Hyperbolic polynomials and majorization

M. Seetharama Gowda
Department of Mathematics and Statistics
University of Maryland, Baltimore County
Baltimore, Maryland 21250, USA
gowda@umbc.edu

Material presented in: UMBC online seminar, Oct. 7 and Oct. 14, 2021

Abstract

On a finite dimensional real vector space $\mathcal{V}$, we consider a real homogeneous polynomial p of degree n that is hyperbolic relative to a vector $e \in \mathcal{V}$. This means that $p(e) \neq 0$ and for any (fixed) $x \in \mathcal{V}$, the roots of the one-variable polynomial $t \mapsto p(te - x)$ are all real. Let $\lambda(x)$ denote the vector in $\mathcal{R}^n$ whose entries are these real roots written in the decreasing order. Relative to the map $\lambda : \mathcal{V} \to \mathcal{R}^n$, we introduce and study automorphisms, majorization, and doubly stochastic transformations.

Key Words:
AMS Subject Classification:

1 Introduction

Hyperbolic polynomials were introduced by Gårding [5] in 1951 in connection with partial differential equations. Since then, they have become important in various areas such as convex analysis, optimization, algebraic geometry, etc. In this talk, we focus on linear algebraic concepts such as automorphisms, majorization and doubly stochastic transformations induced by hyperbolic polynomials.

First, we recall the concept of majorization in $\mathcal{R}^n$ (where vectors are regarded as row or column vectors depending on the context).
For a vector $x = (x_1, x_2, \ldots, x_n)$ in $\mathcal{R}^n$, let $x^\downarrow$ denote the vector obtained from x by rearranging its entries in the decreasing order. Note that $x \geq 0$ in $\mathcal{R}^n$ if and only if $x^\downarrow \geq 0$, so

$$\mathcal{R}^n_+ = \{x \in \mathcal{R}^n : x^\downarrow \geq 0\}.$$

Additionally, $x^{\downarrow} = Px$ for some permutation matrix P (which has exactly one 1 in each row/column and zeros elsewhere). We define, for any $z \in \mathcal{R}^n$, the orbit of z:

$$[z] := \{x \in \mathcal{R}^n : x^{\downarrow} = z^{\downarrow}\}.$$

(For example, if $e_1, e_2, \ldots, e_n$ are the standard coordinate vectors, then $[e_1] = \{e_1, e_2, \ldots, e_n\}$.) Then, a real square matrix A preserves every $[z]$, that is, $(Az)^{\downarrow} = z^{\downarrow}$ for all $z \in \mathcal{R}^n$ if and only if A is a permutation matrix. (To see 'only if' part, observe that A is invertible and $(Ae_k)^{\downarrow} = e_1$ for all k.) Thus, permutation matrices are linear orbit-preservers.

Given $u, v \in \mathcal{R}^n$, we say that u is *majorized* by v and write $u \prec v$ if for all natural numbers k, $1 \leq k \leq n$,

$$\sum_{i=1}^{k} u_i^{\downarrow} \leq \sum_{i=1}^{k} v_i^{\downarrow}$$

with equality for $k = n$. *A result of Hardy, Littlewood, and Pólya says that $u \prec v$ if and only if $u = Dv$, where D is a $n \times n$ doubly stochastic matrix* (i.e., a nonnegative matrix where each row/column sum is one). Additionally, *a result of Birkhoff says that every doubly stochastic matrix D is a convex combination of permutation matrices.*
So, in the setting of $\mathcal{R}^n$,

$$u \prec v \text{ if and only if } u \in conv[v]$$

where 'conv' represents the convex hull.

2 Motivating examples

To motivate hyperbolic polynomials and several concepts that we plan to introduce and study, we begin with some examples.

Example 1: Writing any element of $\mathcal{R}^n$ as $x = (x_1, x_2, \ldots, x_n)$, we consider the homogeneous polynomial of degree n:

$$p(x) = x_1 x_2 \cdots x_n.$$

With $e = (1, 1, \ldots, 1)$, fixing $x \in \mathcal{R}^n$, we consider the one-variable polynomial $t \mapsto p(te - x)$, where

$$p(te - x) = (t - x_1)(t - x_2) \cdots (t - x_n).$$

Clearly, the roots of this polynomial are: $t = x_1, x_2, \ldots, x_n$. From these roots, we form the vector $\lambda(x)$ in $\mathcal{R}^n$ with decreasing components. Note that

$$\lambda(x) = x^{\downarrow}.$$

We regard λ as a map from $\mathcal{R}^n$ to $\mathcal{R}^n$ and define the λ-orbit of any $z \in \mathcal{R}^n$:

$$[z] := \{x \in \mathcal{R}^n : \lambda(x) = \lambda(z)\}.$$

From our observations above,

- $\mathcal{R}^n_+ = \{x \in \mathcal{R}^n : \lambda(x) \geq 0\}$,
- A matrix $A \in \mathcal{R}^{n\times n}$ is a linear preserver of every λ-orbit if and only if A is a permutation matrix,
- $u \in \mathrm{conv}\,[v]$ if and only if $\lambda(u) \prec \lambda(v)$.

Example 2: Now let $\mathcal{V}$ denote either the space $\mathcal{S}^n$ of all $n \times n$ real symmetric matrices or the space $\mathcal{H}^n$ of all $n \times n$ complex Hermitian matrices. For any $X \in \mathcal{V}$, let

$$p(X) := \det(X)$$

which is a homogeneous polynomial of degree n on $\mathcal{V}$. Let e denote the identity matrix I in $\mathcal{V}$; note $p(I) \neq 0$. Then, for any (fixed) $X \in \mathcal{V}$, the roots of the one-variable polynomial

$$t \mapsto p(te - X) = \det(tI - X)$$

are all real, as they are the eigenvalues of X. With these roots (=eigenvalues), we form the vector $\lambda(X)$ in $\mathcal{R}^n$ with decreasing components, thus getting the eigenvalue map $\lambda : \mathcal{V} \to \mathcal{R}^n$. Observe that the set of all positive semidefinite matrices in $\mathcal{V}$ can be expressed as

$$\text{the semidefinite cone} = \{X \in \mathcal{V} : \lambda(X) \geq 0\}.$$

For any $Z \in \mathcal{V}$, we define the λ-orbit of Z:

$$[Z] := \{X \in \mathcal{V} : \lambda(X) = \lambda(Z)\}.$$

We know that $X \in [Z]$ if and only if $X = UZU^*$ for some unitary matrix (orthogonal on the case of $\mathcal{S}^n$). Now consider a linear transformation L on $\mathcal{V}$ that preserves every λ-orbit. Such a transformation is necessarily invertible (and the inverse also preserves every λ-orbit). Also, it and its inverse preserve the so-called semidefinite cone in $\mathcal{V}$. Thus, by a result of Schneider [19], the transformation must be of the form $L(X) = AXA^*$ for some nonsingular matrix A. Since $[I] = I$, we must have $L(I) = I$ implying $AA^* = I$. Hence, L must be given by $L(X) := UXU^*$ for some (fixed) unitary/orthogonal matrix U. Now, one can introduce majorization in $\mathcal{V}$ by saying $X \prec Y$ if and only if $\lambda(X) \prec \lambda(Y)$ in $\mathcal{R}^n$. By a result of Ando [1], this can happen if and only if X can be expressed a convex combination of matrices of the form UYU^* with U unitary/orthogonal. This means that in $\mathcal{V}$, majorization $X \prec Y$ can be defined by

$$X \in \mathrm{conv}\,[Y].$$

Example 3: Consider $\mathcal{R}^n$, where we write any vector in the form $x = (x_0, \bar{x})$ with $x_0 \in \mathcal{R}$ and $\bar{x} = (x_1, x_2, \ldots, x_{n-1}) \in \mathcal{R}^{n-1}$. Then

$$p(x) = x_0^2 - x_1^2 - x_2^2 - \cdots - x_{n-1}^2$$

is a homogeneous polynomial of degree 2. Let $e = (1, 0, 0, \ldots, 0)$. Then $p(e) \neq 0$ and, for any (fixed) $x \in \mathcal{R}^n$, the one-variable polynomial

$$p(te - x) = (t - x_0)^2 - x_1^2 - x_2^2 \cdots - x_{n-1}^2$$

has real roots, namely,

$$x_0 + ||\bar{x}|| \quad \text{and} \quad x_0 - ||\bar{x}||,$$

where $||\bar{x}||$ denotes the (usual) norm of $\bar{x}$ in $\mathcal{R}^{n-1}$. With these roots, we form the vector $\lambda(x)$ in $\mathcal{R}^2$ with decreasing components. Note that

$$\{x = (x_0, \bar{x}) : \lambda(x) \geq 0\} = \{x : x_0 \geq ||\bar{x}||\}$$

is the second-order cone (=Lorentz cone =ice-cream cone). We define the λ-orbit of $z \in \mathcal{R}^n$ by

$$[z] = \{x \in \mathcal{R}^n : \lambda(x) = \lambda(z)\}$$

and observe $A \in \mathcal{R}^{n \times n}$ is a preserver of all λ-orbits if and only if

$$A = \begin{bmatrix} 1 & 0 \\ 0 & U \end{bmatrix},$$

where U is an orthogonal matrix on $\mathcal{R}^{n-1}$. Now, one can introduce majorization in this setting by saying $x \prec y$ if and only if $\lambda(x) \prec \lambda(y)$ in $\mathcal{R}^2$. By a result of Gowda [7], this can happen if and only if x can be expressed a convex combination of Ay, where A is matrix of the above form. This means that majorization $x \prec y$ can be defined by

$$x \in \text{conv}\,[y].$$

Note that in each of the above examples, the starting point was a homogeneous polynomial having the 'real-roots' property. Such polynomials are called *hyperbolic.* Of course, the three examples above are familiar to people who study Euclidean Jordan algebras. The emphasis here is that the examples/results could be explained entirely in terms of hyperbolic polynomials.

3 Hyperbolic polynomials – Definition and more examples

Definition 3.1 ([5, 6]) *Let $\mathcal{V}$ be a finite dimensional real vector space. (Note: We do not assume any other structure on $\mathcal{V}$.) Let p be a real homogeneous polynomial of degree n on $\mathcal{V}$ (relative to some basis of $\mathcal{V}$) and $e \in \mathcal{V}$. Then, p is said to be* **hyperbolic with respect to the vector** *e if*

$p(e) \neq 0$ and for any (fixed) $x \in \mathcal{V}$, the roots of the univariate polynomial $t \mapsto p(te - x)$ from $\mathcal{R}$ to $\mathcal{R}$ are all **real**.

For any $x \in \mathcal{V}$, let $\lambda(x)$ denote the vector in $\mathcal{R}^n$ whose components are the roots of $p(te-x)$ written in the decreasing order. We call the map $\lambda : \mathcal{V} \to \mathcal{R}^n$, the *eigenvalue map* of p. Note that

$$p(te - x) = p(e)\big(t - \lambda_1(x)\big)\big(t - \lambda_2(x)\big) \cdots \big(t - \lambda_n(x)\big)$$

(showing why we need $p(e) \neq 0$). Also, for any $x \in \mathcal{V}$ and $t \in \mathcal{R}$,

$$\lambda(te + x) = t(1, 1, \ldots, 1) + \lambda(x).$$

When $p(e) = 1$, we have

$$p(x) = \lambda_1(x)\lambda_2(x) \cdots \lambda_n(x) \quad (x \in \mathcal{V}).$$

Given a hyperbolic polynomial p on $\mathcal{V}$ (relative to some e) and the induced λ, we call the triple $(\mathcal{V}, p, \lambda)$, *a hyperbolic triple.* In a hyperbolic triple, the λ-orbit of any z is defined by

$$[z] = \{x \in \mathcal{V} : \lambda(x) = \lambda(z)\}.$$

A set E in $\mathcal{V}$ is a *spectral set* if it is a union of λ-orbits, or equivalently, of the form $\lambda^{-1}(Q)$ for some set $Q \subseteq \mathcal{R}^n$.

Here are some properties of λ.

Proposition 3.2 [3] In a hyperbolic triple, the following hold:

- (i) *λ is positively homogeneous, Lipschitz, and continuous (relative to any norm on $\mathcal{V}$).*
- (ii) *For any $w \in \mathcal{R}^n$ with decreasing components, $w^T\lambda(x)$ is sublinear in x, that is,*

$$w^T\lambda(x + y) \leq w^T\lambda(x) + w^T\lambda(y).$$

In particular, for every k, $1 \leq k \leq n$, $\sigma_k(x) := \sum_{i=1}^{k} \lambda_i(x)$ is sublinear in x.

We note, by continuity of λ, that every λ-orbit is closed in $\mathcal{V}$ (with respect to any norm).

A basic result of Gårding [6] is the following:

Proposition 3.3 [6] *Consider a hyperbolic triple $(\mathcal{V}, p, \lambda)$. Then, the set*

$$K := \{x \in \mathcal{V} : \lambda(x) \geq 0\}$$

is a convex cone.

Because of continuity of λ, we see that K is a *closed convex cone with interior* (relative to any normtopology on $\mathcal{V}$). It is called the **hyperbolicity cone** corresponding to p, or simply a *hyperbolic cone.* It is clear from the examples given earlier, the nonnegative orthant, the semidefinite cone (set of all positive semidefinite matrices in $\mathcal{S}^n$ and $\mathcal{H}^n$), and second-order cone are hyperbolic cones. More generally, all symmetric cones are hyperbolic (see below). Guler [11] has shown that all *homogeneous cones are hyperbolic.* Conic optimization on a hyperbolic cone is called *hyperbolic programming.*

Note: In Gårding [6] and other sources, the hyperbolicity cone is defined as the interior of (above cone) K. It is known, see Fact 2.7 in [3], that p is hyperbolic with respect to every c in the interior of K.

Example 4: Let $\mathcal{V}$ be a Euclidean Jordan algebra of rank n [4] with inner product $\langle x, y\rangle$ and Jordan product $x \circ y$. According to the spectral theorem [4], every $x \in \mathcal{V}$ has a decomposition

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n,$$

where $\{e_1, e_2, \ldots, e_n\}$ is a Jordan frame (these elements are primitive idempotents that are mutually orthogonal) and $x_1, x_2, \ldots, x_n$ are the eigenvalues (which depend only on x and not on the Jordan frame). Let $\lambda(x)$ denote the vector of these eigenvalues written in the decreasing order. Then, by definition,

$$tr(x) := \lambda_1(x) + \lambda_2(x) + \cdots + \lambda_n(x) \quad \text{and} \quad \det(x) = \lambda_1(x)\lambda_2(x)\cdots\lambda_n(x).$$

It is known that $(x, y) \mapsto tr(x \circ y)$ defines another inner product on $\mathcal{V}$, called the *trace inner product*, that is compatible with the Jordan product. Henceforth, we assume that $\mathcal{V}$ carries this inner product.

Clearly, $\det(x)$ is homogeneous of degree n. It is known, see [4], Prop. II.2.1 and pages 28-29, that there is a polynomial $a_n(x)$ (homogeneous of degree n) which coincides with $\det(x)$ for all regular (invertible) elements in $\mathcal{V}$. As both $a_n(x)$ and $\det(x)$ are continuous in x, by density of invertible elements we see that the polynomial $a_n(x)$ coincides with $\det(x)$ for all x.

That det is a polynomial can also be seen, perhaps in a more transparent but elaborate way by considering each of the five simple Euclidean Jordan algebras and then quoting the structure theorem that every Euclidean Jordan algebra is a product of of these simple algebras.

- In the case of $\mathcal{S}^n$ or $\mathcal{H}^n$, the determinant of a real/complex Hermitian matrix $A = [a_{ij}]$ is the usual determinant, hence a sum of terms of the form $a_{1i_1}a_{2i_2}\cdots a_{ni_n}$. We can express this as a real homogeneous polynomial.
- In the case of $\mathcal{Q}^n$ (the space of all $n \times n$ quaternion Hermitian matrices), there is the concept of Moore determinant [18]. It is known (see Section 5 and the proof of Theorem 10 in [2]) that this can be expressed in terms of entries of the matrix under consideration (similar to

the previous case, but more delicately defined as quaternions do not commute) and that this determinant is the product of eigenvalues. Since every quaternion is of the form $a+bi+cj+dk$, where a, b, c, d are real, the determinant given as the product of eigenvalues can be expressed as a (homogeneous) polynomial.

- Consider the case of $\mathcal{O}^3$, the space of all 3×3 octonion Hermitian matrices. If $A \in \mathcal{O}^3$ is given by

$$A = \begin{bmatrix} p & a & b \\ \bar{a} & q & \bar{c} \\ \bar{b} & c & r \end{bmatrix},$$

where p, q, r are real and a, b, c are octonions. Then, see [10],

$$\det(A) = pqr + 2Re(\bar{b}(ac)) - r|a|^2 - q|b|^2 - p|c|^2.$$

As each octonion is vector in R^8, by expressing a, b, c in terms of their real components, we see that $\det(A)$ is a homogeneous polynomial of degree 3.

- Consider the case of Jordan spin algebra $\mathcal{L}^n$ which is identified as $\mathcal{R} \times \mathcal{R}^{n-1}$ (as in Example 3). Here, for any $x = (x_0, \bar{x})$,

$$\det(x) = x_0^2 - ||\bar{x}||^2,$$

see [10]. Clearly, this is a homogeneous polynomial.

Having shown that $p(x) := \det(x)$ is a homogeneous polynomial on a general Euclidean Jordan algebra, we now show that it is hyperbolic with respect to the unit element e (which is the sum of all elements in any Jordan frame). From

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n \quad \text{and} \quad e = e_1 + e_2 + \cdots + e_n,$$

we see that for any $t \in \mathcal{R}$,

$$p(te - x) = \det(te - x) = (t - x_1)(t - x_2) \cdots (t - x_n).$$

Hence the roots of this one-variable polynomial are the eigenvalues of x. As these are all real, we see the hyperbolicity of p.

Now consider the Euclidean Jordan algebra $\mathcal{V}$ with (the hyperbolic polynomial) $p(x) = \det(x)$ and the induced λ. In this setting, the hyperbolicity cone is called the *symmetric cone* of $\mathcal{V}$ and is given by

$$\{x \in \mathcal{V} : \lambda(x) \geq 0\} = \{x \circ x : x \in \mathcal{V}\}.$$

Recalling that $\mathcal{V}$ carries the trace inner product, we have the following:

- An invertible linear transformation $A : \mathcal{V} \to \mathcal{V}$ preserves all λ-orbits if and only if A is an algebra automorphism of $\mathcal{V}$, that is, $A(x \circ y) = Ax \circ Ay$ for all $x, y \in \mathcal{V}$, see [9].
- $x \in \text{conv}\,[y]$ in $\mathcal{V}$ if and only if $\lambda(x) \prec \lambda(y)$ in $\mathcal{R}^n$, see Sections 7 and 8 below, and [9].

Example 5: Let $A_1, A_2, \ldots, A_n$ be real/complex Hermitian matrices of the same size with $A_1 = I$ (Identity matrix). For each $x = (x_1, x_2, \ldots, x_n) \in \mathcal{R}^n$, define,

$$p(x) := \det\Big(x_1 A_1 + x_2 A_2 + \cdots + x_n A_n\Big).$$

Clearly, p is a homogeneous polynomial on $\mathcal{R}^n$. We claim that it is hyperbolic with respect to $e = (1, 0, 0, \ldots, 0)$. Fix $x \in \mathcal{R}^n$. Then, $te - x = (t - x_1, -x_2, -x_3, \ldots, -x_n)$ and so (with $A_1 = I$),

$$p(te-x) = \det\Big((t-x_1)I - (x_2A_2 + x_3A_3 + \cdots + x_nA_n)\Big) = \det\Big(tI - (x_1A_1 + x_2A_2 + \cdots + x_nA_n)\Big).$$

We see that the roots of $p(te - x)$ are precisely the eigenvalues of the Hermitian matrix $x_1A_1 + x_2A_2 + \cdots + x_nA_n$, thus real. Hence the claim.

Remark. The hyperbolicity cone of Example 5 is:

$$\Big\{x = (x_1, x_2, \ldots, x_n) : x_1\, I + x_2 A_2 + \cdots + x_n A_n \succeq 0\Big\},$$

where '$\succeq$' denotes positive semidefiniteness. This is an example of a *spectrahedron* – intersection of an affine linear space with the semidefinite cone (in $\mathcal{S}^n$ or $\mathcal{H}^n$). The concept of spectrahedron was introduced by Ramana and Goldman (1995), see an interesting article by Vinzant in AMS Notices, 61 (2014) 492-494. An inequality of the form

$$A_0 + x_1 A_1 + \cdots + x_n A_n \succeq 0$$

is called a *Linear Matrix Inequality.*

Example 6: Let $A_1, A_2, \ldots, A_n$ be real/complex Hermitian matrices of the same size with $A := A_1 + A_2 + \cdots + A_n$ positive definite. Define $C_i := A^{-\frac{1}{2}} A_i A^{-\frac{1}{2}}$ for all i. For any $x = (x_1, x_2, \ldots, x_n) \in \mathcal{R}^n$, let

$$p(x) := \det(x_1 A_1 + x_2 A_2 + \cdots + x_n A_n).$$

Clearly, p is a homogeneous polynomial in x and for $e = (1, 1, \ldots, 1)$, $p(e) = \det(A) > 0$. Also, for any fixed x and any $t \in \mathcal{R}$,

$$p(te - x) = \det\Big((t - x_1)A_1 + \cdots + (t - x_n)A_n\Big) = \det\Big(tA - (x_1A_1 + x_2A_2 + \cdots + x_nA_n)\Big).$$

Writing $A = A^{\frac{1}{2}} A^{\frac{1}{2}}$, we can simplify this to

$$p(te - x) = \det\Big(tI - (x_1 C_1 + \cdots + x_n C_n)\Big)\ \det(A).$$

We see that the roots of this polynomial are all real.

Clearly, product of two hyperbolic polynomials (on an appropriately defined space) is again hyperbolic. We end this section by mentioning some ways of getting new hyperbolic polynomials [3] from any given hyperbolic polynomial. Let p be hyperbolic with respect to e.

- $q(x) := \frac{d}{dt} p(x + td) \mid_{t=0} = \langle \nabla p(x), e \rangle$ *is hyperbolic with respect to* e.
- *For any index* k, $1 \leq k \leq n$, *let* $E_k(x)$ *be the* k*th elementary function in variables* $x_1, x_2, \ldots, x_n$. *So,* $E_1(x) := x_1 + x_2 + \cdots + x_n$, $E_2(x) = \sum_{i<j} x_i x_j$, *more generally,*

$$E_k(x) := \sum_{i_1 < i_2 < \cdots < i_k} x_{i_1} x_{i_2} \cdots x_{i_k};$$

  finally, $E_n(x) = x_1 x_2 \cdots x_n$. *Then, the composition* $E_k(\lambda(x))$ *is a homogeneous polynomial of degree* k *and hyperbolic with respect to* e. *In particular, elementary functions are hyperbolic with respect to* $e = (1, 1, \ldots, 1)$.
- *Let* $q : \mathcal{R}^n \to \mathcal{R}$ *be any permutation invariant function that is hyperbolic with respect to* $(1, 1, \ldots, 1)$ *with eigenvalue map* μ. *Then* $q \circ \lambda$ *is hyperbolic with respect to* e *with eigenvalue map* $\mu \circ \lambda$.

4 Inducing an inner product on the vector space $\mathcal{V}$

Definition 4.1 [3] *Let* p *be hyperbolic on* $\mathcal{V}$. *We say that* p *is complete if*

$$\lambda(x) = 0 \Rightarrow x = 0.$$

Polynomials of Examples 1,2, and 3 are all complete. The polynomial of Example 5 is complete if and only if $A_1, A_2, \ldots, A_n$ are linearly independent. We also note that *when* p *is complete, every* λ*-orbit is compact.*

Proposition 4.2 [6] p *is complete if and only if the hyperbolicity cone of* p *is pointed (equivalently, proper).*

The following result, established in [3], is an elegant way of inducing an inner product on $\mathcal{V}$ as well as proving a 'Fan/Richter-Theobald-von Neumann' type inequality (in $\mathcal{S}^n$ or $\mathcal{H}^n$).

Theorem 4.3 ([3], Theorem 4.2, Proposition 4.4, and Corollary 3.11) *When* p *is complete,* $\mathcal{V}$ *becomes an inner product space under the inner product*

$$\langle x, y \rangle := \frac{1}{4} \Big[||\lambda(x + y)||^2 - ||\lambda(x - y)||^2 \Big]. \tag{1}$$

Additionally,

$$\langle x, y \rangle \leq \langle \lambda(x), \lambda(y) \rangle,$$

with equality if and only if $\lambda(x+y) = \lambda(x) + \lambda(y)$.

Note: If $||x||$ is the norm induced by the above inner product, then, by the positive homogeneity of λ,

$$||x|| = ||\lambda(x)||.$$

The proof (given in [3]) is as follows: We have

$$||\lambda(x)||^2 = \sum_{i=1}^{n} \lambda_i(x)^2 = E_1(\lambda(x))^2 - 2E_2(\lambda(x)).$$

By our observation above, the right-hand side is a homogeneous polynomial of degree 2 and, hence, could be regarded (with respect to some basis in $\mathcal{V}$) as the (usual) inner product $\langle Bx, x\rangle$, where B is a real symmetric matrix. Now, $\langle Bx, x\rangle \geq 0$ for all x and is zero if and only if $x = 0$ (as p is complete). Hence,

$$\langle Bx, y\rangle = \frac{1}{4}\Big[\langle B(x+y), x+y\rangle - \langle B(x-y), x-y\rangle\Big]$$

defines an inner product on $\mathcal{V}$.

Note: In the case of $\mathcal{R}^n$ and $\mathcal{S}^n$ (or $\mathcal{H}^n$), we recover usual inner products.

5 The Lax conjecture

We saw in various examples that hyperbolic polynomials arise in the form of determinants. P. Lax [15] conjectured that in three dimensions, this is the only way of getting hyperbolic polynomials.

Conjecture (Lax, 1958) *A real homogeneous polynomial* p *on* $\mathcal{R}^3$ *is hyperbolic of degree* n *with respect to the vector* $e = (1,0,0)$ *and satisfies* $p(e) = 1$ *if and only if there exist matrices* A, B *in* $\mathcal{S}^n$ *such that*

$$p(x, y, z) = \det(xI + yA + zB).$$

Based on a deep result of Helton and Vinnikov [13] in algebraic geometry, Lewis, Parrilo, and Ramana [17] showed that the *Lax conjecture is true.* Based on the validity of this conjecture, Gurvits [12] proved the following result.

Theorem 5.1 (*a*) *Suppose* $\mathcal{V} = \mathcal{R}^n$ *and* $e = (1, 0, \ldots, 0)$. *Consider a real homogeneous polynomial on* $\mathcal{V}$ *that is hyperbolic relative to* e. *Then, for any* $x, y \in \mathcal{R}^n$, *there exist* $n \times n$ *real* *symmetric matrices A and B such that*

$$p(te + rx + sy) = \det(tI + rA + sB) \quad (\forall\, t, r, s \in \mathcal{R}).$$

In particular, for all $r, s \in \mathcal{R}$,

$$\lambda(rx + sy) = \lambda(rA + sB),$$

where the right-hand side denotes the eigenvalue vector of a symmetric matrix.

(b) *Let $(\mathcal{V}, p, \lambda)$ be a hyperbolic triple. Then, for any $x, y \in \mathcal{V}$, the Lidskii property*

$$\lambda(x) - \lambda(y) \prec \lambda(x - y)$$

holds. In particular, $\lambda(x + y) \prec \lambda(x) + \lambda(y)$.

Proof (from Gurvits [12]) Fix $x, y \in \mathcal{V}$ and define a polynomial $q : \mathcal{R}^3 \to \mathcal{R}$ by

$$q(x_1, x_2, x_3) := p(x_1 e + x_2 x + x_3 y).$$

This is hyperbolic relative to $(1, 0, 0)$ in R^3 and, by the validity of Lax conjecture, could be written as $\det(x_1 I + x_2 A + x_3 B)$ for some real symmetric matrices A and B. By putting $(x_1, x_2, x_3) = (t, r, s)$, we get $p(te + rx + sy) = \det(tI + rA + sB)$. This gives the first item.
We see that for real r, s,

$$\lambda(rx + sy) = \lambda(rA + sB).$$

Since the Lidskii inequality is valid for Hermitian matrices (see Bhatia's book on Matrix Analysis), it is also valid for x and y in $\mathcal{R}^n$. By isomorphism considerations, it is also valid on a general $\mathcal{V}$.

6 Isometric hyperbolic polynomials

The following definition is from Bauschke et al. [3]:

Definition 6.1 *A hyperbolic polynomial p on $\mathcal{V}$ is said to be isometric if for all $y, z \in \mathcal{V}$, there exists an $x \in \mathcal{V}$ such that*

$$\lambda(x) = \lambda(z) \quad \textit{and} \quad \lambda(x + y) = \lambda(x) + \lambda(y).$$

It follows immediately that when p is isometric, the range of λ is a convex cone in $\mathcal{R}^n$. (For a general p, range of λ may not be convex, see Example 5.2 in [3].)

When p is complete, the isometric property can be characterized via the following result.

Proposition 6.2 [3] *Suppose p is complete (which induces an inner product on $\mathcal{V}$). Then p is* *isometric if and only if for all $c, u \in \mathcal{V}$,*

$$\max\{\langle c, x\rangle : x \in [u]\} = \langle \lambda(c), \lambda(u)\rangle.$$

The condition given in the above theorem says that every linear function attains its maximum on any λ-orbit with a specified maximum value. Thus, *every hyperbolic triple $(\mathcal{V}, p, \lambda)$ becomes a Fan-Theobald-von Neumann system [9] when p is complete and isometric.* It is known that in the setting of any Euclidean Jordan algebra, the hyperbolic polynomial det is complete and isometric, see [9].

7 Some new concepts and results

We now introduce some new concepts. In what follows, *we fix a hyperbolic triple $(\mathcal{V}, p, \lambda)$ and denote the corresponding hyperbolicity cone by $\mathcal{V}_+$.* Note that when p is complete, $\mathcal{V}_+$ is a proper cone (that is, it is a closed convex pointed cone with nonempty interior).

Definition 7.1 (a) *Two elements x and y in $\mathcal{V}$ commute in $\mathcal{V}$ if $\lambda(x+y) = \lambda(x) + \lambda(y)$. (When p is complete, this is equivalent to $\langle x, y\rangle = \langle \lambda(x), \lambda(y)\rangle$, see Theorem 4.3.)*

(b) *The center is defined as $C := \{x \in \mathcal{V} : x$ commutes with every $y \in \mathcal{V}\}$.*

(c) *When p is complete, the triple $(\mathcal{V}, p, \lambda)$ is said to be simple if the hyperbolicity cone $\mathcal{V}_+$ is irreducible.*

(d) *An invertible linear transformation $A : \mathcal{V} \to \mathcal{V}$ is an automorphism of $\mathcal{V}$ if $\lambda(Ax) = \lambda(x)$ for all $x \in \mathcal{V}$. The set of all such automorphisms is denoted by $Aut(\mathcal{V}, p, \lambda)$, or just $Aut(\mathcal{V})$.*

(e) *An invertible linear transformation $A : \mathcal{V} \to \mathcal{V}$ is an automorphism of (the cone) $\mathcal{V}_+$ if $A(\mathcal{V}_+) = \mathcal{V}_+$. The set of all automorphisms of $\mathcal{V}_+$ is denoted by $Aut(\mathcal{V}_+)$.*

(f) *Given $x, y \in \mathcal{V}$, x is said to be majorized by y if $x \in \operatorname{conv}[y]$. We then write $x \prec y$.*

(g) *A linear transformation $D : \mathcal{V} \to \mathcal{V}$ is said to be doubly stochastic if $Dx \prec x$ for all $x \in \mathcal{V}$. The set of all doubly stochastic transformations is denoted by $DS(\mathcal{V}, p, \lambda)$, or just $DS(\mathcal{V})$.*

Here are some elementary results.

Proposition 7.2 Let $(\mathcal{V}, p, \lambda)$ be a hyperbolic triple.

(1) *When p is complete, $[e] = \{e\}$.*

(2) *$\mathcal{R}\, e \subseteq C$. The reverse inclusion holds if p is complete.*

(3) *$Aut(\mathcal{V})$ is a group. When p is complete, it is a subgroup of the orthogonal group on the inner*

product space $\mathcal{V}$.

(4) *Every automorphism is doubly stochastic.*

(5) *The set $DS(\mathcal{V})$ of all doubly stochastic transformations on $\mathcal{V}$ is convex.*

(6) *When p is complete, relative to the norm induced by p, $||D|||\leq 1$ for all $D\in DS(\mathcal{V})$. Hence, $DS(\mathcal{V})$ is compact and convex in $\mathcal{V}$.*

Proof. (1): If $\lambda(x)=\lambda(e)=(1,1,\ldots,1)$, then $\lambda(e-x)=0$. By completeness of p, we have $x=e$.
(2): As $\lambda(\alpha e+x)=\alpha(1,1,\ldots,1)+\lambda(x)$, any scalar multiple of e commutes with every x; hence $\mathcal{R}\,e\subseteq C$. To see the reverse inclusion, assume p is complete and let x commute with every element in $\mathcal{V}$. Then x commutes with $-x$ and so $\lambda(x)+\lambda(-x)=\lambda(x+(-x))=\lambda(0)=0$. As the entries of any $\lambda(u)$ are decreasing, $\lambda(x)=-\lambda(-x)$ implies that $\lambda(x)$ is a scalar multiple of $(1,1,\ldots,1)$, say, $\lambda(x)=\alpha(1,1,\ldots,1)$. Then, $\lambda(\alpha e-x)=0$. By completeness, $x=\alpha e$. This shows that $C\subseteq\mathcal{R}\,e$ when p is complete.
(3) Clearly, $\mathrm{Aut}(\mathcal{V})$ is a group. When p is complete, $\mathcal{V}$ carries the inner product induced by p. For any automorphism A, $\lambda(Ax)=\lambda(x)$ for all x. Hence, by (1), $\langle Ax,Ay\rangle=\langle x,y\rangle$ for all x,y. Thus, A is orthogonal.
(4) is obvious.
(5) is obvious.
(6) Fix $D\in\mathrm{DS}(\mathcal{V})$. As $Dx\in\mathrm{conv}\,[x]$, we can write the convex combination $Dx=\sum\alpha_i x_i$, where $x_i\in[x]$ for all i. Then, $||x_i||=||\lambda(x_i)||=||\lambda(x)||=||x||$ for all i and

$$||Dx||\leq\sum\alpha_i||x_i||=\sum\alpha_i||x||=||x||.$$

This implies that $||D||\leq 1$. Since p is complete, each orbit $[x]$ is compact. Now we can use Caratheodory's theorem (to put limitations on the number of elements needed to express any element in $\mathrm{conv}\,[x]$) to show that $\mathrm{DS}(\mathcal{V},p,\lambda)$ is closed. We see that $\mathrm{DS}(\mathcal{V})$ is compact.

We saw in many examples that $x\in\mathrm{conv}\,[y]\Leftrightarrow\lambda(x)\prec\lambda(y)$. Our next result gives one implication in the general case. We do not know if the reverse implication also holds.

Theorem 7.3 *$x\prec y$ in $\mathcal{V}\Rightarrow\lambda(x)\prec\lambda(y)$ in $\mathcal{R}^n$.*

Proof. Suppose $x\prec y$ in $\mathcal{V}$, that is, $x\in\mathrm{conv}\,[y]$. We can write $x=\sum_{k=1}^m t_k y_k$, a convex combination of elements $y_1,y_2,\ldots,y_m$ in $[y]$. From Item (b) in Theorem 5.1, for any $u,v\in\mathcal{V}$,

$$\lambda(u+v)\prec\lambda(u)+\lambda(v).$$

It follows that

$$\lambda(x)=\lambda\Big(\sum_{k=1}^m t_k y_k\Big)\prec\sum_{k=1}^m t_k\lambda(y_k)=\lambda(y).$$

Remark. An alternative proof that avoids Gurvits' result can be given as follows. Let $x \in \text{conv}\,[y]$ and write $x = \sum_{k=1}^{m} t_k y_k$, a convex combination of elements $y_1, y_2, \ldots, y_m$ in $[y]$. Let $p := \lambda(x)$ and $q := \lambda(y)$. Fix any $w \in \mathcal{R}^n$ with decreasing components. By Proposition 3.2, in $\mathcal{R}^n$ we have

$$w^T \lambda(x) \leq \sum t_k w^T \lambda(y_k) = \sum t_k w^T \lambda(y) = w^T \lambda(y).$$

Thus,

for all $w \in \mathcal{R}^n$ with decreasing components, $w^T p \leq w^T q$.

From this, it is easy to show that $p \prec q$ in $\mathcal{R}^n$, proving $\lambda(x) \prec \lambda(y)$.

Remark. When p is complete and isometric, it can be shown that $x \prec y$ in $\mathcal{V} \Leftrightarrow \lambda(x) \prec \lambda(y)$ in $\mathcal{R}^n$, see [9].

Suppose p is complete so that $\mathcal{V}$ carries the induced inner product. Relative to this, for any linear D, we can define the adjoint D^* by $\langle D^* x, y \rangle = \langle x, Dy \rangle$.

Corollary 7.4 *Consider a hyperbolic triple $(\mathcal{V}, p, \lambda)$ with its hyperbolicity cone K. Then, for any doubly stochastic D, we have $D(K) \subseteq K$ and $De = e$. When p is complete, $D^* e = e$.*

Proof. Suppose D is doubly stochastic so $Dx \in \text{conv}\,[x]$ for all x. Since K is a spectral set (it is $\lambda^{-1}(\mathcal{R}^n_+)$) and convex (see Proposition 3.3), we have $D(K) \subseteq K$. As $[e] = e$, we have $De = e$. Now suppose p is complete and $D^* e \neq e$. Then, by a separation theorem (applicable in the Hilbert space $\mathcal{V}$), there is a nonzero $d \in \mathcal{V}$ such that

$$\langle D^* e, d \rangle > \langle e, d \rangle.$$

Since e commutes with all elements in $\mathcal{V}$ and p is complete, this reads:

$$\langle \lambda(Dd), \lambda(e) \rangle = \langle Dd, e \rangle = \langle d, D^* e \rangle > \langle d, e \rangle = \langle \lambda(d), \lambda(e) \rangle.$$

Since $\lambda(e) = (1, 1, \ldots, 1)$, we see that sum of all entries in $\lambda(Dd)$ is greater than the sum of all entries in $\lambda(d)$. Now, an application of Theorem 7.3 gives: $Dd \in \text{conv}\,[d] \Rightarrow \lambda(Dd) \prec \lambda(d)$. It follows, see Section 1, that the sum of all entries in $\lambda(Dd)$ is equal to the sum of all entries in $\lambda(d)$. We reach a contradiction. Hence, $D^* e = e$.

$\square$

8 Some open problems

We conclude our presentation by listing some open problems. Recall that the hyperbolicity cone in the hyperbolic triple $(\mathcal{V}, p, \lambda)$ is denoted by $\mathcal{V}_+$.

(1) *Describe commutativity in the setting of elementary functions/polynomials* E_k*, see end of Section 3.*

(2) *Describe automorphism groups of elementary functions/polynomials.*

(3) *Characterize* p *for which* $(\mathcal{V}, p, \lambda)$ *is simple, that is,* $\mathcal{V}_+$ *is irreducible.*

(4) *Assuming* p *is complete, describe the extreme points of the compact convex set* $DS(\mathcal{V})$*. (Using the strict convexity of the inner product norm, it can be shown that every automorphism is an extreme point.)*

(5) *Assuming* p *is complete, when do we have the equality* $DS(\mathcal{V}) = conv\, Aut(\mathcal{V})$*? (This is true for* p *of Example 3, see [7].) How about the pointwise equality* $DS(\mathcal{V})z = conv\, Aut(\mathcal{V})z$ *holding for all* $z \in \mathcal{V}$*? (This is true for simple Euclidean Jordan algebras, see [7].)*

(6) *Suppose* p *is complete. If* D *is doubly stochastic, do we have* $D^* x \prec x$ *for all* x*? (This is true when* p *is also isometric, more generally in any FTvN system, see [9].)*

(7) *Suppose* p *is complete and* D *is linear on* $\mathcal{V}$*. Does the converse in Corollary 7.4 hold? That is, if* $D(K) \subseteq K$ *with* $De = e$ *and* $D^* e = e$*, can we say that* D *is doubly stochastic? (This is true in the case of Euclidean Jordan algebras, see [9].)*

(8) *What relations exist between* $Aut(\mathcal{V})$*,* $Aut(\mathcal{V}_+)$*, and* $DS(\mathcal{V})$*? For example, (when) can we say* $Aut(\mathcal{V}) = Aut(\mathcal{V}_+) \cap DS(\mathcal{V}) = \{A \in Aut(\mathcal{V}_+) : A(e) = e\}$*. (See [7] for results of this type.)*

References

[1] T. Ando, *Majorization, doubly stochastic matrices, and comparison of eigenvalues*, Linear Alg. Appl., 118 (1989) 163-248.

[2] H. Aslaksen, *Quaternionic determinants*, Math. Intelligencer, 18 (1996) 57-65.

[3] H. H. Bauschke, O. Guler, A. S. Lewis, and H. S. Sendov, *Hyperbolic Polynomials and Convex Analysis*, Canad. J. Math. Vol. 53 (3), 2001 pp. 470-488.

[4] J. Faraut and A. Korányi, *Analysis on Symmetric Cones*, Oxford University Press, Oxford, 1994.

[5] L. Gårding, *Linear hyperbolic differential equations with constant coefficients.* Acta Math. 85(1951), 2-62.

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

[7] M.S. Gowda, *Positive and doubly stochastic maps, and majorization in Euclidean Jordan algebras,* Linear Algebra and its Applications, 528 (2017) 40-61.

[8] M.S. Gowda, *Optimizing certain combinations of spectral and linear/distance functions over spectral sets*, arXiv:1902.06640v2, March 2019.

[9] M.S. Gowda and J. Jeong, *Commutativity, majorization, and reduction in Fan-Theobald-von Neumann systems*, Manuscript, Sept. 30, 2021.

[10] M.S. Gowda and J. Tao, *The Cauchy interlacing theorem in simple Euclidean Jordan algebras and some consequences*, Linear and Multilinear Algebra, 59 (2011) 65-86.

[11] O. Guler, *Hyperbolic polynomials and interior point methods for convex programming*, Math. Oper. Res. 22 (1997) 350-377.

[12] L. Gurvits, *Combinatorics hidden in hyperbolic polynomials and related topics*, arXiv:math/0402088v1 [math.CO], 2004.

[13] J.W. Helton and V. Vinnikov, *Linear matrix inequality representation of sets,* Comm. Pure Appl. Math 60 (2007) 654-674.

[14] G.H. Hardy, J.E. Littlewood, and G. Pólya, *Inequalities*, Cambridge University Press, Cambridge, 1952.

[15] P.D. Lax, *Differential equations, difference equations and matrix theory*, Communications on Pure and Applied Mathematics, 6 (1958) 175-194.

[16] A.S. Lewis, *Group invariance and convex matrix analysis*, SIAM J. Matrix Anal Appl., 17 (1996) 927-949.

[17] A.S. Lewis, P.A. Parrilo, and M.V. Ramana, *The Lax conjecture is true*, Proc. AMS, 133 (2005) 2495-2499.

[18] E. H. Moore, *On the determinant of an hermitian matrix of quaternionic elements,* Bull. Amer. Math. Soc. 28 (1922), 161-162.

[19] H. Schneider, *Positive operators and an inertia theorem,* Numer. Math. 7 (1965), l-17.

[20] C.M. Theobald, *An inequality for the trace of the product of two symmetric matrices*, Math. Proc. Camb. Philos. Soc., 77 (1975) 265-267.

[21] J. von Neumann, *Some matrix inequalities and metrization of matric-spaces*, Tomsk University Rev., 1 (1937) 286-300. In Collected Works, Vol. IV, 205-218, Pergamon, Oxford 1962.