UNIVERSITÉ DE LIMOGES
FACULTÉ DES SCIENCES ET TECHNIQUES

MASTER THESIS OF MASTER ACSYON

Hyperbolic Polynomials and Hyperbolic Programming

Presented by:
Andrii POLIATYKIN

Supervisor:
Simone NALDI

Limoges, September 2020

Acknowledgments

I would like to express my deep gratitude to Professor Simone NALDI, my internship supervisors, for their continued support and thoughtful advice.

In addition, I would like to express my gratitude to the staff of Master 2 ACSYON for their continued help and helpful criticism during my studies at the University of Limoges.

In addition, I am very grateful to the members of the XLIM Research Institute for providing me with a good and friendly work environment that allowed me to stay focused on my studies and overcome obstacles on my way.

Finally, I am very grateful to my parents and friends for their support and sincere encouraging during my work.

Abstract

In this thesis we are studying hyperbolic polynomials which a homogeneous case of the multivariate real-rooted polynomials. Hyperbolicity cones which are related to the hyperbolic programming usually contain singularities, so we propose a variant of relaxations which allow us to connect a regularized feasible set and objective function, what was not done yet.

In this work we also proved that multiplicity of the regularized cone decreases almost everywhere except exact optimal solution. We proposed algorithms which allows approximate the optimal solution what seems to converge, but this fact was not proved.

We completed few numerical examples to show how to apply algorithms and how they work and proposed some options how one can modify proposed algorithms in order to approximate or even reach optimal solution.

Keywords: Hyperbolic Polynomials, Hyperbolic Programming, Optimization, Relaxations, SDP

Contents

Part I

Preliminaries

1 Hyperbolic polynomials

This section gives an introduction to the theory of hyperbolic polynomials. Hyperbolic polynomials first were studied by Lax [4], Gårding [2] in the context of the hyperbolic PDEs.

We recall that the univariate polynomials that have only real roots are called real-rooted polynomials and they can be generalized to the multivariate case in several ways, two of which are: *real zero polynomials* and *hyperbolic polynomials*. Hyperbolic polynomials are the homogeneous version of real zero polynomials and that's why we are interesting to work with this setting.

From now on we consider that there is more than one variable $n > 1$ and that the polynomials are homogeneous.

1.1 Definition and examples

We recall that a multivariate polynomial with all terms having the same degree are called *homogeneous*. Equivalently, $p(\alpha x) = \alpha^d p(x)$ for all $\alpha \in \mathbb{R}$ and $x \in \mathbb{R}^n$, where $d = deg(p)$. We denote by $\mathbb{R}[x]_d$ the set of homogeneous polynomials of degree d with n variables $x = (x_1, \ldots, x_n)$. Hyperbolic polynomials generalize real-rooted polynomials in the sense of the following definition:

Definition 1. A homogeneous polynomial $p \in \mathbb{R}[x]_d$, with $x = (x_1, \ldots, x_n)$ is called *hyperbolic with respect to direction $e \in \mathbb{R}^n$* if:

- $p(e) \neq 0$, and
- for each $a \in \mathbb{R}^n$, the univariate polynomial $t \mapsto p(a - te)$ has only real roots.

Furthermore it is strictly hyperbolic with respect to direction e if it is hyperbolic and the roots of $p(te - a)$ are all distinct, for every $a \in \mathbb{R}^n$.

We denote by $d = \deg(p) \geq 1$ the degree of p, and often assume by

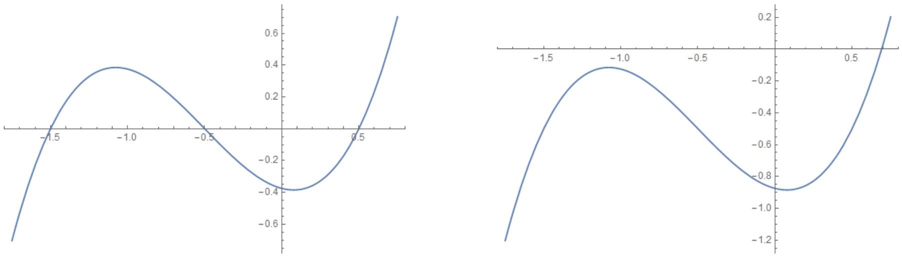

Figure 1.1: (a) plot of the graph of a real rooted cubic polynomial and (b) plot of non real rooted cubic polynomial

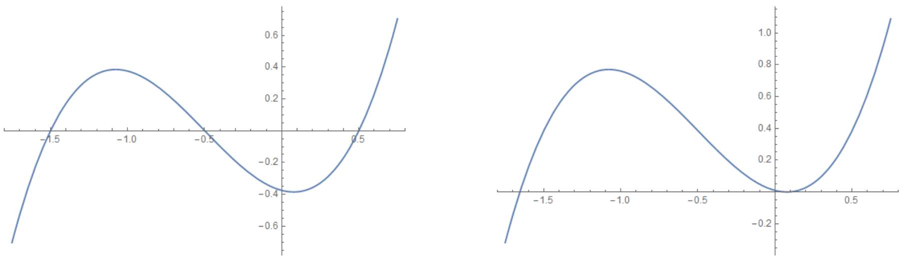

Figure 1.2: (a) plot of the graph of a strict real rooted cubic polynomial and (b) plot of non-strict real rooted cubic polynomial

simplicity that $p(e) > 0$ (we can replace p with $-p$ if necessary). Also, the set of hyperbolic polynomials of degree d, with n variables and with respect to direction e is denoted by $\mathcal{H}_{d,n}(e)$.

Property 1. We could equivalently define a hyperbolic polynomial by replacing "for each $a \in \mathbb{R}^n$, $t \mapsto p(a - te)$ has only real roots" with "for each $a \in \mathbb{R}^n$, $t \mapsto p(e - ta)$ has only real roots". In particular, one can also replace it with $p(a + te)$ or $p(e + ta)$ since the property is for all $a \in \mathbb{R}^n$.

Proof. We want to prove that $p(e - ta)$ has only real roots if and only if $p(a - te)$ has only real roots (with the assumption that p is hyperbolic with respect to e). Since $p(e) \neq 0$, one has that 0 is not a root of $p(e - ta)$. Thus $p(e - ta)$ has only real roots if and only if

$$(-t)^d p(a - \frac{1}{t}e)$$

has only real (non-zero) roots, hence if and only if $p(a - \frac{1}{t}e)$ has only real roots. Since the map $t \mapsto \frac{1}{t}$ preserves reality, we have that this is equivalent to $p(a - te)$ being real rooted. □

Definition 2. The univariate polynomial $t \mapsto p(a - te)$ is called the *characteristic polynomial* of a (with respect to p, in direction e) and we denote it by $\mathrm{ch}_{a;p}(t)$. The roots of $\mathrm{ch}_{a;p}(t)$ are called the eigenvalues of a. Write the eigenvalues of a as $\lambda_1(a) \leq \cdots \leq \lambda_d(a)$ counting multiplicities.

Thus, a hyperbolic polynomial with respect to direction e, is a homogeneous polynomial with the property that $p(e) > 0$ and for each a, all the eigenvalues of a are real.

Example 1. Suppose we have a polynomial $p = det(yI - xA) \in \mathbb{R}[x, y]$ and direction $e = (0, 1)$, where I is identity matrix and A is a symmetric matrix. We are going to check if this polynomial is hyperbolic.

One can check that $p(e) = det(I) = 1 \neq 0$.

Now we need to check if $\mathrm{ch}_{a;p}(t) = p(te - a)$ has only real roots.

$\forall a \in \mathbb{R}^2$ we obtain:

$$\mathrm{ch}_{a;p}(t) = p(te - a) = p(-a_1, t - a_2) = det((t - a_2)I + a_1 A) = det(tI - [a_2 I - a_1 A])$$

Since I and A are symmetric matrices then we have that $det(tI - [a_2 I - a_1 A])$ has only real roots because its roots are the eigenvalues of matrix $[a_2 I - a_1 A]$, which is symmetric.

Two observations of Renegar [8]:

$$\lambda_j(sa - te) = \begin{cases} \lambda_j(sa - te) = s\lambda_j(a) - t & \text{if } s \geq 0, \\ \lambda_j(sa - te) = s\lambda_{d-j}(a) - t & \text{if } s \leq 0, \end{cases}$$

$$p(a) = p(e) \prod_j \lambda_j(a).$$

We can easily see that hyperbolicity is closed under product:

Property 2. If $p_1 \in \mathcal{H}_{d_1,n}(e)$ and $p_2 \in \mathcal{H}_{d_2,n}(e)$, then $p_1 p_2 \in \mathcal{H}_{d_1+d_2,n}(e)$.

Proof. We need to prove that $p_1 p_2$ is hyperbolic in direction e. Indeed, since p_1 and p_2 are hyperbolic then $p_1(e) \neq 0$ and $p_2(e) \neq 0$, then $p_1(e)p_2(e) \neq 0$ as well. And also we know that for each $a \in \mathbb{R}^n$, the univariate polynomials

$t \mapsto \mathrm{ch}_{a;p_1}(t)$ and $t \mapsto \mathrm{ch}_{a;p_2}(t)$ have only real roots, then for each $a \in \mathbb{R}^n$, the univariate polynomial $t \mapsto \mathrm{ch}_{a;p_1}(t)\mathrm{ch}_{a;p_2}(t)$ is also real-rooted. Then we remark that

$$\mathrm{ch}_{a;p_1 p_2}(t) = (p_1 p_2)(te - a) = p_1(te - a)p_2(te - a) = \mathrm{ch}_{a;p_1}(t)\mathrm{ch}_{a;p_2}(t).$$

$\square$

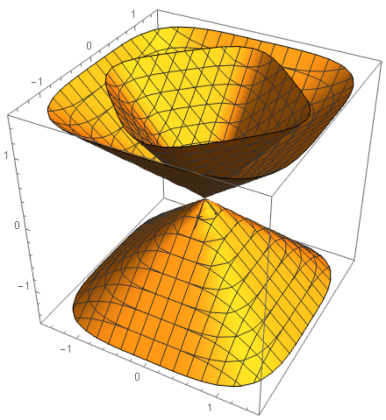

Figure 1.3: quartic hyperbolicity cone in $\mathbb{R}^3$.

Remark 1. The characteristic polynomial of a product is the product of the characteristic polynomials.

Definition 3. The multiplicity of a with respect to p in direction e is the multiplicity of $t = 0$ as an eigenvalue of a, hence as a root of the characteristic polynomial $\mathrm{ch}_{a;p}(t) = p(te - a)$, and is denoted by $mult_p(a)$.

Next we introduce a few fundamental examples of hyperbolic polynomials.

Example 2. A polynomial $p \in \mathbb{R}[x]_d$ that can be expressed as the product of d real linear forms, that is, $p(x) = \ell_1(x)\ell_2(x)\ldots\ell_d(x)$, for some $\ell_i \in \mathbb{R}[x]_1$, is hyperbolic with respect to any vector e such that $\ell_i(e) \neq 0$ for all i. Here we can see that $\mathrm{ch}_{a;p}(t) = \ell_1(a - te)\ell_2(a - te)\cdots\ell_d(a - te)$ and (with assumption that $\ell_i(x) = c_i^T x$) the eigenvalues of a are $\lambda(a) = (\frac{c_1^T a}{c_1^T e}, \cdots, \frac{c_d^T a}{c_d^T e})$

Example 3. The polynomial $p = -x^2 - y^2 + z^2 \in \mathbb{R}[x, y, z]$ is hyperbolic with respect to the vector $e = (0, 0, 1)$, because the characteristic polynomial $\mathrm{ch}_{a;p}(t) = -(a_1)^2 - (a_2)^2 + (t - a_3)^2$ has discriminant $4a_1^2 + 4a_2^2 \geq 0$

and thus has two real roots in t, for all $a = (a_1, a_2, a_3) \in \mathbb{R}^3$. Clearly, $t_{1,2} = a_3 \pm \sqrt{a_1^2 + a_2^2}$. If the discriminant is equal to zero then we get a double root in $t = a_3$.

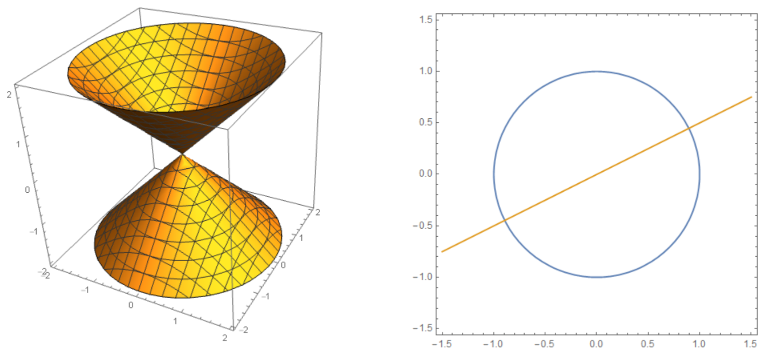

Figure 1.4: plot of the zero set of the polynomial $p = -x^2 - y^2 + z^2$ (a) in $\mathbb{R}^3$ and (b) in projective space $\mathbb{P}^2$.

Example 4. The polynomial $p = -x^4 - y^4 + z^4 \in \mathbb{R}[x, y, z]$ is **not** hyperbolic with respect to any point $e \in \mathbb{R}^3$. In particular, for $e = (0, 0, 1)$, we can check that characteristic polynomial $\text{ch}_{a;p}(t) = p(a - te)$ has two real but also a couple of not real complex-conjugate roots for all $a \in \mathbb{R}^3$.

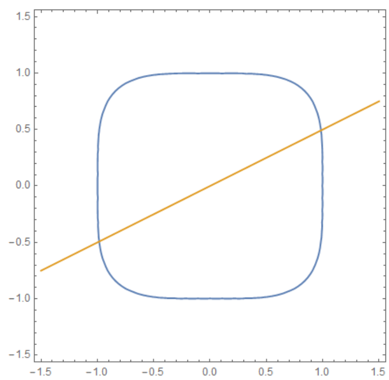

Figure 1.5: plot of the zero set of the polynomial $p = -x^4 - y^4 + z^4$ in the projective space $\mathbb{P}^2$.

Example 5. Perhaps the most important example of hyperbolic polynomials. Let $X \in \mathbb{S}^{d \times d}$ (vector space of $d \times d$ symmetric matrices and thus $n = \frac{d(d+1)}{2}$), $p(X) = det(X)$ is hyperbolic with respect to any symmetric matrix with only positive eigenvalues E. Since E is positive definite we know that $p(E) \neq 0$ and one can replace it as $E = F^2$, where F is a positive definite matrix. Then we obtain:

$$\mathrm{ch}_{a;p}(t) = det(tE - A) = det(F(tI - F^{-1}AF^{-1})F) = det(F)^2 det(tI - F^{-1}AF^{-1})$$

Since $F^{-1}AF^{-1}$ is a symmetric matrix then we obtain that $p(tE - A)$ is a multiple of the characteristic polynomial of the symmetric matrix $F^{-1}AF^{-1}$. Hence the characteristic polynomial is real-rooted and we conclude that p has only real roots.

Instead of the general symmetric determinant one can consider determinant of pencil of symmetric matrices: Let $A = x_1A_1 + \cdots + x_nA_n$ be a homogeneous linear matrix polynomial of size $d \times d$ with assumption that there exists $e \in \mathbb{R}^n$ with $A(e) = I_d$. Then $p = det(A) \in \mathbb{R}[x]$ is homogeneous of degree at most d and hyperbolic with respect to e. Again, this is because $p(te - a) = det(tI_d - A(a))$ is the characteristic polynomial of the symmetric matrix $A(a)$ and therefore has only real roots.

More generally, the same remains true if $A(e)$ is just any positive definite matrix, by considering $\sqrt{A(e)}^{-1}A(x)\sqrt{A(e)}^{-1}$.

Example 6. Let A_1, A_2 be symmetric matrices and let $A = xA_1 + yA_2$ be a homogeneous linear matrix polynomial of size 3×3 and suppose there exists $e \in \mathbb{R}^2$ with positive definite $A(e)$. Taking for instance:

$$A_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 9 \\ 3 & 9 & 8 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 5 & 3 \\ 1 & 3 & 3 \end{pmatrix}$$

we can see that

$$p = \det A = \det(xA_1 + yA_2) = -9x^3 - 38x^2y - 2xy^2 + 4y^3$$

is a hyperbolic cubic in direction $e = (0, 1)$. Indeed, we have that $p(e) = \det(A_2) = 4 \neq 0$ and matrices A_1 and A_2 are symmetric, so according to Example 5, $p = \det A$ is hyperbolic.

For $a = (5, 2)$, the characteristic polynomial is $p(te - a) = det(tA(e) - A(a)) = 4t^3 - 14t^2 - 942t + 3033$. Its roots are approximately:

$$\lambda(a) \approx (-15.23077104, 3.206947332, 15.52382371).$$

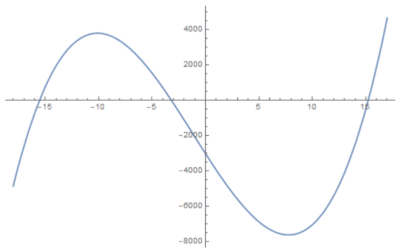

Figure 1.6: (a) plot of the zero set of the polynomial $p = -9x^3 - 38x^2y - 2xy^2 + 4y^3$ in $\mathbb{R}^2$; (b) plot of the characteristic polynomial $\mathrm{ch}_{a;p}(t) = p(te - a) = -3033 - 942t + 14t^2 + 4t^3$ in $\mathbb{R}^2$ with respect to direction $e = (0, 1)$ and $a = (5, 2)$ to have characteristic polynomial $p(te - a)$.

1.2 Hyperbolicity cones

In this subsection we recall some definitions, properties and theorems about hyperbolicity cones, which we are going to use later.

Definition 4. The set $\Lambda_{++}(p, e) := \{a : \lambda_{min}(a) > 0\}$ is called the open hyperbolic cone (for p in direction e). The set $\Lambda_+(p, e) := \{a : \lambda_{min}(a) \geq 0\}$ is called the closed hyperbolic cone (for p in direction e).

It's known that $\Lambda_+(p, e)$ is the closure of $\Lambda_{++}(p, e)$ [8].

Clearly, $e \in \Lambda_{++}(p, e)$. Note, that if $a \in \Lambda_{++}(p, e)$ then $p(a) > 0$ (because $p(a) = p(e) \prod_j \lambda_j(a)$). Also, observe that $\Lambda_{++}(p, e)$ is indeed a cone, i.e., if $a \in \Lambda_{++}(p, e)$, then $\mu a \in \Lambda_{++}(p, e)$ for all $\mu > 0$.

We denote by $V(p)$ the real zero set of the polynomial p:

$$V(p) = \{x \in \mathbb{R}^n : p(x) = 0\}.$$

Theorem 1. *The open hyperbolicity cone is the connected component of the complement of $V(p)$ containing e and hyperbolicity cones are convex.*

Proof. Proof given in [8]. $\square$

Since $\Lambda_+(p, e)$ is the closure of $\Lambda_{++}(p, e)$, it follows that $\Lambda_+(p, e)$ is convex, too.

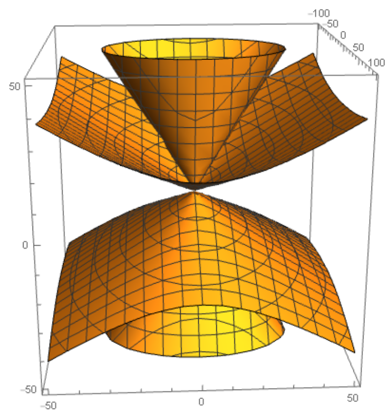

Figure 1.7: quartic hyperbolicity cone of the hyperbolic polynomial $p = 1250000x^4 - 1749500x^3y - 2250800x^2y^2 - 4312500x^2z^2 + 69260xz^3 + 786875xyz^2 + 88176y^4 + 1141000y^2z^2 + 1687500z^4$ in $\mathbb{R}^3$.

Theorem 1 is a corollary of the following result

Theorem 2. *If $a \in \Lambda_{++}(p,e)$, then p is hyperbolic in direction a. Moreover, the hyperbolicity cone in direction a is $\Lambda_{++}(p,e)$ (i.e., the same cone as in direction e), that is*

$$\Lambda_{++}(p,a) = \Lambda_{++}(p,e)$$

Proof. Proof given in [8] □

Later, we consider various directions $e' \in \Lambda_{++}(p,e)$, as is allowed due to Theorem 2. When required for clarity, we make dependence on e explicit, writing, for example, $\lambda_{j,e}(a)$.

Let's recall that a polyhedron is a subset of $\mathbb{R}^n$ described by a finite number of linear inequalities, i.e.

$$P = \{x \in \mathbb{R}^n : \ell_i(x) \geq 0 \text{ for all } i = 1, \ldots, d\}.$$

A spectrahedron is the feasible set of a semidefinite programming (SDP) problem. It is given as the intersection of the cone of positive semidefinite symmetric matrices with a linear subspace, i.e.

$$S = \{x \in \mathbb{R}^n : A_0 + x_1A_1 + \cdots + x_nA_n \succeq 0\}$$

for some symmetric $d \times d$ matrices A_i.

Particularly it's easy to see that polyhedra and spectrahedra are special cases of hyperbolicity cones. More precisely:

Theorem 3. *The following facts hold:*

- *Every polyhedron is a spectrahedron.*
- *Every spectrahedron is a hyperbolicity cone.*

Proof. (1) It was shown in example 2 for polyhedra; (2) and we have shown in example 5 that all polynomials which can be represented as determinants of pencils of symmetric matrices are hyperbolic. $\square$

There exist spectrahedra which are not polyhedral. The easiest example of specrathedron which is not polyhedron is a unit disk $D \subset \mathbb{R}^2$, because it is defining polynomial has symmetric determinantal representation and it can not be represented in the form of finitely many linear inequalities:

$$\det \begin{bmatrix} z-x & y \\ y & z+x \end{bmatrix} = z^2 - x^2 - y^2.$$

and

$$D = \{(x,y,z) \in \mathbb{R}^3 : z = 1, \text{ and } \begin{bmatrix} z-x & y \\ y & z+x \end{bmatrix} \succeq 0\}$$

Lax in his work [4] conjectured that every hyperbolic polynomial has such an explicit representation:

Conjecture 1. Every hyperbolic form have a definite determinantal representation.

But later this conjecture turned out to be false in general. It is still an open question for research whether every hyperbolicity cone is spectrahedral, and called *Generalized Lax Conjecture* which is true for $n = 3$ (see next section).

Next property shows that the cone of the product of hyperbolic polynomials is the intersection of the cones of these polynomials:

Property 3. Let $p \in \mathcal{H}_{d_1,n}(e)$ and $q \in \mathcal{H}_{d_2,n}(e)$. Then $\Lambda_+(pq, e) = \Lambda_+(p, e) \cap \Lambda_+(q, e)$.

Proof. According to Property 2 and Remark 1 we remind that the product of polynomials p and q which are hyperbolic with respect to e is a hyperbolic polynomial with respect to e and characteristic polynomial of the product is a product of characteristic polynomials:

$$\mathrm{ch}_{a;pq}(t) = \mathrm{ch}_{a;p}(t)\mathrm{ch}_{a;q}(t).$$

Then, according to definition and construction we obtain the follows equivalent statements:

$$a \in \Lambda_+(pq, e) \iff \text{roots of } \mathrm{ch}_{a;pq}(t) \text{ are non-negative}$$

$$\iff \text{roots of } \mathrm{ch}_{a;p}(t) \text{ and those of } \mathrm{ch}_{a;q}(t) \text{ are non-negative}$$

$$\iff a \in \Lambda_+(p, e) \cap \Lambda_+(q, e).$$

□

1.3 Determinantal representations

In this subsection we introduce few important results about hyperbolic polynomials which admit a determinantal representation.

Suppose we have polynomial $p \in \mathbb{R}[x]$ and a symmetric linear matrix polynomial $A = x_1A_1 + \cdots + x_nA_n$ such that $p = \det(A)$. We say that A is a symmetric (linear) determinantal representation of p. Moreover, if there exists $e \in \mathbb{R}^n$ such that $A(e)$ is positive definite, then we say that the determinantal representation is definite.

Remark 2. If p has a definite determinantal representation $p = \det(A)$, and e is such that $A(e) \succ 0$, then p is hyperbolic with respect to direction e.

However, in 1957 it was proposed by Lax that every hyperbolic polynomial has a definite determinantal representation and it became known as the Lax Conjecture [4]. It was proved in the work of Helton and Vinnikov [3] for the case of $n = 3$ variables and any degree d.

Theorem 4 (Helton-Vinnikov)**.** *Every hyperbolic polynomial in three variables possesses a definite symmetric determinantal representation.*

This also proves the following result:

Corollary 1. *Every three-dimensional hyperbolicity cone is spectrahedral*

As follows we introduce Generalised Lax Conjecture which says that every hyperbolicity cone is spectrahedral, or every hyperbolicity cone has a spectrahedral representation.

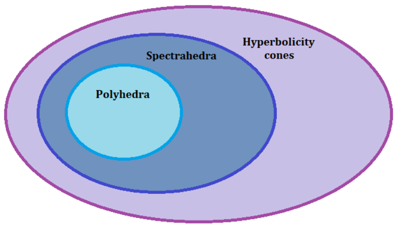

Figure 1.8: The sets of Polyhedra, Spectrahedra and Hyperbolicity Cones

However, it is in fact not hard to see that not every hyperbolic polynomial possesses a symmetric determinantal representation, by counting dimensions:

Proposition 1. *If $n \geq 4$ and $d \geq 7$, there exist hyperbolic polynomials in n variables of degree d, that do not admit a symmetric determinantal representation.*

In 2011 Brändén in his work [1] was studying the case for $n > 3$ and found that there exist hyperbolic polynomials such that they do not have a definite determinantal representation. It was shown that no powers p^k of such polynomials have this representation. It means that one can not prove the Generalised Lax Conjecture by proving that p^k has definite symmetric determinantal representation.

But the Generalised Lax Conjecture is equivalent to the following algebraic statement:

Conjecture 2. (Generalised Lax Conjecture) Let p be a hyperbolic polynomial. Then there exists a hyperbolic polynomial g with respect to direction e such that

- gp has a definite symmetric determinantal representation;
- the hyperbolicity cone of q with respect to direction e contains the hyperbolicity cone of p with respect to direction e:

$$\Lambda_+(g,e) \supseteq \Lambda_+(p,e).$$

The main result of this conjecture is that we can work with hyperbolic programming as with SDP.

1.4 Origin and motivation

Hyperbolic polynomials first were introduced in work of Lax [4] in context of PDEs. Also in this work first was introduced conjecture which was mentioned in previous subsection.

In 2007 original statement was proved in Helton-Vinnikov theorem for case where $n = 3$. However, in general case it is false, according to counterexamples founded by Branden [1]. Branden shows that conjecture can not be proved by proving that $f^k = det(A)$, where f is hyperbolic polynomial and $\det(A)$ is a definite determinantal representation of its polynomial.

But it seemed to be possible to find such representation of the hyperbolicity cone as spectrahedral. Then was introduced generalized Lax conjecture what says that product of hyperbolic polynomials can posses a definite determinantal representation. Lewis in his work [5] tried to prove that generalized Lax conjecture is true, but for the moment it is still an open question if it is true of false.

Garding proved convexity other properties and theorems of the hyperbolicity cone in his work [2].

In 2006 Renegar studied hyperbolic optimization, and some of the most famous results in hyperbolic programming were introduced in his work [8]. He defined derivatives of the hyperbolic polynomials and hyperbolicity cones. We need it to regularize the original cone in optimization problem to avoid computations in singular points in numerical algorithms, but in his work there was no relation to the objective function, so it does not solve the hyperbolic programming and it can be modified at least to get approximation of the solution of the optimization problem.

Naldi and Plauman in their work [6] in computer algebra proposed an approach based on symbolic computation depending on multiplicity of the boundary of the feasible set. Their work gives possibility to certify multiplicity of the solution of the hyperbolic programming and evaluate optimal value of the objective function.

In 2013 Sanyal proposed theorem which says that derivative of the polyhedral cone posses a spectrahedral representation what gives us ability to regularize the cone with relation to the objective fucntion such that at least we are able to approximate the solution of the hyperbolic programming where the feasible set is a polyhedron. However we do not have such kind of determinantal representation for the case of SDP.

2 Hyperbolic programming

2.1 Definition

Hyperbolic programming is a convex optimization problem specified as follows.

Definition 5. Let $p \in \mathbb{R}[x]_d$ with $x = (x_1, \ldots, x_n)$ be hyperbolic with respect to $e \in \mathbb{R}^n$. A hyperbolic program is an optimization instance problem of the form:

$$\min c^T x$$

$$\text{s.t. } Ax = b$$

$$x \in \Lambda_+(p, e)$$

where $x \mapsto c^T x$ is a linear functional, A is a matrix of size $m \times n$ and $Ax = b$ is a system of linear equations.

In other words, a hyperbolic programming asks to minimize a linear function $c^T x$ over the hyperbolicity cone.

2.2 Linear and semidefinite programming

Notice, LP can be rewritten in the sense of hyperbolic programming as follows:

Definition 6. Let $p(x) = x_1 \cdots x_n$ be hyperbolic with respect to any vector $e = (e_1, \cdots, e_n)$ with $e_i > 0$, $\forall i$. Then we have the following optimization problem:

$$\min c^T x$$

$$\text{s.t. } Ax = b$$

$$x \in \mathbb{R}^n_+$$

where $x \mapsto c^T x$ is a linear functional, A is a matrix of size $m \times n$ and $Ax = b$ is a system of linear equations.

The characteristic polynomial corresponding to LP is $ch_a(t) = p(te - a) = \prod_{i=1}^{n}(te_i - a_i)$ and hyperbolicity cone is $\Lambda_+(p,e) = \mathbb{R}_+^n$ and $\Lambda_{++}(p,e) = \mathbb{R}_{++}^n$

Also in the sense of hyperbolic programming we can rewrite SDP as follows:

Definition 7. Semi-Definite Programming (SDP): $M(x) = \sum_{i=1}^{n} x_i M_i$, $p(x) = \det M(x)$ is hyperbolic with respect to $M(e)$, where $e \in \mathbb{R}^n$, such that $M(e) \succ 0$. Then we have the following optimization problem:

$$\min c^T x$$

$$\text{s.t. } x \in \mathcal{S}_M = \{y \in \mathbb{R}^n : M(y) \succeq 0\}$$

where $x \mapsto c^T x$ is a linear functional.

The characteristic polynomial corresponding to SDP is $ch_M(t) = p(te - a) = \det(M(te-a)) = \det(tM(e) - M(a))$ and hyperbolicity cone is $\Lambda_+(p,e) = \mathbb{S}_M$.

2.3 Renegar derivatives

To introduce the derivative cones for multivariate hyperbolic polynomials, we need to recall the univariate case. If $p \in \mathbb{R}[t]$ is real-rooted, then the roots of p' interlace that of p, that is, between any two roots of p, there is a root of p'. Since p' has precisely one less root (counting multiplicities) than does p, by a degree argument, it follows that p having only real roots implies p' has only real roots.

For multivariate hyperbolic polynomials, the situation is the same. If $p \in \mathcal{H}_{d,n}(e)$, we know by definition that for every $a \in \mathbb{R}^n$, the characteristic polynomial $\text{ch}_{a;p}(t)$ is real-rooted. Now it is easy to check (see for instance [8]) that the derivative of $\text{ch}_{a;p}(t)$ is the characteristic polynomial of a with respect to $D_e p$ in direction e, where $D_e p$ is the directional derivative of p in direction e:

$$\text{ch}_{a;D_e p}(t) = \frac{d}{dt}\text{ch}_{a,p}(t), \text{ with } D_e p = \langle \nabla p, e \rangle.$$

Hence this shows that $D_e p$ is again hyperbolic, since $\mathrm{ch}_{a;D_e p}(t)$ is real-rooted.

We denote the hyperbolicity cone of $D_e p$ by $\Lambda'_+(p,e)$ (and by $\Lambda'_{++}(p,e)$ its interior). Remark that the derivative cone depends on the direction e. For example, when we are working in $\mathbb{R}^2$ with $p(x_1,x_2) = x_1 x_2$, the derivative cone of p in direction $e = (e_1, e_2)$ is the closed half-space with boundary orthogonal to $(1/e_1, 1/e_2)$ (recall that the coordinates of e must be nonzero because $p(e) \neq 0$).

By the interlacing property discussed above, we have that the eigenvalues $\lambda'_i(a)$ with respect to $D_e p$ of e, interlace the eigenvalues with respect to p:

$$\lambda_1(a) \leq \lambda'_1(a) \leq \lambda_2(a) \leq \cdots \leq \lambda'_{n-1}(a) \leq \lambda_i(a),$$

where (by [8])

$$(\lambda_j(x) = \lambda'_j(x) \text{ or } \lambda'_j(x) = \lambda_{j+1}(x)) \iff \lambda_j(x) = \lambda'_j(x) = \lambda_{j+1}(x)$$

As a simple consequence of the interlacing we have

$$\Lambda_+(p,e) \subseteq \Lambda'_+(p,e)$$

i.e., the derivative cone is a relaxation of the original cone $\Lambda_+(p,e)$.

We will use the following theorem below, which states that taking derivative of the hyperbolic polynomial reduces the multiplicity, i.e. after taking derivative we obtain the cone which is more regular than without taking derivative but the derivative cone that we obtain is larger and still contains the previous one.

It uses the notation:

$$\partial^m \Lambda_+(p,e) = \{a \in \Lambda_+(p,e) : mult(a) = m\}.$$

Theorem 5 (Theorem 15 in [8]). *For integers $m \geq 2$,*

$$\partial^m \Lambda'_+ = \partial^{m+1} \Lambda_+$$

Also,

$$\partial^1 \Lambda'_+ \cap \Lambda_+ = \partial^2 \Lambda_+$$

Proof. Straightforward consequences of the interlacing of eigenvalues and the equivalence $(\lambda_j(x) = \lambda'_j(x)$ or $\lambda'_j(x) = \lambda_{j+1}(x)) \iff \lambda_j(x) = \lambda'_j(x) = \lambda_{j+1}(x)$. □

Now if we take derivatives repeatedly we obtain a sequence of hyperbolic polynomials $\{p = p^{(0)}, D_e p = p^{(1)}, \ldots, D_e^{(d-1)} p = p^{(d-1)}\}$ with nested hyperbolicity cones:

$$\Lambda_+(p,e) = \Lambda_+^{(0)}(p,e) \subseteq \Lambda'_+(p,e) = \Lambda_+^{(1)}(p,e) \subseteq \ldots \subseteq \Lambda_+^{(d-1)}(p,e)$$

Here the cone $\Lambda_+^{(d-1)}(p,e)$ is a closed half-space, since $D_e^{(d-1)} p$ has degree 1.

Also, one can deduce the following result using Theorem 5 for $m \geq 2$:

$$\partial^m \Lambda_+^{(i)}(p,e) = \partial^{m+1} \Lambda_+^{(i-1)}(p,e) = \ldots = \partial^{m+i} \Lambda_+^{(0)}(p,e).$$

Next we are going to study an example of the chain of derivative cones to show previous statements in the plot.

Example 7. The polynomial $p = x^4 + 3y^4 + 5x^3z + 6x^2z^2 - 6y^2z^2 \in \mathcal{H}_{4,3}(e)$ is hyperbolic with respect to direction $e = (-1,0,1)$. We are going to show that $\Lambda_+(p,e) \subseteq \Lambda'_+(p,e) \subseteq \Lambda_+^{(2)}(p,e) \subseteq \Lambda_+^{(3)}(p,e)$, where $\Lambda_+^{(3)}(p,e)$ is a closed half-space. Since p is hyperbolic in direction $e = (-1,0,1)$, then we are going to take direction derivatives in this direction as follows:

$$p^{(i+1)} = e_1 \partial_x p^{(i)} + e_2 \partial_y p^{(i)} + e_3 \partial_z p^{(i)}, \text{ for all } i = 0,1,2.$$

We obtain:

$$D_e p = x^3 - 3x^2z - 12y^2z - 12xz^2$$
$$D_e^{(2)} p = -6x^2 - 12y^2 - 18xz + 12z^2$$
$$D_e^{(3)} p = -6x + 42z$$

Next we are going to show the hyperbolicity cones related to the derivatives of the hyperbolic polynomials intersected with a plane $L = \{(x,y,z) \in \mathbb{R}^3 : z = 1\}$:

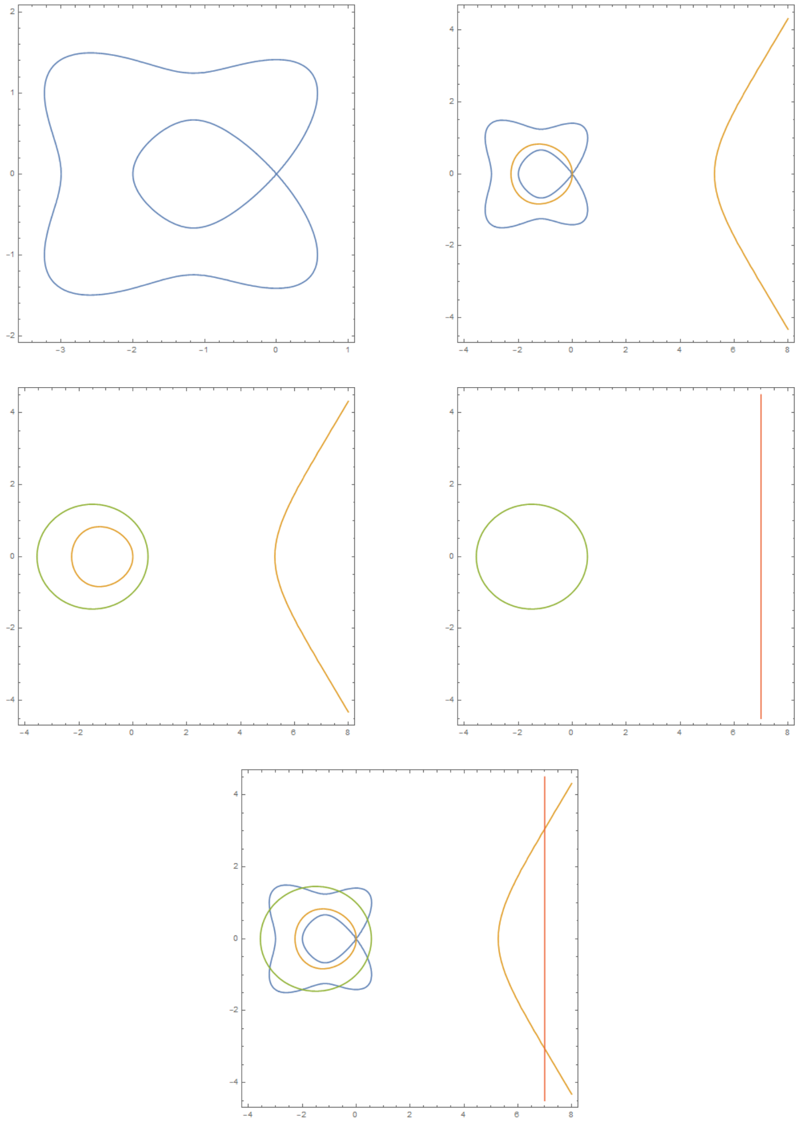

Figure 2.1: The derivative cones of the previously mentioned sequence of the derivatives of the hyperbolic polynomials respectively: (a) $\Lambda_+(p,e)$, (b) $\Lambda_+(p,e) \subseteq \Lambda'_+(p,e)$, (c) $\Lambda'_+(p,e) \subseteq \Lambda^{(2)}_+(p,e)$, (d) $\Lambda^{(2)}_+(p,e) \subseteq \Lambda^{(3)}_+(p,e)$ and (e) all the derivative cones.

2.4 Interior-point methods

Working with hyperbolic programming we can apply interior-point methods. Remind that interior-point methods are class of algorithms that solve linear and nonlinear convex optimization problems. In case of hyperbolic programming applying IPMs methods we can solve the optimization problem.

By Renegar's book [7] for an LP and SDP we have that the main self-concordant functionals are of the following form:

$$\eta c^T x + f|_L(x);$$

where $\eta \geq 0$ is constant, L is affine subspace such that $L = \{x \in \mathbb{R}^n_+ : Ax = b\}$, f is a logarithmic barrier function for:

- the non-negative orthant in case of LP
- the cone of positive definite matrices in case of SDP

In case of hyperbolic programming the barrier function is the following:

$$f = -log\ p$$

For more information about IPM see for instance [7]

Part II

Contributions

3 Modified relaxations

In this section we discuss classical relaxation of Renegar and propose a variant of Renegar derivatives relaxation due to weaknesses of classical Renegar relaxation. With this we propose an algorithm which can approximate the minimum and in some cases can solve the hyperbolic programming. Also it gives an idea that it could be proved in general case.

3.1 Weakness of classical relaxations

Usually a hyperbolic polynomial is singular, and hence the feasible set of the hyperbolic polynomial can contain singularities.

Numerical algorithms typically have better behavior when applied to optimization problem over smooth feasible sets. So the idea of regularizing the feasible set is natural in optimization.

To regularize our objective function and respectively hyperbolicity cone we need to apply relaxation. Renegar proposed to take derivative to obtain something more smooth and regular.

In Renegar's classical relaxation we take a cone (feasible set of the optimization problem) and take derivative of its. It gives us a new hyperbolicity cone which is larger. It regularizes the cone, but at the same time it does not take into account the input objective function. Moreover, Renegar's relaxations give only finitely many lower bounds (that become worse along the relaxation), hence the method can not be used in order to approximate the original minimizer.

This was the main problem we identified, of using this relaxation for hyperbolic programming. So the question now is how to use the derivative relaxation to regularize the original program and make relation to the objective function to make an algorithm which is converging to the minimum.

In the next subsection we propose such modification which at least approximate the minimum, but it was not proved that it converges to the minimum in general.

3.2 Modified Renegar relaxation

Taking derivative we decrease the degree of the function. In general it is nice, but since it is not related to the objective function we can not solve hyperbolic programming. To avoid this weakness we propose to take product of the hyperbolic polynomial p which describes our hyperbolicity cone and something what will give us the same hyperbolicity cone.

Next we focus on the special case of hyperbolic programming where the section is given by the hyperplane $e^T x = 1$:

Let a homogeneous polynomial $p(x) \in \mathbb{R}[x]_d$ with $x = (x_1, \cdots, x_n)$ be hyperbolic with respect to $e \in \mathbb{R}^n$. A hyperbolic program is next

$$c^* = \min c^T x$$
$$\text{s.t. } e^T x = 1$$
$$x \in \Lambda_+(p, e)$$

where $x \mapsto c^T x$ is a linear functional and $e^T x = 1$ is a linear equation.

To approximate (or maybe even solve) this optimization problem we will use next lemma and theorem:

Lemma 1. *Let $p \in \mathcal{H}_{d,n}(e)$, be hyperbolic with respect to a vector e, and let $\Lambda_+(p, e)$ be the hyperbolicity cone of p in direction e. Suppose $\epsilon_0 \leq c^*$, that is, ϵ_0 is a lower bound for the restriction of $c^T x$ on $\Lambda_+(p, e)$. Let $l = c^T x - \epsilon_0 e^T x$ and $q = pl$. Then, the hyperbolicity cone of q in direction e is equal to the hyperbolicity cone of p in direction e:*

$$\Lambda_+(q, e) = \Lambda_+(p, e).$$

Proof. According to property 2 and property 3 we obtain respectively: $q = pl$ is hyperbolic in direction e and $\Lambda_+(q, e) = \Lambda_+(p, e) \cap \Lambda_+(l, e)$. The linear function $l = c^T x - \epsilon_0 e^T x$ by construction does not cross the interior of the $\Lambda_+(p, e)$, thus, since the $\epsilon_0 \leq c^*$, then the hyperbolic polynomial l is positive definite over $\Lambda_+(p, e)$. Thus, we obtain:

$$\Lambda_+(p, e) \cap \Lambda_+(l, e) = \Lambda_+(p, e)$$

and hence

$$\Lambda_+(q,e) = \Lambda_+(p,e).$$

$\square$

Next modification proposes a variant of regularizing of the feasible set which does not change the multiplicity and make the feasible set smoother.

Theorem 6. *Let $p \in \mathcal{H}_{d,n}(e)$, be hyperbolic with respect to a vector e, and let $\Lambda_+(p,e)$ be the hyperbolicity cone of p in direction e. Suppose $\epsilon_0 \le c^*$, that is, ϵ_0 is a lower bound for the restriction of $c^T x$ on $\Lambda_+(p,e) \cap L$, where $L = \{x \in \mathbb{R}^n : x_n = 1\}$. Let $l = c^T x - \epsilon_0 e^T x$. The following holds (for $a \in \partial\Lambda_+(p,e)$):*

1. *Assume $\epsilon_0 = c^*$. If $c^T a = c^*$, then $mult_{D_e q}(a) = mult_p(a)$; otherwise, $mult_{D_e q}(a) = mult_p(a) - 1$.*
2. *Otherwise, if $\epsilon_0 < c^*$, then $mult_{D_e q}(a) = mult_p(a) - 1$.*

Proof.

1. If $\epsilon_0 = c^*$ we obtain that for $a = a^*$: $l(a^*) = 0$ and $mult_l(a^*) = 1$. Then $mult_{D_e q}(a) = mult_{D_e(pl)}(a) = mult_{pl}(a) - 1 = mult_p(a) + mult_l(a) - 1 = mult_p(a)$. Otherwise if $a \neq a^*$, then $mult_l(a) = 0$ and $mult_{D_e q}(a) = mult_{D_e(pl)}(a) = mult_{pl}(a) - 1 = mult_p(a) + mult_l(a) - 1 = mult_p(a) - 1$.
2. If $\epsilon_0 < c^*$ we obtain $\forall a$: $l(a) \neq 0$ and $mult_l(a) = 0$. Then $mult_{D_e q}(a) = mult_{D_e(pl)}(a) = mult_{pl}(a) - 1 = mult_p(a) + mult_l(a) - 1 = mult_p(a) - 1$.

$\square$

This theorem shows that we can regularize the hyperbolicity cone with relation to the objective function.

Working with previously described optimization problem we need next information as input of the algorithm:

- $p \in \mathcal{H}_{d,n}(e)$ is hyperbolic with respect to a vector e;
- $\Lambda_+(p,e)$ is the hyperbolicity cone of p in direction e;

- $c^T x$ is objective function;
- $\epsilon_0 \leq c^*$ is a known lower bound for $c^T x$ over the feasible set $\Lambda_+(p, e) \cap L$, where $L = \{x \in \mathbb{R}^n : x_n = 1\}$;

Knowing a lower bound is a strong assumption, but in some cases (for example working with compact sections) we can assume it.

We construct a linear function $l = c^T x - \epsilon_0 e^T x$ which gives us relation to the objective function. Clearly, by construction the function $l(x)$ is positive definite over the feasible set $\Lambda_+(p, e)$.

Then we construct new polynomial $q = p \cdot l$. This polynomial is hyperbolic by Property 2 and, by Lemma 1, the hyperbolicity cone of q in direction e is equal to the hyperbolicity cone of p in direction e:

$$\Lambda_+(q, e) = \Lambda_+(p, e).$$

Taking the derivative $D_e q$ of the hyperbolic polynomial q we will obtain a new hyperbolic polynomial with $deg(D_e q) = deg(p)$.

Now, according to the property of interlacers and theorem 6, we will obtain more regular hyperbolicity cone and we can find a lower bound which is closer to the minimum of the objective function. Also, by property of interlacers, we can approximate the minimum.

Then we use interior point method to find new lower bound and repeating the algorithm we will approximate the minimum of the original program.

But it is better to find a definite determinantal representation of the feasible set. By Helton-Vinnikov theorem there is such representation and we will try to avoid taking derivative. Also, when we find a definite determinantal representation we have a feasible set as a spectrahedron. It means that our program is converted to SDP.

3.3 The case of linear programming

Remind that a polyhedron is a hyperbolicity cone. We are working with the hyperbolic programming given in Section 3.2. Let feasible set $\Lambda_+(p,e)(x)$ be a polyhedron.

Sanyal in his work [9] proposed the following theorem:

Theorem 7. *Let $P = \{x \in \mathbb{R}^n : \ell_i(x) \geq 0, \text{for } i \in 1,\dots,d\}$ be a full-dimensional polyhedral cone. Let $e \in$ int (P) and assume that $\ell_i(e) = 1$ for all i. Then the first derivative cone is given by all $x \in \mathbb{R}^n$ satisfying*

$$\begin{bmatrix} \ell_1(x)+\ell_n(x) & \ell_n(x) & \cdots & \ell_n(x) \\ \ell_n(x) & \ell_2(x)+\ell_n(x) & \cdots & \ell_n(x) \\ \vdots & \vdots & \ddots & \vdots \\ \ell_n(x) & \ell_n(x) & \cdots & \ell_{n-1}(x)+\ell_n(x) \end{bmatrix} \succeq 0.$$

Previous theorem shows that the derivative of a polyhedron is a spectrahedron. It means that we have the spectrahedral representation of the derivative cone. So, it solves problem of finding the definite determinantal representation and we can work with hyperbolic programming as with SDP.

Now, taking feasible set $\Lambda_+(p,e) = \{x \in \mathbb{R}^n : \ell_i(x) \geq 0, \text{for } i \in 1,\dots,d\}$ and all needed information of input given in Section 3.2 we can apply the following algorithm to solve hyperbolic program:

1. We construct the boundary constraint $\ell^{(1)}_{n+1}(x) = c^T x - \epsilon_0 e^T x$, where ϵ_0 is a known lower bound;
2. We normalize linear constraints as follows:
$$\ell_i^{(1)}(x) := \frac{\ell_i(x)}{\ell_i(e)} \text{ for all } i = 1,\dots,n+1,$$
such that by construction $\ell_i^{(1)}(e) = 1$.

   We obtain the hyperbolic polynomial:
$$p^{(1)}(x) = \prod_{i=1}^{n} \ell_i^{(1)}(x)$$

3. We construct then the polynomial function:

$$q^{(1)}(x) = p^{(1)}(x)\ell^{(1)}_{n+1}(x) = \prod_{i=1}^{n+1} \ell^{(1)}_i(x).$$

4. We construct a new feasible set $\Lambda'^{(1)}_+(q, e)$ using theorem by Sanyal.

5. Apply interior point method over the new feasible set $\Lambda'^{(1)}_+(q, e)$ to find new lower bound ϵ_1 and the minimizer $x^{*(1)}$

6. Update the lower bound $\epsilon_0 \leftarrow \epsilon_1$ and if $|\epsilon_0 - \epsilon_1|$ is smaller than some tolerance, we return $x^{*(1)}$, otherwise we go back to step 1.

3.4 The case of semidefinite programming

Unfortunately we do not have such kind of theorem proposed by Sanyal for the case of semidefinite programming. It means that we will need to find determinantal representation of the hyperbolic polynomial, but also we are able to take derivative and our algorithm will be useful in this case. But sometimes we know the definite determinantal representation as will be shown in example of the disk.

So algorithm of approximation will be a little different from previously proposed for the linear case and we propose a new algorithm as follows:

1. We find a definite determinantal representation of the polynomial p:

$$p = \det(A)$$

where $A = x_1 A_1 + \cdots + x_n A_n$ is a symmetric linear matrix polynomial and e is such that $A(e) \succ 0$.

2. We construct the boundary constraint $\ell^{(1)}(x) = c^T x - \epsilon_0 e^T x$, where ϵ_0 is a known lower bound;

3. We construct then the polynomial function:

$$q(x) = D_e(p\ell)$$

and find a definite determinantal representation of the polynomial q as follows:

$$q^{(1)}(x) = \det(A^{(1)}) = D_e(p\ell)$$

where $A^{(1)} = x_1 A_1^{(1)} + \cdots + x_n A_n^{(1)}$ is a symmetric linear matrix polynomial and e is such that $A^{(1)}(e) \succ 0$.

4. We construct a new feasible set $\Lambda_+^{(1)}(q,e)$ which is spectrahedron by construction.

5. Apply interior point method over the new feasible set $\Lambda_+^{(1)}(q,e)$ to find new lower bound ϵ_1 and the minimizer $x^{*(1)}$

6. Update the lower bound $\epsilon_0 \leftarrow \epsilon_1$ and if $|\epsilon_0 - \epsilon_1|$ is smaller than some tolerance, we return $x^{*(1)}$, otherwise we go back to step 1.

4 Numerical examples

In this section we are going to approximate the solution in the following examples of polyhedron and disk. Also we propose some ways of approximation by modifications of the direction e.

4.1 Example of polyhedron

Example 8. Let $p = (z+x)(z+y)(z-x)(z-y) \in \mathbb{R}[x,y,z]_4$. Note that p is hyperbolic in the direction of z, that is, in direction $e = (0,0,1)$. The picture of its hyperbolicity cone is the green square Figure 4.1, seen in the hyperplane $L = \{(x,y,z) \in \mathbb{R}^3 : z = 1\}$.

The hyperbolic problem is the following:

$$\begin{gathered} c^* = \min c^T x \\ \text{s.t. } e^T x = 1 \\ x \in \Lambda_+(p,e) \end{gathered}$$

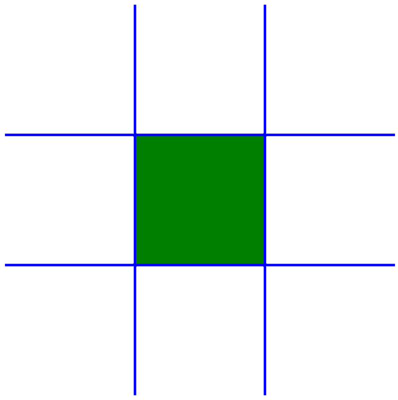

Figure 4.1: The feasible set of the original hyperbolic program

where $x \mapsto c^T x$ is a linear functional with $c = (1, 2, 0)$ and $e^T x = 1$ is a plane L.

We assume that we know the lower bound $\epsilon_0 = -5$

Solution. We are going to approximate the minimum of given hyperbolic program in 2 ways:

1. Without changing of direction $e = (0, 0, 1) = const$.
2. Updating the direction $e^{(i)}$ at each step. We take some $0 \leq \lambda^{(i)} \leq 1$ and updating direction as follows:

$$e^{(i+1)} = \lambda^{(i)} e^{(i)} + (1 - \lambda^{(i)}) x^{*(i)},$$

where $x^{*(i)}$ is the minimizer computed in step i.

1. Approximation of the minimum of the hyperbolic programming without changing of direction.

Taking $\ell_1 = z + x$, $\ell_2 = z + y$, $\ell_3 = z - x$, $\ell_4 = z - y$ we have $p = \ell_1 \ell_2 \ell_3 \ell_4$.

We construct the hyperbolic polynomial $\ell_5 = c^T x - \epsilon_0 e^T x$ and this polynomial is positive definite over the feasible set.

Now we construct the polynomial $q = \ell_5 p$. The polynomial q is still hyperbolic with respect to $e = (0, 0, 1)$.

Then we normalize linear constraints as follows:

$$\ell_1^{(1)}(x) = z + x;$$

$$\ell_2^{(1)}(x) = z + y;$$

$$\ell_3^{(1)}(x) = z - x;$$

$$\ell_4^{(1)}(x) = z - y;$$

$$\ell_5^{(1)}(x) = 0.2(x + 2y + 5z);$$

So we obtain normalized constructed hyperbolic polynomial:

$$q^{(1)} = 0.2 \prod_{i=1}^{5} \ell_i^{(1)}.$$

We construct a new feasible set $\Lambda_+'^{(1)}(q^{(1)}, e)$ applying the theorem of Sanyal, where $D_e q^{(1)} = x^2y^2 - 0.4(x+2y)(x^2+y^2)z - 3(x^2+y^2)z^2 + 0.8(x+2y)z^3 + 5z^4$.

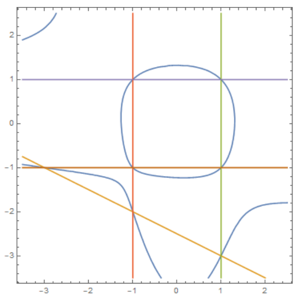

Figure 4.2: The feasible set of the original program and the first modified relaxation

To compute the lower bound for this cone we need to represent this function in the form of SDP to find minimizer. Since we are looking for minimizer in L where $z = 1$ we need to represent our cone in the form $\Lambda_+'^{(1)}(q^{(1)}, e) = zA_0^{(1)} + xA_1^{(1)} + yA_2^{(1)}$:

$$A_0 = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix} \quad A_1 = \begin{pmatrix} 1.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & -0.8 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 \end{pmatrix} \quad A_2 = \begin{pmatrix} 0.4 & 0.4 & 0.4 & 0.4 \\ 0.4 & 1.4 & 0.4 & 0.4 \\ 0.4 & 0.4 & 0.4 & 0.4 \\ 0.4 & 0.4 & 0.4 & -0.6 \end{pmatrix}$$

Now we obtained next SDP optimization problem:

$$\begin{gathered} \min x + 2y \\ \text{s.t. } zA_0^{(1)} + xA_1^{(1)} + yA_2^{(1)} \succeq 0 \\ (x, y, z) \in \mathbb{R}^3 \\ z = 1 \end{gathered}$$

Using SeDuMi in MATLAB we obtain new minimizer $x^{*(1)} = (-0.884487, -1.08177, 1)$ and a new lower bound -3.04803.

Repeating the algorithm we obtain next table of minimizers and lower bounds:

N	x	y	z	Lower bound
1	-0.884487	-1.08177	1	-3.04803
2	-0.993911	-1.00431	1	-3.00253
3	-0.999675	-1.00024	1	-3.00015
4	-0.999981	-1.00001	1	-3.00001
5	-0.999999	-1	1	-3
6	-1	-1	1	-3
7	-1	-1	1	-3

As we can see Algorithm approximated the minimum of the hyperbolic programming in 7 repetitions.

2. Approximation of the minimum of the hyperbolic programming with changing of the direction $e^{(i)}$

1st step is the same with previous case, because we didn't obtain minimizer $x^{*(i)}$. So, we start from step 2 since we have already computed the minimizer $x^{*(1)} = (-0.884487, -1.08177, 1)$.

Let $\lambda^{(i)} = 0.5^i$, then we obtain updated direction $e^{(2)}$:

$$e^{(2)} = \lambda^{(1)} e^{(1)} + (1 - \lambda^{(1)}) x^{*(1)} = (-0.442244, -0.540885, 1),$$

Then we normalize linear constraints as follows:

$$\ell_1^{(2)}(x) = 1.7929(z + x);$$

$$\ell_2^{(2)}(x) = 2.1781(z + y);$$

$$\ell_3^{(2)}(x) = 0.693364(z - x);$$

$$\ell_4^{(2)}(x) = 0.624102(z - y);$$

$$\ell_5^{(2)}(x) = -0.115533x + 0.116662y + 1.01201z;$$

So we obtain normalized constructed hyperbolic polynomial:

$$q^{(2)} = 1.75721 \prod_{i=1}^{5} \ell_i^{(2)}.$$

We construct a new feasible set $\Lambda_+^{\prime(2)}(q^{(2)}, e)$ applying the theorem of Sanyal, where $D_e q^{(2)} = 0.219617x^3y + 1.71502x^2y^2 - 0.18132xy^3 + 0.406033x^3z - 2.33373x^2yz - 1.16686xy^2z - 0.410001y^3z - 5.49341x^2z^2 - 0.0382966xyz^2 - 5.09208y^2z^2 + 0.76083xz^3 + 2.74373yz^3 + 8.87047z^4$.

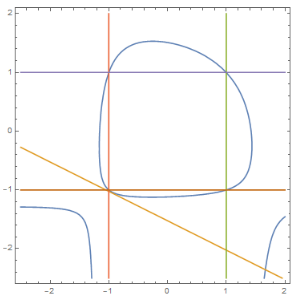

Figure 4.3: The feasible set of the original program and of the first modified relaxation

To compute the lower bound for this cone we need to represent this function in the form of SDP to find minimizer. Since we are looking for minimizer in L where $z = 1$ we need to represent our cone in the form $\Lambda'^{(2)}_{+}(q^{(2)}, e) = zA_0^{(1)} + xA_1^{(1)} + yA_2^{(1)}$:

$$A_0 = \begin{pmatrix} 2.8049 & 1.01201 & 1.01201 & 1.01201 \\ 1.01201 & 3.19011 & 1.01201 & 1.01201 \\ 1.01201 & 1.01201 & 1.70537 & 1.01201 \\ 1.01201 & 1.01201 & 1.01201 & 1.66098 \end{pmatrix}$$

$$A_1 = \begin{pmatrix} 1.67736 & -0.115533 & -0.115533 & -0.115533 \\ -0.115533 & -0.115533 & -0.115533 & -0.115533 \\ -0.115533 & -0.115533 & -0.808897 & -0.115533 \\ -0.115533 & -0.115533 & -0.115533 & -0.115533 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} 0.116662 & 0.116662 & 0.116662 & 0.116662 \\ 0.116662 & 2.29476 & 0.116662 & 0.116662 \\ 0.116662 & 0.116662 & 0.116662 & 0.116662 \\ 0.116662 & 0.116662 & 0.116662 & -0.532316 \end{pmatrix}$$

Now we obtained next SDP optimization problem:

$$\begin{gathered} \min x + 2y \\ \text{s.t. } zA_0^{(1)} + xA_1^{(1)} + yA_2^{(1)} \succeq 0 \\ (x, y, z) \in \mathbb{R}^3 \\ z = 1 \end{gathered}$$

Using SeDuMi in MATLAB we obtain new minimizer $x^{*(2)} = (-0.937232, -1.04032, 1)$ and a new lower bound -3.01787.

Repeating the algorithm we obtain next table of minimizers and lower bounds:

N	x	y	z	Lower bound
1	-0.884487	-1.08177	1	-3.04803
2	-0.937232	-1.04032	1	-3.01787
3	-1.0032	-0.998478	1	-3.00015
4	-0.999472	-1.00029	1	-3.00004
6	-1.00016	-0.999941	1	-3.00004

As we can see this method allows us only approximate the minimum and some times not that good as we want

Note, that we can change the direction e but we need to be careful because if e is not inside the original cone then the polynomial p is not hyperbolic polynomial.

4.2 Example of disc

Example 9 (Plane circle, as in Example 3)**.** As already remarked $p = z^2 - x^2 - y^2 \in \mathbb{R}[x, y, z]_2$ is hyperbolic with respect to the direction $e = (0, 0, 1)$. Moreover, coherently with Helton-Vinnikov theorem [3], p has the following determinantal representation:

$$p = \det \begin{bmatrix} z - x & y \\ y & z + x \end{bmatrix}$$

The picture of its hyperbolicity cone is the green square Figure 4.4, seen in the hyperplane $L = \{(x, y, z) \in \mathbb{R}^3 : z = 1\}$.

We assume that we know the lower bound $\epsilon_0 = -2$

The hyperbolic problem is the following:

$$c^* = \min c^T x$$
$$\text{s.t. } e^T x = 1$$
$$x \in \Lambda_+(p, e)$$

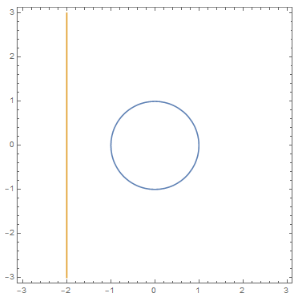

Figure 4.4: The feasible set of the original program and given lower bound $\epsilon_0 = -2$

where $x \mapsto c^T x$ is a linear functional with $c = (1, 0, 0)$ and $e^T x = 1$ is a plane L.

Solution. The hyperbolic polynomial p is given by

$$p = \det \begin{bmatrix} z - x & y \\ y & z + x \end{bmatrix} = z^2 - x^2 - y^2.$$

Now we construct the polynomial $q^{(1)} = D_e(p\ell^{(1)})$ and obtain a definite determinantal representation:

$$\begin{aligned} q^{(1)} &= D_e(p\ell^{(1)}) \\ &= \ell^{(1)} D_e(p) + p D_e(\ell^{(1)}) \\ &= (x - \epsilon_0 z) D_e(z^2 - x^2 - y^2) + (z^2 - x^2 - y^2) D_e(x - \epsilon_0 z) \\ &= 2xz - 2\epsilon_0 z^2 - \epsilon_0 z^2 + \epsilon_0 x^2 + \epsilon_0 y^2 \\ &= 2xz - 3\epsilon_0 z^2 + \epsilon_0 x^2 + \epsilon_0 y^2 \\ &= \det \begin{bmatrix} (x - \frac{z}{\epsilon_0}(-1 + \sqrt{1 + 3\epsilon_0^2})) & \sqrt{-\epsilon_0} y \\ \sqrt{-\epsilon_0} y & \epsilon_0 (x - \frac{z}{\epsilon_0}(-1 - \sqrt{1 + 3\epsilon_0^2})) \end{bmatrix}. \end{aligned}$$

Remind that by assumption the lower bound $\epsilon_0 < 0$ and we also can deduce the following statement for $\epsilon_k \leq c^* < 0$ (as it was done with ϵ_0):

$$\begin{aligned} q^{(k)} &= -3\epsilon_k z^2 + 2xz + \epsilon_k x^2 + \epsilon_k y^2 \\ &= \det \begin{bmatrix} (x - \frac{z}{\epsilon_k}(-1 + \sqrt{1 + 3\epsilon_k^2})) & \sqrt{-\epsilon_k} y \\ \sqrt{-\epsilon_k} y & \epsilon_k (x - \frac{z}{\epsilon_k}(-1 - \sqrt{1 + 3\epsilon_k^2})) \end{bmatrix}. \end{aligned}$$

The polynomial $p\ell^{(1)}$ is hyperbolic with respect to e and hence $q^{(1)} = D_e(p\ell^{(1)})$ is hyperbolic with respect to e as well.

Next we construct a new feasible set $\Lambda_+^{(1)}(q^{(1)}, e)$ using obtained definite determinantal representation $q^{(1)}$.

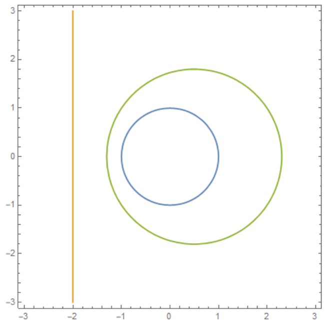

Figure 4.5: The feasible set of the original program, given lower bound $\epsilon_0 = -2$ and the first modified relaxation

To compute the lower bound of this relaxation of the cone we need to represent this function in the form of SDP to find minimizer. Since we are looking for minimizer in L where $z = 1$ we need to represent our cone in the form $\Lambda_+^{(1)}(q^{(1)}, e) = zA_0^{(1)} + xA_1^{(1)} + yA_2^{(1)}$:

$$A_0^{(1)} = \begin{pmatrix} 1.30278 & 00 & 4.60555 \end{pmatrix} \quad A_1^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix} \quad A_2^{(1)} = \begin{pmatrix} 0 & 1.41421 \\ 1.41421 & 0 \end{pmatrix}$$

Now we obtained next SDP optimization problem:

$$\begin{gathered} \min x \\ \text{s.t. } zA_0^{(1)} + xA_1^{(1)} + yA_2^{(1)} \succeq 0 \\ (x, y, z) \in \mathbb{R}^3 \\ z = 1 \end{gathered}$$

Using SeDuMi in MATLAB we obtain new minimizer $x^{*(1)} = (-1.30278, 0, 1)$ and a new lower bound -1.30278.

Note, that since objective function $c^T x = x$ it is clear that solution could be obtained as the smallest root of the equation $q^{(1)}(x, 0, 1) = 0$, or $6 + 2x - 2x^2 = 0$.

Repeating the algorithm we obtain next table of minimizers and lower bounds:

N	x	y	z	Lower bound
1	-1.30278	0	1	-1.30278
2	-1.12693	0	1	-1.12693
3	-1.05876	0	1	-1.05876
4	-1.02834	0	1	-1.02834
5	-1.01392	0	1	-1.01392
6	-1.0069	0	1	-1.0069
7	-1.00344	0	1	-1.00344
8	-1.00171	0	1	-1.00171
9	-1.00086	0	1	-1.00086
10	-1.00043	0	1	-1.00043
11	-1.00021	0	1	-1.00021
12	-1.00011	0	1	-1.00011
13	-1.00005	0	1	-1.00005
14	-1.00003	0	1	-1.00003
15	-1.00001	0	1	-1.00001
16	-1.00001	0	1	-1.00001
17	-1	0	1	-1
18	-1	0	1	-1

So we can approximate the solution of this hyperbolic programming.

Remark 3. As we noticed during solution for this hyperbolic programming we can compute minimizers as the smallest roots of of the polynomials $q^{(k)}(x, 0, 1)$ for all k.

Since the smallest root of the equation $q^{(k)}(x, 0, 1) = 0$ we obtain a sequence of lower bounds.

We want to prove that this sequence converges to the minimum for this hyperbolic programming.

Proof. Roots of the equation $q^{(k)}(x,0,1) = 0$ or $-3\epsilon_k + 2x + \epsilon_k x^2 = 0$ are following:

$$q_{1,2}^{(k)}(x,0,1) = \frac{-1 \pm \sqrt{1+3\epsilon_k^2}}{\epsilon_k}$$

Since we assume that $\epsilon_k \leq c^* < 0$, then $q_{min}^{(k)}(x,0,1) = \frac{-1+\sqrt{1+3\epsilon_k^2}}{\epsilon_k}$ because this root is negative and another root is positive. So, we obtain a sequence of ϵ_k:

$$q_{min}^{(k)}(x,0,1) = \epsilon_{k+1} = \frac{-1 + \sqrt{1+3\epsilon_k^2}}{\epsilon_k}$$

Let's check 3 cases of behavior of the ϵ_k:

1. If $\epsilon_{k+1} = \epsilon_k = -\infty$ when $k \to \infty$.
2. If $\epsilon_k = \epsilon_{k+1} = v$.

1. If $\epsilon_k = -\infty$ then we obtain:

$$\lim_{k\to\infty}\epsilon_{k+1} = \lim_{k\to\infty}\left(\frac{-1}{\epsilon_k} + \frac{\sqrt{1+3\epsilon_k^2}}{\epsilon_k}\right)$$
$$= \lim_{k\to\infty}\left(\frac{-1}{\epsilon_k} + \sqrt{\frac{1}{\epsilon_k^2} + 3\frac{\epsilon_k^2}{\epsilon_k^2}}\right)$$
$$= \sqrt{3} \neq -\infty.$$

So the first statement is false and this sequence is not divergent;

2. If $\epsilon_k = \epsilon_{k+1} = v$ where $v \leq c^* < 0$. Then we obtain:

$$v = \frac{-1}{v} + \frac{\sqrt{1+3v^2}}{v}$$
$$\Leftrightarrow \frac{v^2+1}{v} = \frac{\sqrt{1+3v^2}}{v}$$
$$\Leftrightarrow \frac{v^4+2v^2+1}{v^2} = \frac{1+3v^2}{v^2}$$
$$\Leftrightarrow \frac{v^4-v^2}{v^2} = 0$$

$$\Leftrightarrow \frac{v^2(v-1)(v+1)}{v^2} = 0$$
$$\Rightarrow v = -1 \text{ (by assumption that } v \leq c^* < 0).$$

So we obtained that if this sequence converge to some value then it converge to the optimal solution of previous hyperbolic programming. □

5 Software implementation

Working with computation on this work we used next software:

- Wolfram Mathematica 11.1;
- MATLAB R2018b.

During computation there we needed to solve an SDP over the obtained relaxations of the feasible sets. It was done using SeDuMi, special package in MATLAB which can solve an SDP. That's how minimizers were obtained.

Unfortunately in using Wolfram Mathematica there is no possibility to solve an SDP numerically as MATLAB does. We used SeDuMi, special MATLAB package which allows us to solve optimization problems over symmetric cones.

But the drawback of MATLAB is that we can apply only numerical algorithms without possibility to work with symbolic computations.

So combination those software allowed us to work with all the algorithms what we wanted to use and helped to compute results of proposed algorithms.

6 Conclusions and open questions

In this thesis we were studying hyperbolic polynomials in context of optimization problems, provided approach to regularize the hyperbolicity cone with such kind of modification that we are able to keep relation between the regularized cone and objective function. With this we proposed two algorithms for LP and SDP in context of hyperbolic programming, which seem to converge to the minimum, but since it was not proved we can say that it is only approximate to optimal solution.

According to proposed method of relaxations we provided small proof of the fact that regularized cone becomes smoother and in regularized cone singularities are reduced.

Also we showed how to apply algorithms in practical numerical experiments for polyhedron, according to results in the table, was approximating minimizers quite close to the optimal solution just in few cycles of the algorithm.

Working on numerical computation we noticed that sometimes we can achieve the minimum after one cycle of the algorithms, for instance, in the case of disk, if we apply our method in direction of the optimal solution or in direction $e = (0.5, 0, 1)$ with given lower bound $\epsilon_0 = -2$ then we obtain an optimal solution of the hyperbolic programming, so it gives an idea that there is a relation between known lower bound and hyperbolic direction and maybe it's valuable to study this question.

One of the main open question for the moment it is if such kind of modified relaxation allows algorithms converge to the optimal solution, since it was not proved.

However, in some specific cases, for instance in the example of the disk, proving this fact can be easier according to symmetry. But proving convergence in general case becomes complicated.

References

[1] P. Brändén. Obstructions to determinantal representability. *Adv. Math.*, 226(2):1202–1212, 2011.

[2] L. Gårding. An inequality for hyperbolic polynomials. *J. Math. Mech.*, 8(6):957–965, 1959.

[3] J. Helton and V. Vinnikov. Linear matrix inequality representation of sets. *Communications on Pure and Applied Mathematics*, 60(5):654–674, 2007.

[4] P. Lax. Differential equations, difference equations and matrix theory. *Communications on pure and applied mathematics*, 11(2):175–194, 1958.

[5] A. Lewis, P. Parrilo, and M. Ramana. The Lax conjecture is true. *Proceedings of the American Mathematical Society*, 133(9):2495–2499, 2005.

[6] S. Naldi and D. Plaumann. Symbolic computation in hyperbolic programming. *J. Alg. Appl.*, 17(10):1850192, 2018.

[7] J. Renegar. *A mathematical view of interior-point methods in convex optimization*. SIAM, 2001.

[8] J. Renegar. Hyperbolic programs, and their derivative relaxations. *Found. Comput. Math.*, 6(1):59–79, 2006.

[9] R. Sanyal. On the derivative cones of polyhedral cones. *Advances in Geometry*, 13(2):315–321, 2013.