderson's Paving Conjecture '79

For all $\epsilon > 0$ there is a k so that for every n-by-n symmetric matrix A with zero diagonals,

there is a partition of $\{1, ..., n\}$ into $S_1, ..., S_k$

$$\|A(S_j, S_j)\| \leq \epsilon \|A\| \quad \text{for} \quad j = 1, \ldots, k$$

Is equivalent if restrict to projection matrices.
[Casazza, Edidin, Kalra, Paulsen '07]

Anderson's Paving Conjecture '79

Equivalent to [Harvey '13]:

There exist an $\epsilon > 0$ and a k so that for $v_1, ..., v_n \in \mathbb{C}^d$

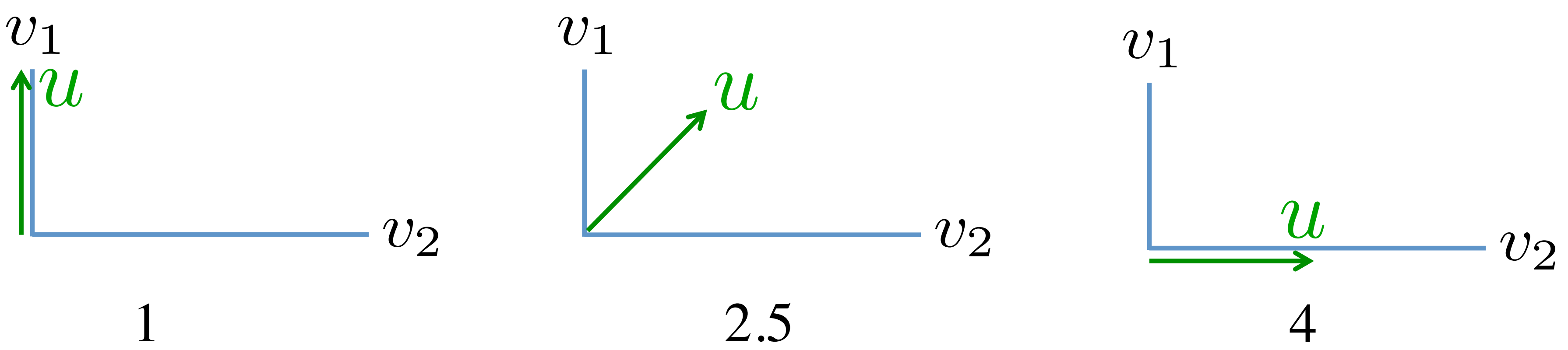
such that $\|v_i\|^2 \leq 1/2$ and $\sum v_i v_i^* = I$

then exists a partition of $\{1, \ldots, n\}$ into k parts s.t.

$$\left\| \sum_{i \in S_j} v_i v_i^* \right\| \leq 1 - \epsilon$$

Moments of Vectors

The moment of vectors $v_1, ..., v_n$ in the direction of a unit vector u is $$\sum_i (v_i^T u)^2$$



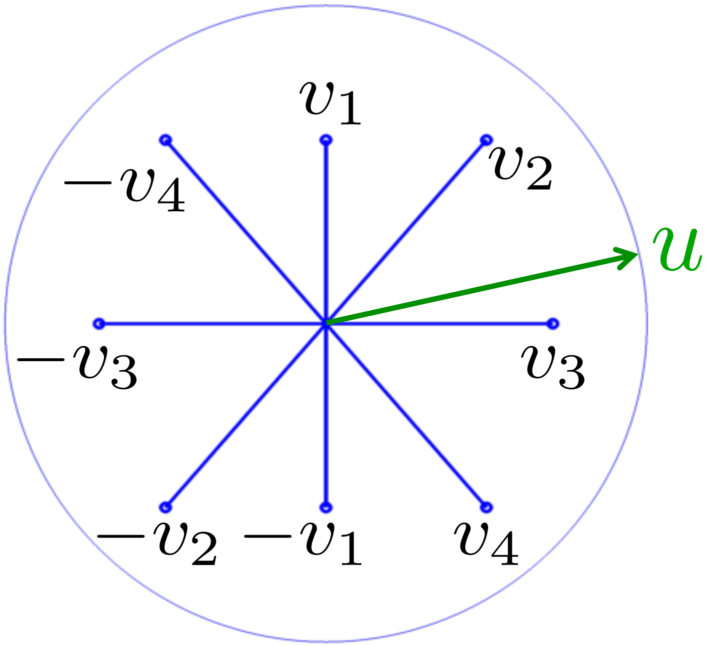
Moments of Vectors

The moment of vectors $v_1, ..., v_n$
in the direction of a unit vector u is

$$\sum_i (v_i^T u)^2 = \sum_i u^T (v_i v_i^T) u$$

$$= u^T \left(\sum_i v_i v_i^T \right) u$$

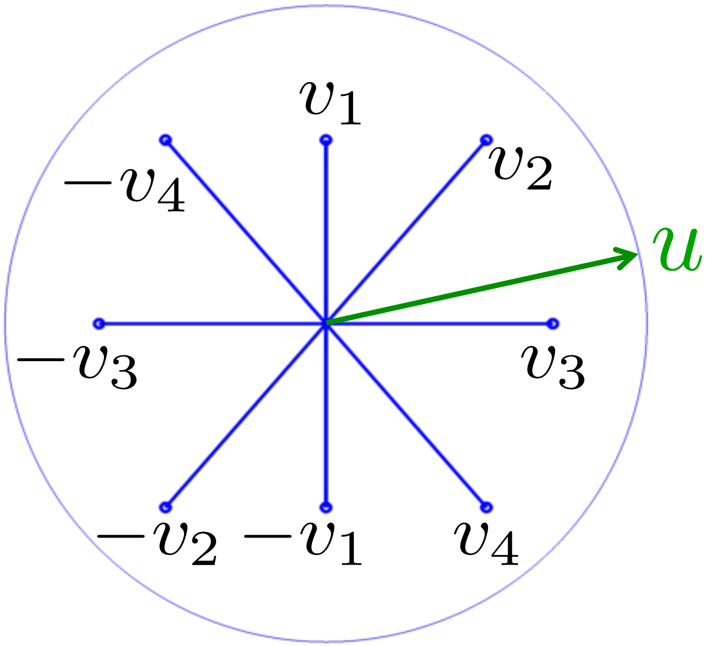
Vectors with Spherical Moments



For every unit vector u

$$\sum_i (v_i^T u)^2 = 1$$

Vectors with Spherical Moments



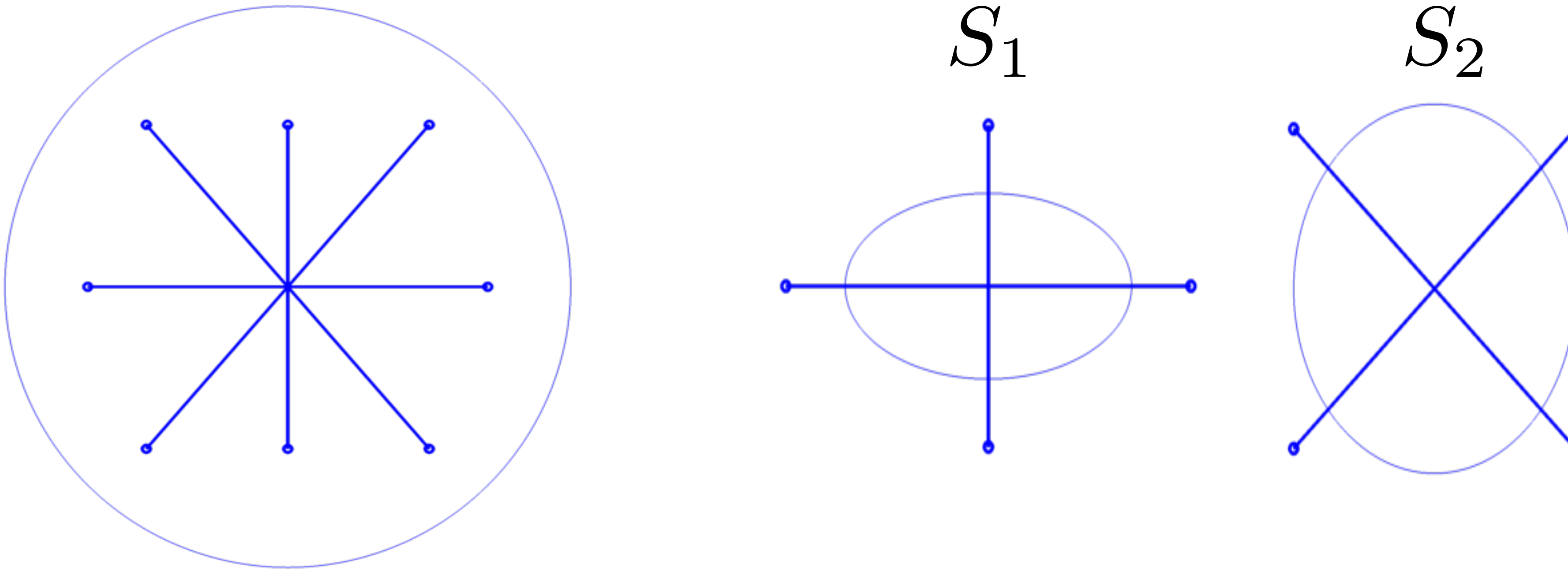
For every unit vector u

$$\sum_i (v_i^T u)^2 = 1$$

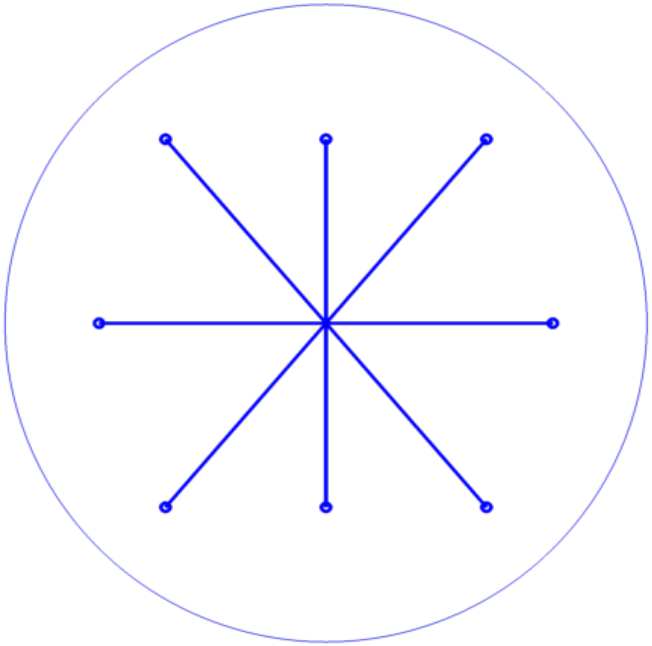
$$\sum_i v_i v_i^T = I$$

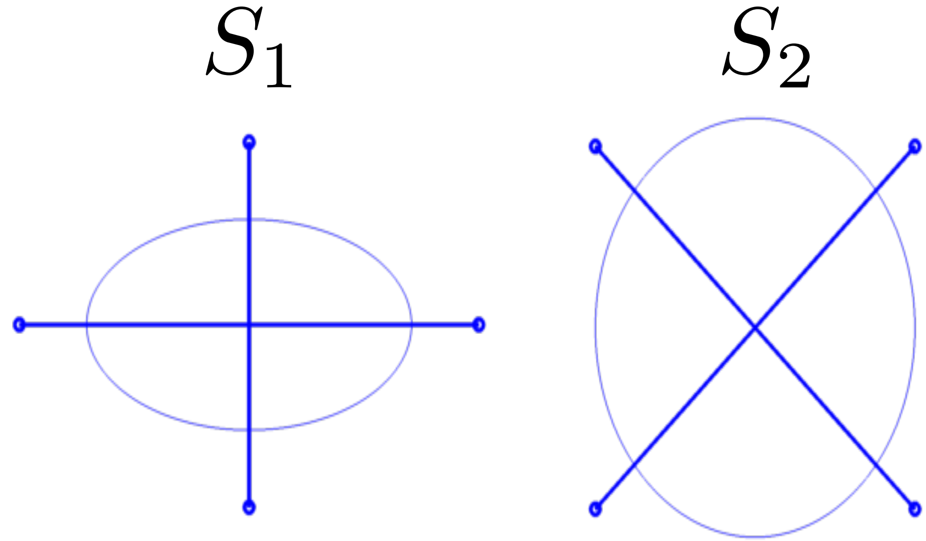
Also called *isotropic position*

Partition into Approximately ½-Spherical Sets



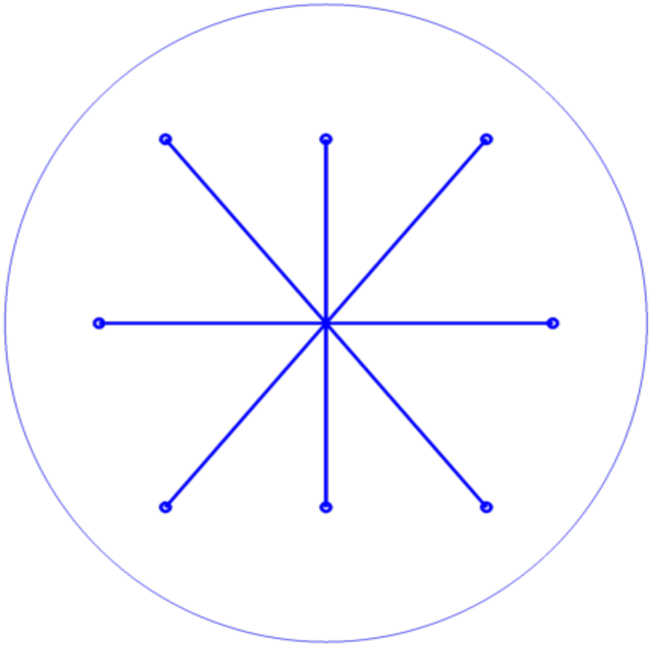
Partition into Approximately ½-Spherical Sets

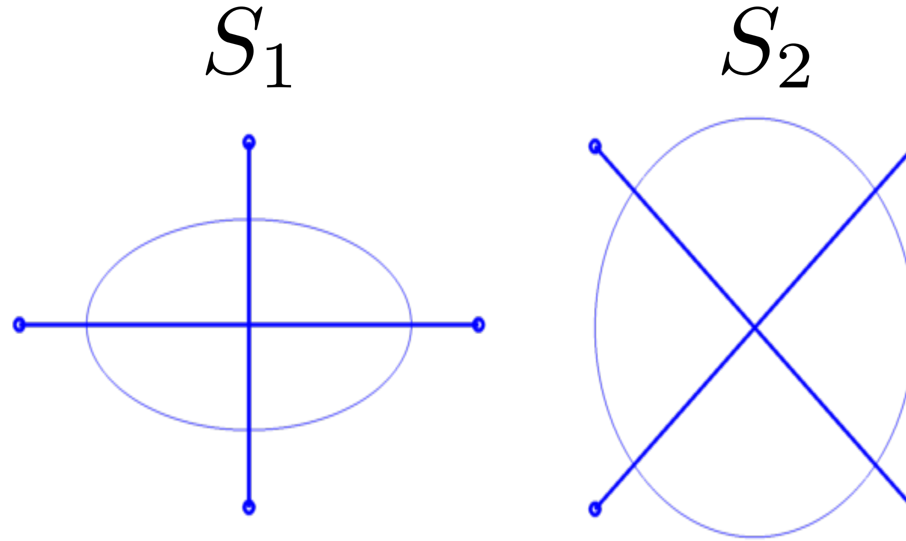




$$1/4 \leq \sum_{i \in S_j} (v_i^T u)^2 \leq 3/4$$

Partition into Approximately ½-Spherical Sets

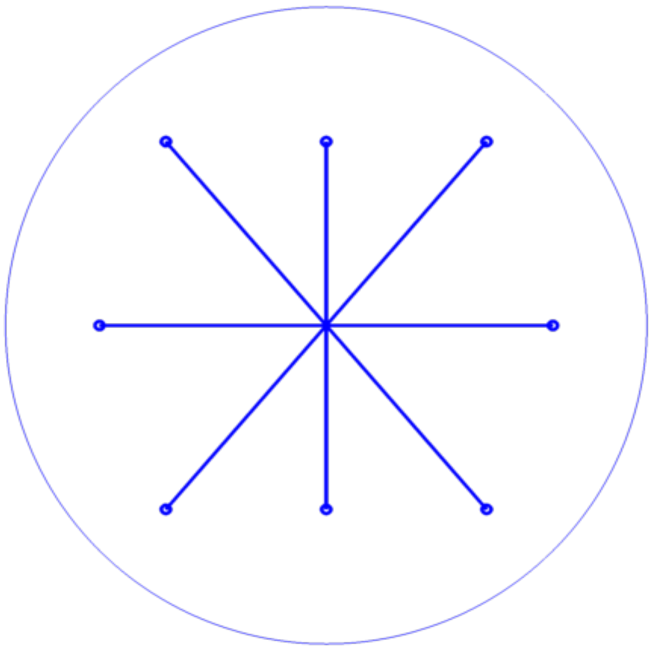


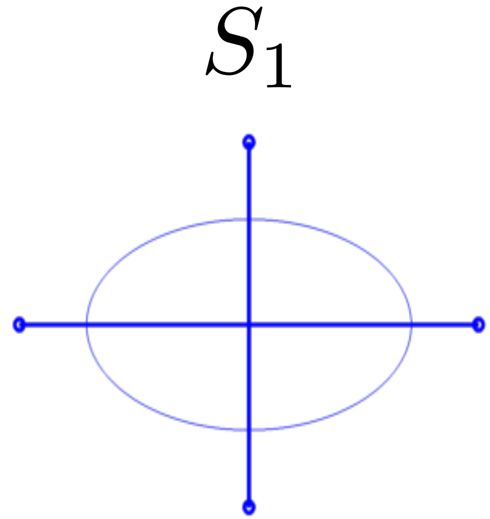


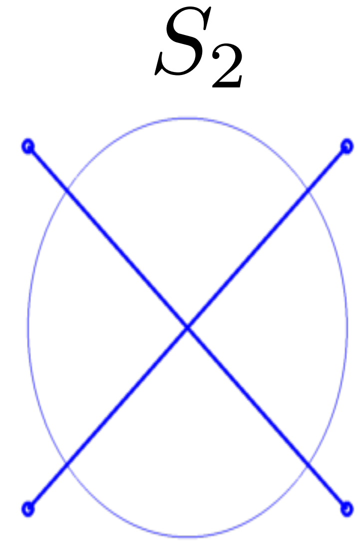
$$1/4 \leq \sum_{i \in S_j} (v_i^T \textcolor{green}{u})^2 \leq 3/4$$

$$1/4 \leq \mathrm{eigs}(\sum_{i \in S_j} v_i v_i^T) \leq 3/4$$

Partition into Approximately ½-Spherical Sets







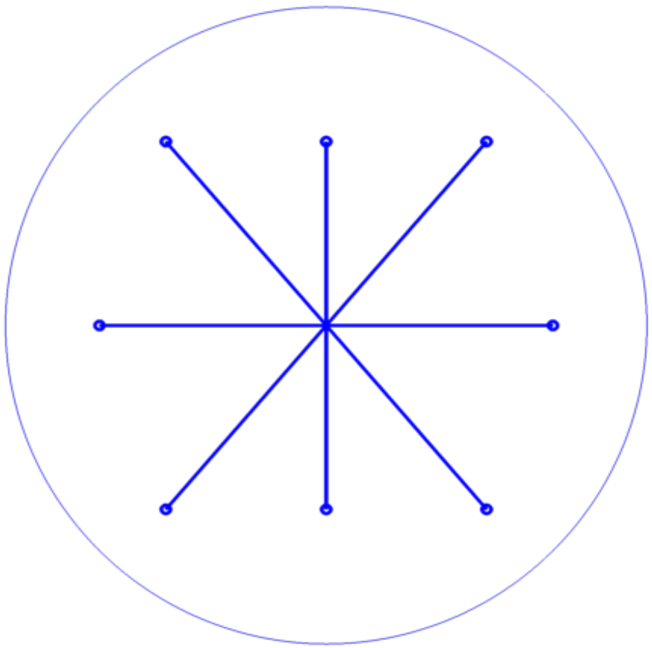
$$1/4 \leq \sum_{i \in S_j} (v_i^T u)^2 \leq 3/4$$

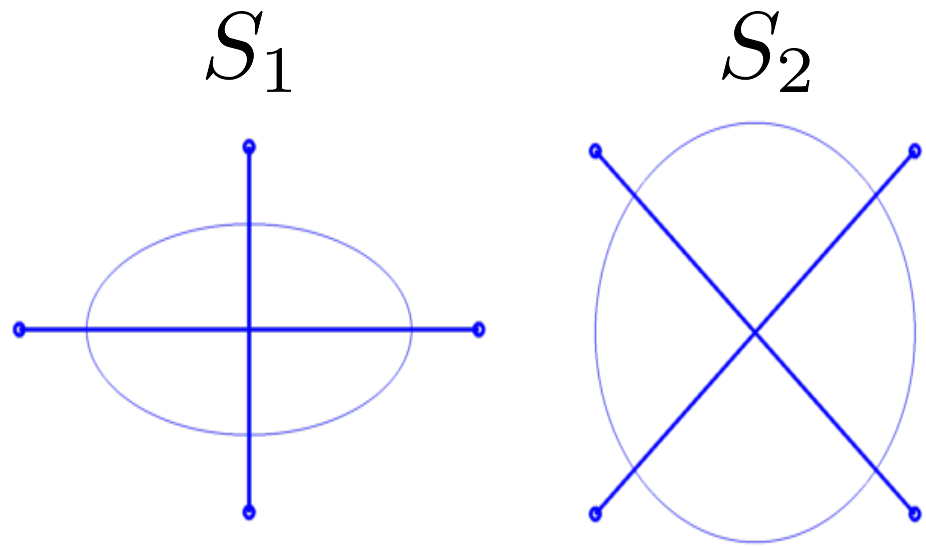
$$1/4 \leq \text{eigs}\left(\sum_{i \in S_j} v_i v_i^T\right) \leq 3/4$$

$$\iff \quad \text{eigs}\left(\sum_{i \in S_j} v_i v_i^T\right) \leq 3/4$$

because $\displaystyle \sum_{i \in S_1} v_i v_i^T = I - \sum_{i \in S_2} v_i v_i^T$

Partition into Approximately ½-Spherical Sets





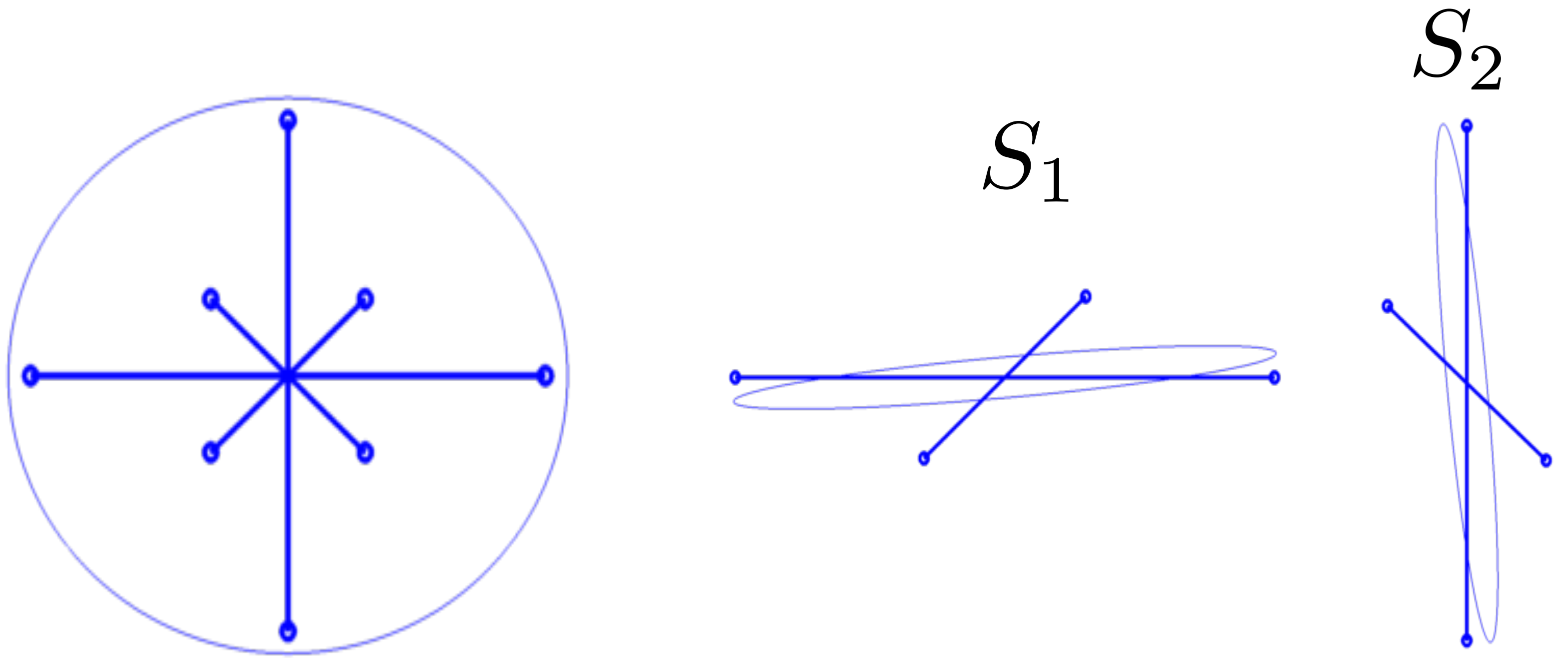
$$1/4 \leq \sum_{i \in S_j} (v_i^T \color{green}{u}\color{black})^2 \leq 3/4$$

$$1/4 \leq \operatorname{eigs}\left(\sum_{i \in S_j} v_i v_i^T\right) \leq 3/4$$

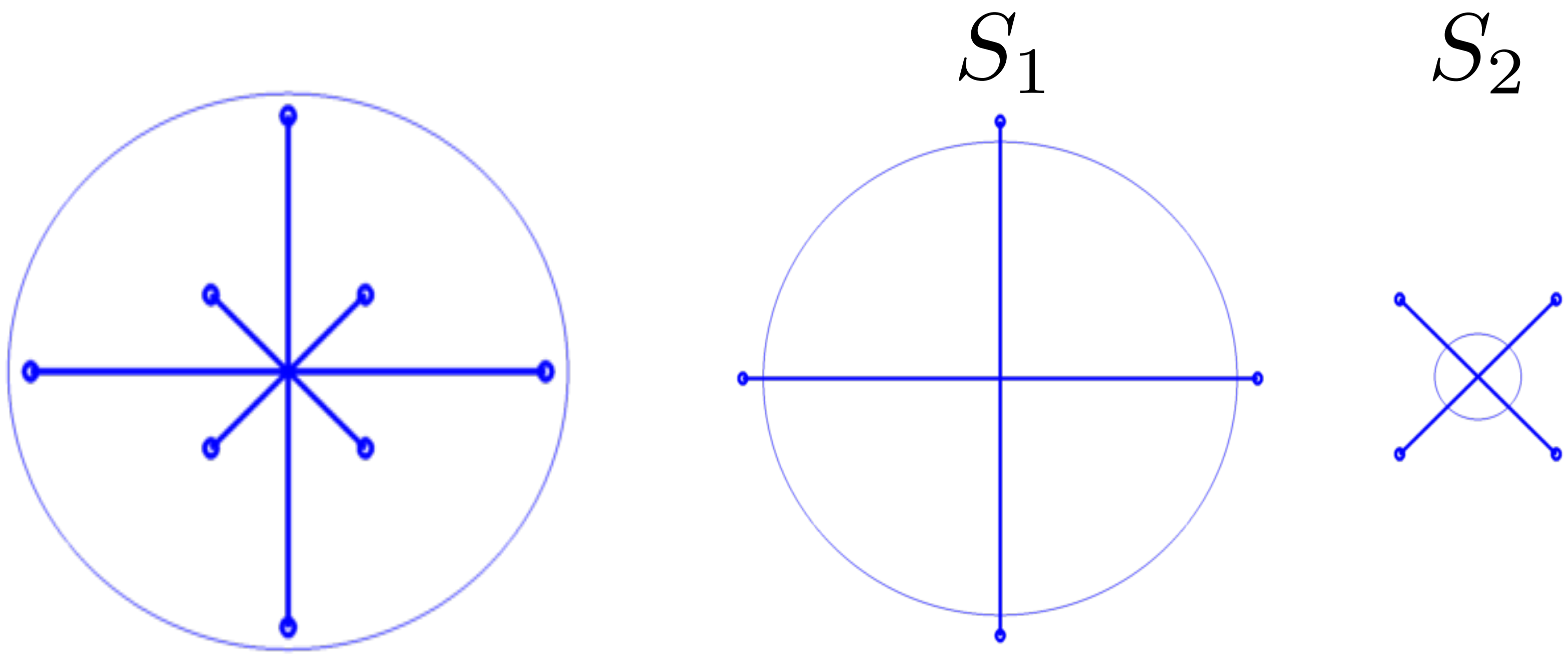
$$\iff \quad \left\| \sum_{i \in S_j} v_i v_i^T \right\| \leq 3/4$$

because $$\sum_{i \in S_1} v_i v_i^T = I - \sum_{i \in S_2} v_i v_i^T$$

Big vectors make this difficult



Big vectors make this difficult



Weaver's Conjecture KS_2

There exist positive constants α and ϵ so that

if all $\|v_i\| \leq \alpha$

then exists a partition into S_1 and S_2 with

$$\text{eigs}(\sum_{i \in S_j} v_i v_i^*) \leq 1 - \epsilon$$

Weaver's Conjecture KS_2

There exist positive constants α and ϵ so that

if all $\|v_i\| \leq \alpha$

then exists a partition into S_1 and S_2 with

$$\mathrm{eigs}(\sum_{i \in S_j} v_i v_i^*) \leq 1 - \epsilon$$

Implies Akemann-Anderson Paving Conjecture, which implies Kadison-Singer

Main Theorem

For all $\alpha > 0$

if all $\|v_i\| \leq \alpha$

then exists a partition into S_1 and S_2 with

$$\text{eigs}(\sum_{i \in S_j} v_i v_i^*) \leq \frac{1}{2} + 3\alpha$$

Implies Akemann-Anderson Paving Conjecture, which implies Kadison-Singer

A Random Partition?

Works with high probability if all $\|v_i\|^2 \leq O(1/\log d)$

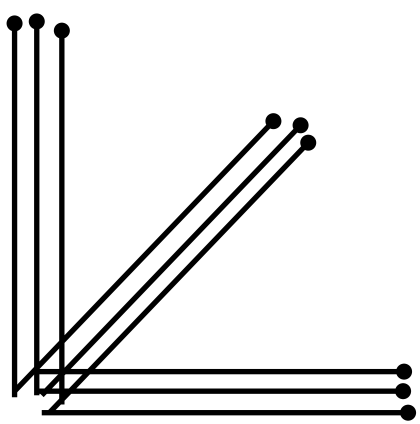
(by Tropp '11, variant of Matrix Chernoff, Rudelson)

A Random Partition?

Works with high probability if all $\|v_i\|^2 \leq O(1/\log d)$
(by Tropp '11, variant of Matrix Chernoff, Rudelson)

Troublesome case: each $\|v_i\| = \alpha$
is a scaled axis vector

are $1/\alpha^2$ of each

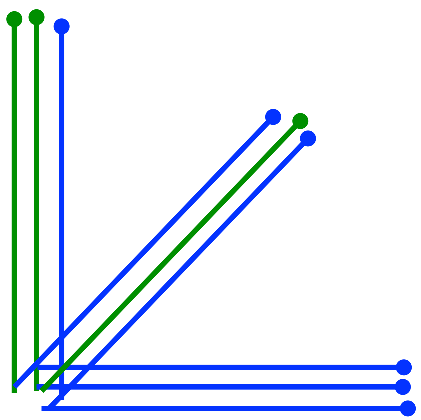


A Random Partition?

Works with high probability if all $\|v_i\|^2 \leq O(1/\log d)$
(by Tropp '11, variant of Matrix Chernoff, Rudelson)

Troublesome case: each $\|v_i\| = \alpha$ is a scaled axis vector

are $1/\alpha^2$ of each

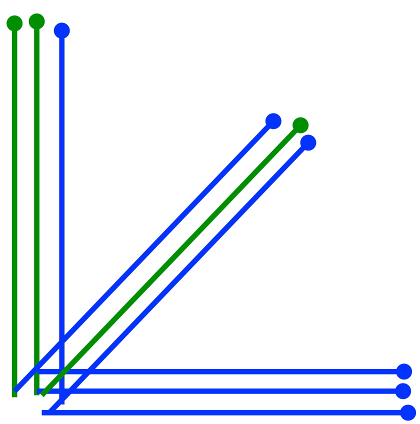


chance that all in one direction land in same set is $2^{-1/\alpha^2}$

A Random Partition?

Works with high probability if all $\|v_i\|^2 \leq O(1/\log d)$
(by Tropp '11, variant of Matrix Chernoff, Rudelson)

Troublesome case: each $\|v_i\| = \alpha$
is a scaled axis vector



are $1/\alpha^2$ of each

chance that all in one direction
land in same set is $2^{-1/\alpha^2}$

Chance there exists a direction in which
all land in same set is

$$1 - \left(1 - 2^{-1/\alpha^2}\right)^d \rightarrow 1$$

The Graphical Case

From a graph G = (V,E) with |V| = n and |E| = m
Create m vectors in n dimensions:

$$v_{a,b}(c) = \begin{cases} 1 & \text{if } c = a \\ -1 & \text{if } c = b \\ 0 & \text{otherwise} \end{cases}$$

$$\sum_{(a,b)\in E} v_{a,b} v_{a,b}^T = L_G$$

If G is a good d-regular expander, all eigs close to d

very close to spherical

Partitioning Expanders

Can partition the edges of a good expander to obtain two expanders.

Broder-Frieze-Upfal '94:
construct random partition guaranteeing degree at least d/4, some expansion

Frieze-Molloy '99: Lovász Local Lemma,
good expander

Probability is works is low, but can prove non-zero

Interlacing Families of Polynomials

A new technique for proving existence
from very low probabilities

Restricted Invertibility (Bourgain-Tzafriri)

Special case:

For $v_1, ..., v_n \in \mathbb{C}^d$ with $\sum_i v_i v_i^* = I$

for every $k \leq d$ there is a $S \subset \{1, ..., n\}$, $|S| = k$

so that

$$\lambda_k \left(\sum_{i \in S} v_i v_i^* \right) \geq \left(1 - \sqrt{\frac{k}{d}} \right)^2 \frac{d}{n}$$

Restricted Invertibility (Bourgain-Tzafriri)

Special case:

For $v_1, ..., v_n \in \mathbb{C}^d$ with $\sum_i v_i v_i^* = I$

for every $k \leq d$ there is a $S \subset \{1, ..., n\}$, $|S| = k$

so that

$$\lambda_k \left(\sum_{i \in S} v_i v_i^* \right) \geq \left(1 - \sqrt{\frac{k}{d}} \right)^2 \frac{d}{n}$$

Is far from singular on the span of $\{v_i\}_{i \in S}$

Restricted Invertibility (Bourgain-Tzafriri)

For $v_1, ..., v_n \in \mathbb{C}^d$ with $\sum_i v_i v_i^* = I$

for every $k \leq d$ there is a $S \subset \{1, ..., n\}, |S| = k$

so that $\lambda_k \left(\sum_{i \in S} v_i v_i^*\right) \geq \left(1 - \sqrt{\frac{k}{d}}\right)^2 \frac{d}{n}$

For $k = 1$ says $\lambda_1(vv^*) \gtrsim \frac{d}{n}$,

while $\lambda_1(vv^*) = v^* v = \|v\|^2 \approx \frac{d}{n}$

Similar bound for k a constant fraction of d!

Method of proof

Let $r_1, \ldots, r_k$ be chosen uniformly from $\{v_1, \ldots, v_n\}$

1. $\mathbb{E}\,\chi\left[\sum r_j r_j^*\right](x)$ is real rooted

the characteristic polynomial in the variable x of the matrix inside the brackets

Method of proof

Let $r_1, \ldots, r_k$ be chosen uniformly from $\{v_1, \ldots, v_n\}$

1. $\mathbb{E}\,\chi\left[\sum r_j r_j^*\right](x)$ is real rooted

2. $\lambda_k\left(\mathbb{E}\,\chi\left[\sum r_j r_j^*\right](x)\right) \geq \left(1 - \sqrt{\frac{k}{d}}\right)^2 \frac{d}{n}$

the k-th root of the polynomial

Method of proof

Let $r_1, ..., r_k$ be chosen uniformly from $\{v_1, ..., v_n\}$

1. $\mathbb{E}\,\chi\left[\sum r_j r_j^*\right](x)$ is real rooted

2. $\lambda_k\left(\mathbb{E}\,\chi\left[\sum r_j r_j^*\right](x)\right) \geq \left(1 - \sqrt{\frac{k}{d}}\right)^2 \frac{d}{n}$

3. With non-zero probability

$$\lambda_k\left(\chi\left[\sum r_j r_j^*\right](x)\right) \geq \lambda_k\left(\mathbb{E}\,\chi\left[\sum r_j r_j^*\right](x)\right)$$

Because is an interlacing family of polynomials

Method of proof

Let $r_1, ..., r_k$ be chosen uniformly from $\{v_1, ..., v_n\}$

1. $\mathbb{E}\chi\left[\sum r_j r_j^*\right](x)$ is real rooted

2. $\lambda_k\left(\mathbb{E}\chi\left[\sum r_j r_j^*\right](x)\right) \geq \left(1 - \sqrt{\frac{k}{d}}\right)^2 \frac{d}{n}$

3. With non-zero probability

$$\lambda_k\left(\chi\left[\sum r_j r_j^*\right](x)\right) \geq \lambda_k\left(\mathbb{E}\chi\left[\sum r_j r_j^*\right](x)\right)$$

Because is an interlacing family of polynomials

Rank-1 updates of characteristic polynomials

As $\sum_i v_i v_i^* = I$, $\mathbb{E}\, r_j r_j^* = \frac{1}{n} I$

Lemma: For a symmetric matrix A,

$$\mathbb{E}\, \chi \left[A + r_j r_j^* \right] (x) = (1 - \tfrac{1}{n} \partial_x) \chi[A](x)$$

Rank-1 updates of characteristic polynomials

As $\sum_i v_i v_i^* = I$, $\mathbb{E}\, r_j r_j^* = \frac{1}{n} I$

Lemma: For a symmetric matrix A,

$$\mathbb{E}\, \chi\left[A + r_j r_j^*\right](x) = (1 - \tfrac{1}{n}\partial_x)\chi[A](x)$$

Proof: follows from rank-1 update for determinants:

$$\det(A + uu^*) = \det(A)(1 + u^* A^{-1} u)$$

The expected characteristic polynomial

Lemma: For a symmetric matrix A,

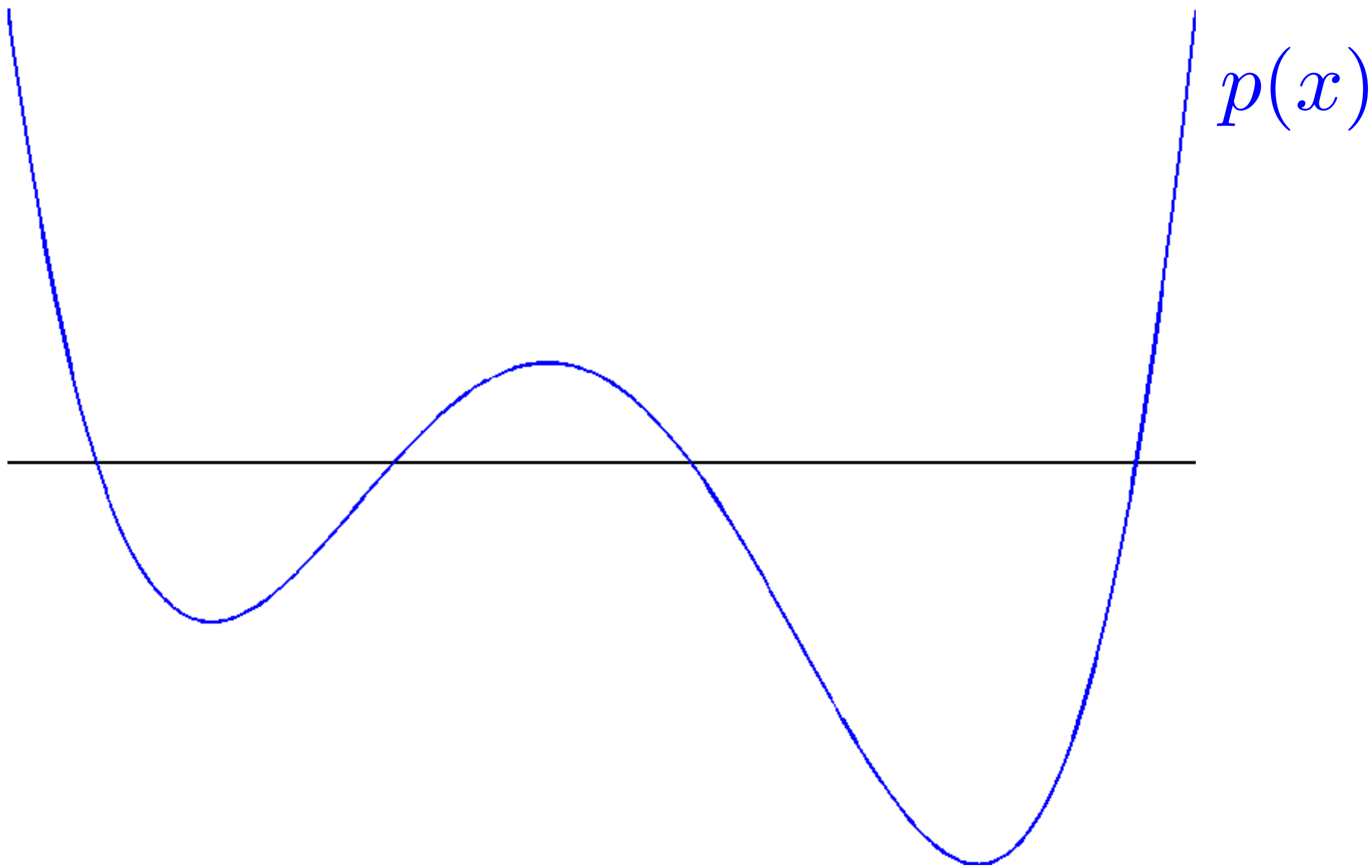
$$\mathbb{E}\,\chi\left[A + r_j r_j^*\right](x) = (1 - \tfrac{1}{n}\partial_x)\chi[A](x)$$

Corollary:

$$\mathbb{E}\,\chi\left[\sum_{j=1}^{k} r_j r_j^*\right](x) = (1 - \tfrac{1}{n}\partial_x)^k x^d$$

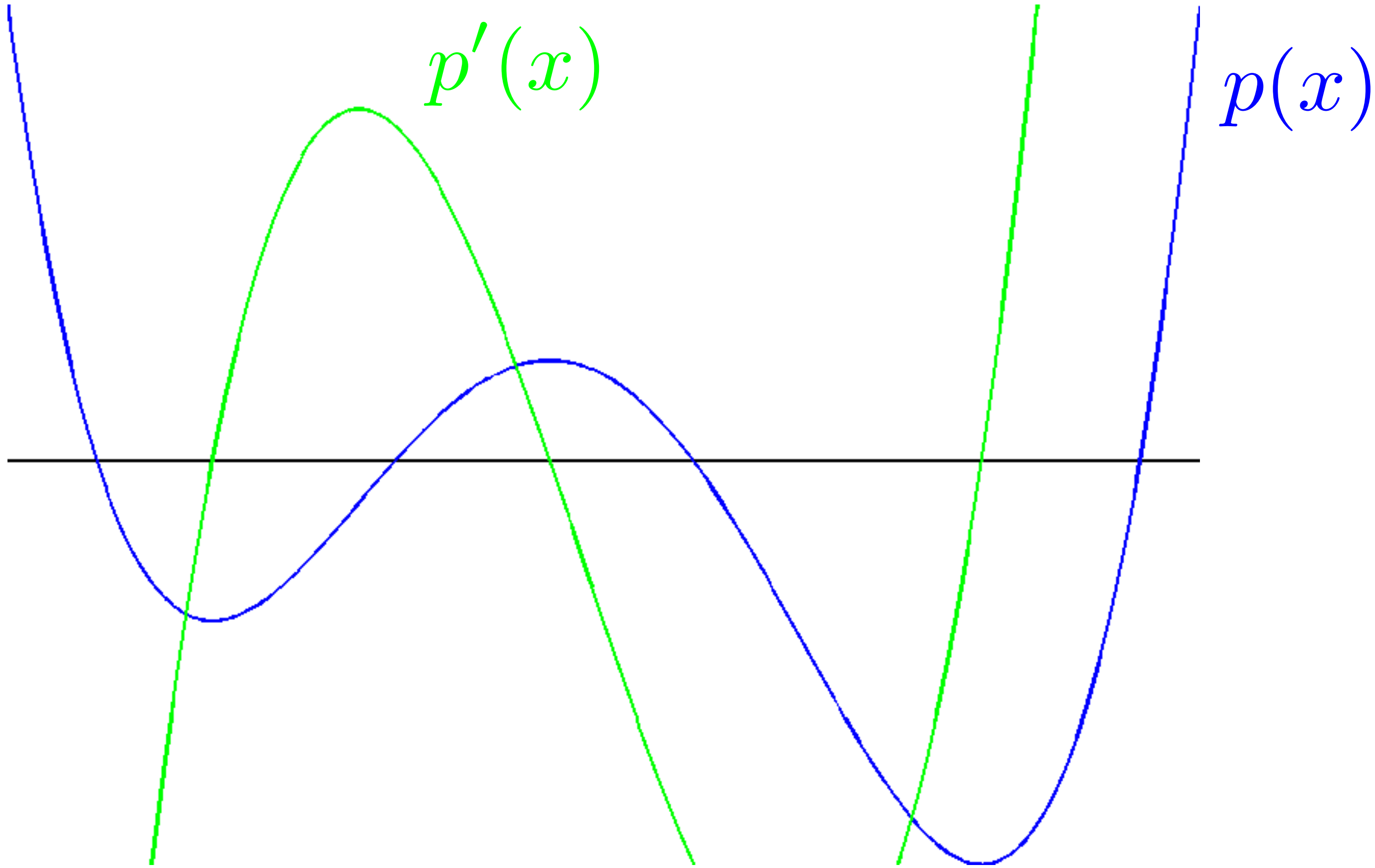
Real Roots

Lemma: if $p(x)$ is real rooted, so is $(1 - c\partial_x)p(x)$



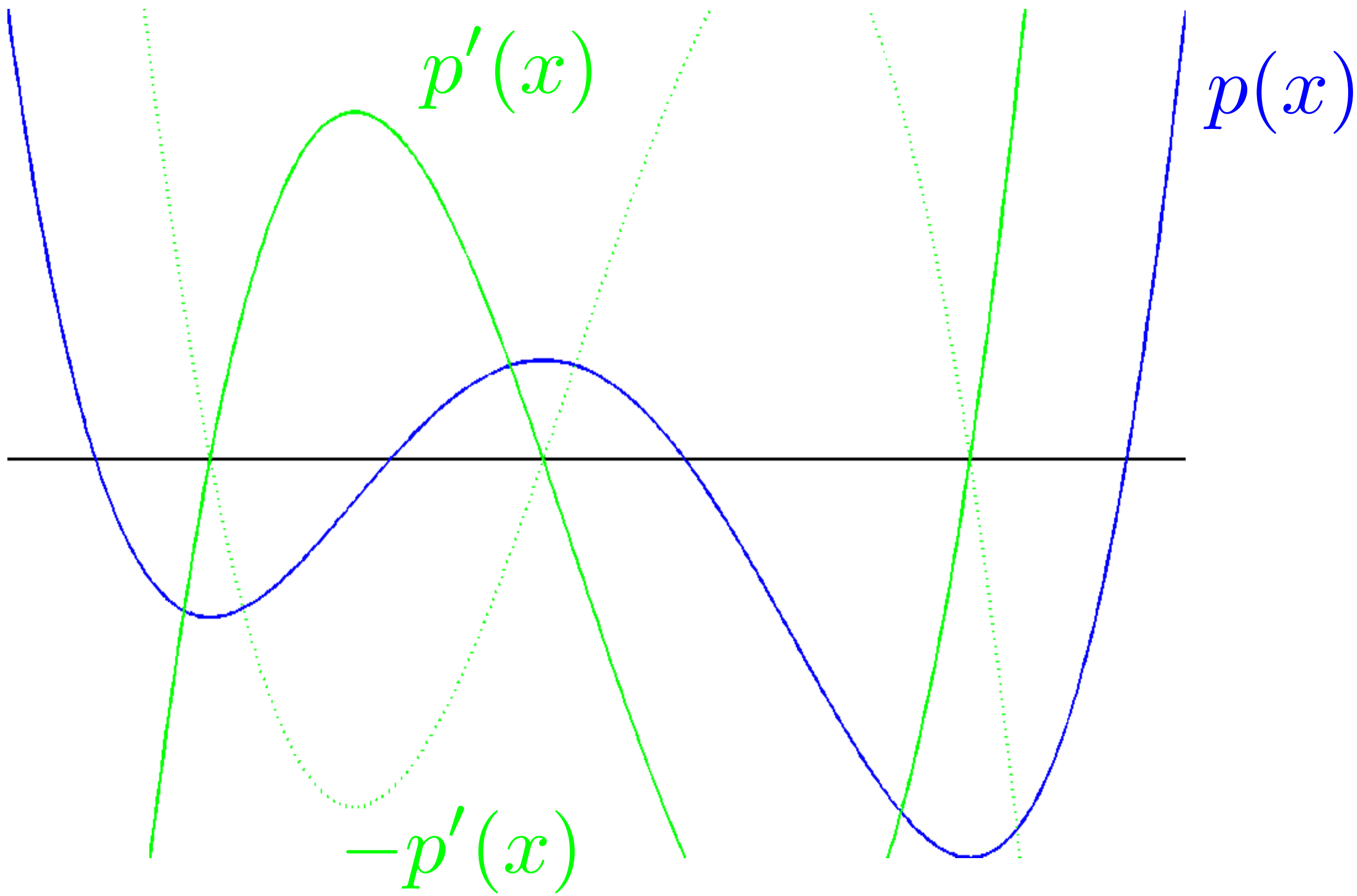
Real Roots

Lemma: if $p(x)$ is real rooted, so is $(1 - c\partial_x)p(x)$



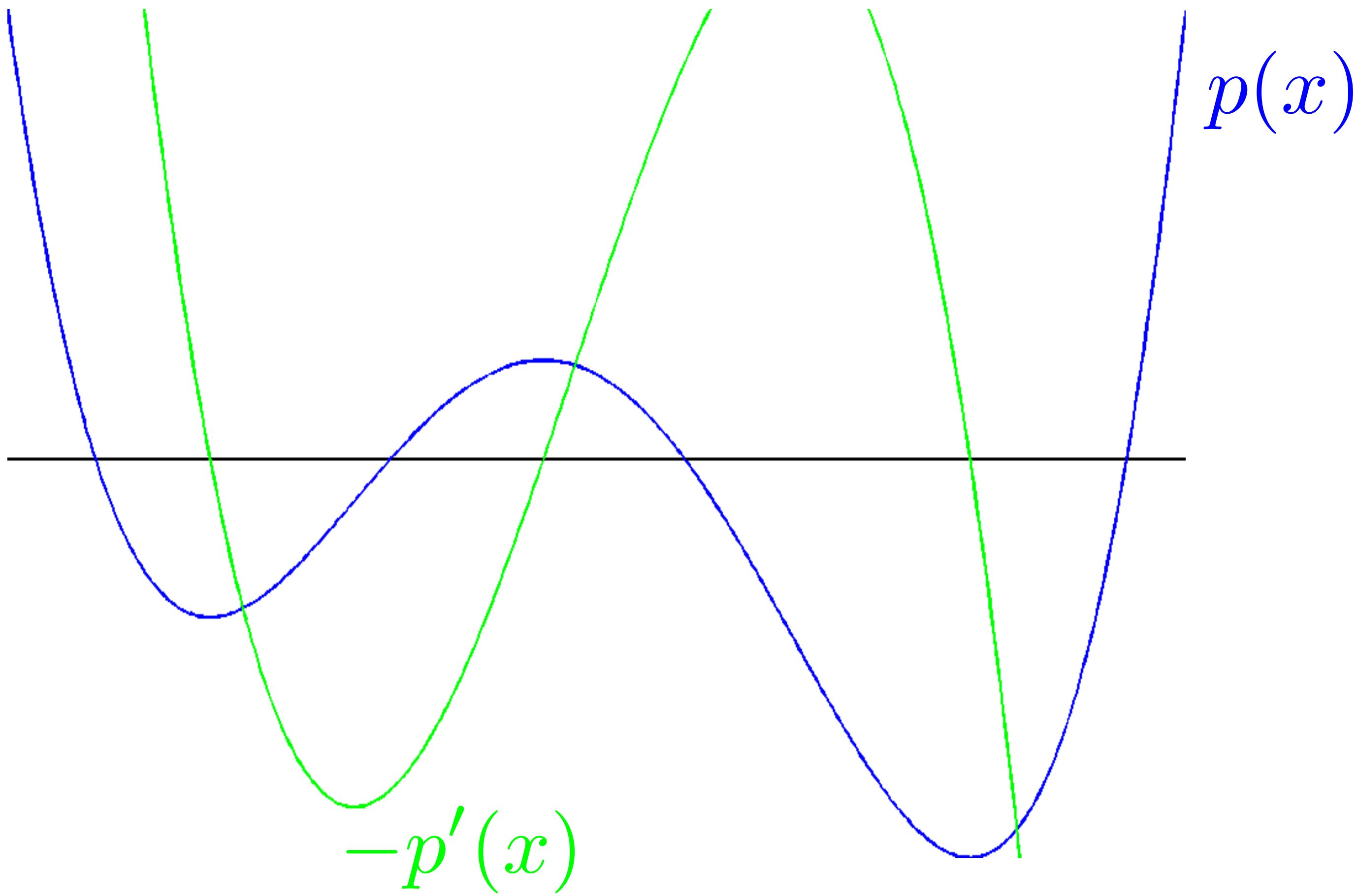
Real Roots

Lemma: if $p(x)$ is real rooted, so is $(1 - c\partial_x)p(x)$



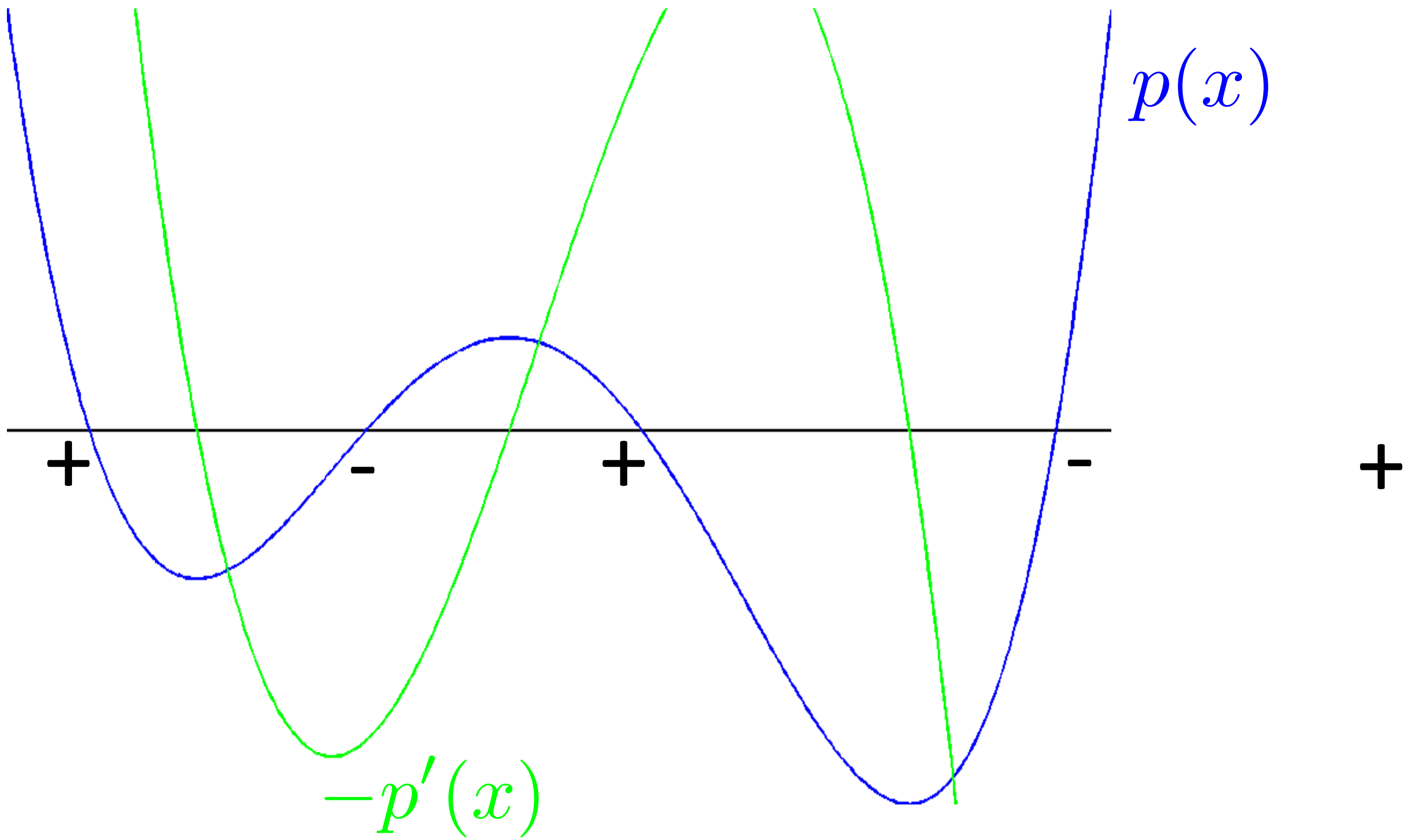
Real Roots

Lemma: if $p(x)$ is real rooted, so is $(1 - c\partial_x)p(x)$



Real Roots

Lemma: if $p(x)$ is real rooted, so is $(1 - c\partial_x)p(x)$



Real Roots

Lemma: if $p(x)$ is real rooted, so is $(1 - c\partial_x)p(x)$

$$\mathbb{E}\,\chi\left[\sum_{j=1}^{k} r_j r_j^*\right](x) = (1 - \tfrac{1}{n}\partial_x)^k x^d$$

So, $\mathbb{E}\,\chi\left[\sum_{j=1}^{k} r_j r_j^*\right](x)$ is real rooted

Lower bound on the kth root

$$\mathbb{E}\,\chi\left[\sum_{j=1}^{k} r_j r_j^*\right](x) = (1 - \tfrac{1}{n}\partial_x)^k x^d$$

$$= x^{d-k}(1 - \tfrac{1}{n}\partial_x)^d x^k$$

$$= x^{d-k} L_k^{d-k}(nx)$$

a scaled associated Laguerre polynomial

Lower bound on the kth root

$$\mathbb{E}\,\chi\left[\sum_{j=1}^{k} r_j r_j^*\right](x) = x^{d-k} L_k^{d-k}(nx)$$

a scaled associated Laguerre polynomial

Let R be a k-by-d matrix of independent $\mathcal{N}(0, 1/n)$

$$\mathbb{E}\,\chi[RR^T] = L_k^{d-k}(nx)$$

$$\mathbb{E}\,\chi[R^T R] = x^{d-k} L_k^{d-k}(nx)$$

Lower bound on the kth root

$$\mathbb{E}\,\chi\left[\sum_{j=1}^{k} r_j r_j^*\right](x) = x^{d-k} L_k^{d-k}(nx)$$

a scaled associated Laguerre polynomial

$$\lambda_k\left(\mathbb{E}\,\chi\left[\sum r_j r_j^*\right](x)\right) \geq \left(1 - \sqrt{\frac{k}{d}}\right)^2 \frac{d}{n}$$

(Krasikov '06)

3. With non-zero probability

$$\lambda_k \left(\chi \left[\sum r_j r_j^* \right] (x) \right) \geq \lambda_k \left(\mathbb{E} \, \chi \left[\sum r_j r_j^* \right] (x) \right)$$

Proof: the polynomials

$$p_{i_1, i_2, \ldots, i_k}(x) = \chi \left[v_{i_1} v_{i_1}^* + \cdots + v_{i_k} v_{i_k}^* \right] (x)$$

form an interlacing family.

Interlacing

Polynomial $p(x) = \prod_{i=1}^{d}(x - \alpha_i)$

interlaces $q(x) = \prod_{i=1}^{d-1}(x - \beta_i)$

if $\alpha_1 \leq \beta_1 \leq \alpha_2 \leq \cdots \alpha_{d-1} \leq \beta_{d-1} \leq \alpha_d$

Interlacing

Polynomial $p(x) = \prod_{i=1}^{d}(x - \alpha_i)$

interlaces $q(x) = \prod_{i=1}^{d-1}(x - \beta_i)$

if $\alpha_1 \leq \beta_1 \leq \alpha_2 \leq \cdots \alpha_{d-1} \leq \beta_{d-1} \leq \alpha_d$

For example, $p(x)$ interlaces $\partial_x p(x)$

Interlacing

Polynomial $p(x) = \prod_{i=1}^{d} (x - \alpha_i)$

interlaces $q(x) = \prod_{i=1}^{d-1} (x - \beta_i)$

if $\alpha_1 \leq \beta_1 \leq \alpha_2 \leq \cdots \alpha_{d-1} \leq \beta_{d-1} \leq \alpha_d \leq \beta_d$

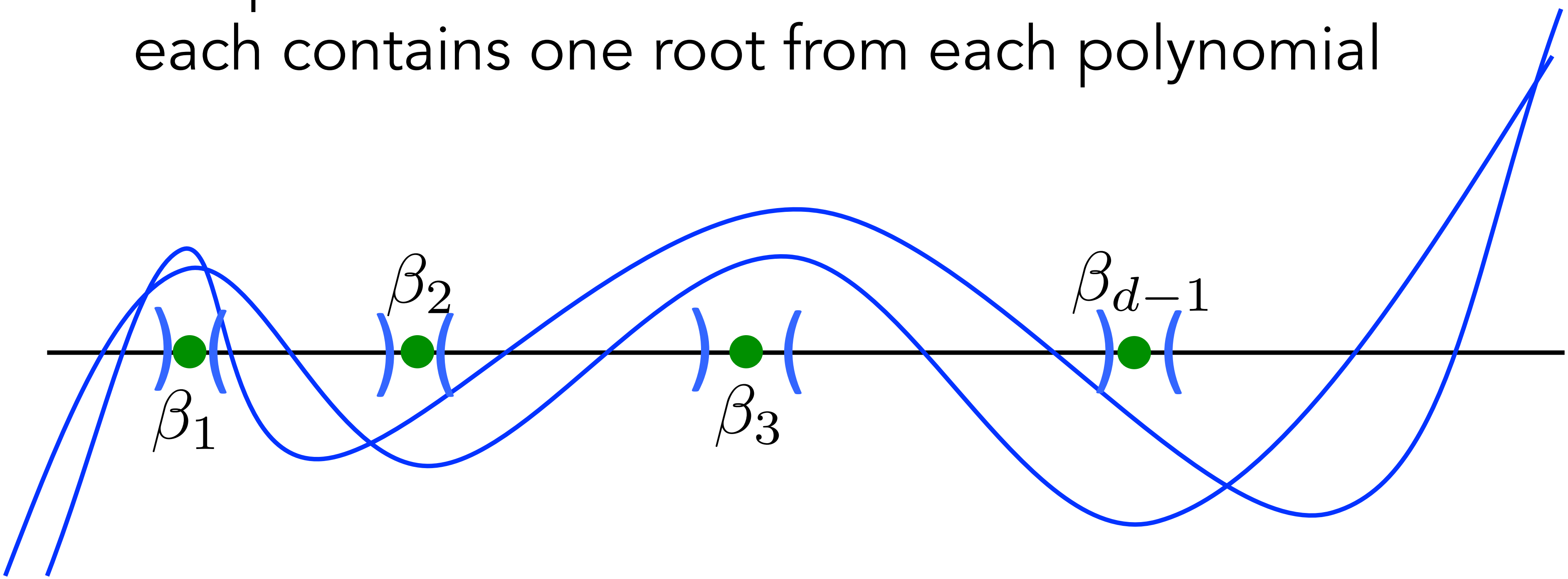
If generalize to allow same degree,

Cauchy's interlacing theorem says

$\chi[A](x)$ interlaces $\chi[A + vv^*](x)$

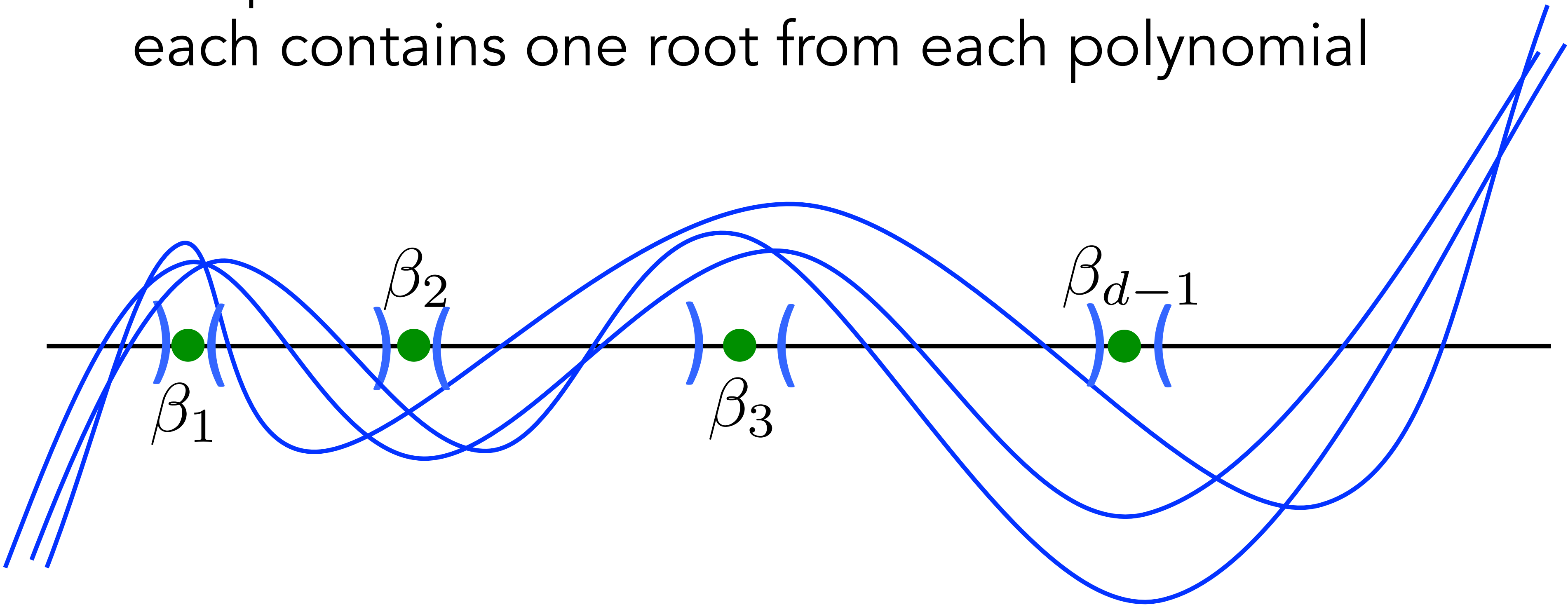
Common Interlacing

$p_1(x)$ and $p_2(x)$ have a common interlacing if can partition the line into intervals so that each contains one root from each polynomial



Common Interlacing

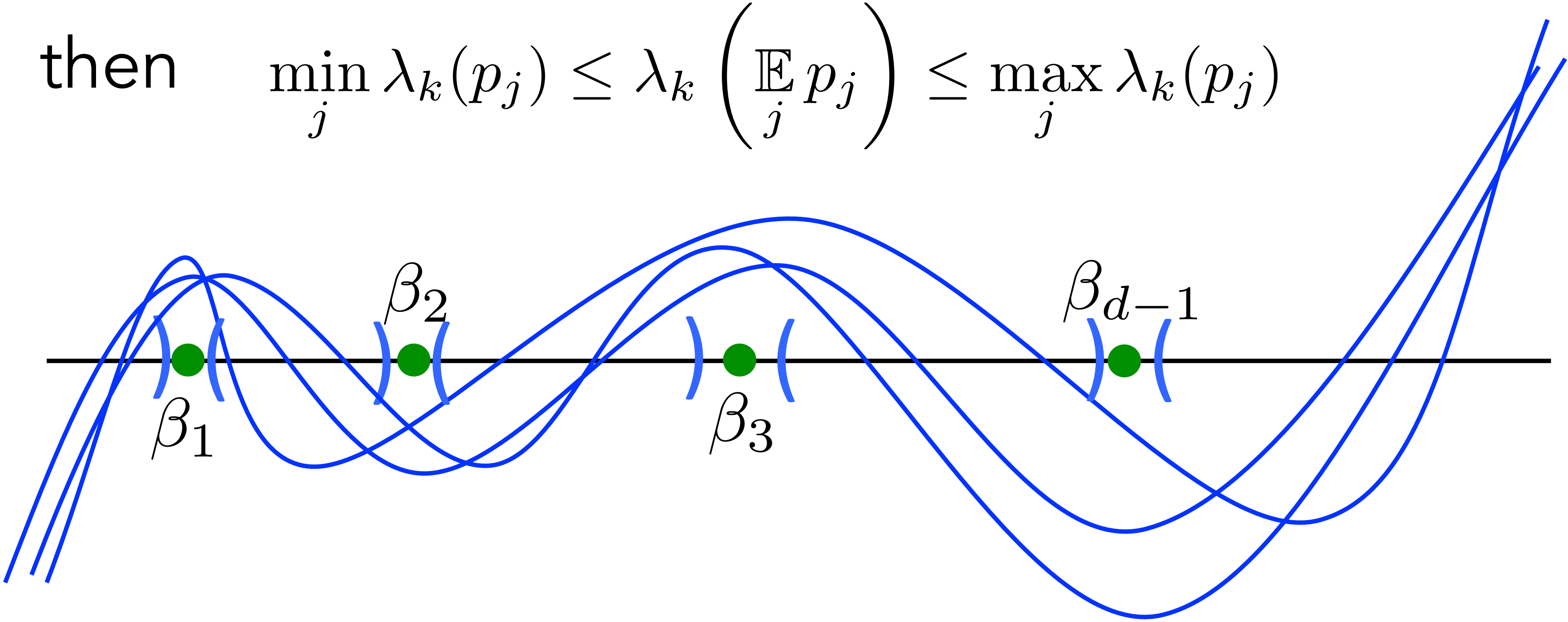
$p_1(x)$ and $p_2(x)$ have a common interlacing if can partition the line into intervals so that each contains one root from each polynomial



Common Interlacing

If $p_1(x), p_2(x), \ldots, p_n(x)$ have a common interlacing,

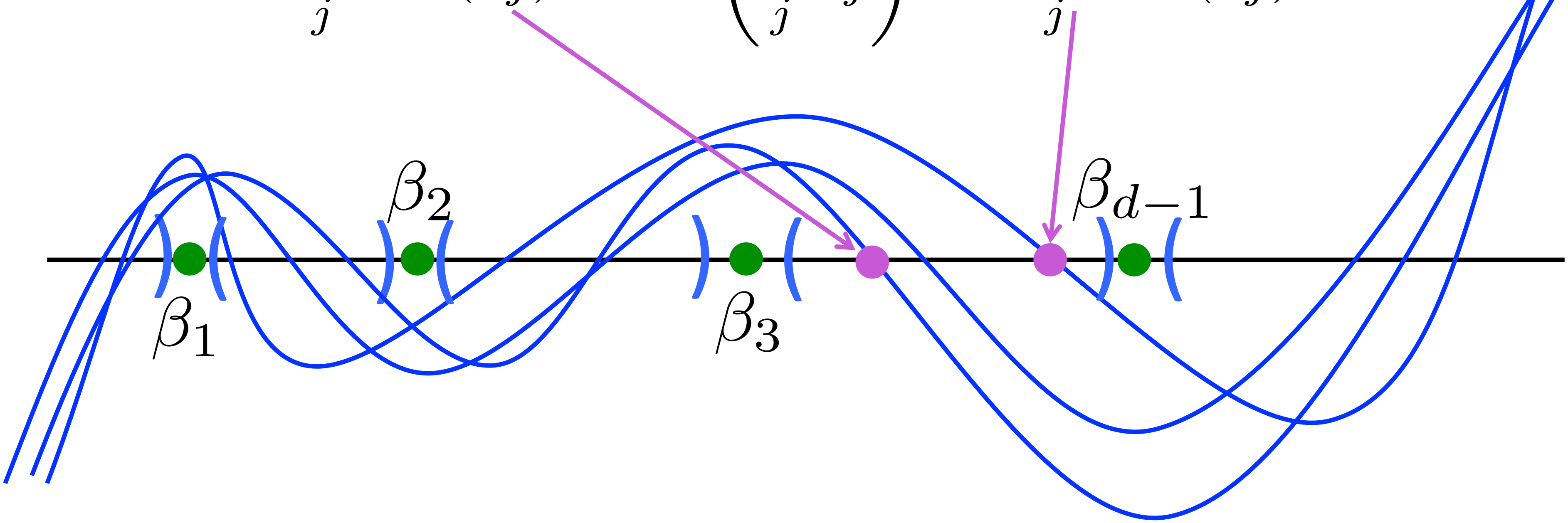
then $\min_j \lambda_k(p_j) \leq \lambda_k \left(\mathbb{E}_j p_j \right) \leq \max_j \lambda_k(p_j)$



Common Interlacing

If $p_1(x), p_2(x), \ldots, p_n(x)$ have a common interlacing,

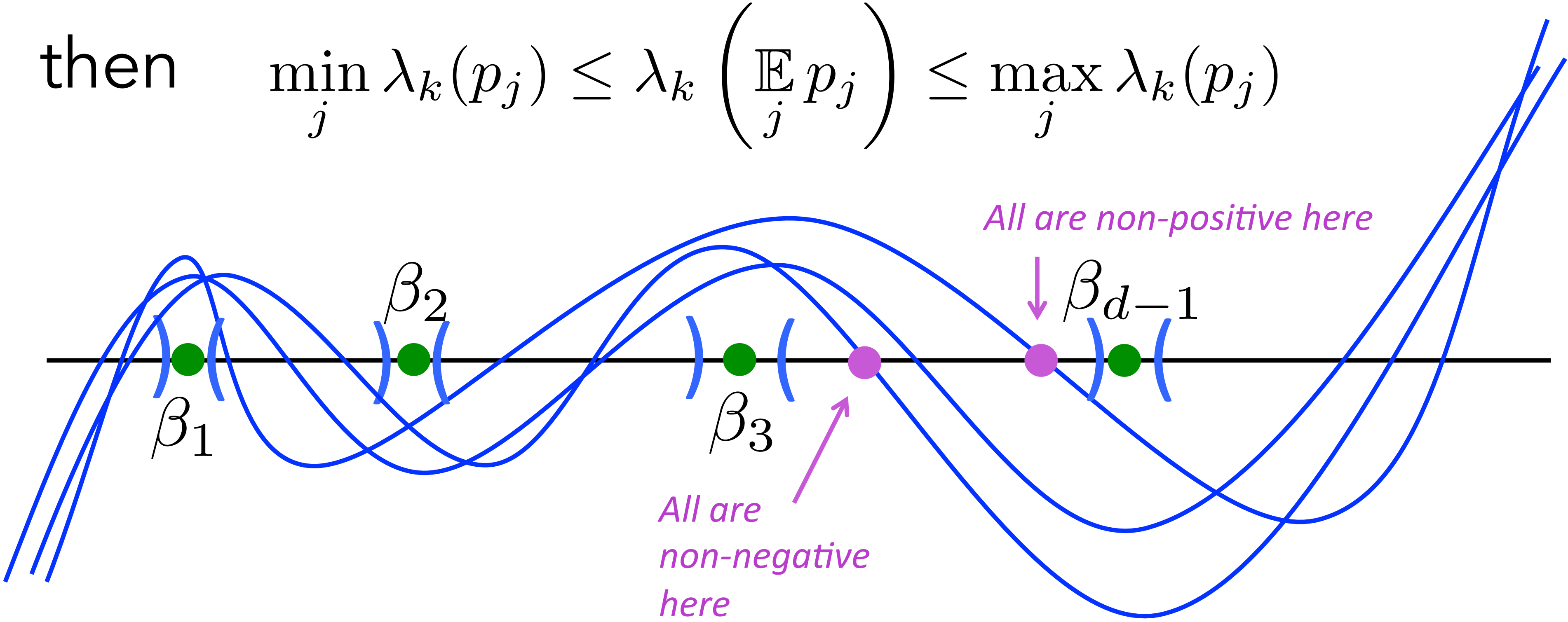
then $$\min_j \lambda_k(p_j) \leq \lambda_k\left(\mathbb{E}_j\, p_j\right) \leq \max_j \lambda_k(p_j)$$



Common Interlacing

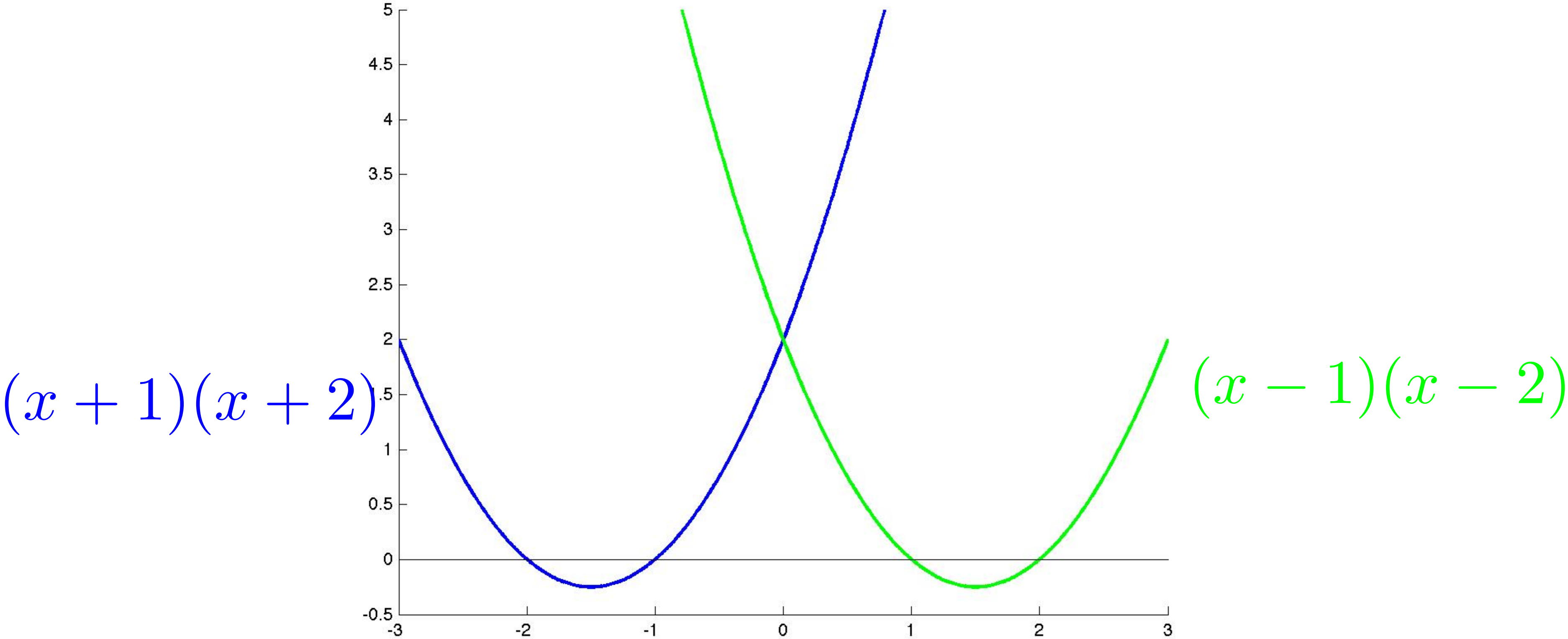
If $p_1(x), p_2(x), \ldots, p_n(x)$ have a common interlacing,

then $$\min_j \lambda_k(p_j) \leq \lambda_k\left(\mathbb{E}_j p_j\right) \leq \max_j \lambda_k(p_j)$$

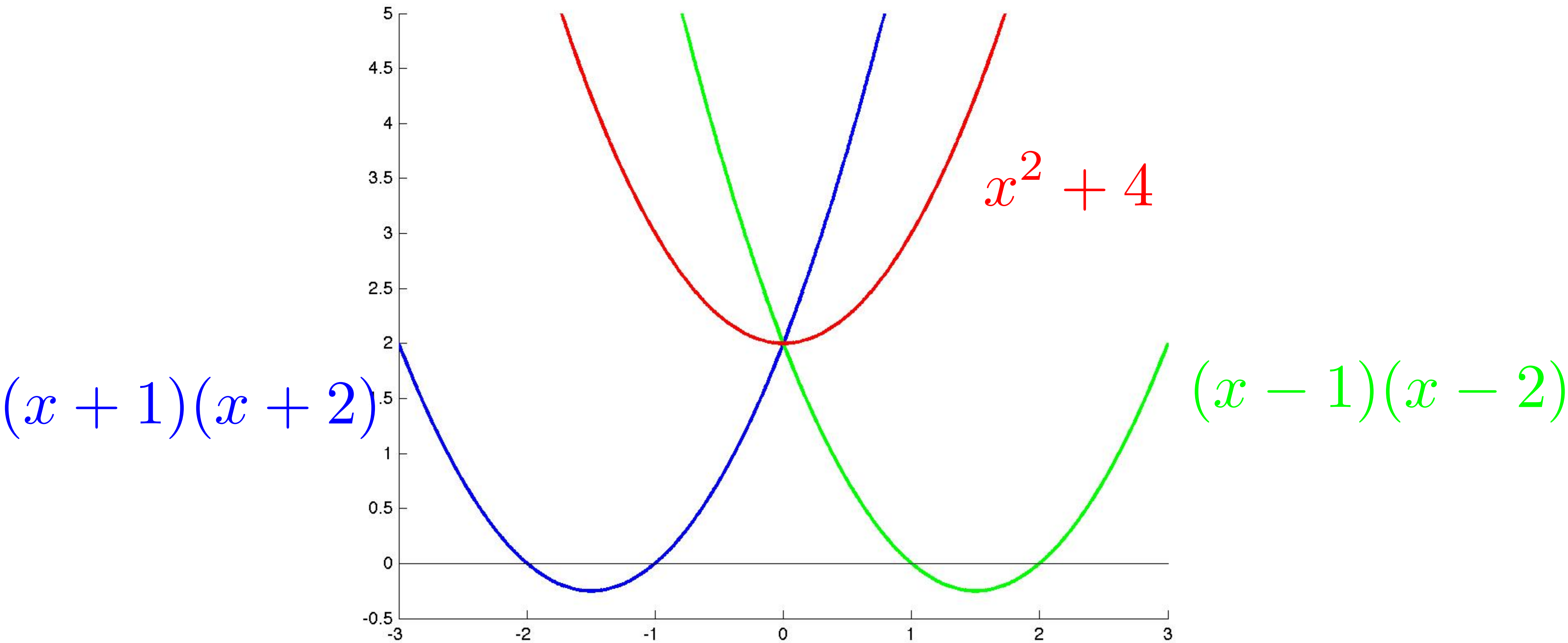


So, the average has a root between the smallest and largest kth roots

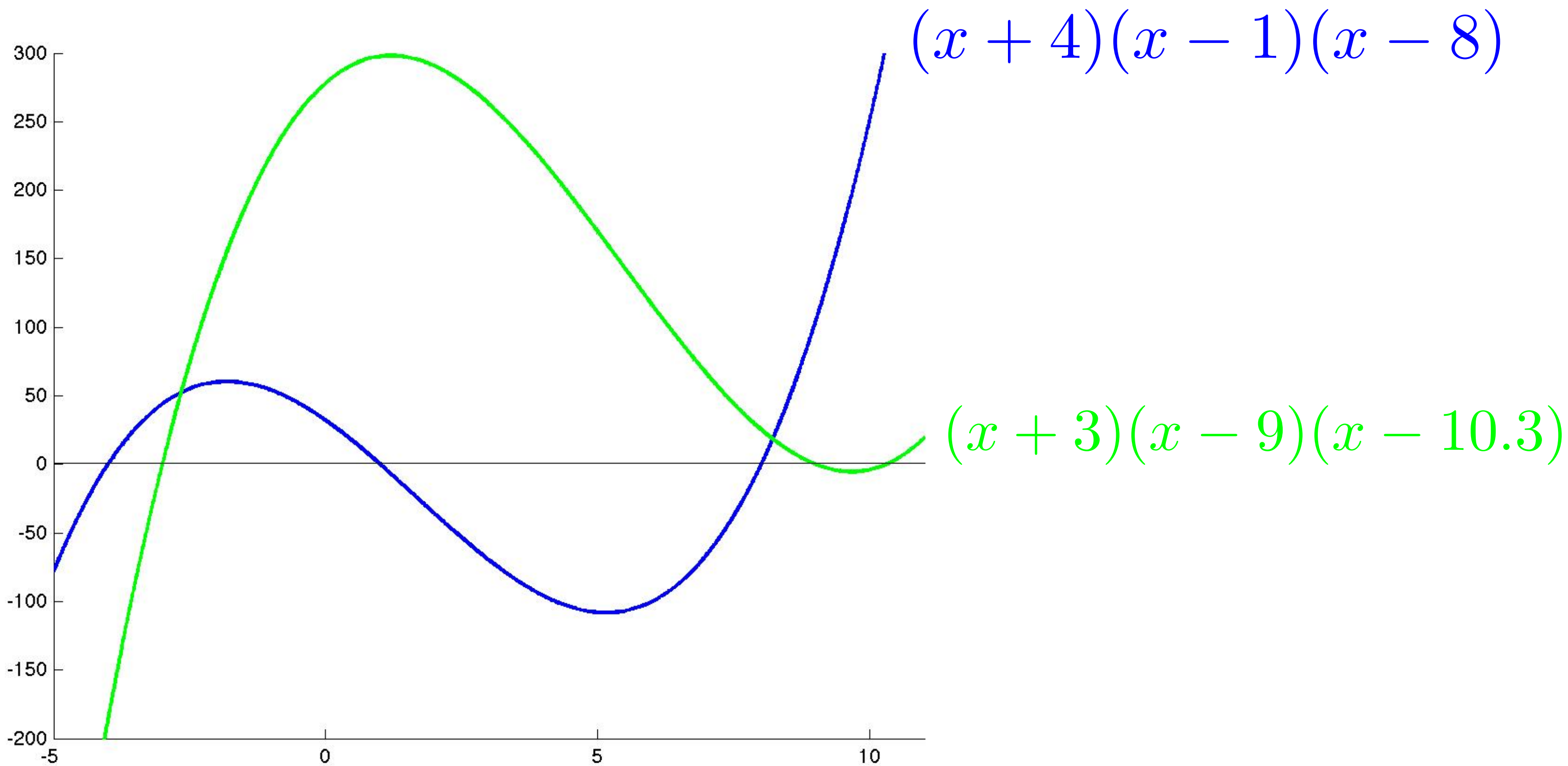
Without a common interlacing



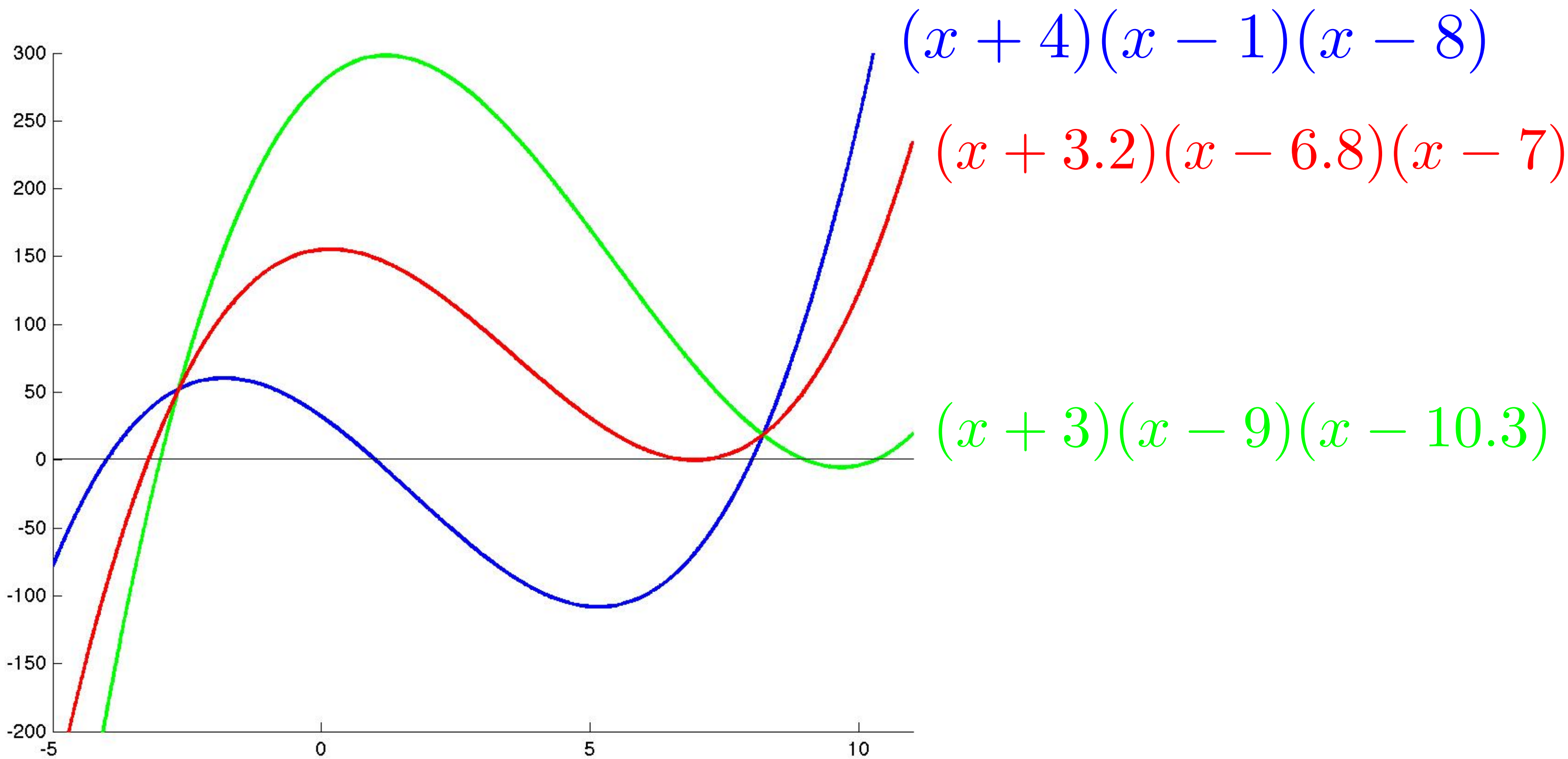
Without a common interlacing



Without a common interlacing



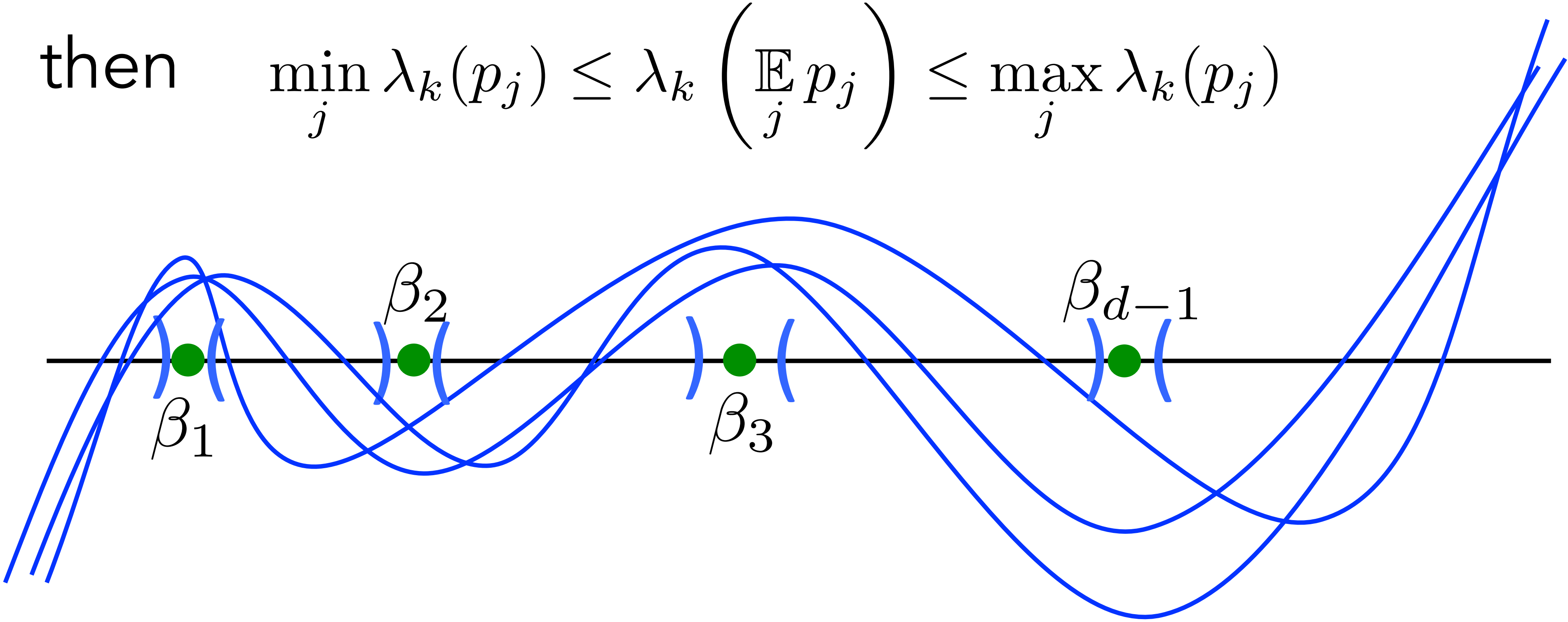
Without a common interlacing



Common Interlacing

If $p_1(x), p_2(x), \ldots, p_n(x)$ have a common interlacing,

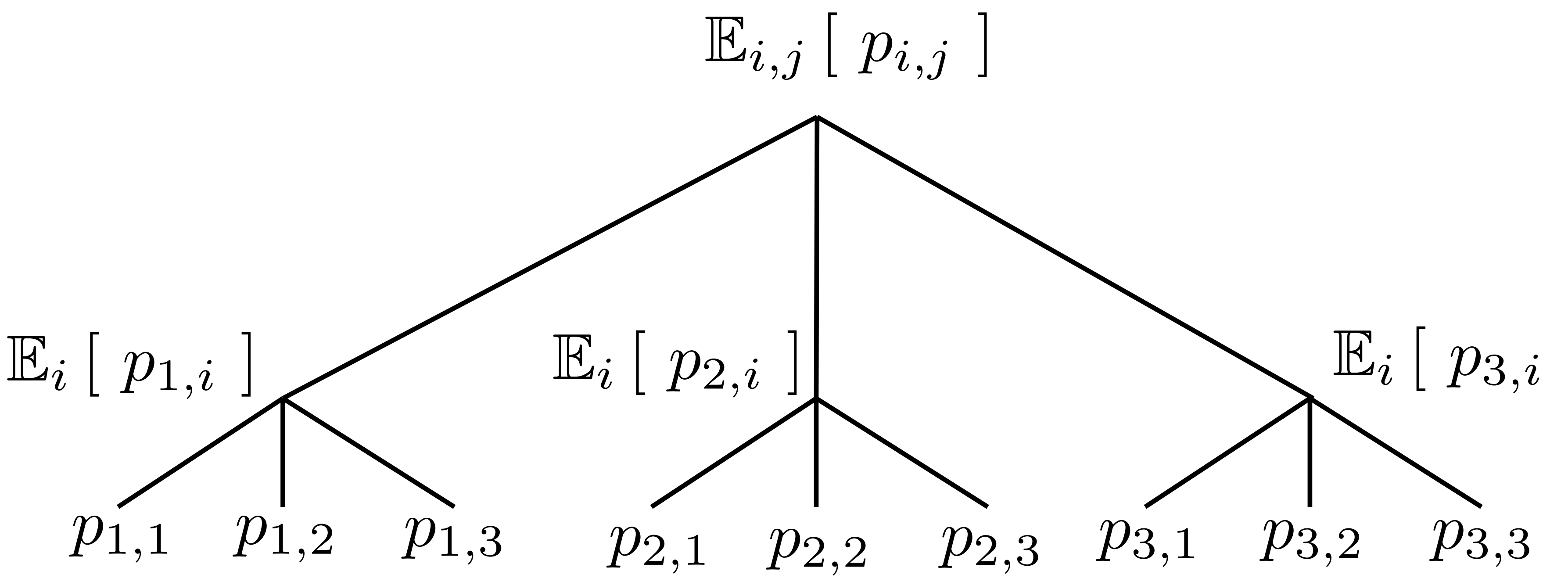
then $$\min_j \lambda_k(p_j) \leq \lambda_k\left(\mathbb{E}_j\, p_j\right) \leq \max_j \lambda_k(p_j)$$



Interlacing Family of Polynomials

$\{p_\sigma(x)\}_\sigma$ is an interlacing family

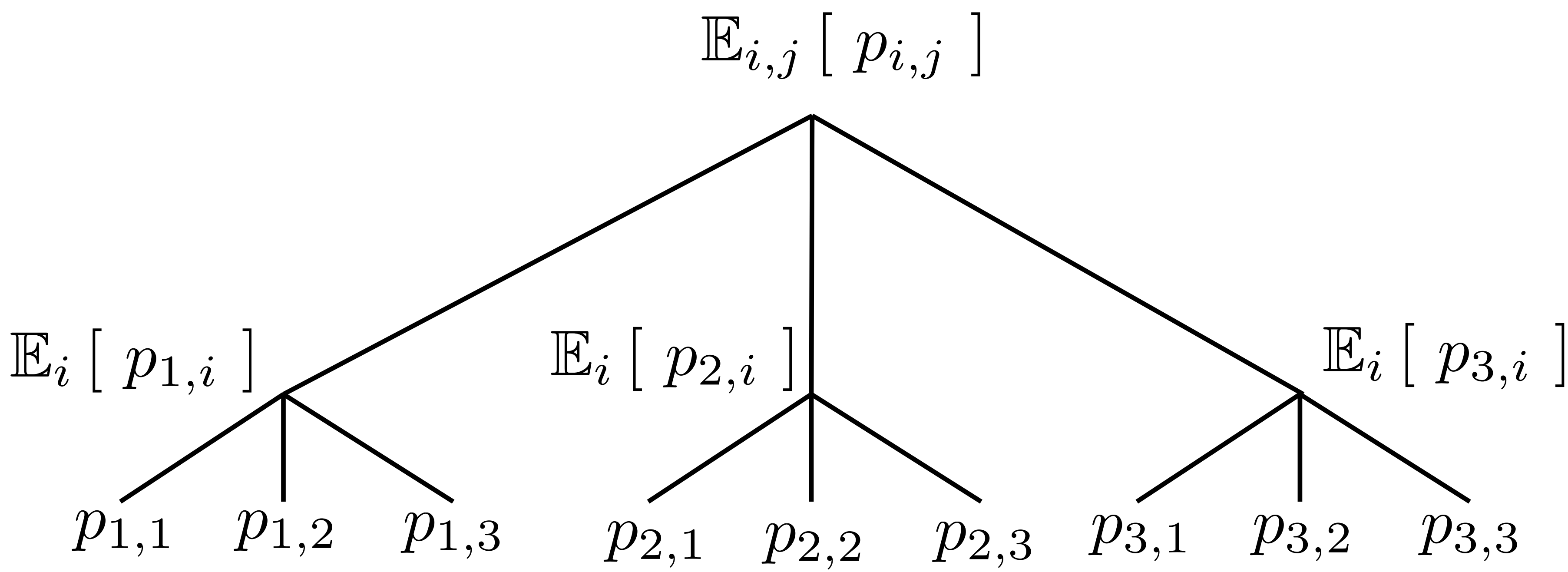
if its members can be placed on the leaves of a tree so that when every node is labeled with the average of leaves below, siblings have common interlacings



Interlacing Family of Polynomials

$\{p_\sigma(x)\}_\sigma$ is an interlacing family

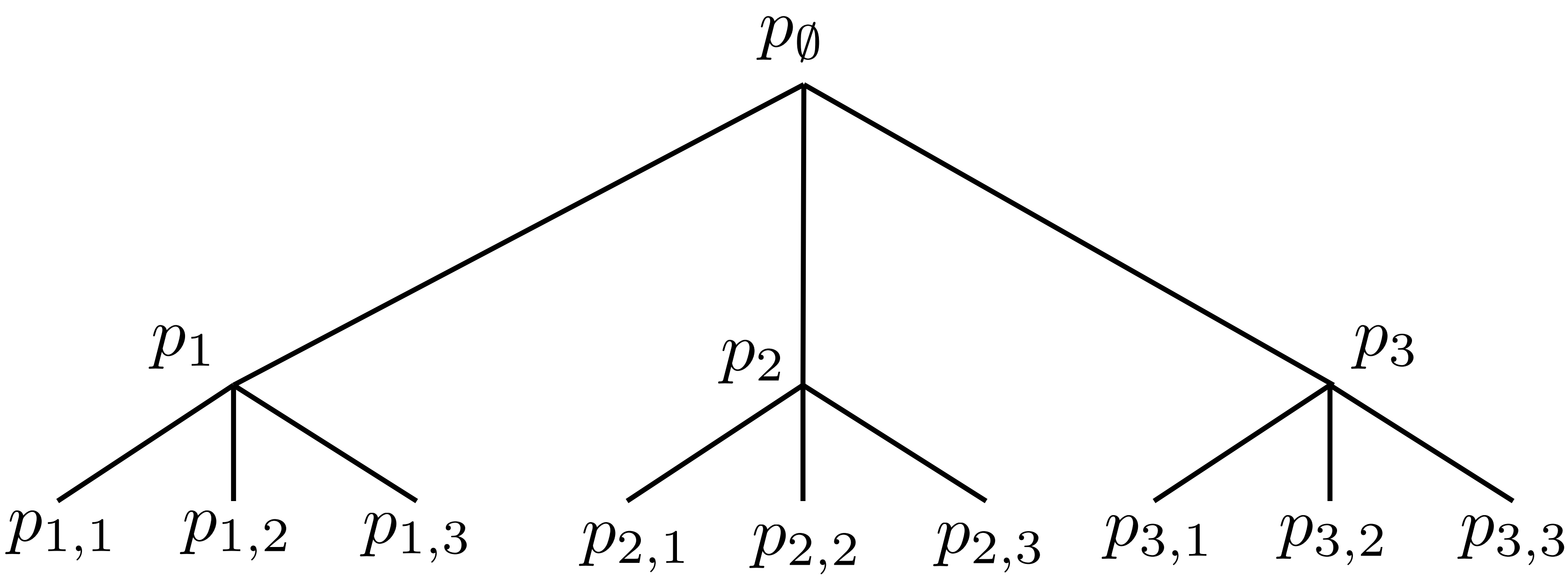
For $\sigma \in \{1, .., n\}^k$, set $p_{i_1,\ldots,i_h} = \mathbb{E}_{i_{h+1},\ldots,i_k}\, p_{i_1,\ldots,i_k}$



Interlacing Family of Polynomials

$\{p_\sigma(x)\}_\sigma$ is an interlacing family

For $\sigma \in \{1, .., n\}^k$, set $p_{i_1,\ldots,i_h} = \mathbb{E}_{i_{h+1},\ldots,i_k}\, p_{i_1,\ldots,i_k}$

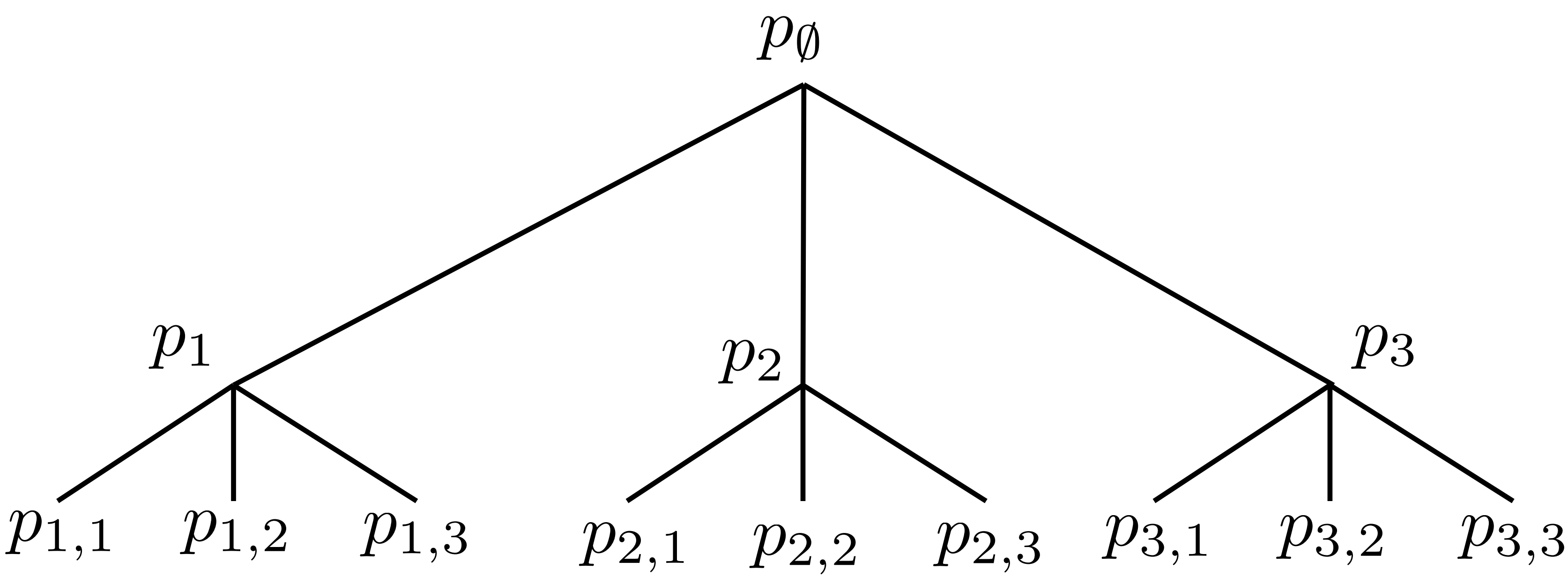


Interlacing Family of Polynomials

Theorem:

There is an $i_1, ..., i_k$ so that

$$\lambda_k(p_{i_1,...,i_k}) \geq \lambda_k\left(p_\emptyset\right)$$



Interlacing Family of Polynomials

It remains to prove that the polynomials

$$p_{i_1,i_2,\ldots,i_k}(x) = \chi\left[v_{i_1}v_{i_1}^* + \cdots + v_{i_k}v_{i_k}^*\right](x)$$

form an interlacing family.

Will imply that with non-zero probability

$$\lambda_k\left(\chi\left[\sum r_j r_j^*\right](x)\right) \geq \lambda_k\left(\mathbb{E}\,\chi\left[\sum r_j r_j^*\right](x)\right)$$

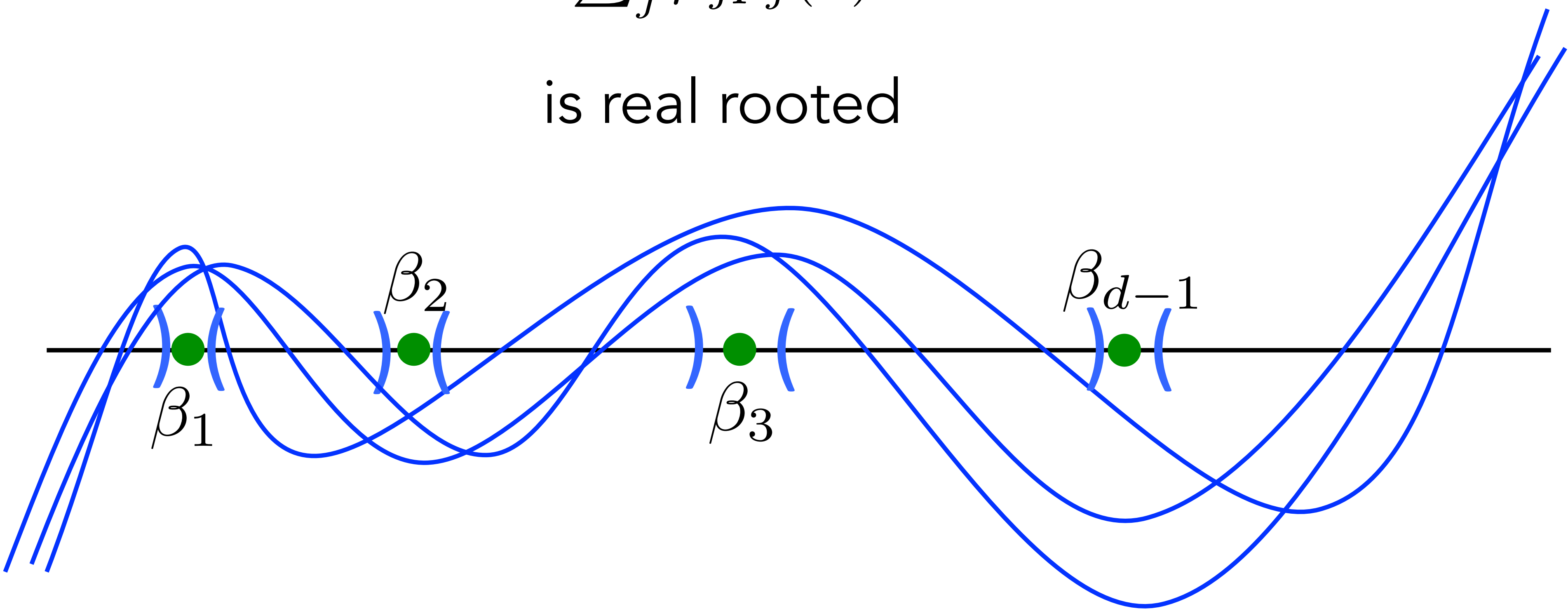
Common interlacings

$p_1(x), p_2(x), \ldots, p_n(x)$ have a common interlacing iff

for every convex combination $\sum_j \mu_j = 1, \mu_j \geq 0$

$$\sum_j \mu_j p_j(x)$$

is real rooted



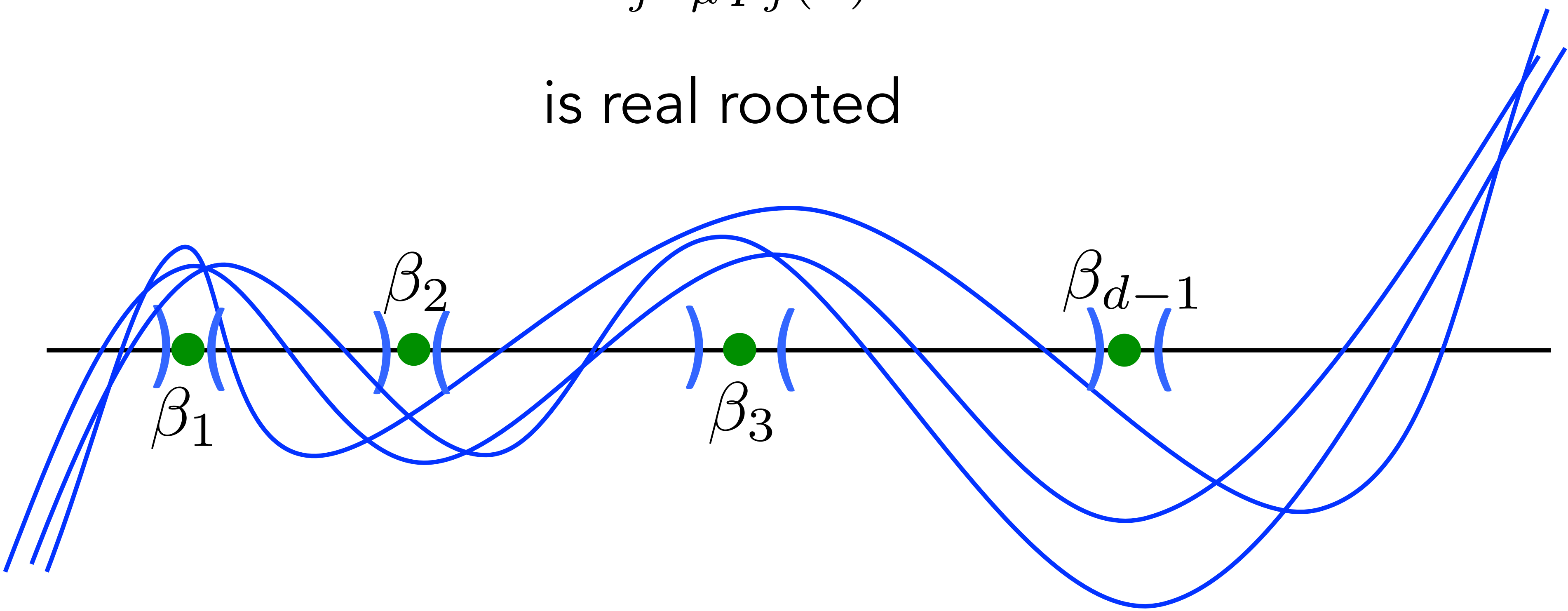
Common interlacings

$p_1(x), p_2(x), \ldots, p_n(x)$ have a common interlacing iff

for every distribution μ on $\{1, ..., n\}$

$$\mathbb{E}_{j\sim\mu}\, p_j(x)$$

is real rooted



Common interlacings

$p_1(x), p_2(x), \ldots, p_n(x)$ have a common interlacing iff

for every distribution μ on $\{1, ..., n\}$

$$\mathbb{E}_{j\sim\mu}\, p_j(x)$$

is real rooted

Proof: by similar picture.
(Dedieu '80, Fell '92, Chudnovsky-Seymour '07)

An interlacing family

$$p_{i_1,i_2,\ldots,i_k}(x) = \chi\left[v_{i_1}v_{i_1}^* + \cdots + v_{i_k}v_{i_k}^*\right](x)$$

Nodes on the tree are labeled with

$$p_{i_1,\ldots,i_h}(x) = \mathbb{E}_{i_{h+1},\ldots,i_k}\, p_{i_1,\ldots,i_k}(x)$$

We need to show that for each $i_1, \ldots, i_h$ the polynomials

$p_{i_1,\ldots,i_h,\underline{j}}(x)$ have a common interlacing

An interlacing family

$$p_{i_1,i_2,\ldots,i_k}(x) = \chi\left[v_{i_1}v_{i_1}^* + \cdots + v_{i_k}v_{i_k}^*\right](x)$$

Nodes on the tree are labeled with

$$p_{i_1,\ldots,i_h}(x) = \mathbb{E}_{i_{h+1},\ldots,i_k}\, p_{i_1,\ldots,i_k}(x)$$

We need to show that for each $i_1, ..., i_h$ the polynomials

$p_{i_1,\ldots,i_h,j}(x)$ have a common interlacing

Proof:

$$= (1 - \tfrac{1}{n}\partial_x)^{k-h-1}\chi\left[v_{i_1}v_{i_1}^* + \cdots + v_{i_h}v_{i_h}^* + v_j v_j^*\right](x)$$

An interlacing family

We need to show that for each $i_1, ..., i_h$ the polynomials

$p_{i_1,\ldots,i_h,j}(x)$ have a common interlacing

Proof:

$$= (1 - \tfrac{1}{n}\partial_x)^{k-h-1} \chi\left[v_{i_1}v_{i_1}^* + \cdots + v_{i_h}v_{i_h}^* + v_j v_j^*\right](x)$$

Cauchy's interlacing theorem implies

$$\chi\left[v_{i_1}v_{i_1}^* + \cdots + v_{i_h}v_{i_h}^* + v_j v_j^*\right](x)$$

All interlace

$$\chi\left[v_{i_1}v_{i_1}^* + \cdots + v_{i_h}v_{i_h}^*\right](x)$$

An interlacing family

We need to show that for each $i_1, ..., i_h$ the polynomials

$p_{i_1,\ldots,i_h,j}(x)$ have a common interlacing

Proof:

$$= (1 - \tfrac{1}{n}\partial_x)^{k-h-1}\chi\left[v_{i_1}v_{i_1}^* + \cdots + v_{i_h}v_{i_h}^* + v_j v_j^*\right](x)$$

Cauchy's interlacing theorem implies

$$\chi\left[v_{i_1}v_{i_1}^* + \cdots + v_{i_h}v_{i_h}^* + v_j v_j^*\right](x)$$

have a common interlacing

An interlacing family

Proof:

$$= (1 - \tfrac{1}{n}\partial_x)^{k-h-1} \chi \left[v_{i_1} v_{i_1}^* + \cdots + v_{i_h} v_{i_h}^* + v_j v_j^* \right] (x)$$

Cauchy's interlacing theorem implies

$$\chi \left[v_{i_1} v_{i_1}^* + \cdots + v_{i_h} v_{i_h}^* + v_j v_j^* \right] (x)$$

have a common interlacing. So, for all distributions μ

$$\mathbb{E}_{j \sim \mu} \, \chi \left[v_{i_1} v_{i_1}^* + \cdots + v_{i_h} v_{i_h}^* + v_j v_j^* \right] (x)$$

is real rooted.

An interlacing family

Proof:

$$= (1 - \tfrac{1}{n}\partial_x)^{k-h-1} \chi \left[v_{i_1} v_{i_1}^* + \cdots + v_{i_h} v_{i_h}^* + v_j v_j^* \right](x)$$

Cauchy's interlacing theorem implies

$$\chi \left[v_{i_1} v_{i_1}^* + \cdots + v_{i_h} v_{i_h}^* + v_j v_j^* \right](x)$$

have a common interlacing. So, for all distributions μ

$$\mathbb{E}_{j \sim \mu}\, \chi \left[v_{i_1} v_{i_1}^* + \cdots + v_{i_h} v_{i_h}^* + v_j v_j^* \right](x)$$

is real rooted. And,

$$\mathbb{E}_{j \sim \mu} (1 - \tfrac{1}{n}\partial_x)^{k-h-1} \chi \left[v_{i_1} v_{i_1}^* + \cdots + v_{i_h} v_{i_h}^* + v_j v_j^* \right](x)$$

is also real rooted for every distribution μ.

QED

Method of proof

Let $r_1, \ldots, r_k$ be chosen uniformly from $\{v_1, \ldots, v_n\}$

1. $\mathbb{E}\, \chi \left[\sum r_j r_j^* \right] (x)$ is real rooted

2. $\lambda_k \left(\mathbb{E}\, \chi \left[\sum r_j r_j^* \right] (x) \right) \geq \left(1 - \sqrt{\frac{k}{d}} \right)^2 \frac{d}{n}$

3. With non-zero probability

$$\lambda_k \left(\chi \left[\sum r_j r_j^* \right] (x) \right) \geq \lambda_k \left(\mathbb{E}\, \chi \left[\sum r_j r_j^* \right] (x) \right)$$

Because is an interlacing family of polynomials

In part 2

Will use the same approach to prove
Weaver's conjecture, and thereby Kadison-Singer

But, employ multivariate analogs of these arguments

and a direct bound on the roots of the polynomials.

Main Theorem

For all $\alpha > 0$

if all $\|v_i\| \leq \alpha$ and $\sum v_i v_i^* = I$

then exists a partition into S_1 and S_2 with

$$\text{eigs}(\sum_{i \in S_j} v_i v_i^*) \leq \frac{1}{2} + 3\alpha$$

Implies Akemann-Anderson Paving Conjecture, which implies Kadison-Singer

We want

$$\operatorname{eigs}\begin{pmatrix} \sum_{i\in S_1} v_i v_i^* & 0 \\ 0 & \sum_{i\in S_2} v_i v_i^* \end{pmatrix} \le \frac{1}{2} + 3\alpha$$

We want

$$\mathrm{roots}\left(\chi\begin{bmatrix}\sum_{i\in S_1} v_i v_i^* & 0 \\ 0 & \sum_{i\in S_2} v_i v_i^*\end{bmatrix}(x)\right) \leq \frac{1}{2} + 3\alpha$$

We want

$$\mathrm{roots}\left(\chi\begin{bmatrix}\sum_{i\in S_1} v_i v_i^* & 0 \\ 0 & \sum_{i\in S_2} v_i v_i^*\end{bmatrix}(x)\right) \leq \frac{1}{2} + 3\alpha$$

Consider expected polynomial with a random partition.

Method of proof

1. Prove expected characteristic polynomial has real roots

2. Prove its largest root is at most $1/2 + 3\alpha$

3. Prove is an interlacing family, so exists a partition whose polynomial has largest root at most $1/2 + 3\alpha$

The Expected Polynomial

Indicate choices by $\sigma_1, ..., \sigma_n : i \in S_{\sigma_i}$

$$p_{\sigma_1,\ldots,\sigma_n}(x) = \chi \begin{bmatrix} \sum_{i:\sigma_i=1} v_i v_i^* & 0 \\ 0 & \sum_{i:\sigma_i=2} v_i v_i^* \end{bmatrix} (x)$$

The Expected Polynomial

$$a_i = \begin{pmatrix} v_i \\ 0 \end{pmatrix} \text{ for } i \in S_1 \qquad a_i = \begin{pmatrix} 0 \\ v_i \end{pmatrix} \text{ for } i \in S_2$$

$$\sigma_i = 1 \qquad \sigma_i = 2$$

$$\begin{pmatrix} \sum_{i \in S_1} v_i v_i^* & 0 \\ 0 & \sum_{i \in S_2} v_i v_i^* \end{pmatrix} = \sum_i a_i a_i^*$$

Mixed Characteristic Polynomials

For $a_1, ..., a_n$ independently chosen random vectors

$$\mathbb{E}\,\chi[\sum_i a_i a_i^*]$$

is their *mixed characteristic polynomial*.

Theorem: It only depends on $A_i = \mathbb{E}\, a_i a_i^*$
and, is real-rooted

Mixed Characteristic Polynomials

For $a_1, ..., a_n$ independently chosen random vectors

$$\mathbb{E}\,\chi[\sum_i a_i a_i^*] = \mu(A_1, ..., A_n)$$

is their *mixed characteristic polynomial.*

Theorem: It only depends on $A_i = \mathbb{E}\, a_i a_i^*$
and, is real-rooted

$$\mathrm{Tr}\,(A_i) = \mathrm{Tr}\,(\mathbb{E}\, a_i a_i^*) = \mathbb{E}\,\mathrm{Tr}\,(a_i a_i^*) = \mathbb{E}\,\|a_i\|^2$$

Mixed Characteristic Polynomials

For $a_1, ..., a_n$ independently chosen random vectors

$$\mathbb{E}\,\chi[\textstyle\sum_i a_i a_i^*] = \mu(A_1, ..., A_n)$$

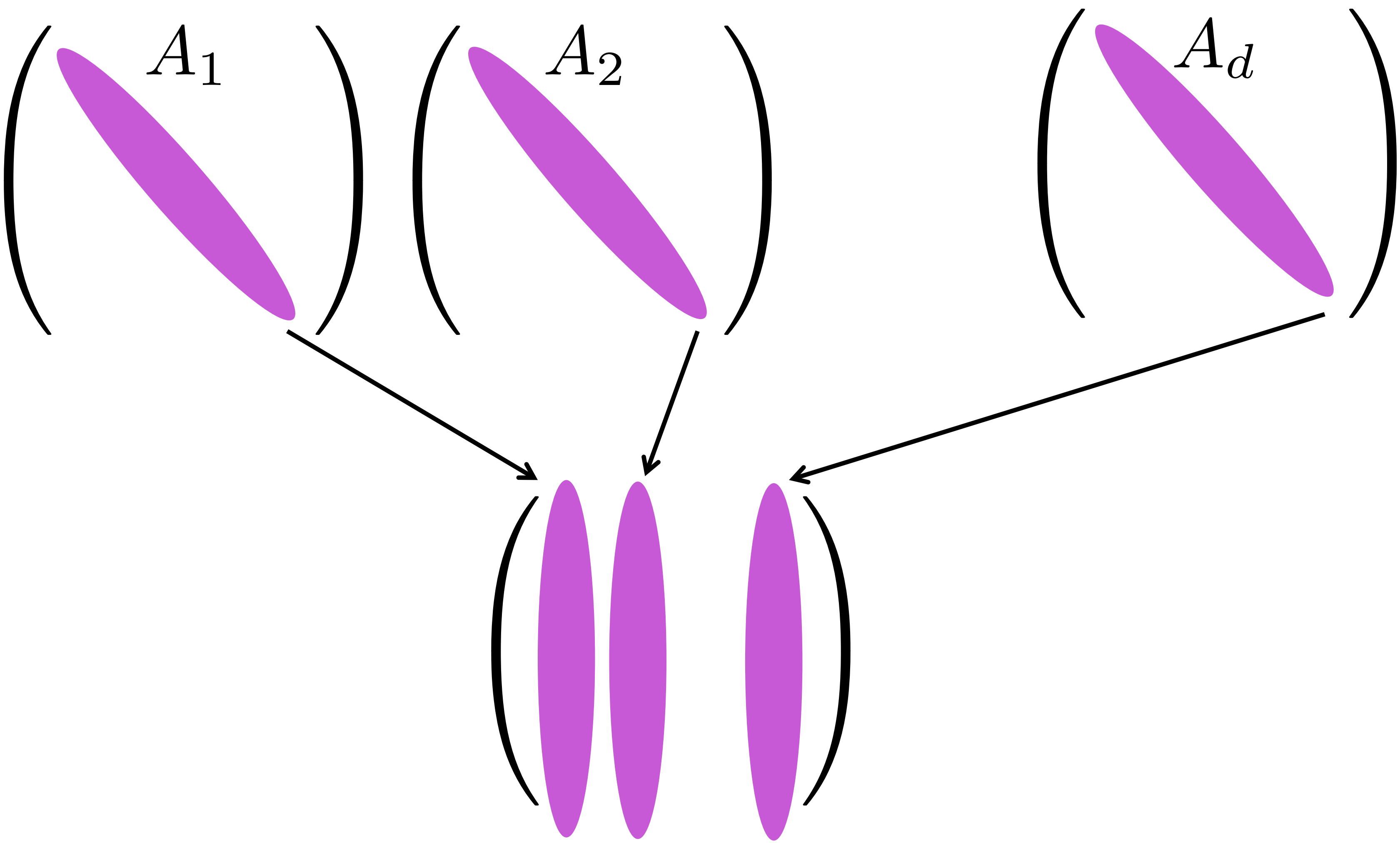
is their *mixed characteristic polynomial*.

The constant term is the *mixed discriminant* of

$$A_1, \ldots, A_n$$

The constant term

When diagonal and $d = n$, c_d is a matrix permanent.



The constant term

When diagonal and $d = n$, c_d is a matrix permanent.
Van der Waerden's Conjecture becomes

If $\sum A_i = I$ and $\mathrm{Tr}(A_i) = 1$

c_d is minimized when $A_i = \frac{1}{n} I$

Proved by Egorychev and Falikman '81.
Simpler proof by Gurvits (see Laurent-Schrijver)

The constant term

For Hermitian matrices, c_d is the mixed discriminant
Gurvits proved a lower bound on c_d:

If $\sum A_i = I$ and $\mathrm{Tr}(A_i) = 1$

c_d is minimized when $A_i = \frac{1}{n} I$

This was a conjecture of Bapat.

Other coefficients

One can generalize Gurvits's results
to prove lower bounds on all the coefficients.

If $\sum A_i = I$ and $\mathrm{Tr}(A_i) = 1$

$\left| c_k \right|$ is minimized when $A_i = \frac{1}{n} I$

But, this does not imply useful bounds on the roots

Real Stable Polynomials

A multivariate generalization of real rootedness.

Complex roots of $p \in \mathbb{R}[z]$ come in conjugate pairs.

So, real rooted iff no roots with positive complex part.

Real Stable Polynomials

$p \in \mathbb{R}[z_1, \ldots, z_n]$

is real stable if $\mathrm{imag}(z_i) > 0$ for all i
implies $p(z_1, \ldots, z_n) \neq 0$

it has no roots in the upper half-plane

Isomorphic to Gårding's hyperbolic polynomials

Used by Gurvits (in his second proof)

Real Stable Polynomials

$p \in \mathbb{R}[z_1, \ldots, z_n]$

is real stable if $\mathrm{imag}(z_i) > 0$ for all i
implies $p(z_1, \ldots, z_n) \neq 0$

it has no roots in the upper half-plane

Isomorphic to Gårding's hyperbolic polynomials

Used by Gurvits (in his second proof)

See surveys of Pemantle and Wagner

Real Stable Polynomials

Borcea-Brändén '08:

For PSD matrices $A_1, \ldots, A_n$

$$\det[z_1 A_1 + \cdots + z_n A_n]$$

is real stable

Real Stable Polynomials

$p(z_1, ..., z_n)$ real stable

implies $(1 - \partial_{z_i})\, p(z_1, ..., z_n)$ is real stable

(Lieb Sokal '81)

$p(z_1, ..., z_n)$ real stable

implies $p(x, x, ..., x)$ is real rooted

Real Roots

$$\mu(A_1, \ldots, A_n)(x) =$$

$$\left(\prod_{i=1}^{n} 1 - \partial_{z_i} \right) \det \left(\sum_{i=1}^{n} z_i A_i \right) \Big|_{z_1 = \cdots = z_n = x}$$

So, every mixed characteristic polynomial is real rooted.

Our Interlacing Family

Indicate choices by $\sigma_1, ..., \sigma_n : i \in S_{\sigma_i}$

$$p_{\sigma_1,\ldots,\sigma_n}(x) = \chi \begin{bmatrix} \sum_{i:\sigma_i=1} v_i v_i^* & 0 \\ 0 & \sum_{i:\sigma_i=2} v_i v_i^* \end{bmatrix} (x)$$

$$p_{\sigma_1,\ldots,\sigma_k}(x) = \mathop{\mathbb{E}}_{\sigma_{k+1},\ldots,\sigma_n} \left[\, p_{\sigma_1,\ldots,\sigma_n} \, \right](x)$$

Interlacing

$p_1(x)$ and $p_2(x)$ have a common interlacing iff
$\lambda p_1(x) + (1-\lambda)p_2(x)$ is real rooted for all $0 \leq \lambda \leq 1$

We need to show that

$$\lambda p_{\sigma_1,\ldots,\sigma_k,1}(x) + (1-\lambda)p_{\sigma_1,\ldots\sigma_k,2}(x)$$

is real rooted.

Interlacing

$p_1(x)$ and $p_2(x)$ have a common interlacing iff
$\lambda p_1(x) + (1-\lambda)p_2(x)$ is real rooted for all $0 \leq \lambda \leq 1$

We need to show that

$$\lambda p_{\sigma_1,\ldots,\sigma_k,1}(x) + (1-\lambda)p_{\sigma_1,\ldots\sigma_k,2}(x)$$

is real rooted.

It is a mixed characteristic polynomial, so is real-rooted.

Set $\sigma_{k+1} = 1$ *with probability* λ
Keep σ_i *uniform for* $i > k+1$

An upper bound on the roots

Theorem: If $\sum A_i = I$ and $\mathrm{Tr}(A_i) \leq \epsilon$ then

$$\text{max-root}\,(\mu(A_1, ..., A_n)(x)) \leq (1 + \sqrt{\epsilon})^2$$

An upper bound on the roots

Theorem: If $\sum A_i = I$ and $\mathrm{Tr}(A_i) \leq \epsilon$ then

$$\text{max-root}\left(\mu(A_1, ..., A_n)(x)\right) \leq (1 + \sqrt{\epsilon})^2$$

An upper bound of 2 is trivial (in our special case).

Need any constant strictly less than 2.

An upper bound on the roots

Theorem: If $\sum A_i = I$ and $\mathrm{Tr}(A_i) \leq \epsilon$ then

max-root $(\mu(A_1, ..., A_n)(x)) \leq (1+\sqrt{\epsilon})^2$

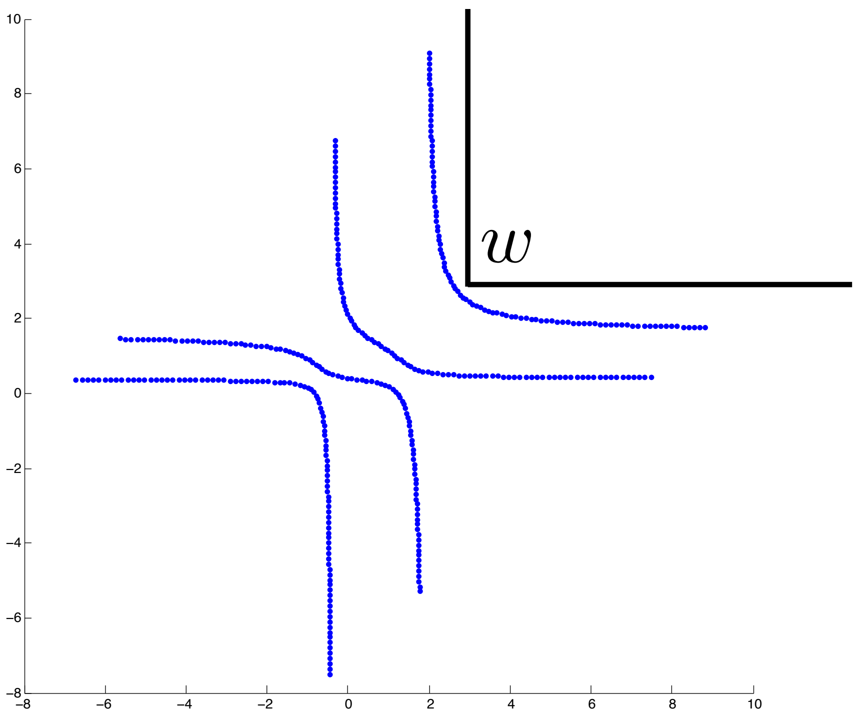
$$\mu(A_1, \ldots, A_n)(x) =$$

$$\left(\prod_{i=1}^{n} 1 - \partial_{z_i}\right) \det\left(\sum_{i=1}^{n} z_i A_i\right)\Big|_{z_1 = \cdots = z_n = x}$$

An upper bound on the roots

Define: $(w_1, ..., w_n)$ is an *upper bound* on the roots of $p(z_1, ..., z_n)$ if

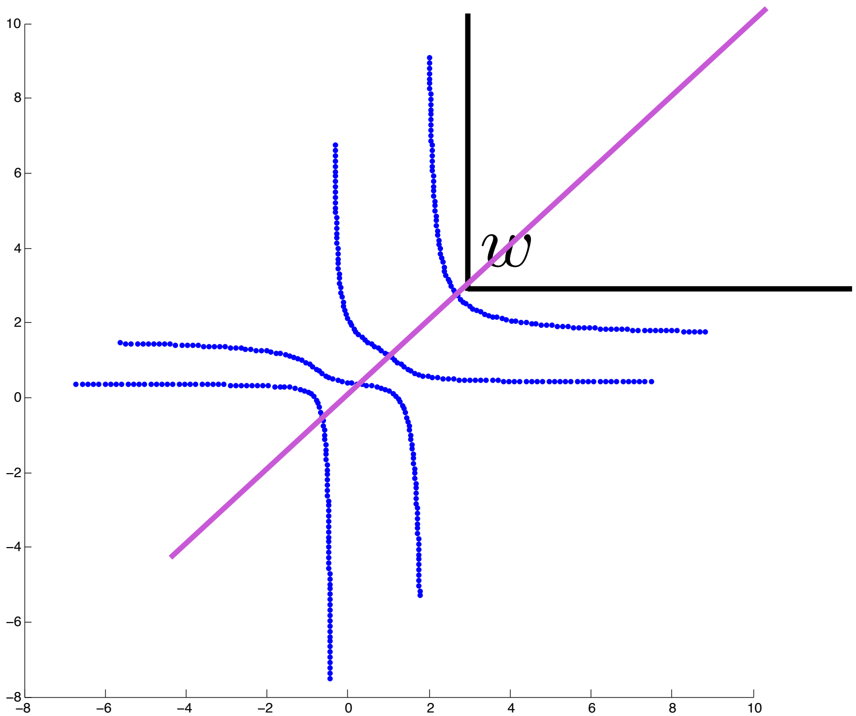
$$p(z_1, ..., z_n) > 0 \text{ for } (z_1, ..., z_n) \geq (w_1, ..., w_n)$$



An upper bound on the roots

Define: $(w_1, ..., w_n)$ is an *upper bound* on the roots of $p(z_1, ..., z_n)$ if

$$p(z_1, ..., z_n) > 0 \text{ for } (z_1, ..., z_n) \geq (w_1, ..., w_n)$$



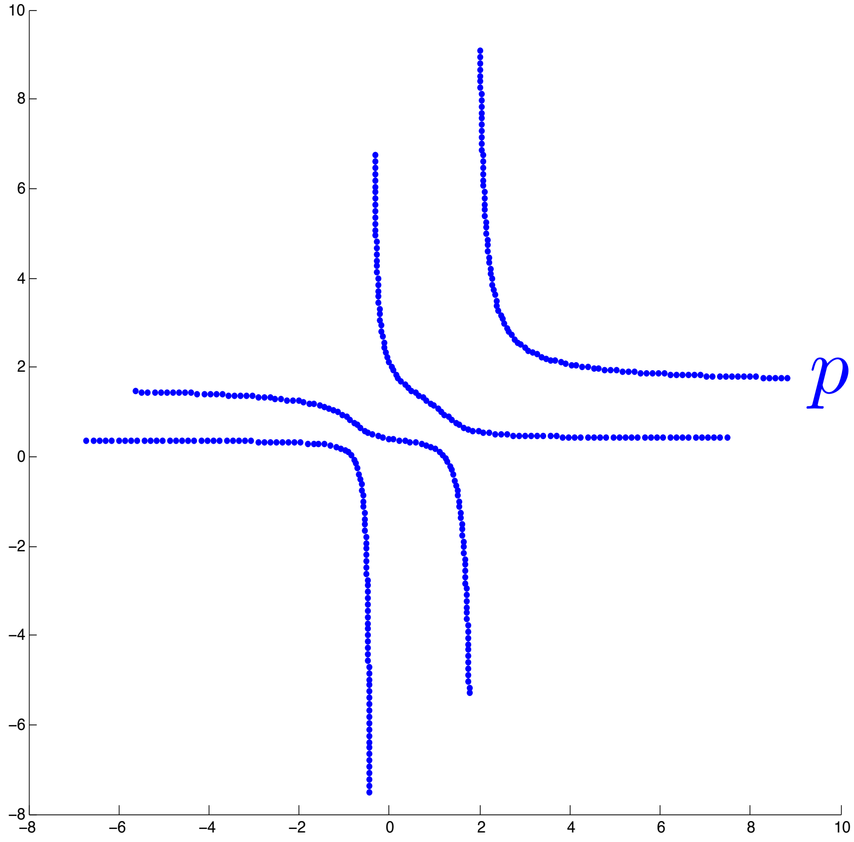
Eventually set

$$z_1, ..., z_n = x$$

so want

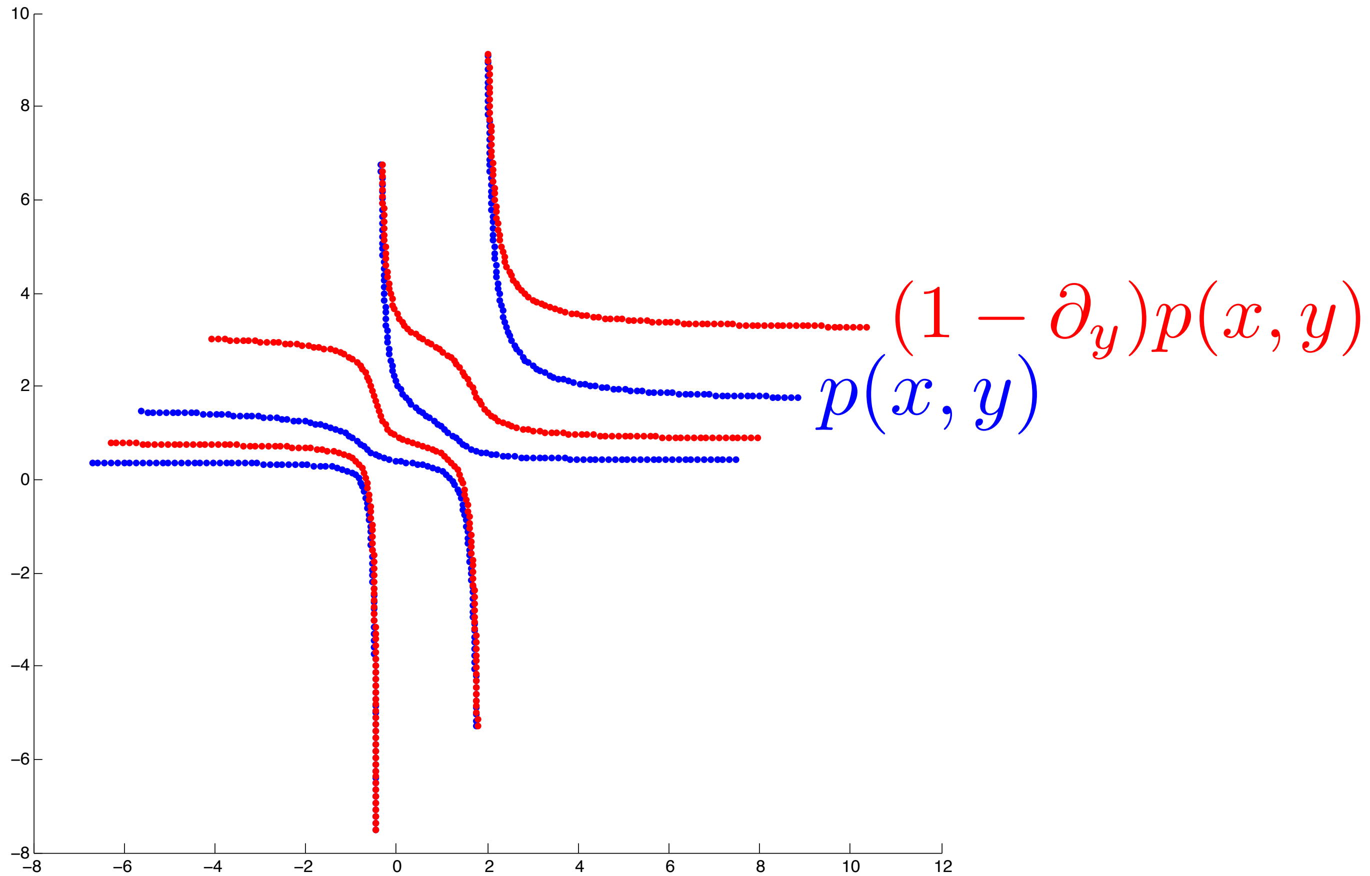
$$w_1 = \cdots = w_n$$

Action of the operators

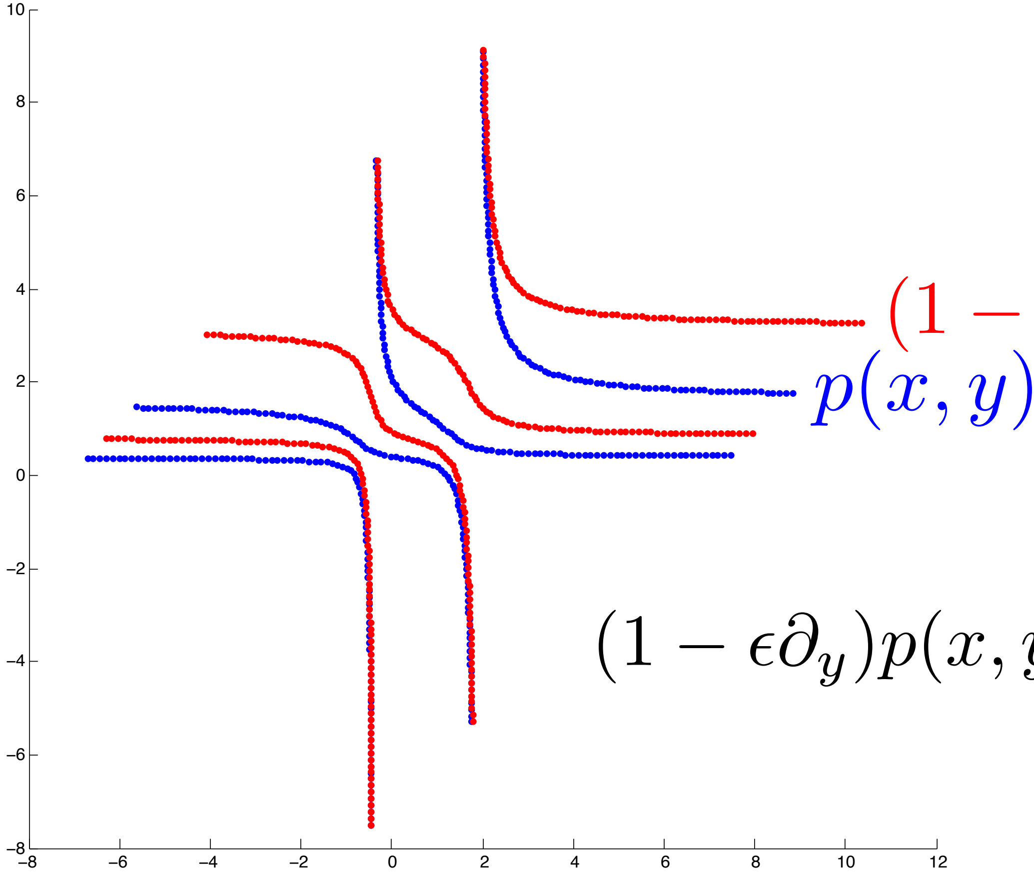


$p(x, y)$

Action of the operators



Action of the operators



$$(1-\epsilon\partial_y)p(x,y) \approx p(x,y-\epsilon)$$

The roots of $(1-\partial_x)p(x)$

Define:

$$\alpha\text{-}\max(\lambda_1, ..., \lambda_n) = \max\left\{u : \sum \tfrac{1}{u-\lambda_i} = \alpha\right\}$$

$$\alpha\text{-}\max(p(x)) = \alpha\text{-}\max(\mathrm{roots}(p))$$

Theorem (Batson-S-Srivastava):
If $p(x)$ is real rooted and $\alpha > 0$

$$\alpha\text{-}\max((1-\partial_x)p(x)) \leq \alpha\text{-}\max(p(x)) + \frac{1}{1-\alpha}$$

The roots of $(1-\partial_x)p(x)$

Theorem (Batson-S-Srivastava):
If $p(x)$ is real rooted and $\alpha > 0$

$$\alpha\text{-max}((1-\partial_x)p(x)) \leq \alpha\text{-max}(p(x)) + \frac{1}{1-\alpha}$$

Proof: Define $\Phi_p(u) = \dfrac{p'(u)}{p(u)} = \displaystyle\sum_i \frac{1}{u-\lambda_i} = \partial_u \log p(u)$

Set $u = \alpha\text{-max}(p(x))$, so $\Phi_p(u) = \alpha$

Suffices to show for all $\delta \geq \frac{1}{1-\alpha}$

$$\Phi_{p-p'}(u+\delta) \leq \Phi_p(u)$$

The roots of $(1-\partial_x)p(x)$

Define $\Phi_p(u) = \frac{p'(u)}{p(u)} = \sum_i \frac{1}{u-\lambda_i}$

Set $u = \alpha\text{-}\max(p(x))$, so $\Phi_p(u) = \alpha$

Suffices to show for all $\delta \geq \frac{1}{1-\alpha}$

$$\Phi_{p-p'}(u+\delta) \leq \Phi_p(u)$$

⇕ *(algebra)*

$$\Phi_p(u) - \Phi_p(u+\delta) \geq \frac{-\Phi'_p(u+\delta)}{1-\Phi_p(u+\delta)}$$

The roots of $(1 - \partial_x)p(x)$

Define $\Phi_p(u) = \frac{p'(u)}{p(u)} = \sum_i \frac{1}{u - \lambda_i}$

$$\Phi_p(u) - \Phi_p(u+\delta) \geq \frac{-\Phi_p'(u+\delta)}{1 - \Phi_p(u+\delta)}$$

$\Phi_p(u)$ convex for $u > \max(p(x))$ implies

$$\Phi_p(u) - \Phi_p(u+\delta) \geq \delta(-\Phi_p'(u+\delta))$$

Monotone decreasing implies only need

$$\delta \geq \frac{1}{1 - \Phi_p(u+\delta)}$$

The roots of $(1 - \partial_x)p(x)$

$\Phi_p(u)$ convex for $u > \max(p(x))$ implies

$$\Phi_p(u) - \Phi_p(u + \delta) \geq \delta(-\Phi_p'(u + \delta))$$

Monotone decreasing implies only need

$$\delta \geq \frac{1}{1 - \Phi_p(u + \delta)}$$

and that $\delta \geq \frac{1}{1 - \alpha} = \frac{1}{1 - \Phi_p(u)}$ suffices.

The roots of $(1 - \partial_x)p(x)$

$\Phi_p(u)$ convex for $u > \max(p(x))$ implies

$$\Phi_p(u) - \Phi_p(u + \delta) \geq \delta(-\Phi_p'(u + \delta))$$

Monotone decreasing implies only need

$$\delta \geq \frac{1}{1 - \Phi_p(u + \delta)}$$

and that $\delta \geq \dfrac{1}{1 - \alpha} = \dfrac{1}{1 - \Phi_p(u)}$ suffices.

The roots of $(1-\partial_x)p(x)$

Theorem (Batson-S-Srivastava):
If $p(x)$ is real rooted and $\alpha > 0$

$$\alpha\text{-max}((1-\partial_x)p(x)) \leq \alpha\text{-max}(p(x)) + \frac{1}{1-\alpha}$$

Gives a sharp upper bound on the roots of associated Laguerre polynomials.

The analogous argument with the min gives the lower bound that we claimed before.

An upper bound on the roots

Theorem: If $\sum A_i = I$ and $\mathrm{Tr}(A_i) \leq \epsilon$ then

max-root $(\mu(A_1, ..., A_n)(x)) \leq (1 + \sqrt{\epsilon})^2$

$$\mu(A_1, \ldots, A_n)(x) =$$

$$\left(\prod_{i=1}^{n} 1 - \partial_{z_i} \right) \det \left(\sum_{i=1}^{n} z_i A_i \right) \Big|_{z_1 = \cdots = z_n = x}$$

A robust upper bound

Define: $(w_1, ..., w_n)$ is an α-*upper bound* on
$$p(z_1, ..., z_n)$$
if it is an α_i-max, in each z_i , and $\alpha_i \leq \alpha$

Theorem:

If w is an α-upper bound on p, then

$w + \delta e_j$ is an α-upper bound on $p - \partial_{z_j} p$,

for $\delta \geq \frac{1}{1-\alpha}$

A robust upper bound

Theorem:

If w is an α-upper bound on p, and $\delta \geq \frac{1}{1-\alpha}$

$w + \delta e_j$ is an α-upper bound on $p - \partial_{z_j} p$,

Proof:

Same as before, but need to know that

$$\frac{\partial_{z_i} p(z_1, \ldots, z_n)}{p(z_1, \ldots, z_n)}$$

is decreasing and convex in z_j, above the roots

A robust upper bound

Proof:

Same as before, but need to know that

$$\frac{\partial_{z_i} p(z_1, \ldots, z_n)}{p(z_1, \ldots, z_n)}$$

is decreasing and convex in z_j, above the roots

Follows from a theorem of Helton and Vinnikov '07:

Every bivariate real stable polynomial can be written

$$\det(A + Bx + Cy)$$

A robust upper bound

Proof:

Same as before, but need to know that

$$\frac{\partial_{z_i} p(z_1, \ldots, z_n)}{p(z_1, \ldots, z_n)}$$

is decreasing and convex in z_j, above the roots

Or, as pointed out by Renegar,
from a theorem Bauschke, Güler, Lewis, and Sendov '01

Or, by a theorem of Brändén '07.

Or, see Terry Tao's blog for a (mostly) self-contained proof

An upper bound on the roots

Theorem: If $\sum A_i = I$ and $\mathrm{Tr}(A_i) \leq \epsilon$ then

$$\text{max-root}\left(\mu(A_1, ..., A_n)(x)\right) \leq (1+\sqrt{\epsilon})^2$$

A probabilistic interpretation

For $a_1, ..., a_n$ independently chosen random vectors with finite support

such that $\mathbb{E}\left[\ \sum_i a_i a_i^T\ \right] = I$ and $\left\| \mathbb{E}\left[\ a_i a_i^T\ \right] \right\| \leq \epsilon$

then $$\Pr\left[\left\| \sum_i a_i a_i^T \right\| \leq (1+\sqrt{\epsilon})^2 \right] > 0$$

Main Theorem

For all $\alpha > 0$

if all $\|v_i\| \leq \alpha$

then exists a partition into S_1 and S_2 with

$$\text{eigs}(\sum_{i \in S_j} v_i v_i^T) \leq \frac{1}{2} + 3\alpha$$

Implies Akemann-Anderson Paving Conjecture, which implies Kadison-Singer

Anderson's Paving Conjecture '79

Reduction by Casazza-Edidin-Kalra-Paulsen '07 and Harvey '13:

There exist an $\epsilon > 0$ and a k so that

if all $\|v_i\|^2 \leq 1/2$ and $\sum v_i v_i^T = I$

then exists a partition of $\{1, \ldots, n\}$ into k parts s.t.

$$\text{eigs}(\textstyle\sum_{i \in S_j} v_i v_i^T) \leq 1 - \epsilon$$

Can prove using the same technique

A conjecture

If $\sum A_i = I$ and $\mathrm{Tr}(A_i) \leq \epsilon$ then

max-root $(\mu(A_1, ..., A_n)(x))$

is largest when $A_i = \frac{\epsilon}{d} I$

Questions

Can the partition be found in polynomial time?

What else can one construct this way?

How do operations that preserve real rootedness move the roots and the Stieltjes transform?