# Hyperbolic Polynomials and SOS

Inofficial lecture notes
for the lecture held by Cynthia Vinzant, Rainer Sinn, Greg Blekherman, Daniel Plaumann, Summer 2018

Henning Seidler

## Contents

# 1 Hyperbolic Polynomials

We consider objects in $\mathbb{R}[x_1, \ldots, x_n]_d$, of homogeneous polynomials in $n$ variables of degree $d$. These can be written as

$$\left\{ \sum_\alpha c_\alpha x^\alpha : c_\alpha \mathbb{R}, \alpha \in \mathbb{N}^d, |\alpha|_1 = d \right\}$$

They have the property $f(\lambda x) = \lambda^d f(x)$.

**1.1 Definition.** $f \in \mathbb{R}[x]_d$ is hyperbolic with respect to $e \in \mathbb{R}^n$ if $f(e) \neq 0$ and for every $v \in \mathbb{R}^n$, all of the roots of the univariate polynomial $f(te+v) \in \mathbb{R}[t]$ are real.

Alternatively all roots of $f(e+sv) \in \mathbb{R}[s]$ are real.
The definition takes direction $e$, the alternative takes lines through $e$.

**1.2 Example.** Take $f = x_1^2 - x_2^2 - x_3^2$ and $e = (1, 0, 0)$. Then the roots of $f$ are a double cone . pictur
Each vertical line intersects the cones in exactly 2 points (or 2 double root at $\mathbf{0}$).

**1.3 Example (Non-example).** $f = x_1^4 - x_2^4 - x_3^4$ is not hyperbolic wrt. to any point in $\mathbb{R}^3$. In each case we only get 2 real roots.

**Exercise (Hyperbolic quartic).** *The shape would have to be some cone where an inner cone is missing.*

**1.4 Example (Important).** Let $A_1, \ldots, A_n \in \mathbb{R}^{d \times d}$ symmetric. Put $A(x) = \sum x_i A_i$. If $A(e)$ is positive definite, then $f := \det(A(x))$ is hyperbolic wrt $e$.

*Proof.* Assume $A(e) = I_d$. Then $f(te - v) = \det(tI_d - A(v))$. Its roots are the eigenvalues of $A(v)$ and these are real, since $A(v)$ is symmetric. □

Going back to our first example we get

$$x_1^2 - x_2^2 - x_3^2 = \det \begin{pmatrix} x_1 + x_2 & x_3 \\ x_3 & x_1 - x_2 \end{pmatrix}$$

**1.5 Definition.** If $f$ is hyperbolic wrt $e$, for $x \in \mathbb{R}^n$ we call the roots $\lambda_1(x) \geq \ldots \geq \lambda_d(x)$ for $f(t - ev)$ the eigenvalues of $x$. The *rank* of $x$ is its number of non-zero eigenvalues.

Take $f = x_1^2 - x_2^2 - x_3^2$ and $e = (1, 0, 0)$. The eigenavlues of $x \in \mathbb{R}^3$ are the roots of $(t - x_1)^2 - x_2^2 - x_3^2$, which are $(x_1 \pm \sqrt{x_2^2 + x_3^2}$.

**1.6 Example.** Take $f = \prod x_1$ and $e = (1, \ldots, 1)$. The eigenvalues of $x$ are the roots of $f(te - x) = \prod(t - x_i)$. So $\lambda_1 = \max\{x_1, \ldots, x_n\}$, $\lambda_2 = \max\{x_1, \ldots, x_n\} \setminus \{\lambda_1\}$ and so on.

**1.7 Example.** Let $X \in \mathbb{R}^{d \times d}$ symmetric. Say $f = \det X$ and $e = I_d$, shaped into $\mathbb{R}^{d \cdot d}$. Then the eigenvalues of $X$ in terms of hyperbolic polynomial are the roots of $f(te - X) = \det(tI_d - X)$ which are the eigenvalues of $X$ as a real matrix.

**1.8 Definition.** The *hyperbolicity cone* of $f$ wrt $e$ is

$$C_e f = \{x \in \mathbb{R}^n : \text{ roots of } f(te - x) \geq 0\}$$

**1.9 Example.** For our previous three examples we get

**Example 1.2** $C_e f = \left\{x \in \mathbb{R}^3 : x_1 \geq \sqrt{x_2^2 + x_3^2}\right\}$

**Example 1.6** $C_e f = (\mathbb{R}_{\geq 0})^n$

**Example 1.7** $C_e f$ are the positive semidefinite matrices.

**1.10 Theorem (Gårding, 1959).** *Let $f \in \mathbb{R}[x]_d$ be hyperbolic wrt $e$. Then $C_e f$ is a convex cone and $f$ is hyperbolic wrt to any points in its interior.*

**1.11 Lemma.** *Let $\vec{a} \in C_{\vec{e}} f, \vec{b} \in \mathbb{R}^n, s \geq 0$. Then the foots of*

$$f\left(is\vec{e} + t\vec{a} + \vec{b}\right) \in \mathbb{C}[t]$$

*have $\leq 0$ imaginary part. Write $t = x - iy$ for $x, y \in \mathbb{R}$ for some root $t$. Rewriting, we get*

$$f(is\vec{e} + (x - iy)\vec{a} + \vec{b}) = 0 \implies f\left(s\vec{e} - ix\vec{a} - y\vec{a} - i\vec{b}\right) = 0$$

*By assumption hyperbolic, any lines through vece yields real points, to the imaginary part must cancel.*

*homework*

$$f(s\vec{e} - y\vec{a}) = 0 \xRightarrow{y \geq 0} f\left(\frac{s}{y}\vec{e} - \vec{a}\right) = 0$$

*Proof of Theorem 1.10.* Taking $s \to 0$, all roots of $f(t\vec{a} + \vec{b}) \in \mathbb{R}[t]$ have $\leq 0$ imaginary part. Hence all roots are real.

convex cone is homework

□

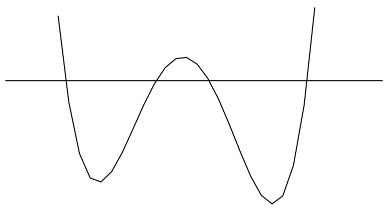


Figure 1: Alternating roots of $f$ and $f'$

Another way to see the hyperbolicity cone $C_e f$ is the closure of the connected component of $e$ in $\mathbb{R}^n \setminus \{x : f(x) = 0\}$.

**1.12 Lemma.** *If $p(t) \in \mathbb{R}[t]$ is real rooted, then so is $p'(t)$.*

**1.13 Lemma.** *If $f \in \mathbb{R}[x]_d$ is hyperbolic wrt e, then so is the derivative*

$$D_{\vec{e}} f = \sum_{i=1}^{n} (\vec{e})_i \frac{\partial f}{\partial x_i}$$

*Proof.* Chain rule

$$\frac{d}{dt} f(te + x) = D_{\vec{e}} f(t\vec{e} + x)$$

□

**1.14 Example.** Put $f = \prod x_i$ and $e = \mathbf{1}$. Then

$$D_{\vec{e}} f = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} = e_{n-1}(x_1, \dots, x_n) = \sum_i \prod_{j \neq i} x_j$$

which is the second elementary symmetric polynomial.

# 2 Convexity

**2.1 Definition.** $C \subseteq \mathbb{R}^n$ is convex if

$$\forall x, y \in C. \forall \lambda \in [0, 1]. \lambda x + (1 - \lambda) y \in C$$

**2.2 Example.**

1. polyhedra: intersections of finitely many half-spaces

$$\bigcap_{i=1}^{m} \{x \in \mathbb{R}^n : l_i(x) \leq z_i\} \qquad l_i \text{ linear}$$

   sub-examples:

   - $\mathbb{R}^n_{\geq 0}$,
   - Birkhoff polytope: take non-negative $m \times m$-matrices, where row- and column-sums are 1.

2. positive semi-definite matrices

$$\mathbb{S}^d_+ = \left\{A \in \mathbb{S}^d : \forall x \in \mathbb{R}^n . x^T A x \geq 0\right\}$$

**2.3 Definition.** The *convex hull* of $S$ is the smallest convex set containing $S$. Equivalently the set of finite convex combinations

$$\operatorname{conv}(S) = \left\{ \sum_{i=1}^{r} \lambda_i x_i : \sum \lambda_i = 1, \lambda \geq 0, x_i \in S \right\}$$

**2.4 Theorem (Caratheodory).** *Every $x \in \operatorname{conv}(S) \subseteq \mathbb{R}^d$ can be written as a convex combinations of at most $(d+1)$ points in $S$.*

**2.5 Corollary.** *If $S$ is compact, then* $\operatorname{conv}(S)$ *is closed.*

*Proof.* Regard the map

$$\underbrace{S \times \ldots \times S}_{d+1} \times \Delta_d \to \mathbb{R}^d$$
$$(x_1, \ldots, x_{d+1}, \lambda) \mapsto \sum \lambda_i x_i$$

The left hand side is closed, so the image is closed as well. But due to Theorem 2.4 the image is the convex hull. $\square$

**2.6 Theorem.** *Let $A \subseteq \mathbb{R}^d$ convex,* $\operatorname{int}(A) = \emptyset$*. Then there exists a proper affine subspace $L \subset \mathbb{R}^d$ with $A \subseteq L$.*

**2.7 Definition.** The *dimension* of a convex set $C \subseteq \mathbb{R}^d$ is the dimension of its affine span.

## 2.1 Isolation Theorem

**2.8 Theorem.** *Suppose $C \subseteq \mathbb{R}^d$ convex, closed set, $u \notin C$. Then there exists an affine hyperplane $H = \left\{x \in \mathbb{R}^d : l(x) = z\right\}$ such that*

$$C \subseteq H^+ = \left\{x \in \mathbb{R}^d : l(x) > z\right\}$$
$$u \in H^- = \left\{x \in \mathbb{R}^d : l(x) < z\right\}$$

*Proof.* We take the distance function

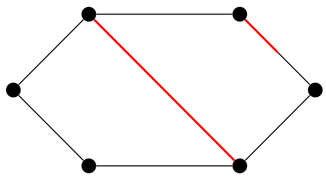
$$\min\left\{\operatorname{dist}(u, x) : x \in C\right\} > 0$$

Since $C$ is closed, the minimum is attained at some $x_0$. Since $C$ is convex, the minimum is unique (triangle inequality). So take the hyperplane perpendicular to $u - x_0$, acros half the distance. $\square$

**2.9 Theorem (Farkas Lemma).** *Let $A \in \mathbb{R}^{m \times d}, z \in \mathbb{R}^m$. Either there exists $x \in \mathbb{R}^d_{\geq 0}$ such that $Ax = z$ or there exists $c \in (\mathbb{R}^m) \setminus \{\mathbf{0}\}$ such that $cA \geq 0, c \cdot z < 0$.*

**2.10 Theorem.**

1. *Let $C \subseteq \mathbb{R}^d$ open, convex, $u \notin C$. Then there exists a hyperplane $H$ such that $u \in H$ and $C \subseteq H^+$.*
2. *Let $C \subseteq \mathbb{R}^d$ convex,* $\operatorname{int}(X) \neq \emptyset$*, $u \in \partial C$. Then there exists a hyperplane $H$ such that $u \in H$ and $C \subseteq \overline{H^+} = \left\{x \in \mathbb{R}^d : l(x) \geq z\right\}$.*



## 2.2 Faces

**2.11 Definition.** A *face* $F$ of a convex set $C \subseteq \mathbb{R}^d$ is a convex subset of $C$ such that

$$\forall x, y \in C. \frac{1}{2}(x+y) \in F \implies x, y \in F$$

Red lines are no faces.

**2.12 Definition.** An *exposed face* of a convex set $C \subseteq \mathbb{R}^d$ is the intersection of $C$ with a *supporting hyperplane* $H$, i.e. $C \subseteq \overline{H^+}$.

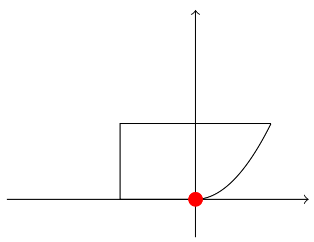
**2.13 Lemma.** *Every exposed face actually is a face.*

*Proof.* Let $x, y \in C$. Then

$$\frac{1}{2}(x+y) \in C \cap H \implies \frac{1}{2}l(x) + \frac{1}{2}l(y) = z \implies l(x) = l(y) = z \implies x, y \in C \cap H \qquad \square$$

The converse is true for polytopes, but not in general.

**2.14 Example.** picture



Here the origin is a face, but not exposed.

**2.15 Corollary (to is Isolation Theorem).** *Let $C \subseteq \mathbb{R}^d$ convex, closed, $\operatorname{int}(X) \neq \emptyset$ and $u \in \partial C$. Then $u$ is contained in a proper exposed face $F$ of $C$ (proper means $F \neq \emptyset, C$).*

**2.16 Definition.** An *extreme point* of a convex set $C \subseteq \mathbb{R}^d$ is a 0-dimensional face. We denote it with $\operatorname{ex}(C)$.

**2.17 Theorem.** *If $C \subseteq \mathbb{R}^d$ is convex and compact, then $C = \operatorname{conv}(\operatorname{ex}(C))$.*

Hence every bounded polyhedron is a polytope (defined as convex hull of a finite set of points).

## 2.3 Duality/Polarity

**2.18 Definition.** Let $C \subseteq \mathbb{R}^d$. The *polar* of $C$ is

$$C^\circ = \left\{ l \in \mathbb{R}^d : l \neq \mathbf{0}, \forall x \in C. l(x) \leq 1 \right\}$$

The *dual* cone of $C$ is

$$C^\vee = \left\{ l \in \mathbb{R}^d : l \neq \mathbf{0}, \forall x \in C. l(x) \geq 0 \right\}$$

**2.19 Example.**

- $\left(\mathbb{R}^d_{\geq 0}\right)^\vee = \mathbb{R}^d_{\geq 0}$
- $\left(\mathbb{S}^m_+\right)^\vee = \mathbb{S}^m_+$

**2.20 Theorem (Biduality, Bipolarity).** *For any $C \subseteq \mathbb{R}^d$ we have*

$$(C^\circ)^\circ = \mathrm{cl}\left(\mathrm{conv}\left(C \cup \{0\}\right)\right)$$

**2.21 Remark.** Suppose $C$ is closed, convex and $0 \in \mathrm{int}(C)$. Then the extreme points of $C^\circ$ correspond almost to irredundant linear inequalities defining $C$.

## 2.4 Homogenisation of Convex Sets

**2.22 Definition.** Let $C \subseteq \mathbb{R}^d$. Put the *homogenisation*

$$\widehat{C} = \mathrm{conv}(C \times \{1\}) = \mathrm{cone}\left(\left\{(x, 1) : \mathbb{R}^{d+1} : x \in C\right\}\right)$$

Then affine combinations somehow correspond to linear combinations.

**2.23 Theorem.** $\widehat{C^\circ} = -\left(\widehat{C}\right)^\vee$

**2.24 Lemma.** *If $C$ is compact, closed, then $\widehat{C}$ is closed and* pointed *(meaning $\widehat{C} \cap -\widehat{C} = \{0\}$). Up to change of coordinates, the converse is also true.*

# 3 Non-negative Polynomials

**3.1 Definition.** A real polynomial $p \in \mathbb{R}[x_1, \ldots, x_n]$ is called non-negative if $\forall x \in \mathbb{R}^n . p(x) \geq 0$. A polynomial is called sum of squares if it can be written as $p = \sum q_i^2$ for $q_i \in \mathbb{R}[x]$.

**3.2 Example.**

- Obviously $p = 1$ is non-negative.
- $p = 1 + x^2 = 1^2 + x^2$ is SOS.
- $p = 2x^4 - 2x^2 + 1 = x^4 + (x^2 - 1)^2$ is SOS.

**3.3 Theorem (Hilbert, 1888).** *Non-negative polynomials = SOS only in the following 3 cases*

1. *univariate, $n = 1$*
2. *quadratic, “$2d = 2$*
3. *bivariate of degree 4, $(n, 2d) = (2, 4)$*

**3.4 Example (Motzkin Polynomial).** The first known explicit example for non-equality is

$$M(x, y) = x^4y^2 + x^2y^4 + 1 - 3x^2y^2$$

It is non-negative by AM-GM-inequality.

So we expand the question, what happens for rational functions. Three equivalent formulations are

$$p = \sum \left(\frac{f_i}{g_i}\right)^2 \qquad\qquad p \cdot r^2 = \sum f_i^2 \qquad\qquad p \cdot \sum h_i^2 = \sum f_i^2$$

Hilbert showed "yes" for $n = 2$. In particular $M(x,y)(1+x^2+y^2)$ is SOS. Even more, for $2d = 6$, then quadratic multipliers of degree 2 suffice.
This became Hilbert's 17th problem: What about $n \geq 3$? It was solved in the affirmative by Artin-Schreier in 1928.
There still remain the question how to find such a decomposition. In particular we need a bound on the degree of the $h_i$. The known bounds greatly differ (linear versus exponential tower).
So far we only regarded global non-negativity. But what if we restrict ourselves to some set defined by polynomial, inequalities?
Say $A = \{x \in \mathbb{R}^n : f(x) \geq 0\}$. Then obviously $p = f \cdot \mathrm{SOS} + \mathrm{SOS} \geq 0$ on $A$. For further constraints $A = \{x \in \mathbb{R}^n : g_1(x) \geq 0, g_2(x) \geq 0\}$, as obvious non-negative polynomials we have

$$\mathrm{SOS} + g_1 \cdot \mathrm{SOS} + g_2 \cdot \mathrm{SOS} + g_1 g_2 \cdot \mathrm{SOS}$$

which can be expanded to arbitrary many constraints.

## 3.1 Positivstellensätze

**3.5 Theorem (Krivine, Stengel).** *Assume $f \geq 0$ on a closed semialgebraic set, defined by polynomial inequalities $g_i(x) \geq 0$. Then $f \cdot (1 + \mathrm{SOS})$ is the set of obviously non-negative polynomials.*

**3.6 Theorem (Schmüdgen).** *If $f > 0$ on a compact semialgebraic set, then $f$ is obviously non-negative.*

**Exercise.** *If you look at the cusped cubic $A : y^2 - x^3 = 0$, then $f = x$ is non-negative on $A$, but $f$ is nor obviously non-negative in any degree. If we take $f + \varepsilon$, then certificates exist, but degree $\to \infty$ as $\varepsilon \to 0$.*

**3.7 Theorem (Putinar).** *If $f > 0$ on a compact semialgebraic set, and a small extra condition, we have*

$$f = \mathrm{SOS} + \sum g_i \cdot \mathrm{SOS}$$

*which means, we can avoid the combinatorial blow up.*

## 3.2 Computationally Find SOS Certificates

Go back to our example $f = 2x^4 - 2x^2 + 1$.
Each summand is of type $(cx^2 + bx + a)^2$, so write $\alpha = (c, b, a)$ and $\vec{x} = (x^2, x, 1)$.

finish

Applied to the example this means

$$2x^4 - 2x^2 + 1 = (1, x, x^2) \begin{pmatrix} \alpha_0 & \alpha_1\alpha_2 \\ \alpha_1 & \alpha_3\alpha_4 \\ \alpha_2 & \alpha_4\alpha_5 \end{pmatrix} \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix}$$

Comparing the coefficients, we get

$$\alpha_0 = 1 \qquad 2\alpha_1 = 1 \qquad 2\alpha_3 + \alpha_4 = -2 \qquad 2\alpha_4 = 0 \qquad \alpha_5 = 2$$

and the above matrix has to be positive semidefinite. Solving this kind of problems can be done, although we suffer from a serious blow up when constructing problem, where both $n$ and $2d$ become larger. (10 already is a large number in this case.)

This can be applied for optimisation problem. A general optimisation problem is

$$\min \{f(x) : x \in K\} = \max \{\gamma : \forall x \in K. f(x) - \gamma \geq 0\}$$

This we relax to $f(x) - \gamma$ is obviously non-negative on $K$ and apply our previous theory. The method is called Lasserre relaxation.

**3.8 Example (Max-Cut).** Given a graph $G = (V, E)$ we want to find the maximal cut. Our variables are $x_i \in \{-1, 1\}$ given by equations $x_i^2 - 1 = 0$. So we have the problem

$$\max \left\{ \frac{1}{2} \left( |E| - \sum_{i,j \in V} x_i x_j \right) : \forall i. x_u^2 - 1 = 0 \right\}$$

The degree 2 SOS relaxation is the Goemans-Williamson algorithm.

# 4 Conic Programming

The lecture will follow the book of Barvinok.

picture

To be more precise, we have the following setup: Domain $D$ is a section of a cone. Let $K \subseteq \mathbb{R}^l$ a closed convex cone, and $\varphi : \mathbb{R}^l \to \mathbb{R}^m$ linear, with some point $b \in \mathbb{R}^m$. Then $D = K \cap \varphi^{-1}(b)$ is called *set of feasible points*. $\lambda(x) = \langle x, c \rangle$ for some $c \in \mathbb{R}^l$ is the *target function*. The task is to find

$$\begin{tikzcd} \mathbb{R}^l \arrow[r, "\varphi"] \arrow[d, "\lambda"] & \mathbb{R}^m \ni b \\ \mathbb{R} & \end{tikzcd}$$

$$\gamma = \inf \{\langle x, c \rangle : \varphi(x) = b, x \in K\}$$

Any $x \in D$ with $\gamma = \langle x, c \rangle$ is an *optimal point.*

## 4.1 Duality

$\varphi^\vee : \mathbb{R}^m \to \mathbb{R}^l$ such that

$$\forall x \in \mathbb{R}^l, y \in \mathbb{R}^m. \langle \varphi(x), y \rangle = \langle x, \varphi^\vee(y) \rangle$$

is the *dual linear map* and

$$K^\vee = \left\{ a \in \mathbb{R}^l : \forall x \in K. \langle x, a \rangle \geq 0 \right\}$$

is the *dual cone.*
Let $K_2 \subseteq \mathbb{R}^m$ be a cone. The *primal problem* is

$$\gamma = \inf \{\langle x, c \rangle : \varphi(x) - b \in K_2, \varphi(x) = b, x \in K\}$$

The corresponding *dual problem* is

$$\beta = \sup \{\langle y, b \rangle : c - \varphi^\vee(y) \in K^\vee, y \in K_2^\vee\}$$

In practice we usually put $K_2 = \{0\}$, which yields $K_2^\vee \mathbb{R}^m$. In red we have the original condition, in blue the simplified one.

**4.1 Theorem (Weak Duality).** *$\gamma \geq \beta$.*

The difference $\gamma - \beta$ is called *duality gap*.
As further optimality criteria we have the following.

**4.2 Lemma.** *Assume $\gamma = \beta$. If $x, y$ are feasible, then the following are equivalent*

- *$x, y$ are optimal*
- *$\langle x, c - \varphi^\vee(b)\rangle = 0$ and $\langle y, \varphi(x) - b\rangle = 0$*

**LP** In this setting we have $K = \mathbb{R}^l_+$ the positive orthant (though strictly speaking it is the non-negative orthant). $D = \mathbb{R}^l_+ \cap \varphi^{-1}(b)$ is a *polyhedron*. Note that $K = K^\vee$ is *self-dual*.

**SDP** Our cone is

$$K = \mathbb{S}^n_+ \subseteq \mathbb{S}^n \cong \mathbb{R}^{\binom{n+1}{2}}$$

the cone of positive semi-definite $n \times n$-matrices and our product is $\langle A, B\rangle := \operatorname{Tr}(AB)$. Again we have $K = K^\vee$. The linear function $\varphi$ has the shape

$$\begin{aligned} \varphi : \mathbb{S}^n &\to \mathbb{R}^m \\ X &\mapsto (\langle X, A_1\rangle, \dots, \langle X, A_m\rangle) \end{aligned}$$

for some $A_i \in \mathbb{S}^n$. For short we write $X \succeq 0$ for $X \in K$ and $X \succeq Y$ for $X - Y \succeq 0$. The domain $D = \mathbb{S}^n_+ \cap \varphi^{-1}(b)$ is called a *spectrahedron*.

The primal problem is

$$\gamma = \inf \{\langle X, C\rangle : \forall i. \langle X, A_i\rangle = b_i, X \succeq 0\}$$

and its corresponding dual is

$$\beta = \sup \left\{ \langle b, y\rangle : C - \sum_{i=1}^m y_i A_i \succeq 0 \right\}$$

**4.3 Example.** 1. Compute the Lovasz-Theta-number for graphs (lies between clique-number and chromatic number).

2. Correlation matrices

$$\begin{pmatrix} 1 & x_{12} & x_{13} \\ x_{12} & 1 & x_{23} \\ x_{13} & x_{23} & 1 \end{pmatrix}$$

Here we want we find

$$\gamma = \inf \{x_{13} : x_{11} = x_{22} = x_{33} = 1, X \succeq 0\}$$

In practice, we usually have further inequalities on the variables.

3. Let $f \in \mathbb{R}[x_1, \dots, x_n]_{\leq 2d}$ of even degree. Take the vector of all monomials $m = (x^\alpha)_{|\alpha| \leq d}$. Then

$$\mathscr{G} := \left\{A : m^T A m = f\right\}$$

is an affine subspace of $\mathbb{S}^N_+$ where $N = \binom{n+d}{d}$. Now $f$ is a sum of squares iff $\mathscr{G}$ contains a psd-matrix. check
$\mathscr{G} \cap \mathbb{S}^N_+$ is a spectrahedron, over which we are optimising.

**HYP** Hyperbolic Programming: This is a conic programme for $K = C_e(f)$ the *hyperbolicity cone* of some polynomial $f$ hyperbolic wrt $e$. Here our dual problem will involve $K^\vee \neq K$. In general, $K^\vee$ will not even be a hyperbolicity cone.

## 4.2 Interior Point Methods

Take (D) the dual of an SDP. Let $D^*$ denote the domain of the dual problem.

**4.4 Lemma.** *Assume $D, D^*$ have interior points and $A_1, \ldots, A_m$ are linearly independent. Then $\gamma = \beta$, i.e. no duality gap.*

First note $\det\left(C - \sum y_i A_i\right) = 0$ on $\partial D^*$, and the optimum is attained at the boundary. But since the determinant is no convex, we use an alternative.

**4.5 Lemma.** *The function $X \mapsto -\log(\det(X))$ is strictly convex on $\mathbb{S}^n_{++}$.*

**4.6 Definition.** The function

$$B_\lambda(y) := \langle b, y\rangle + \lambda \cdot \log\left(\det\left(C - \sum y_i A_i\right)\right)$$

is called the logarithmic barrier function of (D) with parameter $\lambda$.

**4.7 Theorem.** *Let $y(\lambda)$ be the unique maximiser of $B_\lambda(y)$ on $D^*$. Then $\lim_{\lambda\to 0} y(\lambda)$ is an optimal point.*

The path $\{y(\lambda) : \lambda > 0\}$ is called the *central path.*
For HYP we use $\log(f)$, reasonably restricted.

# 5 Geometry of Hyperbolicity Cones

The lecture follows the paper “Hyperbolic Programmes and their Derivative Relaxations” by Hames Renegar.
Fix some hyperbolic polynomial $f \in \mathbb{R}[x_1, \ldots, x_n]_d$ hyperbolic wrt $e \in \mathbb{R}^n$.

1. $f(e) > 0$
2. $\forall x \in \mathbb{R}^n . f(te - x) \in \mathbb{R}[t]$ is real rooted

We always order the eigenvalues $\lambda_1(x) \leq \ldots \leq \lambda_d(x)$. Then the hyperbolicity cone is

$$C_e(f) = \{x \in \mathbb{R}^n : \lambda_1(x)\}$$

which means all eigenvalues are non-negative.

**5.1 Remark.** Observe that

$$\lambda_j(sx + te) = \begin{cases} s\lambda_j(x) + t & : s \geq 0 \\ s\lambda_{d-j}(x) + t & : s \leq 0 \end{cases}$$

and

$$f(x) = f(e) \cdot \prod_{j=1}^{d} \lambda_j(x)$$

**5.2 Proposition.** *$C_e(f)$ is the closure of the connected component $S$ of $\{x \in \mathbb{R}^n : f(x) \neq 0\}$ containing $e$.*

$$\begin{tikzcd}
 & x+\bar{t}e \arrow[rr] & & e+\bar{t}e \arrow[dl] \\
x \arrow[ur] \arrow[rr, dash] & & e &
\end{tikzcd}$$

*Proof.* Show mutual inclusion

**$S \subseteq C_e(f)$:** $S$ is connected and $\lambda_1(e) = \lambda_d(e) = 1$ (the latter since $f(te - e) = (t-1)^d f(e)$). Since $\lambda_1(x)$ is a continuous function of $x$ and $f(x) = 0 \Leftrightarrow \lambda_1(x) = 0$.

**$C_e(f) \subseteq S$:** We will walk along a path

For all sufficiently large $\bar{t}$, all $<\in [x, e]$ satisfy

$$0 < f\left(y + \bar{t}e\right) = \bar{t}^d f\left(\frac{1}{\bar{t}}y + e\right)$$

□

## 5.1 Boundary Basics

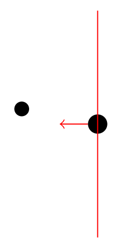
**5.3 Definition.** The *multiplicity* of $x \in \mathbb{R}^n$ wrt $f$ the multiplicity of $0$ as an eigenvalue of $x$.

**5.4 Remark.**
- Note $\mathrm{mult}(x) > 0 \Leftrightarrow f(x) = 0$.
- $\mathrm{mult}(x) = d - \mathrm{rank}(x)$

**5.5 Theorem.** *The set $\{x \in \mathbb{R}^n : \mathrm{mult}(x) = 1\}$ is (if non-empty) a codimension 1 analytic submanifold.*

**5.6 Lemma.** *The gradient at these points does not vanish, i.e.* $\mathrm{mult}(x) = 1 \Leftrightarrow f(x) = 0 \wedge \nabla f(x) \neq 0$.

*Proof.* Observe $\frac{d}{dx} f(te - x) = (\nabla f(te - x)) \cdot e$. Assume $f(x) = 0$. If $\nabla f(x) = (-1)^{d-1} \nabla f(-x)$, then $\mathrm{mult}(x) = 1$. If $\nabla f(x) \neq 0$, then $\{y \in \mathbb{R}^n : \nabla f(x) \cdot y = 0\}$ is the supporting hyperplane to $C_e(f)$ at $x$.



picture

□

## 5.2 Curvature of the Boundary

**5.7 Proposition.** *Let $f(x) = 0$, $\nabla f(x) \neq 0$ and $x \in C_e(f)$. If $\nabla f(x) \cdot v = 0$, then $v^t f''(x) v \leq 0$.*

Under the above assumptions, we also have

$$v^t f''(x) v = \frac{d^2}{dt^2} f(x + tv)_{|t=0}$$

**5.8 Theorem.** *If $x \in \partial C_e(f)$, $\operatorname{mult}(x) = 1$ and $\nabla f(x) \cdot v = 0$, then one of the following holds*

1. $\forall t \in \mathbb{R}. f(x + tv) = 0$ *and* $\exists \varepsilon > 0. \forall t \in (-\varepsilon, \varepsilon). x + tv \in C_e(f)$
2. $v^t f(x) v < 0$

So if the curvature is not negative, then locally we have a flat face.

## 5.3 Derivative Cones

**5.9 Claim.** $D_e f = \nabla f \cdot e = \sum e_i \frac{\partial f}{\partial x_i}$ *is hyperbolic wrt $e$.*

picture

*Derivative Cone: $C_e(D_e f) \subseteq C_e(f)$.*

**5.10 Theorem.** *For integers $m \geq 2$, the multiplicity of $x$ wrt $D_e f$ is one less than the multiplicity of $x$ wrt $f$, i.e.* $\operatorname{mult}'(x) = \operatorname{mult}(x) - 1$*. Also if* $\operatorname{mult}'(x) = 1$ *and* $\operatorname{mult}(x) > 0$*, then* $\operatorname{mult}(x) = 2$.

**5.11 Theorem.** *Suppose $C_e(f)$ is pointed, $d \geq 3$. Let $x \in C_e(D_e f) \setminus C_e(f)$,* $\operatorname{mult}'(x) = 1$*, $v \in T'_x$ some tangent vector to the derivative cone. If $v \notin \mathbb{R}x$, then $v^t (D_e f)'' v < 0$.*

**5.12 Corollary.** *So $x$ is an exposed extreme direction of $C_e(D_e f)$.*

## 5.4 Higher Dimensions and Faces Exposed

$$C_e(f) \subseteq C_e(D_e f) \subseteq \dots C_e\left(D_e^{d-1} f\right)$$

Then the above Theorem 5.11 translates to

**5.13 Theorem.** *Suppose $C_e(f)$ is pointed, $d \geq 3$. Let $x \in C_e(D_e^k f) \setminus C_e(f)$,* $\operatorname{mult}^{(k)}(x) = 1$*, $v \in T_x^{(k)}$ some tangent vector to the derivative cone. If $v \notin \mathbb{R}x$, then $v^t (D_e^k f)'' v < 0$.*

So $x$ is an exposed extreme direction of $C_e(D_e^k f)$.

**5.14 Theorem.** *All faces of $C_e(f)$ are exposed.*

The proof consists of showing the following two propositions.

**5.15 Proposition.** *For $k = 0, 1, \dots, d-2$ each proper face of $C_e\left(D_e^k f\right)$ either is a face of $C_e(f)$ or it is an exposed extreme ray not in $C_e(f)$.*

*Proof.* Just a rephrasing of Theorem 5.13. □

**5.16 Proposition.** *Let $F$ be a proper face of $C_e(f)$ and let $x \in$* $\operatorname{relint}(F)$*. Set $m =$* $\operatorname{mult}(x)$*. Then $F$ is a proper face of $C_e\left(D_e^{m-1} f\right)$.*

# 6 Sums of Squares in Extremal Combinatorics

I will diverge from the notation on the board.
We want to tackle some problems in graph theory. So we index our variables by the edges, or have them double-indexed by the vertices.
Let $G = (V, E)$ be a simple graph. We have 0/1-problems, so we include constraints $x_{ij}^2 = x_{ij}$ for all $\{i, j\} \in E$, or for all $i, j \in V$.

**6.1 Example.** We want to minimise the density of triangles. To check whether $1, 2, 3$ forms a triangle, we use $x_{12}x_{23}x_{13}$. The density function can then be written as

$$\mathrm{Sym}_n\left(x_{12}x_{23}x_{13}\right) = \frac{1}{n!}\sum_{\sigma\in S_n}\sigma(x_{12}x_{23}x_{13})$$

In general, we want to introduce notation.

- $\{1,2\} \cong x_{12}$
- $P_3 \cong x_{12}x_{23}$
- $\{1,*\} \cong \mathrm{Sym}_{n-1}(\{1,2\}) = \frac{1}{n-1}\sum_{i\geq 2} x_{1i}$

What are inequalities for subgraph-densities, e.g. for $P_2$ or $C_3$? We will abuse notation and identify a graph with its subgraph-density. Trivially we have $0 \leq H \leq 1$ for any subgraph. But mainly we are interested in asymptotic behaviour, i.e. inequalities that are valid on accumulation points.
A small miracle

$$\{1,2\}\times\{1,3\} = \{\{1,2\},\{1,3\}\}$$
$$x_{12}\times x_{13} = x_{12}x_{13}$$

which means, we take fully labelled graphs and glue them together on the common labels. Thanks to our constraints we can eliminate squares as in

$$(1-2-3)\times(2-3-4) = 1-2-3-4$$
$$x_{12}x_{23}\times x_{23}x_{34} = x_{12}x_{23}x_{34}$$

For unlabelled graphs, this becomes tricky

$$\begin{aligned}(1-*)\times(1-*) &\cong \left(\frac{1}{n-1}\sum_{i\geq 2}x_{1i}\right)^2 = \underbrace{\frac{1}{(n-1)^2}\sum_{i\geq 2}x_{1i}}_{\to 0} + \frac{2}{(n-1)^2}\sum_{i>j\geq 2}x_{1i}x_{1j}\\ &\approx \frac{1}{\binom{n-1}{2}}\sum_{i>j\geq 2}x_{1i}x_{1j} \cong (*-1-*)\end{aligned}$$

which is the graph we expected. Note how we needed the asymptotic behaviour here.

**6.2 Remark.** Full symmetrisation just removes all labels (just a big average).

We can allow forbid edges by using $(1-x_{ij})$. This allows us to find densities of induced subgraphs. To regard something mildly non-trivial, we take

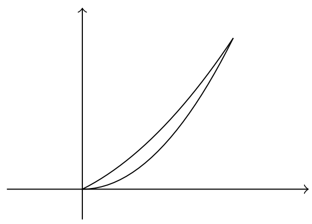
$$\begin{aligned}\mathrm{Sym}_n\left(((1-*)-(2-*))^2\right) &= \mathrm{Sym}_n\left((1-*)^2+(2-*)^2-2((1-*)\times(2-*))\right)\\ &= \mathrm{Sym}_n\left((*-1-*)+(*-2-*)-2(1-*,2-*)\right)\\ &= 2(*-*-*)-2(*-*)^2\end{aligned}$$

In terms of graphs, this means $P_3 - P_2^2 \geq 0$.

**Exercise.** *Show that this inequality is tight on regular graphs (all vertices same degree). This means: Take sequence $G_1,\dots,G_k,\dots$ of regular graphs. If $P_2(G_i)\to d$, then $P_3(G_i)-P_2(G_i)^2\to 0$ as $k\to\infty$.*

$$\mathrm{Sym}_n\left((12-23+34-14)^2\right) = 4(*-*) + 8(*-*)^2 - 8(*-*-*)$$

which shows $P_2 + P_2^2 - 2P_3 \geq 0$.
So with this little bit of effort, we showed that our densities lie in the small area given by:



# 7 Determinantal Representations

Suppose $f(x) = \det(x_1 A_1 + \ldots + x_n A_n)$ where $A_i \in \mathbb{S}^d(\mathbb{R})$, so $\deg(f) = d$. The term $\sum x_i A_i$ is called *real symmetric matrix pencil of size $d \times d$*. If $A(e) \succ 0$, then call this *definite determinantal representation* of $f$. This implies $f$ is hyperbolic wrt $e$.

*Proof.* Wlog we restrict to $A(e) = I_d$. Then $f(te - v)$ is the characteristic polynomial of $A(v)$, which is real rooted. $\square$

Furthermore $C_e(F) = \{v \in \mathbb{R}^n : A(V) \succeq 0\}$ is a spectrahedron. So this is a certificate for hyperbolicity (see: SOS as certificate for non-negativity).

**7.1 Lemma.** *Not every hyperbolic polynomial has a (definite) determinantel representation.*

*Proof.* For $n, d$ large, we simply count the dimension. $\square$

For smaller parameters, however, thing look better. Regard $n = 2$, i.e. $f(x_1, x_2)$ homogeneous of degree $d$. If $x_2 \nmid f$, then

$$f(x_1, 1) = c \cdot \prod_{j=1}^{d} (x_1 - \alpha_j) = c \cdot \det\left(x_1 I_d - \mathrm{diag}(\alpha_1, \ldots, \alpha_d)\right)$$

Note that all $\alpha_j$ are real.
For $n = 3$, things are more difficult.

**7.2 Theorem (Helton-Vinnikov, 2004/Lax-conjecture).** *If $f \in \mathbb{R}[x, y, z]$ is hyperbolic wrt $e$, then $f$ has a real symmetric determinantal representation at $e$.*

The same is not true for $n > 3$. Furthermore the representations are hard to compute, but they are very useful.

**7.3 Example.** Consider the cubic

$$f = x^3 - x^2 z - x z^2 - y^2 z + z^3$$

To show hyperbolicity we take

$$A = x \begin{pmatrix} -2 & 0 & -1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \end{pmatrix} + y \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + z \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -2x+2z & y & -x+z \\ y & -x+z & 0 \\ -x+z & 0 & z \end{pmatrix}$$

then $f = \det(A)$, so $f$ is hyperbolic (found by "trial and error").

*Proof of Theorem 7.2, general idea.* Suppose $\det(A) = f$. Define the *adjugate matrix*

$$A^{\mathrm{adj}} := \left( (-1)^{i+j} \underbrace{\det\left(A'_{j,i}\right)}_{(d-1)\text{-minors}} \right)_{i,j}$$

Then

$$A \cdot A^{\mathrm{adj}} = A^{\mathrm{adj}} \cdot A = \det(A) \cdot I_d$$

Let $p \in \mathbb{R}^n$ with $f(p) = 0$. If $(\nabla f)(p) \neq 0$, we have

$$\underbrace{A(p)}_{\mathrm{rk}=d-1} \cdot \underbrace{A^{\mathrm{adj}}(p)}_{\mathrm{rk}=1} = 0$$

so $\operatorname{Ker} A(p)$ is 1-dimensional. The map $p \mapsto \operatorname{Ker} A(p)$ is called *line bundle* on $\{f = 0, \nabla f \neq 0\}$. This is parametrised by any one column of $A^{\mathrm{adj}}$. ☐

This building of the determinantal representation is called “Dixon process”.

## 7.1 Generalised Lax Conjecture

**7.4 Claim.** *Every hyperbolicity cone is a spectrahedron.*

Given $f$ irreducible, hyperbolic wrt $e$, there exists $A$ such that $G = \det(A)$, $A(e) \succ 0$ and $C_e(f) = C_e(g)$. If $\deg g \geq \deg f$, this means $f \mid g$.
It was shown, that taking $g$ as a power of $f$ does not suffice, as shown by Brändén.
Equivalently: Given such $f$, there exists $h$ hyperbolic wrt $e$ such that $C_e(f) \subseteq C_e(h)$ and $f \cdot h$ has a determinantal representation at $e$.

**7.5 Theorem (Mario Kummer).** *This is true, up to the inclusion.*

This approach is similar to Hilbert’s 17th problem. We cannot have SOS, but some multiples has an SOS-representation.

## 7.2 Hermite Method

Suppose $H \in \mathbb{R}[t]$ is monic, $\deg h = d$. Then we can write

$$h = \sum_{j=0}^{d} a_j t^{d-j} = \prod_{j=1}^{d} (t - \alpha_j)$$

To count the nuber of real roots, there is a method by Sturm, but here we want to focus on another one by Hermite.

**7.6 Definition.** The *power sum* is $\omega_k := \sum_{j=1}^{d} \alpha_j^k$.

For these we have the Newton identities, which express $\omega_k$ in the coefficients $a_i$, e.g.

$$\begin{aligned}
\omega_0 &= d \\
\omega_1 &= -a_1 \\
\omega_2 &= a_1^2 - 2a_2 \\
\omega_3 &= -a^3 + 3a_1 a_2 - 3a_3
\end{aligned}$$

These we put in a matrix

$$H(k) := (\omega_{j+k-2})_{1\le j,k\le d}$$

**7.7 Theorem.** *$h$ is real rooted iff $H(k) \succeq 0$.*

Now suppose $f \in \mathbb{R}[x_1,\ldots,x_n]$ homogeneous of degree $d$, $e = (0,\ldots,0,1)$ and $f(e) = 1$. Then

$$f = \sum_{j=0}^{d} f_j(x_1,\ldots,x_{n-1}) \cdot x_n^{d-j}$$

**7.8 Corollary.** *The following are equivalent*

- *$f$ is hyperbolic wrt $e$*
- *for all $a \in \mathbb{R}^{n-1}$ the univariate $f(a,x_n) \in \mathbb{R}[x_n]$ is real rooted*
- *$H_{x_n}(f)(a) \succeq 0$ for all $a \in \mathbb{R}^{n-1}$, where $H_{x_n}$ is a symmetric matrix with entries in $x_1,\ldots,x_{n-1}$*

So we rephrased the question "hyperbolic" to "psd" (aka "non-negative"). This can be further translated to non-negativity of polynomials.

$$\forall a \in \mathbb{R}^{n-1}. M(a) \succeq 0 \Leftrightarrow 0 \le (y_1,\ldots,y_d)\cdot M \cdot (y_1,\ldots,y_d)^T \in \mathbb{R}[x,y]$$

**7.9 Theorem.** *If $f^r = \det(A)$, $A(e) = I_d$ for some $r \ge 1$, then $H(f)$ is SOS.*

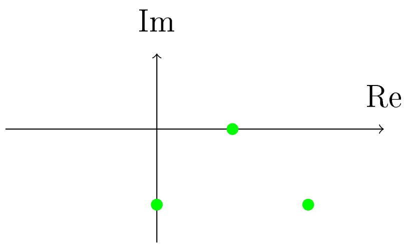
**7.10 Remark.** A polynomial matrix $M$ is SOS as above iff $M$ is a sum of matrix squares $M = \sum Q_i^T Q_i$.

# 8 Stable Polynomials

We follow Wagner: Multivariate stable polynomials and their applications.

**8.1 Definition.** A polynomial $f \in \mathbb{C}[x_1,\ldots,x_n]$ is *stable* if $f(z) \neq 0$ for all points $z \in \mathbb{C}^n$ with $\mathrm{Im}(z) \in \mathbb{R}^n_{>0}$ and *real stable* if $f$ is stable and $f \in \mathbb{R}[x_1,\ldots,x_n]$.

**8.2 Example.** Take $n = 1$ and $f = (x+i)(x-(2-i))(x-1)$.




Another example, this time real stable, is $f = (x-1)(x+1)(x-2)$.

**8.3 Proposition.** *Polynomial $f \in \mathbb{R}[x_1,\ldots,x_n]$ is real stable iff for all $a \in \mathbb{R}^n_{>0}$ and $b \in \mathbb{R}^n$, the polynomial $f(at+b) \in \mathbb{R}[t]$ is real rooted.*

*Proof.* **⇒:** Assume $f$ is not real rooted. Consider

$$f(\underbrace{(\alpha + i\beta)a + b}_{z}) = 0$$

where $a \in \mathbb{R}^n_{>0}$ and take $\beta > 0$. Then

$$\mathrm{Im}(z) = b \cdot a \in \mathbb{R}^n_{>0}$$

so $f$ is not stable.

**⇐:** Assume $f$ is not stable. Take $a \in \mathbb{R}^n_{>0}$ and $b \in \mathbb{R}^n$ such that $f(ia + b) = 0$. Then $f(ta + b)$ has root $t = i$, so it is not real rooted. □

**8.4 Corollary.** *For $f \in \mathbb{R}[x_1, \ldots, x_n]_d$ we have: $f$ is stable iff $f$ is hyperbolic wrt every $a \in \mathbb{R}^n_{>0}$.*

**8.5 Remark.** The following are stable

- $\prod_{i=1}^n x_i$
- $D_a f$ for stable $f$ and $a \in \mathbb{R}^n_{\geq 0}$
- elementary-symmetric polynomials $e_k(x_1, \ldots, x_n)$
- $\det\left(\sum x_i A_i + B\right)$ for $A_i \in \mathbb{S}^d_+$ and $B \in \mathbb{S}^d$

**8.6 Example.** Consider

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \qquad A_3 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Then we get the stable polynomial

$$\det\left(\sum x_i A_i\right) = \det\begin{pmatrix} x_1 + x_3 & x_3 \\ x_3 & x_2 + x_3 \end{pmatrix} = x_1x_2 + x_1x_3 + x_2x_3$$

More generally, gien $v_1, \ldots, v_n \in \mathbb{R}^d$, take $A_i = v_i v_i^T$. Then

$$\det\left(\sum x_i v_i v_i^T\right) \sum_{I \subseteq [n], |I| = d} \det\left(v_i : i \in I\right)^2 \cdot \prod_{i \in I} x_i$$

**8.7 Theorem (COSW,2004).** *If we have a stable polynomial of the form*

$$f = \sum_{I \subseteq [n], |I| = d} c_I \prod_{i \in I} x_i \in \mathbb{R}[x]$$

*then $\{I : c_I \neq 0\}$ are the bases of a matroid. The matroid is called* hyperbolic matroid.

In Example 8.6 we have $v_1 = (1, 0)$, $v_2 = (0, 1)$ and $v_3 = (1, 1)$. The set of bases is $\{\{1, 2\}, \{1, 3\}, \{2, 3\}\}$ and we clearly see that any pair of the vectors in linearly independent.
On the other hand, there are no $a, b \in \mathbb{R}^*$ such that $ax_1x_2 + bx_3x_4$ is stable.
More generally $e_k(x_1, \ldots, x_n)$ corresponds to the uniform matroid of rank $k$ on $n$ elements.
Now consider graphs $G(V, E)$ where $|V| = d + 1$ and $|E| = n$. For each edge $e = ij \in E$ define

$$v_{ij} := \begin{cases} e_i - e_j & : i < j \leq d \\ e_i & : j = d + 1 \end{cases}$$

| Operations preserving stability | Matroid operations | Operations on $\{v_1, \ldots, v_n\} \subseteq \mathbb{R}^d$ |
|---|---|---|
| $f \mapsto f_{\|x_i=0}$ | Deletion $M \mapsto M - i$, new bases $B(M-i) = \{b \in B : i \notin b\}$ | drop $v_i$ |
| $F \mapsto \frac{\partial f}{\partial x_i}$ | Contraction $M \mapsto M/i$, $B(M/i) = \{B \setminus \{i\} : i \in B\}$ | Project $v_j$ for $j \neq i$ onto $v_i^{\perp}$ |
| $f \mapsto \prod x_i \cdot f\left(\frac{1}{x_1}, \ldots, \frac{1}{x_n}\right)$ | Dual $M \mapsto M^*$, where $B(M^*) = \{[n] \setminus B : B \in B(M)\}$ | Columns of matrix, whose rows span the orthogonal complement of rowspan$(v_1, \ldots, v_n)$ |

**8.8 Example.** Take $G = K_4$, which means $d = 3, n = 6$. This yields a matrix

$$v = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{pmatrix} 12 & 13 & 23 & 14 & 24 & 34 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 & 0 & 1 \end{pmatrix}$$

**8.9 Theorem.** *We have*

$$F_G(x) = \det\left(\sum_{ij \in E} x_{ij} v_{ij} v_{ij}^T\right) = \sum_{T \text{ spanning tree of } G} \prod_{ij \in T} x_{ij}$$

**8.10 Example.** Take $G = K_3$.

$$\det\begin{pmatrix} x_{12} + x_{13} & -x_{12} \\ -x_{12} & x_{12} + x_{23} \end{pmatrix} = x_{12}x_{13} + x_{12}x_{23} + x_{13}x_{23}$$

For $F = K_4$ we get

$$f_{K_4} = \det\left(\sum_{ij \in E} x_{ij} v_{ij} v_{ij}^T\right) = 12P_4 + 4S_3$$

using our previous notation for polynomials.

**Remark (continuing Remark 8.5).**

- *graphical matroids*
- *matroid represented by* $v_1, \ldots, v_n$: $f = \det\left(\sum x_i v_i v_i^T\right)$.

## 8.1 Operations preserving stability

It turns out that representable matroids are a proper subset of hyperbolic matroids, which are a proper subset of all matroids. The first one is shown by "Vamos matroid", $(d = 4, n = 8)$; the other by "Fano matroid" $(d = 3, n = 7)$.

## 8.2 Reduce to Multiaffine Polynomials via Polarisation

Assume we have $f \in \mathbb{R}[x_1, \ldots, x_n]$, degree $d_i$ in variable $x_i$ (write $\deg_i f = d_i$).
The *polarisation* of $f$, written $P(f) \in \mathbb{R}[x_{1,1}, \ldots, x_{1,d_1}, \ldots, x_{n,1}, \ldots, x_{n,d_n}]$ is the unique multiaffine polynomial such that

- $P(f)$ is symmetric in $x_{j,1}, \ldots, x_{j,d_j}$ for all $j$
- we have

$$P(f)\left(\underbrace{x_1,\ldots,x_1}_{d_1},\ldots,\underbrace{x_n,\ldots,x_n}_{d_n}\right) = f$$

**8.11 Theorem.** *$F$ is stable iff $P(f)$ is stable.*

This means, we can restrict to multiaffine polynomials.

**8.12 Example.** The polynomial $f = \sum_{k=0}^{n} a_k x^k$ is real rooted iff

$$P(f) = \sum \frac{a_k}{\binom{n}{k}} e_k(x_1,\ldots,x_n)$$

is stable.

# 9 Interlacers

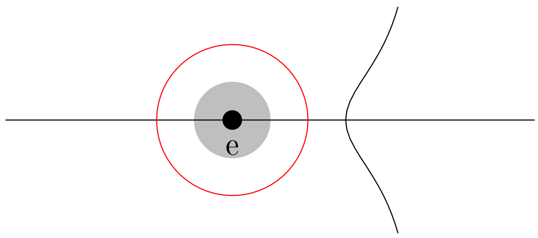
**9.1 Definition.** Let $f, g \in \mathbb{R}[t]$ be real rooted, $d = \deg f = \deg g + 1$. Suppose $\alpha_1, \ldots, \alpha_d$ are the roots of $f$, $\beta_1, \ldots, \beta_{d-1}$ are the roots of $g$, both including multiplicities. Then we say $g$ *interlaces* $f$, written $g \ll f$, if the roots of $g$ sit between the roots of $f$, i.e. $\alpha_i \leq \beta_i \leq \alpha_{i+1}$ for $1 \leq i < d$. See figure 1. We say $g$ *strictly interlaces* $f$ if all inequalities are strict.

**9.2 Example.** If $f$ is real rooted, then $f' \ll f$.

**9.3 Definition.** Let $f, g \in \mathbb{R}[x_1, \ldots, x_n]$ hyperbolic wrt $e$, and $\deg f = \deg g + 1$. Then $g$ *interlaces* $f$ if

$$\forall v \in \mathbb{R}^n . g(te + v) \ll f(te + v)$$

**9.4 Example.** Let $F, g \in \mathbb{R}[x, y, z]$, $e = (1, 0, 0)$ and fix $z = 1$ (dehomogenise).




**9.5 Example.**

1. Since $\frac{d}{dt} d(te + v) \ll f(te + v)$ for all $v$, we have

$$D_e f = \sum_{j=1}^{n} e_j \frac{\partial f}{\partial x_j} \ll f$$

   More generally, we have $D_a f \ll f$ for all $a \in C_e(f)$.

2. Let $f \det X$ for $X \in \mathbb{S}^n$, and $E \succeq 0$. Then $D_E(\det X) := \operatorname{tr}\left(E \cdot X^{\text{adj}}\right) \ll \det X$. More generally, if $f = \det\left(\sum x_i A_i\right)$ and $A(e) \succ 0$, then $\operatorname{tr}\left(E \cdot A^{\text{adj}}\right) \ll f$ (wrt $e$). In particular, we can pick $E = e_1 \cdot e_1^T$ (just single 1 in corner), then for the $d-1$-minor we have $\det(A'_{1,1}) \ll f$, which means the eigenvalues of $A'_{1,1}$ interlace the eigenvalues of $A$.

## 9.1 The Interlacer Cone

For simplicity assume $f$ is irreducible and $f(e) > 0$. Denote $Z(f) = \{a \in \mathbb{R}^n : f(a) = 0\}$. Now we are interested in

$$\mathrm{Int}_e(f) = \{g \in \mathbb{R}[x_1, \dots, x_n]_{d-1} : g \ll f, g(e) > 0\}$$

If we take the intervals, given by the roots of $f$ and look at the possible sign of $g$, we gat

$$\forall g_1, g_2 \in \mathrm{Int}_e(F). \forall a \in \mathbb{Z}(f). g_1(a) \cdot g_2(a) \geq 0$$

which means, both $g_i$ have the same sign. Furthermore

$$g_1 \ll f \wedge g_1(e) > 0 \wedge \forall a \in \mathbb{Z}(f). g_1(a) g_2(a) \geq 0 \implies g_2 \in \mathrm{Int}_e(f)$$

**9.6 Theorem.** *For $g \in \mathbb{R}[x_1, \dots, x_n]_{d-1}$ with $g(e) > 0$ the following are equivalent*

1. $g \in \mathrm{Int}_e(f)$
2. $D_e f \cdot g \geq 0$ *on* $Z(f)$
3. $D_e f \cdot g - f \cdot D_e g \geq 0$ *on* $\mathbb{R}^n$.

*Proof.* (1) $\Leftrightarrow$ (2) and (3) $\Rightarrow$ (2) we have basically done.
For (2) $\implies$ (3) regard the univariate case: $f, g \in \mathbb{R}[t]$ monic and real rooted. Then, if $g \ll f$, for the Wronskian we have

$$W(f, g) = f'g - g'f = g^2 \left(\frac{f}{g}\right)' \geq 0$$

□

**9.7 Corollary.** $\mathrm{Int}_e(f)$ *is a convex cone.*

**9.8 Theorem.** *For any $a \in \mathbb{R}^n$, we have $a \in C_e(f) \Leftrightarrow D_a f \ll f$.*

**9.9 Theorem.** *Define the* generalised Wronskian $\Delta_{e,a} = D_e f \cdot D_a f - f \cdot D_e D_a f$. *Then*

$$C_e(f) = \{a \in \mathbb{R}^n : D_a f \in \mathrm{Int}_e(f)\} = \{a \in \mathbb{R}^n : \forall x \in \mathbb{R}^n. \Delta_{e,a}(x) \geq 0\}$$

*So this is a slice of the cone of non.negative polynomials.*

This allows for an SOS-relaxation of $C_e(f)$, as $\{a \in \mathbb{R}^n : \Delta_{e,a} \in \mathrm{SOS}\}$.

**9.10 Theorem.** *If $f$ is determinantal, then $\Delta_{e,a}$ is SOS for all $a \in \mathbb{R}^n$.*

# 10 Greg III

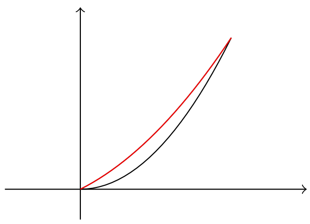
My notes are not useful. The board was mainly a collection of pictures.

**10.1 Remark.** The density of any tree in a regular graph is asymptotically $P_2^k$ where $k$ is the number of edges in the tree.

*Proof.* By induction on $k$, the base is $T = P_2$.
Now let $S = T + e$ some new tree, where we added a single edge (with its end as new leaf). Then $S(G) \approx P_2(G) \times T(G)$ if $G$ is a regular graph.

$$S(G) = \mathrm{Sym}\left((1 - *) \times T_1\right)(G) = \mathrm{Avg}\left((1 - *)(G) \times T_1(G)\right) = P_2(G) \times \mathrm{Sym}\left(T_1\right)(G)$$

where $T_1$ is $T$ with the label 1 at the appropriate place. □



To extend the picture form the last time, we add another result.

fix red line

The red line is due to Ahlswede-Katone. The extremal cases for the second half are Complete graphs + empty vertices (quasi-clique). For the first half it is the complement of a quasi-clique. For tree versus edge, we have the curve $y = x^{E(T)}$ as lower bound (due to Sidorenko). Reiher-wagner: $S_k$ is similar to $P_3$ and $P_5$ is similar to $P_3$. If $T$ has a perfect matching, then clique always wins.

**10.2 Definition.** A *moment curve* is the curve $t \mapsto (1, t, t^2, \dots, t^k) =: C_k$, where $t \in [0, 1]$.

# 11 Stable Polynomials II

See e.g. "Hyperbolic and stable polynomials in combinatorics and probability" by Pemantle, around 2000.

**11.1 Example (Newton's Inequalities).** Take $f = \sum_{k=0}^{d} a_k x^k \in \mathbb{R}[x]$ real rooted. Then

$$\left(\frac{a_k}{\binom{d}{k}}\right)^2 \geq \frac{a_{k+1}}{\binom{d}{k+1}} \cdot \frac{a_{k-1}}{\binom{d}{k-1}}$$

If $a_k \geq 0$, then $a_k^2 \geq a_{k+1} a_{k-1}$. Hence

$$\log(a_k) \geq \frac{\log(a_{k+1}) + \log(a_{k-1})}{2}$$

Therefore the roots are unimodal (on some concave curve).

## 11.1 Application: Graph Matching

A matching on a graph $G = (V, E)$ is a subset of disjoint edges. Let $m_k$ be the number of matchings with $k$ edges.

**11.2 Theorem.** *$p(x) := \sum_k m_k x^k$ is real rooted.*

*Proof.* Let $n := |V|$. Then

$$M_G(x_1, \dots, x_n) := \prod_{ij \in E} (1 - x_i x_j)$$

is stable. Consider

$$T_{MA} : \mathbb{R}[x_1, \dots, x_n] \to \mathbb{R}[x_1, \dots, x_n]$$
$$T_{MA}(x^\alpha) = \begin{cases} x^\alpha & : \forall i. \alpha_i \leq 1 \\ 0 & : \text{else} \end{cases}$$

By some theorem of Borcea and Brändén and some exercise, $T_{MA}$ preserves stability.

$$T_{MA}(M_G) = T_{MA}\left(\sum_{S\subseteq E}(-1)^{|S|}\prod_{ij\in S}x_ix_j\right) = \sum_{\substack{M\subseteq E\\ \text{matching}}}(-1)^{|M|}\prod_{ij\in M}x_ix_j$$

again is stable. Consider

$$T_{MA}(M_G)(x,\dots,x) = \sum_k(-1)^k m_k x^{2k} = p(-x^2)$$

is stable, thus also real rooted. By another exercise, $p(x)$ thus is real rooted. $\square$

## 11.2 Multiaffine Polynomials → SUBMODULAR

**11.3 Theorem (Brändén).** *If $f\in\mathbb{R}[x_1,\dots,x_n]$ is stable and $i,j\in[n]$, then*

$$\Delta_{ij} = \frac{\partial f}{\partial x_i}\cdot\frac{\partial f}{\partial x_j} - f\cdot\frac{\partial^2 f}{\partial x_i\partial x_j} \geq 0$$

*for all $x\in\mathbb{R}^n$.*

**11.4 Definition.** A function $F: 2^{[n]}\to\mathbb{R}\cup\{-\infty\}$ is *submodular* if for all $S,T\subseteq[n]$ we have

$$F(S)+F(T)\geq F(S\cap T)+F(S\cup T)$$

Equivalently for all $S\subseteq[n]$ and $i,j\notin S$ we have

$$F(S\cup\{i\})+F(S\cup\{j\})\geq F(S)+F(S\cup\{i,j\})$$

**11.5 Proposition.** *If we have a stable function*

$$f=\sum_{S\subseteq[n]}c_S\prod_{i\in S}x_i$$

*with $c_S\geq 0$, then $F(S):=\log(c_S)$ is submodular.*

*Proof.* We have

$$0\leq\Delta_{ij}f(0)=c_{\{i\}}c_{\{j\}}-c_\emptyset c_{\{i,j\}}$$

Therefore $F(\{i\})+F(\{j\})\geq F(\emptyset)+F(\{i,j\})$.
Note that the polynomial

$$\prod_{k\in S}\frac{\partial}{\partial x_k}f=\sum_{S\subseteq T\subseteq[n]}c_T\prod_{i\in T\setminus S}x_i$$

is stable.

$$\left(\prod_{k\in S}\frac{\partial}{\partial x_k}f\right)_{x=\mathbf{0}}=c_S$$

and taking the Wronskian yields

$$\Delta_{ij}\left(\prod_{k\in S}\frac{\partial}{\partial x_k}f\right)_{x=\mathbf{0}}=c_{S\cup\{i\}}c_{S\cup\{j\}}-C_Sc_{S\cup\{i,j\}}\geq 0$$

Applying log yields the result. $\square$

**11.6 Remark (Application).** Suppose $A \in \mathbb{S}_+^n$. Then we get a stable polynomial with non-negative coefficients by

$$f(x) := \det\left(\operatorname{diag}(x_1, \ldots, x_n) + A\right) = \sum_{S \subseteq [n]} \det\left(A[S^c]\right) \prod_{i \in S} x_i$$

where $A[S^c]$ denotes the principal minor of $A$, whose rows/columns are not in $S$. Then $F(S) := \log\det(A[S^c])$ is submodular.

## 11.3 Probability Distributions

Let $\mu : 2^{[n]} \to \mathbb{R}_{\geq 0}$ be a probability distribution. This means

$$\mu(S) \geq 0 \qquad\qquad \sum_{S \subseteq [n]} \mu(S) = 1$$

Define

$$f_\mu(x_1, \ldots, x_n) := \sum_{S \subseteq [n]} \mu(S) \prod_{i \in S} x_i$$

**11.7 Definition.** If $f_\mu$ is stable, then call $\mu$ *strongly Rayleigh*.

$$\frac{\partial f_\mu}{\partial x_i} = \sum_{i \in S \subseteq [n]} \mu(S) \prod_{j \in S - i} x_j$$

$$\frac{\partial f_\mu}{\partial x_i}(\mathbf{1}) = \operatorname{Prob}_\mu(i \in S)$$

Recall $\Delta_{ij}(f_\mu) \geq 0$, so in particular

$$0 \leq \Delta_{ij}(f_\mu)(\mathbf{1}) = \frac{\partial f_\mu}{\partial x_i}(\mathbf{1}) \cdot \frac{\partial f_\mu}{\partial x_j}(\mathbf{1}) - f(\mathbf{1}) \cdot \frac{\partial^2 f_\mu}{\partial x_i \partial x_j}(\mathbf{1})$$

Translating back to probability, we get

$$\operatorname{Prob}(i \in S) \cdot \operatorname{Prob}(j \in S) \geq \operatorname{Prob}(i, j \in S)$$

which means the events $i \in S$ and $j \in S$ are negatively correlated.

**11.8 Remark (Application: Spanning Trees).** Given graph $G$ with $t$ spanning trees. Define measure $\mu : 2^E \to \mathbb{R}_{\geq 0}$ via

$$\mu(S) = \begin{cases} \frac{1}{t} & : S \text{ is a spanning tree} \\ 0 & : \text{else} \end{cases}$$

define the (stable) polynomial

$$g_\mu := \frac{1}{t} \sum_{\substack{T \subseteq E \\ \text{spanning tree}}} \prod_{e \in T} x_e$$

Then for all $e, e' \in E$ we have

$$\operatorname{Prob}(e \in T) \cdot \operatorname{Prob}(e' \in T) \geq \operatorname{Prob}(e, e' \in T)$$

# 12 Ranks on (the Boundary of) Spectrahedra

**12.1 Definition.** A *spectrahedron* $S = \mathscr{A} \cap \mathbb{S}^N_+$ is the solution space of an SDP, i.e. $\mathscr{A} \subseteq \mathbb{R}^N$ is an affine subspace.

**12.2 Theorem (Face Lattice of $\mathbb{S}^N_+$).** *Suppose $L \subseteq \mathbb{R}^N$ is a linear subspace. Define $F_L := \{X \in \mathbb{S}^N_+ : L \subseteq \ker X\}$.*

1. *For every $X \in \mathbb{S}^N_+$, there is a particular subspace $F_{\ker X}$, which is the unique face of $\mathbb{S}^N_+$ that contains $X$ in its relative interior.*
2. *Assume* $\operatorname{codim}(L) = r$. *Then* $\dim(F_L) = \binom{r+1}{2}$ *and there exists $O \in O(N)$ such that*
$$O^T F_L O = \left\{ \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} : B \in \mathbb{S}^r_+ \right\}$$
3. *$L \mapsto F_L$ is an anti-isomorphism of lattices, i.e.*
$$L \subseteq L' \Leftrightarrow F_L \supseteq F_{L'} \qquad F_{L+L'} = F_L \sqcap F_{L'} \qquad F_{L\cap L'} = F_L \sqcup F_{L'}$$

*Proof ingredients.* (a) $\mathbb{S}+^N$ is invariant under $X \mapsto O^T X O$ for all $O \in O(N)$.

(b) $L = \{x \in \mathbb{R}^N : x_1, \dots, x_r = 0\}$ for nice choice of coordinates. □

## 12.1 Pataki Interval

**12.3 Definition (Rank Varieties).** Let $V_r \subseteq \mathbb{S}^N$ be the set of all $X \in \mathbb{S}^N$ with $\operatorname{rk}(X) \leq r$.

**12.4 Remark.** If $X, Y \in V_r$ and $\ker(X \subseteq \ker(Y)$, then $\forall s, t \in \mathbb{R}. sX + tY \in V_r$.

The variety $V_r$ is *ruled* by linear spaces, i.e.
$$\begin{aligned} U_L &:= \{X \in \mathbb{S}^N : L \subseteq \ker(X)\} \\ V_r &= \bigcup_{\substack{L \subseteq \mathbb{R}^N \\ \operatorname{codim}(L) = r}} U_L \end{aligned}$$

In fact $U_L = \operatorname{span}(F_L)$.

## 12.2 Lower Bound of the Pataki Interval

Next we want to find $\dim(V_r)$. We can do anything in the $r \times r$-part, and then in the remaining $r \times (N-r)$-part. The rest of the matrix is then determined. This gives us $\dim(V_r) = \binom{r+1}{2} + r(N-r)$.

**12.5 Remark.** Let $X \subseteq \mathbb{P}^{N-1}$ be an irreducible algebraic variety of dimension $k$. For a generic linear space $L \subseteq \mathbb{P}^{N-1}$, we have $X \cap L \neq \emptyset \Leftrightarrow \operatorname{codim}(L) \leq \dim(X)$.

**12.6 Proposition.** *Let $\mathscr{A} \subseteq \mathbb{S}^N$ be a generic affine subspace of dimension $m$. Then rank $r$ of an extreme point of $\mathscr{A} \cap \mathbb{S}^N_+$ satisfies*
$$rN - \binom{r}{2} \geq \binom{N+1}{2} - M \iff m \geq \binom{N-r+1}{2}$$
*This gives a lower bound on $r$.*

## 12.3 Upper Bound of the Pataki Interval

Let $S = \mathscr{A} \cap \mathbb{S}^N_+ \neq \emptyset$, with $\dim(\mathscr{A}) = m$ suppose $X \in \mathscr{A} \cap \mathbb{S}^N_+$ is of rank $r$. Then

$$X \in F_{\ker(X)} \cap \mathbb{S}^N_+ = \left\{ Y \in \mathbb{S}^N_+ : \ker(X) \subseteq \ker(Y) \right\} = O^T \left\{ \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} : B \in \mathbb{S}^r_+ \right\} O$$

**12.7 Remark.** If $X \in \mathrm{ex}(\mathbb{S})$, then $\mathscr{A} \cap F_{\ker(X)} = \{X\}$, so $\mathscr{A} \cap U_{\ker(X)} = \{X\}$. Therefore

$$\begin{aligned} &\underbrace{\dim\left(\operatorname{span} \mathscr{A}\right)}_{=m+1} + \dim\left(U_{\ker(X)}\right) \\ &= \dim\left(\operatorname{span} \mathscr{A} + U_{\ker(X)}\right) + \underbrace{\dim\left(\operatorname{span} \mathscr{A} \cap U_{\ker(A)}\right)}_{=1} \\ &\leq \dim\left(\mathbb{S}^N\right) + 1 = \binom{N+1}{2} + 1 \end{aligned}$$

This gives us an upper bound on $r$ by

$$\binom{r+1}{2} \leq \binom{N+1}{2} - m$$

**12.8 Proposition (Pataki Interval).** *Let $\mathscr{A} \subseteq \mathbb{S}^N$ be an affine space of dimension $m$. The rank $r$ of an extreme point of $\mathscr{A} \cap \mathbb{S}^N_+$ satisfies*

$$\binom{r+1}{2} \leq \binom{N+1}{2} - m$$

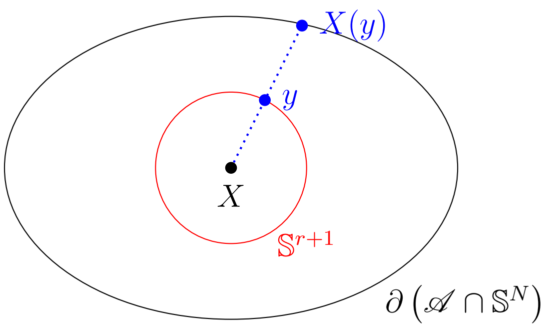
*If $\mathscr{A}$ is* generic, *it also satisfies*

$$m \geq \binom{N-r+1}{2}$$

**12.9 Theorem.** *Let $\mathscr{A} \subseteq \mathbb{S}^N$ be an affine subspace, and assume $\emptyset \neq \mathscr{A} \cap \mathbb{S}^N_+$ is bounded. Suppose* $\operatorname{codim}(\mathscr{A}) = \binom{r+1}{2}$ *for some $r \in \mathbb{N}$ and $N \geq r+2$. Then there exists a matrix $X \in \mathscr{A} \cap \mathbb{S}^N_+$ with* $\operatorname{rk}(X) \leq r$*.*

*Proof.* 1. Reduce to $N = r+2$, $\emptyset \neq \mathscr{A} \cap \mathbb{S}^{r+2}_{++}$.

2. Proof by contradiction: Suppose all matrices in $\partial\left(\mathscr{A} \cap \mathbb{S}^{r+2}_+\right)$ have rank $\geq r+1$.




The map

$$\phi : \mathbb{S}^{r+1} \to \mathbb{RP}^{r+1} \qquad\qquad y \mapsto \ker(X(y))$$

is continuous, injective. $\square$

**12.10 Example.** Take the map

$$\operatorname{Gr}\mathbb{S}^6 \to \mathbb{R}[x_1, x_2, \ldots]_{\leq 4} \qquad\qquad X \mapsto \vec{m}^T X \vec{m}$$

where $\vec{m}$ is the vector of all monomials. we have $\dim(\mathbb{S}^6) = 6$ and $\dim(\mathbb{R}[x_1, x_2]_{\leq 4}) = 15$, hence $m = 6$. We have

$$\forall f \in \mathbb{R}[x_1, x_2]_{\leq 4}. f \geq 0 \Leftrightarrow \operatorname{Gr}^{-1}(f) \cap \mathbb{S}^6_+ \neq \emptyset$$

Our upper bound by Pataki gives us $r \leq 5$, with the additional Theorem 12.9 $r \leq 4$, but Hilbert's Theorem tells us $r = 3$, so there is a gap.

# 13 Symmetry Reductions for Sums of Squares

We are given a quadratic form $Q$ in $n$ variables. To check for SOS, we can solve an $n \times n$-SDP. But we might be able to do better, if $Q$ is symmetric.
In the end, we want to have $Q = \sum l_i^2$ where the $l_i$ are linear forms. $S_n$ acts on linear forms, by permuting the variables. First note that $\sum x_i$ is fixed and so is its complement $\{\sum \alpha_i x_i : \sum \alpha_i = 0\}$. Assume we have

$$l \in \mathbb{R}[x_1, \ldots, x_n] = S_1 \oplus S_2$$

Then we have a unique decomposition $l = h_1 + h_2$ such that $h_i \in S_i$.

$$Q = \sum l_i^2 \implies Q = \operatorname{Sym} Q = \operatorname{Sym}(\sum l_i^2) = \sum \operatorname{Sym}(l_i^2)$$

When focusing on a single linear form, we have

$$\operatorname{Sym}(l^2) = \operatorname{Sym}((h_1 + h_2)^2) = \operatorname{Sym}(h_1)^2 + \operatorname{Sym}(h_2^2) + 2\underbrace{\operatorname{Sym}(h_1 h_2)}_{=0}$$

$$\operatorname{Sym}(x_1 - x_2)^2 = \operatorname{Sym}(x_1^2 + x_2^2) - 2\operatorname{Sym}(x_1 x_2) = \frac{2}{n}\left(x_1^2 + \ldots + x_n^2\right) - \frac{2}{\binom{n}{2}} \sum_{i<j} x_i x_j$$

As another example, we have

$$\begin{aligned}\operatorname{Sym}(x_2 - 2x_3 + x_4)^2 &= \operatorname{Sym}\left(x_2^2 + x_4^2 + 4x_3^2\right) + \operatorname{Sym}\left(-4x_2x_3 - 4x_3x_4 + 2x_2x_4\right)\\ &= \frac{6}{n}\left(x_1^2 + \ldots + x_n^2\right) - \frac{6}{\binom{n}{2}} \sum_{i<j} x_i x_j\end{aligned}$$

Something more general (symmetric quartics). Symmetric polynomials on $\{0,1\}^n$, which means, we are module $\langle x_i^2 - x_i : i = 1, \ldots, n\rangle$.
Consider the space of polynomials, we want to squares and decompose into isotopic components

$$V = W_1 + \ldots + W_k$$

where each $W_i$ is a direct sum of isomorphic irreducibles. Then

$$\operatorname{Sym} p^2 = \operatorname{Sym}\left((h_1 + \ldots + h_k)^2\right) = \sum \operatorname{Sym} h_i^2 + 2\sum \operatorname{Sym}(h_i h_j) = \sum \operatorname{Sym} h_i^2$$

where $h_i \in W_i$. Note that cross-products form non-isomorphic irreducibles always vanish.

No we restrict to a single isotypic component $W = V_1 + \ldots + V_m$ and from each $V_i$ we choose a representative $f_i$. Then we can define a matrix

$$F = (\mathrm{Sym}(f_i f_j) : 1 \leq i, j \leq m)$$

Then $Q$ is a sum of squares from $W$ iff there exists $A \in \mathbb{S}_+^m$ such that $Q = \langle A, F \rangle$.
Symmetric sum of squares on polynomials of degree $\leq d \leq \frac{n}{2}$ on $\{0,1\}^n$.

$$\begin{aligned}(0,\ldots,0) &: 1\\ (1,0,\ldots,0) &: \sum x_i, x_1 - x_2\\ (1,1,0,\ldots,0) &: \left(\sum x_i\right)^2, (x_1-x_2)(x_3-x_4)\end{aligned}$$

Again, I could not make sense of the board, so the notes stop.

# 14 Sums of Squares and Determinantal Representations

This lecture is based on "Hyperbolic polynomials, interlacers and SOS" by Kummer/Plaumann/Vinzant and "Non-representable hyperbolic matroid" by Brändén.
Recall for any $f \in \mathbb{R}[x_1, \ldots, x_n]$ and $\vec{e}, \vec{a} \in \mathbb{R}^n$ we have

$$\Delta_{\vec{e},\vec{a}}(f) = D_{\vec{e}} f \cdot D_{\vec{a}} f - f \cdot D_{\vec{e}} D_{\vec{a}}(f)$$

If $f$ is hyperbolic and $\vec{a} \in C(f, \vec{e})$, then $\Delta_{\vec{e},\vec{a}} f \geq 0$ on $\mathbb{R}^n$.
Now we focus on functions of the form $f = \det(X)$, with $X = (x_{ij} : 1 \leq i, j \leq n)$.

## 14.1 Dodgson condensation

Let $A \in \mathbb{R}^{n \times n}$. We define $A_{S,T}$ as the matrix obtained by removing rows $S$ and columns $T$. Note that we have

$$|A_{1,1}| \cdot |A_{n,n}| - |A_{1,n}| \cdot |A_{n,1}| = |A_{1n,1n}| \cdot |A|$$

**14.1 Example.** Assume we have a symmetric matrix, then we have

$$A = \begin{pmatrix} 1 & b & c \\ b & d & e \\ c & e & f \end{pmatrix} \implies \det(A) = \frac{(ad - b^2)(df - e^2) - (be - cd)^2}{d}$$

**14.2 Corollary.** *Let $f = \det(X)$ for $X^T = X$. Then*

$$\Delta_{e_1 e_1^T, e_n e_n^T} f = \frac{\partial f}{\partial x_{11}} \cdot \frac{\partial f}{\partial x_{nn}} - f \cdot \frac{\partial f}{\partial x_{11} \partial x_{nn}} = |X_{1,n}|^2$$

**14.3 Corollary.** *Instead of the unit vectors, we can take any vector.*

$$\Delta_{vv^T, ww^T}(f) = \left(v^T X^{\mathrm{adj}} w\right)^2$$

**14.4 Theorem.** *Suppose, we have a multiaffine polynomial*

$$f = \sum_{S \subseteq [n]} c_S \prod_{i \in S} x_i \in \mathbb{R}[x_1, \ldots, x_n]$$

*Then $f$ has a definite determinantal representation iff for all $i, j \in [n]$ the polynomial $\Delta_{ij}(f)$ is a square in $\mathbb{R}[\vec{x}]$.*

*Proof.* Assume $f = \det(A(x))$ where $A(x) = \sum x_i A_i + b$ with $A_i \succeq 0$. Then $A_i = v_i v_i^T$. Thus $\Delta_{ij}(f) = \left(v_i A(x)^{\mathrm{adj}} v_j\right)^2$. The converse direction is harder. □

More generally

**14.5 Theorem.** *If $f^r = \det(A(x))$, then $\Delta_{ij}(f)$ is SOS for any $i, j \in [n]$.*

*Proof sketch.* 1. Assume $f = \det(X)$ and $A, B \succeq 0$, so we can write them as $A = \sum v_i v_i^T$ and $B = \sum w_j w_j^T$. Check that $\Delta_{\cdot,\cdot}$ is bilinear. Then

$$\Delta_{A,B} f = \sum_{i,j} \Delta_{v_i v_i^T, w_j w_j^T} f = \sum_{i,j} \left(v_i^T X^{\mathrm{adj}} w_j\right)^2$$

2. $\Delta_{ij}(f^r) = r \cdot f^{2(r-1)} \Delta_{ij}(f)$, so it suffices to find an SOS decomposition for the latter. □

**14.6 Theorem (Wagner, Wei).** *Put*

$$B = \binom{[8]}{4} \setminus \{1234, 1256, 3456, 3478, 5678\}$$

*Then define the polynomial*

$$f = \sum_{I \in B} \prod_{i \in I} x_i \in \mathbb{R}[x_1, \ldots, x_8]$$

*Then $f$ is stable.*

**14.7 Theorem (Brändén).** *Let $f$ as in Theorem 14.6, and $r \in \mathbb{N}$. Then $f^r$ does not have a definite determinantal representation.*

*Proof.* $\Delta_{78}(f)$ is not SOS. □

**14.8 Lemma (Kummer).** *$C(f, \mathbf{1})$ is spectradral.*

**14.9 Theorem (Brändén).** *For any graph $G = ([n], E)$, the polynomial*

$$f_G := e_4(x_1, \ldots, x_n, y_1, \ldots, y_n) - \sum_{ij \in E} x_i x_j y_i y_j$$

*is stable.*

In particular, this yields the result for the Vámos-matroid, for the graph

$G_0$ :

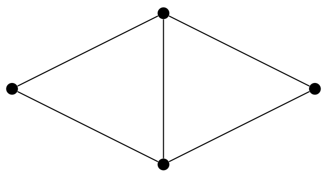


Figure 2: The graph for the Vámos matroid

Furthermore, if $G$ has this $G_0$ as a subgraph, then $(f_G)^r$ does not have a definite determinantel representation.

**14.10 Remark.** An open question is: Is the hyperbolicity cone of $f_G$ spectrahedral? Rephrased, we ask whether there is some $q(x)$ such that

$$q(x) f_G = \det\left(\sum x_i A_i\right) \qquad A_i \succeq 0$$

$$C_{\mathbf{1}}(Q \cdot f_G) = C(f_G) \iff C_{\mathbf{1}}(f_G) \subseteq C_{\mathbf{1}}(q)$$