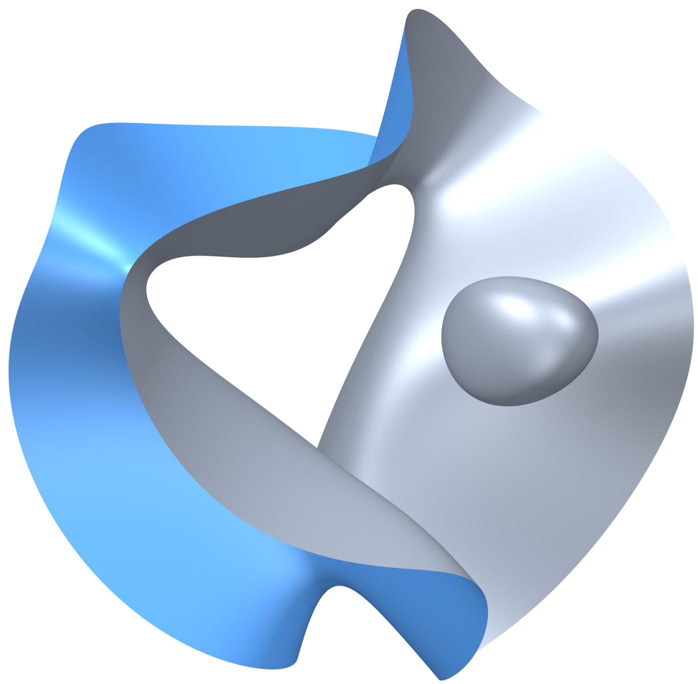
Hyperbolic polynomials and optimization

Simone Naldi

Séminaire Parisien d'Optimisation

Institut Henri Poincaré, Paris – 3 Février 2020



Introduction

Spectrahedral representations (j/w M. Kummer and D. Plaumann)

Local relaxations (work in progress)

Introduction

Historical motivations and definitions

COMMUNICATIONS ON PURE AND APPLIED MATHEMATICS, VOL. XI, 175–194 (1958)

Differential Equations, Difference Equations and Matrix Theory*†

P. D. LAX

The paper studies the *sets of real matrices all of whose linear combinations with real coefficients have only real eigenvalues*.

Motivated by understanding uniqueness/existence of solutions to

$$u_t = \sum_{k=1}^{m} A^k u_{x^k} = A u_x$$

where here A^k are real matrices, and $x = (x^1, x^2, \ldots, x^m)$.

Given a polynomial $f \in \mathbb{R}[x_1, \ldots, x_n]_{\leq d}$, one can associate to f a differential operator $\partial_f = f(d/dx_1, \ldots, d/dx_n)$.

D'alembertien : $f = t^2 - \sum_{i=1}^n x_i^2 \quad \mapsto \quad \partial_f = d/dt^2 - \sum_i d/dx_i^2$.

Lax, Mizohata ('60). Let $f = \sum_{i \leq d} f_i$ with $f_i \in \mathbb{R}[x]_i$. If the Cauchy problem $\partial_f(u) = p$ is well-posed (existence/uniqueness of solutions) then f_d is hyperbolic (definition later).

Hence hyperbolicity deals with necessary conditions for existence/uniqueness of the Cauchy problem.

How to check hyperbolicity ?

The Lax conjecture (is true)

Let $p(\xi, \eta, \lambda)$ be a form of degree n in ξ, η, λ, with the following property: The coefficient of λ^n is one, and for each fixed real choice for ξ and η, $p(\xi, \eta, \lambda)$ has only real zeros in λ.

We shall call such a form hyperbolic. If A and B are symmetric matrices, the characteristic polynomial of $\xi A + \eta B$, $\det\,[\xi A + \eta B - \lambda \mathrm{I}]$, is hyperbolic. Conversely, I

CONJECTURE: *Every hyperbolic form can be so represented.*

Helton-Vinnikov (2007): Let $f \in \mathbb{R}[x, y, z]$ be a homogeneous hyperbolic polynomial of degree d. Then there exist $d \times d$ symmetric matrices A, B, C such that:

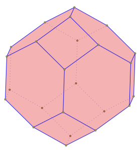
$$\boxed{f = \det(xA + yB + zC)}$$

Hyperbolic polynomials

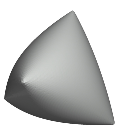
$f \in \mathbb{R}[x]_d$ is *hyperbolic w.r.t.* $e = (e_1, \ldots, e_n) \in \mathbb{R}^n$ if

- $f(e) \neq 0$
- $\forall\, a \in \mathbb{R}^n \;\; t \mapsto ch_a(t) := f(t\,e - a)$ has only real roots

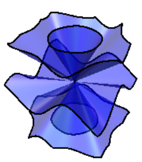
Examples :



$f = x_1 \cdots x_d$, in which case $ch_a(t) = \prod_i (te_i - a_i)$
or more generally $f = \ell_1(x) \cdots \ell_d(x)$, ℓ_i linear forms



$f = \det(X)$, $X = (x_{ij})$ symmetric, and $ch_A(t) = \det(tE - A)$
or more generally $f = \det(x_1 A_1 + \cdots + x_n A_n)$



$\leftarrow$ the general shape of a hyperbolic hypersurface

Non-homogeneous case : real zero polynomials

In the univariate case (so, non-homogeneous) there is only one direction, hence not so much choice of e :

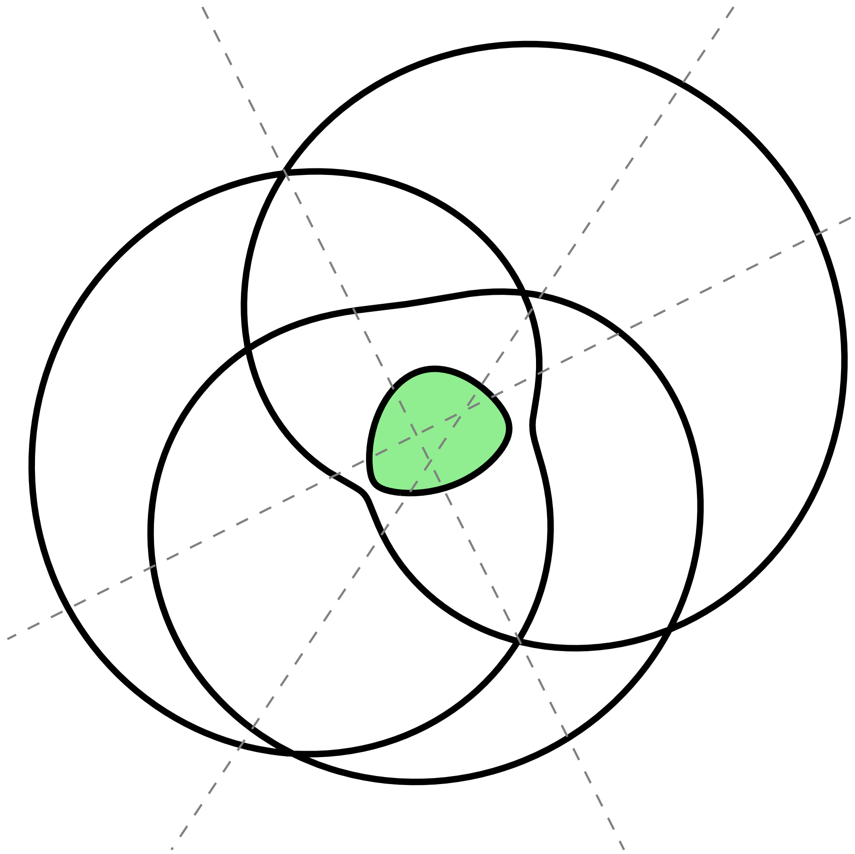
$$f \in \mathbb{R}[x_1] \text{ \textit{real-rooted if} } \quad f(a) = 0 \;\Rightarrow\; a \in \mathbb{R}.$$

The definition extends to a property for non-homogeneous multivariate polynomials, classically called *real zero (RZ)* :

$$f \in \mathbb{R}[x] \text{ RZ if } \quad \forall\, \lambda \in \mathbb{C}, a \in \mathbb{R}^n \;\; f(\lambda a) = 0 \;\Rightarrow\; \lambda \in \mathbb{R}.$$

In this talk I consider hyperbolicity in the homogeneous setting.

Example in degree 8



Hyperbolicity cones

$f \in \mathbb{R}[x]_d$ hyperbolic w.r.t. $e \in \mathbb{R}^n$. Then the set

$$\Lambda_+(f,e) = \{a \in \mathbb{R}^n : ch_a(t) = 0 \Rightarrow t \geq 0\}$$

is called the *hyperbolicity cone* of f in direction e.

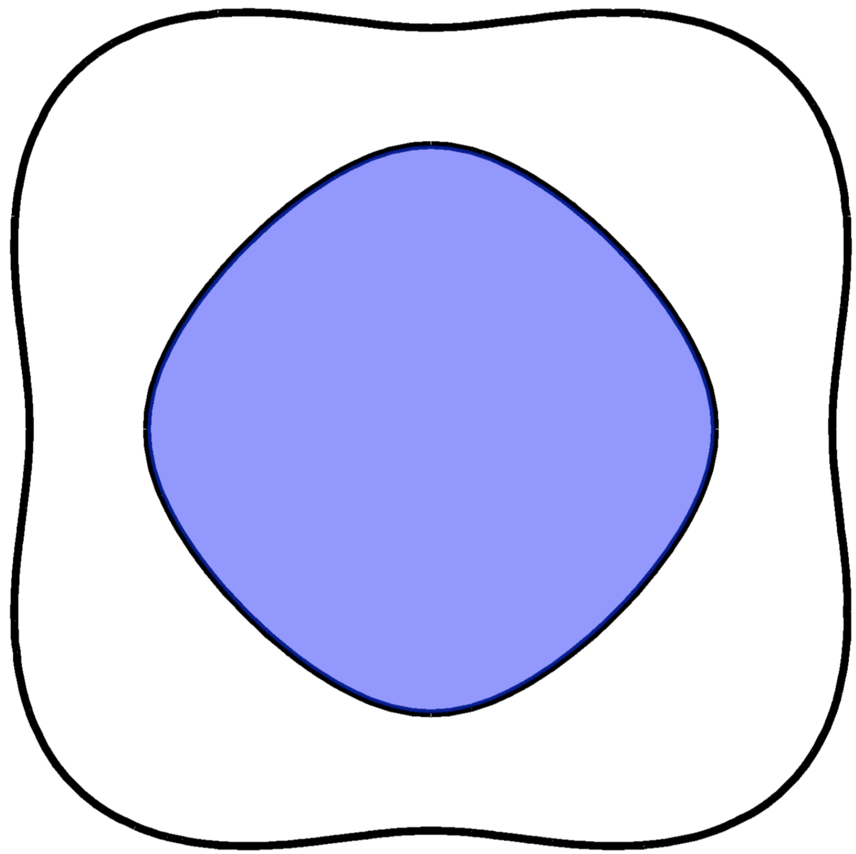
Clearly, it is not unique, it comes in pairs : $\Lambda_+(f,-e) = -\Lambda_+(f,e)$.

The dependence on e is weak: $a \in \Lambda_+(f,e) \Rightarrow \Lambda_+(f,a) = \Lambda_+(f,e)$

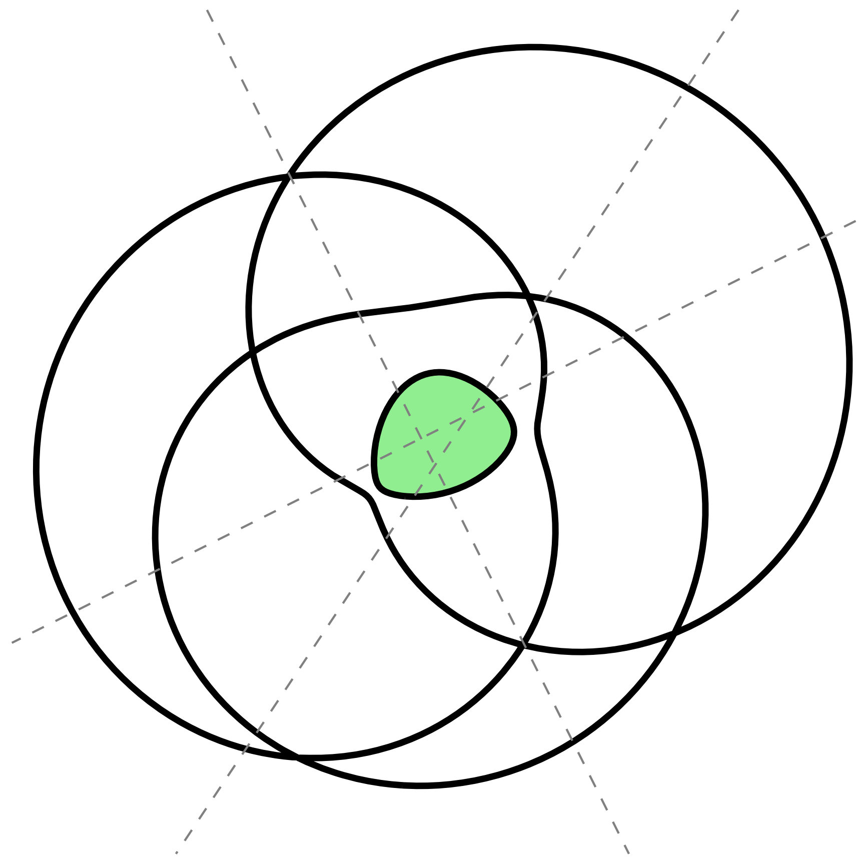
If f is irreducible, there is only one pair (*Kummer*, 2018). But in the polyhedral case there are exponentially many cones.

Jörgens, Theobald (2008). The number of hyperbolicity cones is at most 2^d for $d \leq n$ and at most $2\sum_{k=0}^{n} \binom{d-1}{k}$ otherwise. The borne is sharp for polyhedra.

A hyperbolic plane quartic



The three-ellipse is a hyperbolicity cone

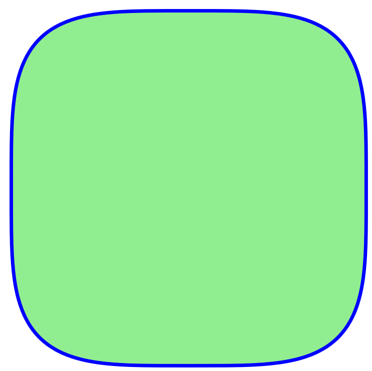


A non-hyperbolic polynomial

Consider the quartic $f = x^4 + y^4 - z^4$, and the semi-algebraic set defined by $f \leq 0$:

$$S = \{(x, y, z) \in \mathbb{R}^3 : x^4 + y^4 \leq z^4\}.$$

Its viewpoint from the section $z = 1$ is pictured here, and we see that f cannot be hyperbolic (every lines meets the curve in only 2 real points):



Hence S is not a hyperbolicity cone, hence not a spectrahedron. It is a *projected* spectrahedron ($\pi : \mathbb{R}^6 \to \mathbb{R}^2$) : optimization over S is still a SDP, up to slack.

Hyperbolic Programming (HP)

Hyperbolic programming asks to minimize a linear function to affine sections of hyperbolicity cones:

$$\begin{array}{lll} c^* := & \inf & c^T x \\ & \text{s.t.} & Ax = b \\ & & \boxed{x \in \Lambda_+(f,e)} \end{array}$$

$\Lambda_+(f,e)$ is convex (Garding), closed basic semi-algebraic.

Logarithmic barrier functions are known for HP, for instance $-log(f)$ which is (strictly) convex in the interior of $\Lambda_+(f,e)$

Spectrahedra are hyperbolicity cones

For $f = \det(X)$, and $E \succ 0$ (say $E = I_d$) one has that

$$\Lambda_+(\det, I_d) = \{A \in \mathcal{S}_d : \det(tI_d - A) = 0 \Rightarrow t \geq 0\} = \mathcal{S}_d^+$$

In particular :

Every spectrahedron is a hyperbolicity cone

Every SDP problem is a HP problem

The big question in the theory of hyperbolic polynomials is whether the converse of these implications holds.

Few hyperbolic polynomials are determinants

Counting dimensions, the Helton-Vinnikov theorem cannot directly extend to high n, d. Indeed, the set $H_{n,d}$ of hyperbolic polynomials is full-dimensional in $\mathbb{R}[x]_d$ (Nuij, 1968), that is

$$\dim(H_{n,d}) = \dim(\mathbb{R}[x]_d) = \binom{n+d-1}{d}$$

But the image of the determinant map has lower dimension.

Generalized representations. The d'Alembertien is hyperbolic, there are determinantal formulae, up to factor :

$$t^2 \cdot (t^2 - x^2 - y^2 - z^2) = \det\left(\begin{bmatrix} t & x & y & z \\ x & t & & \\ y & & t & \\ z & & & t \end{bmatrix}\right)$$

Derivative relaxations *à la* Renegar

f hyperbolic w.r. to e $\quad\Rightarrow\quad$ $D_e f = \sum_i e_i \frac{\partial f}{\partial x_i}$ still hyperbolic

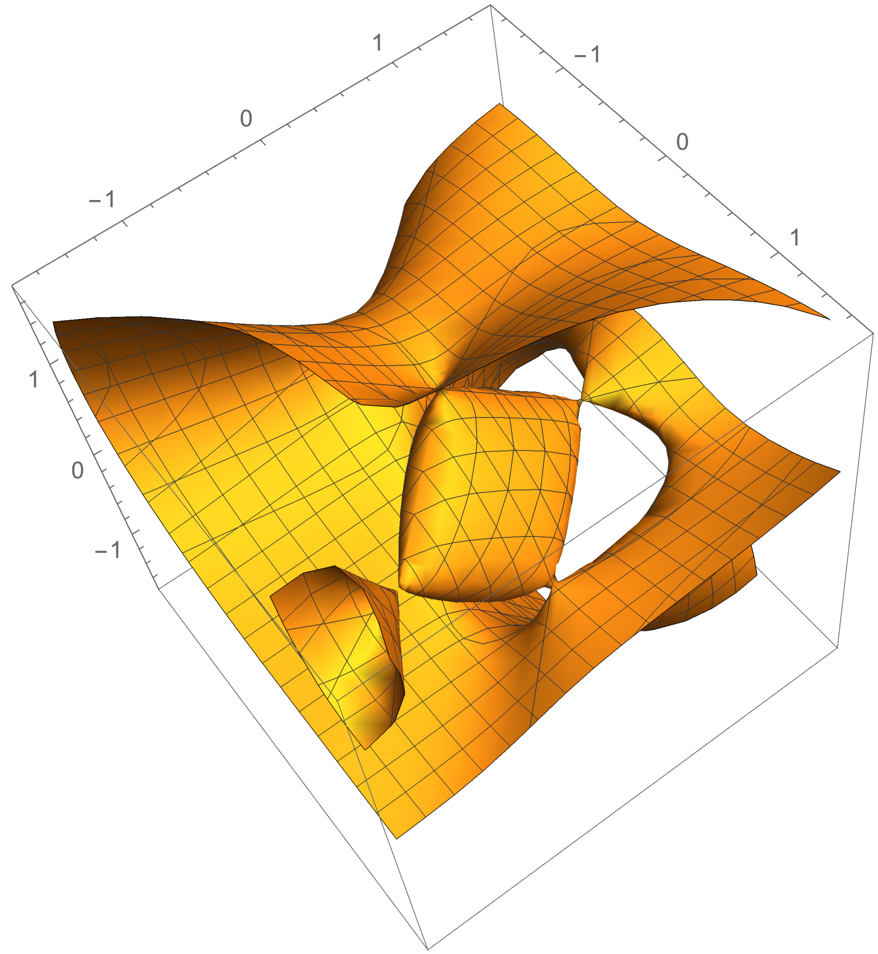
This gives a nested sequence of convex hyperbolicity cones:

$$\Lambda_+(f, e) \subset \Lambda_+(D_e f, e) \subset \cdots \subset \Lambda_+(D_e^{(d-1)} f, e)$$

(the last one is a half-space) giving a sequence of *lower bounds* for the linear function to optimize:

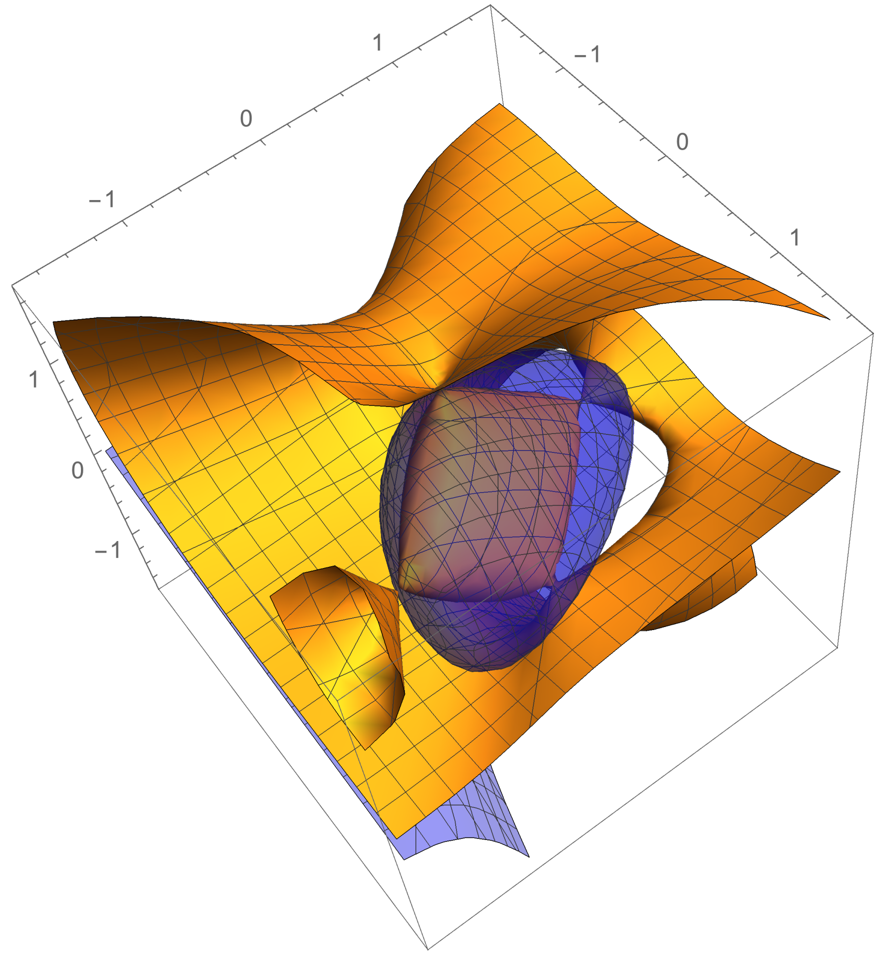
$$\inf_{\Lambda_+(f,e)} c^T x \geq \inf_{\Lambda_+(D_e f,e)} c^T x \geq \cdots \geq \inf_{\Lambda_+(D_e^{(d-1)} f,e)} c^T x$$

Example of Renegar derivative



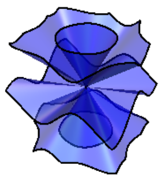
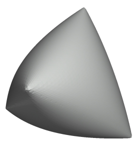
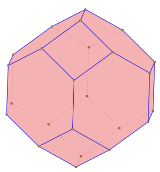
Quartic hyperbolic polynomial

Example of Renegar derivative



+ First derivative

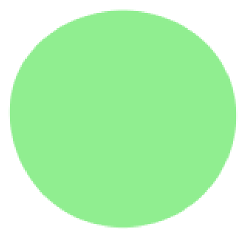
Optimization viewpoint

Feasible set	*name*	**Optimization**	**Polynomial**
	Hyperbolicity Cone	HP	Hyperbolic polynomial
	Spectrahedron	SDP	$f = \det A(x)$
	Polyhedron	LP	$f = \prod \ell_i(x)$

Spectrahedral representations of hyperbolic plane curves

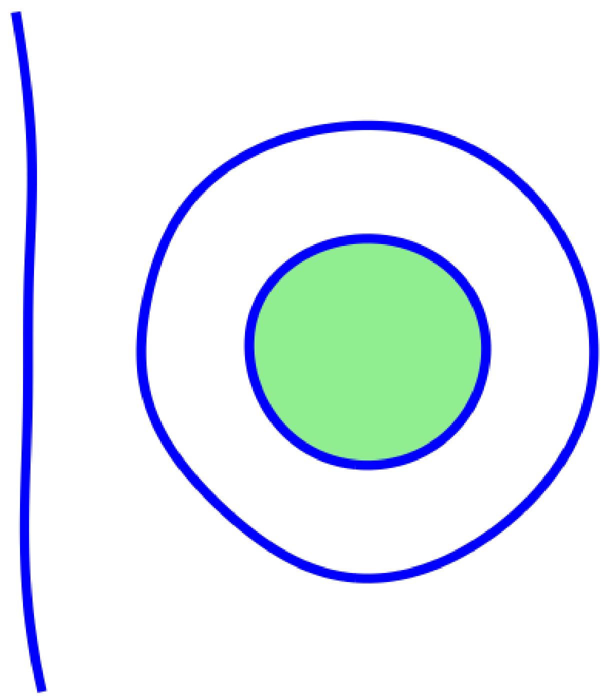
joint work with D. Plaumann and M. Kummer

Convex semialgebraic set : $\Lambda_+(f, e) = \{a \in \mathbb{R}^n : f(a) \geq 0, \ldots\}$



Convex semialgebraic set : $\Lambda_+(f,e) = \{a \in \mathbb{R}^n : f(a) \geq 0, \ldots\}$

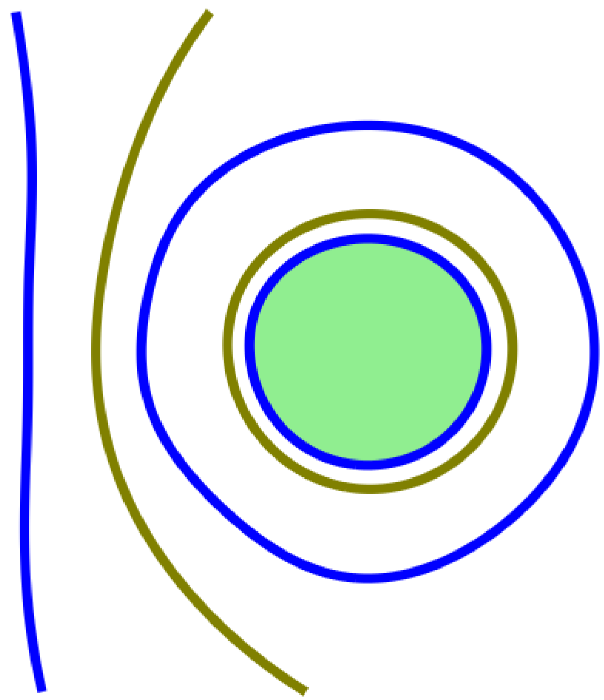
Hyperbolic curve : $V(f) = \{a \in \mathbb{R}^n : f(a) = 0\}$



Convex semialgebraic set : $\Lambda_+(f,e) = \{a \in \mathbb{R}^n : f(a) \geq 0, \ldots\}$

Hyperbolic curve : $V(f) = \{a \in \mathbb{R}^n : f(a) = 0\}$

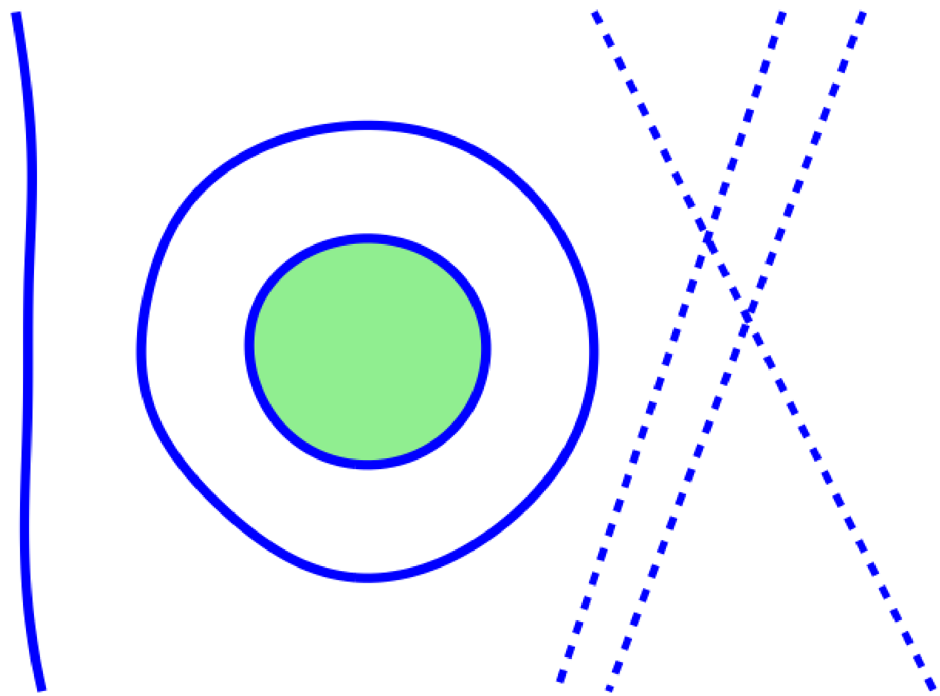
Interlacer : $g \in \mathbb{R}[x]_{\deg f - 1}$.



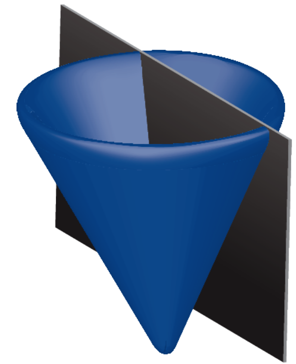
Convex semialgebraic set : $\Lambda_+(f, e) = \{a \in \mathbb{R}^n : f(a) \geq 0, \ldots\}$

Hyperbolic curve : $V(f) = \{a \in \mathbb{R}^n : f(a) = 0\}$

Interlacer : $g \in \mathbb{R}[x]_{\deg f - 1}$. Extra-factor : $\ell_1 \cdot \ell_2 \cdots \ell_s$



(Generalized) Lax Conjecture



Geometric version

Every hyperbolicity cone is a spectrahedron

$$\Lambda_+(f,e) = \{a \in \mathbb{R}^n : A_1 a_1 + \ldots + A_n a_n \succeq 0\}$$

Helton-Vinnikov – true for curves (even stronger : $f = \det A$)

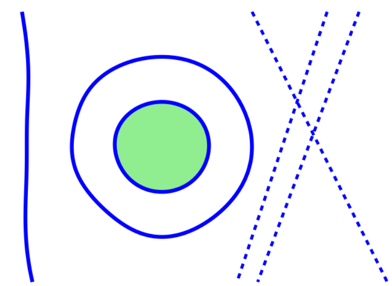
Brändén – GLC cannot be proved by means of det. represent. of f

Algebraic version

Let f be hyperbolic with respect to e. Then there exist a poly. q and matrices A_i such that

(1) $q \cdot f = \det(A(x))$

(2) $\Lambda_+(q,e) \supset \Lambda_+(f,e)$



$f =$ blue curve

$q =$ dashed curve

Kummer – (1) is true for real-smooth hyperbolic polynomials

Spectrahedral vs Determinantal representations

Given

A hyperbolic polynomial $f \in \mathbb{R}[x]_d$

A direction of hyperbolicity $e \in \mathbb{R}^n$

one would like to compute a spectrahedral representation

$$\Lambda_+(f, e) = \{a \in \mathbb{R}^n : A_1 a_1 + \ldots + A_n a_n \succeq 0\}$$

Can be achieved by computing determinantal representations

$f = \det A$	not necessary (Brändén counterexamples)
$f^k = \det A$	not necessary (Brändén counterexamples)
$q \cdot f = \det A$	open (equivalent to GLC)

Contact curves and Interlacers ($n = 3$)

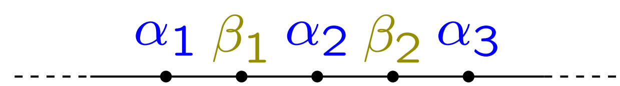
Let $f, g \in \mathbb{R}[x, y, z]$ be coprime. We say that

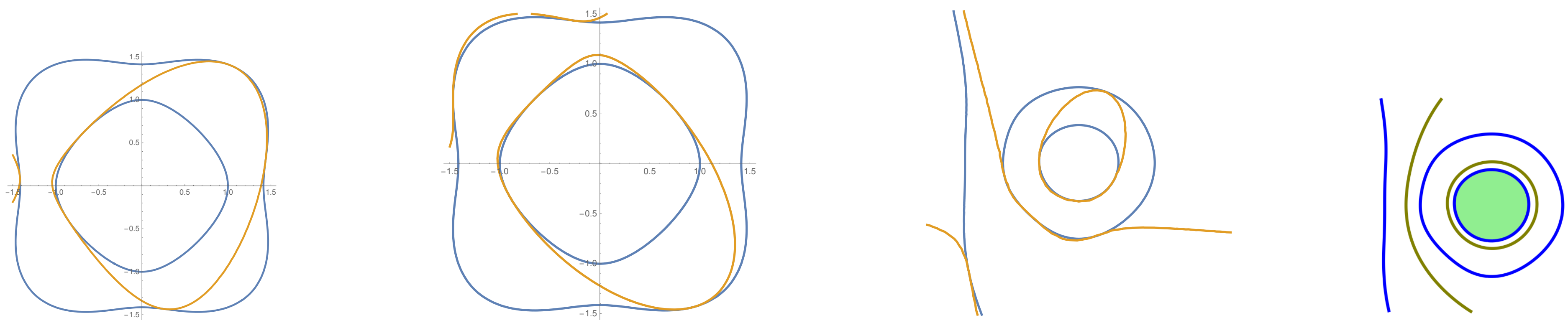
$p \in V_{\mathbb{C}}(f) \cap V_{\mathbb{C}}(g)$ is a *contact point* if $\mathrm{mult}_p(f, g)$ even

g is a *contact curve* (resp. *real contact curve*) if every $p \in V_{\mathbb{C}}(f) \cap V_{\mathbb{C}}(g)$ (resp. $p \in V_{\mathbb{R}}(f) \cap V_{\mathbb{R}}(g)$) is contact

g is an *interlacer* if

$t \mapsto g(te + a)$ interlaces $t \mapsto f(te + a)$ for all a :





contact curve >> real contact curve << interlacers

Dixon process (Hermitian version with interlacer)

Essentially based on the property $\boxed{A \cdot A^{adj} = \det A \cdot \mathrm{Id}_d}$

Remark : if $f = \det A$ and $V_{\mathbb{C}}(f)$ is smooth, then

$$\text{co-rank}(A) = \text{rank}(A^{adj}) = 1 \mod \det A.$$

INPUT

f hyperbolic with resp. to e

Sketch of the PROCEDURE :

$m_{11} \leftarrow D_e f := e_1 \frac{\partial f}{\partial x} + e_2 \frac{\partial f}{\partial y} + e_3 \frac{\partial f}{\partial z}$ # interlacer

split $S \cup \overline{S} = V_{\mathbb{C}}(f) \cap V_{\mathbb{C}}(D_e f)$

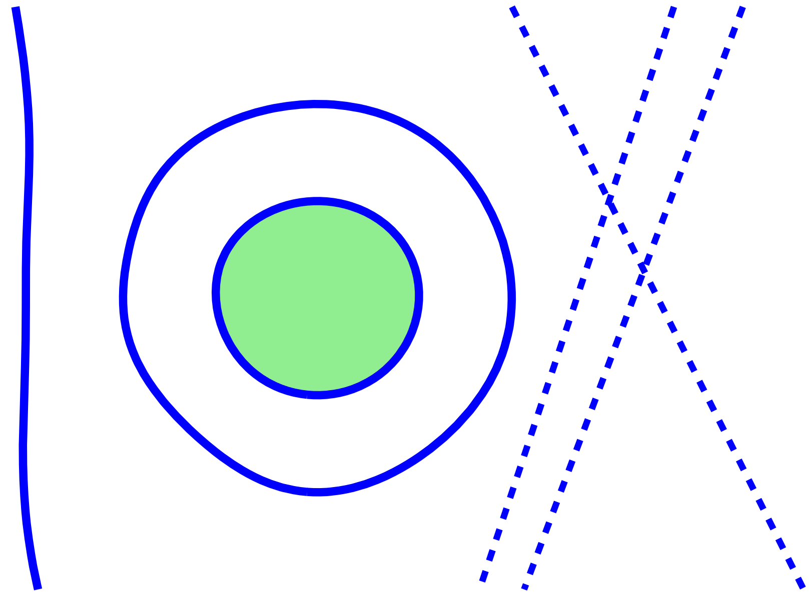
extend m_{11} to basis $\langle m_{11} \ldots m_{1d} \rangle = V(\langle S \rangle)$

$m_{jk} \leftarrow$ solve $a_{11}a_{jk} - a_{1j}a_{1k} \in \langle f \rangle$ for $j \leq k$ # rank $= 1$

$M \leftarrow (m_{jk})$

$A \leftarrow M^{adj}/f^{d-2}$

OUTPUT A (satisfying $f = \det A$ and $A(e) \succ 0$)



Positive aspects of our contribution

The size of the det. repr. depends on r = real contact points

There are (poss. large) spectrahedral representations over $\mathbb{Q}$

The multiplier is the simplest : $q = \ell_1 \cdot \ell_2 \cdots \ell_s$

For $r = d(d-1)/2$ (maximal) we get Helton-Vinnikov repr.

Extremal interlacers

We say that an interlacer is **extremal** if it is an extreme point of the set of interlacers (that is a cone). If f is smooth, any extremal interlacer has at least

$$\left\lceil \frac{(d+1)d-2}{4} \right\rceil$$

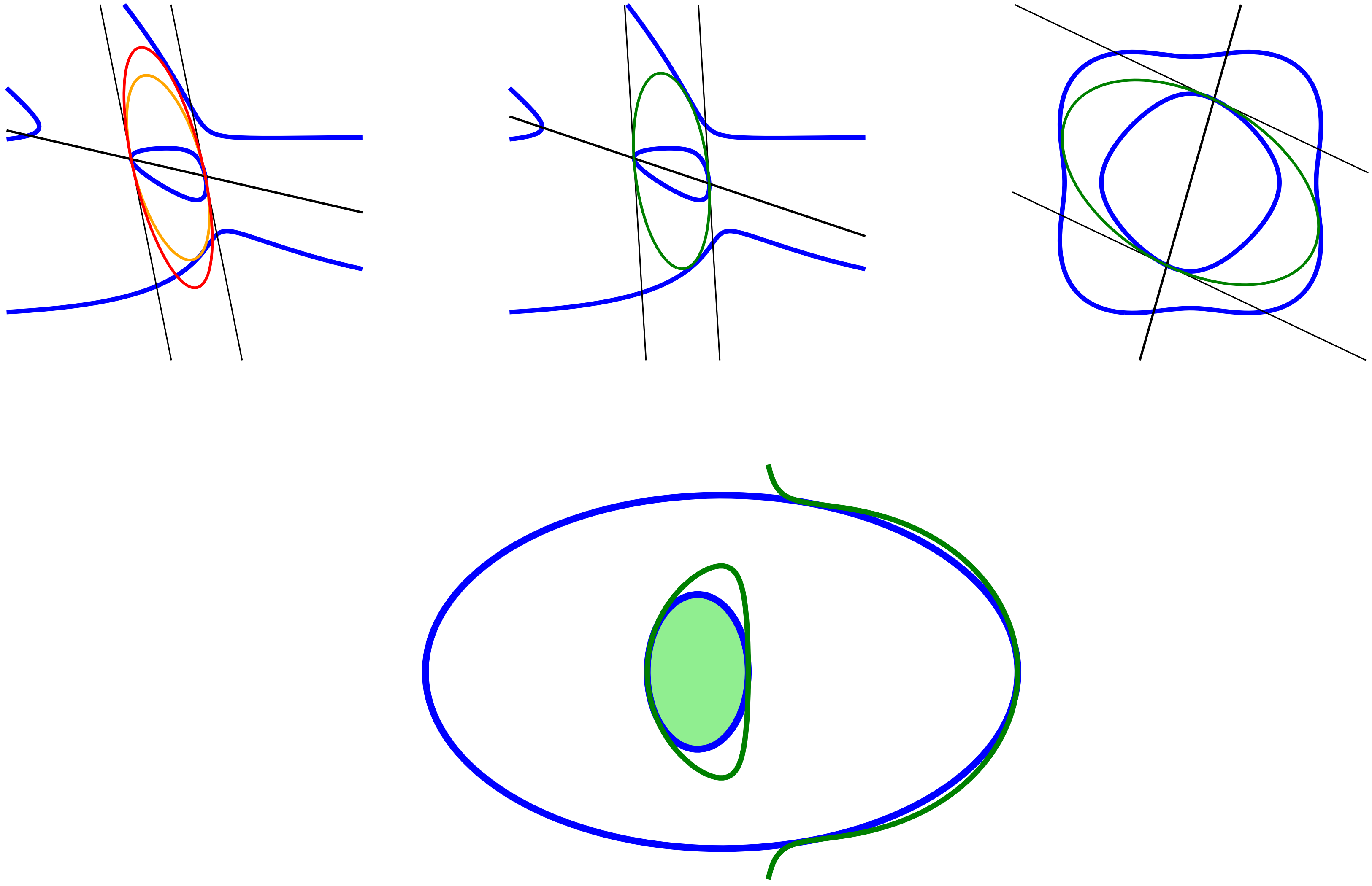
contact points (counted multiplicities).

Expected number of real contact points compared with the number of points for a full contact curve:

d	2	3	4	5	6	$\cdots$
$\left\lceil \frac{(d+1)d-2}{4} \right\rceil$	1	3	5	7	10	$\cdots$
$\frac{d(d-1)}{2}$	1	3	6	10	15	$\cdots$

Not clear whether there always exist interlacers with $\frac{d(d-1)}{2}$ real contact points.

The case of quartics



Main result

Let f be hyperbolic with respect to e, and let g be an interlacer of f with r real contact points.

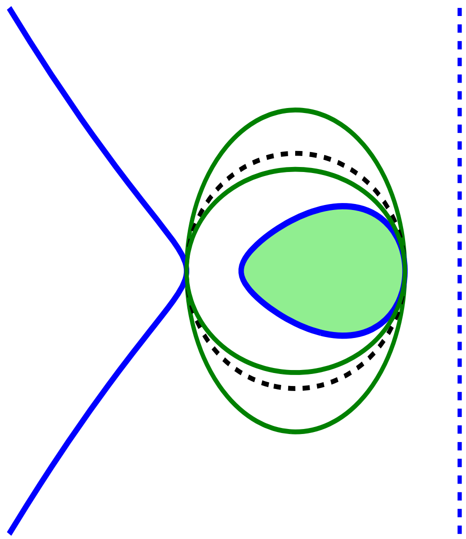
Idea. Fix the lack of g being a contact curve by increasing the multiplicity of the complex intersections of f with g. Let ℓ_i be the line through the conjugate points $p_i, \overline{p}_i \in V_{\mathbb{C}}(f) \cap V_{\mathbb{C}}(g)$.

Theorem. If f defines a smooth curve, and up to genericity assumptions on g there are real matrices A, B, C of size $m = (d^2+d-2r)/2$ such that $\Lambda_+(f,e) = \{xA+yB+zC \succeq 0\}$. Moreover

$$f \cdot \ell_1 \ell_2 \cdots \ell_s = \det(xA + yB + zC).$$

One example

The cubic $f = x^3 + 2x^2y - xy^2 - 2y^3 - xz^2$ is hyperbolic with respect to $e = (1, 0, 0)$.



By exact computation one gets :

Two interlacers living in $K[x, y, z]$ with $|K : \mathbb{Q}| = 4$ (green curves)

One rational interlacer (dashed green)

Corresponding spectrahedral representation:

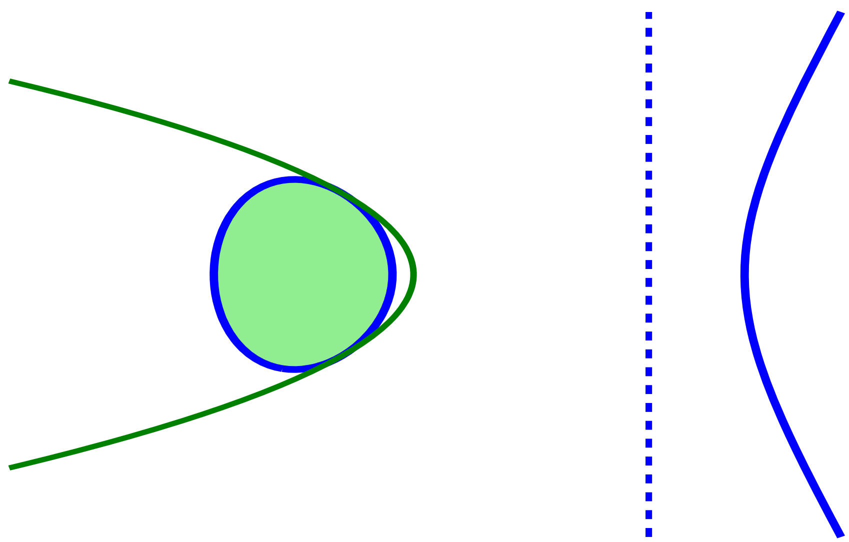
$$\frac{24}{125} f \cdot (2x - y) = \det \begin{pmatrix} 5x + 10y & -x - 2y & -4z & 2z \\ -x - 2y & x & 0 & 0 \\ -4z & 0 & 4x + 2y & -2x - 4y \\ 2z & 0 & -2x - 4y & 4x + 2y \end{pmatrix}.$$

Rational spectrahedral representation

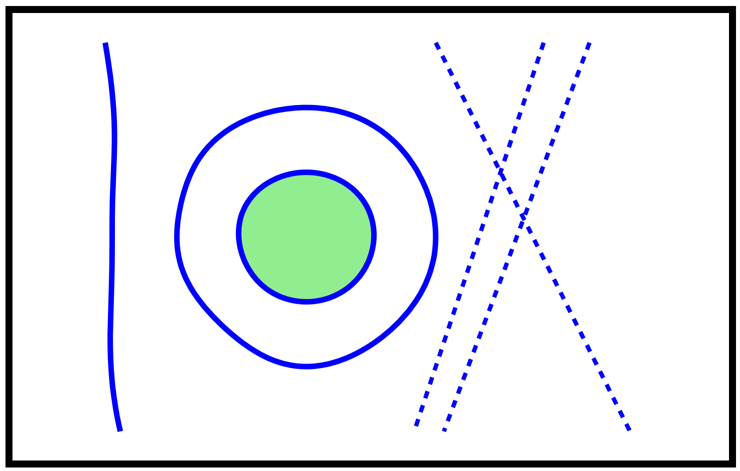
$f = y^2z - (x^3 - 6x - 3)$ has no rational 3×3 determinantal representation (because $x^3 - 6x - 3$ is irreducible). But :

$$\Lambda_+(f,e) = \left\{ \begin{pmatrix} 3z & y & -x-z & -3x+z \\ y & -x+2z & 0 & -y \\ -x-z & 0 & z & x+4z \\ -3x+z & -y & x+4z & -x+18z \end{pmatrix} \succeq 0 \right\}.$$

Extra-factor is a line and the interlacer has two contact points:



Spectrahedral representations of plane hyperbolic curves

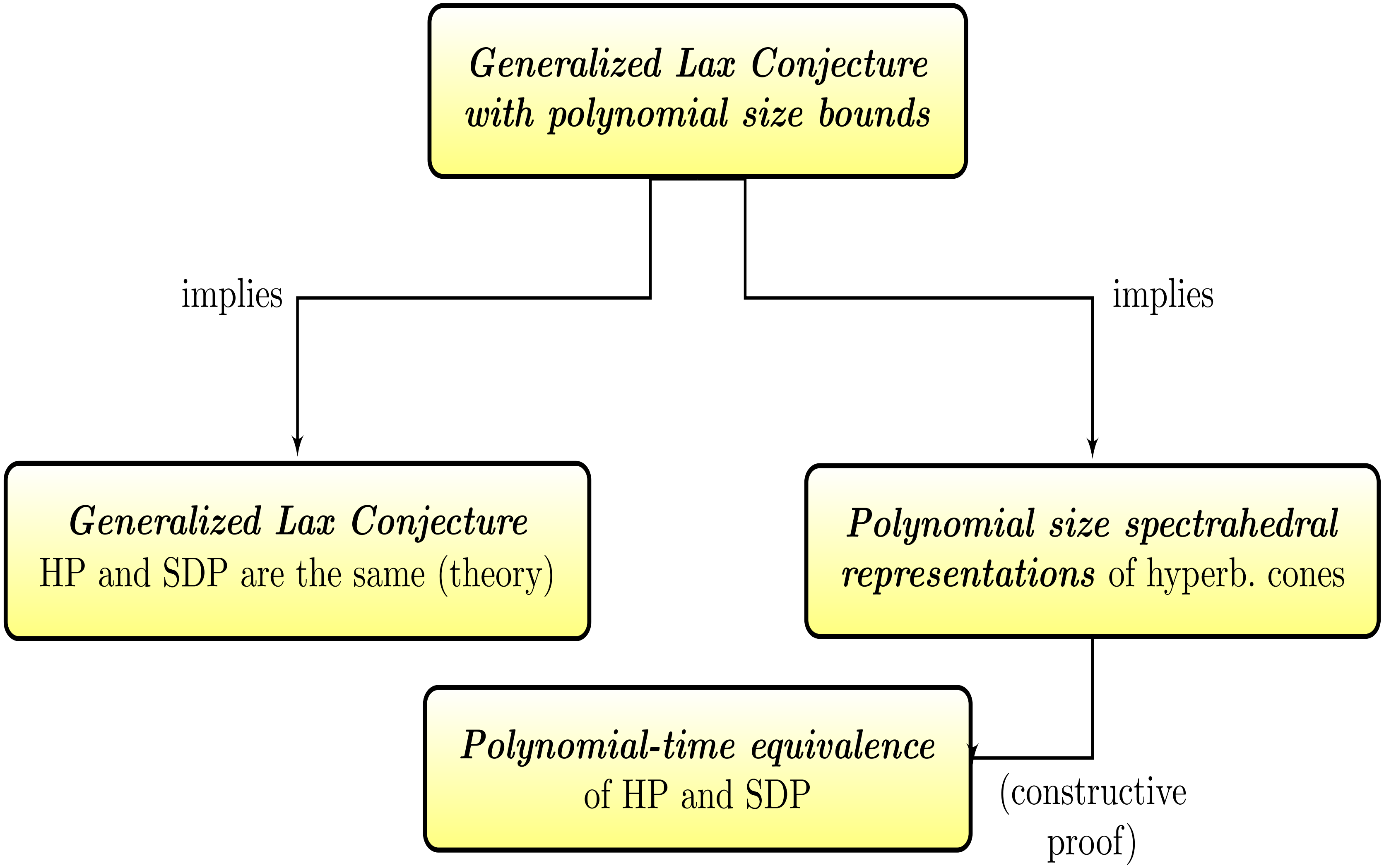


on the arXiv : `https://arxiv.org/abs/1807.10901`

Pac. J. Math. – in press

Equivalence of HP and SDP

A work in progress



Generalized Lax Conjecture
with polynomial size bounds
implies
implies
Generalized Lax Conjecture
HP and SDP are the same (theory)
Polynomial size spectrahedral
representations of hyperb. cones
Polynomial-time equivalence
of HP and SDP
(constructive proof)

Obstructions

The polynomial-time equivalence might not be provable via the GLC. Indeed recently lower bounds for the size of the determinantal representations have been computed :

Raghavendra et al. (2018)
There exists $\kappa > 0$ such that for $n, d \gg 0$, there exists a $f \in H_{n,d}$ s.t. $\Lambda_+(f, e)$ does not admit an $(1/n^{4nd})$-approximate spectrahedral representation of dimension $\leq (n/d)\kappa d$.

The examples by Raghavendra and co-authors are essentially perturbations $\Lambda_+(\sigma_k + q, e)$ of the hyperbolicity cones of the elementary symmetric polynomials

$$\sigma_k(x_1, \ldots, x_n) = \sum_{i_1,\ldots,i_k} x_{i_1} x_{i_2} \cdots x_{i_k}$$

Local hyperbolic relaxations of HP

Recall that a HP has the form

$$\begin{array}{rcll} c^* & := & \text{inf} & c^T x \\ & & \text{s.t.} & x \in L \\ & & & x \in \Lambda_+(f,e) \end{array}$$

Assume that a lower bound θ for $c^t x$ on $\Lambda_+(f,e)$ is known.

Then $c^t x - \theta \geq 0$ on $\Lambda_+(f,e)$, in other words, $c^t x - \theta$ does not vanish on the interior of $\Lambda_+(f,e)$, by convexity.

Then one can locally relax the original program :

$$\begin{array}{rcll} c_k^* & := & \text{inf} & c^T x \\ & & \text{s.t.} & x \in L \\ & & & (x_0, x) \in \Lambda_+(f_k, \epsilon) \end{array}$$

with $f_k(x_0, x) = D^k_{\epsilon^k}(f \cdot (c^T x - x_0)^k)$ and $\epsilon = (\theta, e)$.