Hyperbolic Polynomials and Convex Analysis

Heinz H. Bauschke[*], **Osman Güler**[†], **Adrian S. Lewis**[‡], and **Hristo S. Sendov**[§]

June 9, 1998

Abstract

A homogeneous polynomial $p(x)$ is *hyperbolic* with respect to a given vector d if the real polynomial $t \mapsto p(x + td)$ has all real roots for all vectors x. We show that any symmetric convex function of these roots is a convex function of x, generalizing a fundamental result of Gårding. Consequently we are able to prove a number of deep results about hyperbolic polynomials with ease. In particular, our result subsumes von Neumann's characterization of unitarily invariant matrix norms, and Davis's characterization of convex functions of the eigenvalues of Hermitian matrices. We then develop various convex-analytic tools for such symmetric functions, of interest in interior-point methods for optimization problems posed over related cones.

[*]Department of Combinatorics & Optimization, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada. Email: bauschke@cecm.sfu.ca. Research supported by an NSERC Postdoctoral Fellowship and by the Department of Combinatorics & Optimization, University of Waterloo.

[†]Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, Maryland 21228, USA. Email: guler@math.umbc.edu. Research partially supported by the National Science Foundation under grant DMS-9623135

[‡]Department of Combinatorics & Optimization, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada. Email: aslewis@math.uwaterloo.ca. Research supported by NSERC.

[§]Department of Combinatorics & Optimization, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada. Email: hssendov@barrow.uwaterloo.ca. Research supported by NSERC.

1991 M.R. Subject Classification.
Primary 90C25; Secondary 15A45, 52A41.
Key words and phrases. convex analysis, eigenvalue, Gårding's inequality, hyperbolic barrier function, hyperbolic polynomial, hyperbolicity cone, interior-point method, semidefinite program, singular value, symmetric function, unitarily invariant norm.

1 Introduction

A famous result of von Neumann [35] states that for any symmetric gauge g on $\mathbb{R}^n$, the function

$$Z \in M^n \mapsto g(\sigma(Z)) \tag{1}$$

is a norm (and so in particular is convex). By a *symmetric gauge* we mean a norm which is invariant under sign changes and permutations of its arguments. For any matrix Z in M^n, the space of $n \times n$ real matrices, the vector $\sigma(Z)$ has components the singular values of Z arranged in decreasing order. Rather less well-known is a parallel result of Davis [5] stating that for any symmetric convex function f on $\mathbb{R}^n$, the function

$$Z \in S^n \mapsto f(\lambda(Z)) \tag{2}$$

is convex. Here, for any matrix Z in S^n, the space of $n \times n$ real symmetric matrices, the vector $\lambda(Z)$ has components the eigenvalues of Z, arranged in decreasing order.
Norms of the form (1) are important in matrix approximation [14]. Analogously, functions of the form (2) are fundamental in eigenvalue optimization and semidefinite programming [23]. In both areas there is an attractive duality theory: the Fenchel conjugate of the function (2) is described elegantly by the formula $(f \circ \lambda)^* = f^* \circ \lambda$ [19], and von Neumann showed a parallel result for the norm (1).
The analogies between these two families of results are not accidental. The paper [20] develops an axiomatic framework subsuming both models. At a more sophisticated level, both convexity results follow quickly from the Kostant convexity theorem in semisimple Lie theory [21].
The work we describe in this current paper also concerns the above type of convexity result, but with a very different and remarkably simple approach.

To illustrate the key idea, consider the determinant as a function on S^n. This function is a homogeneous polynomial which is *hyperbolic* with respect to the identity matrix I: that is, the real polynomial

$$t \in \mathbb{R} \mapsto \det(Z - tI)$$

has all real roots, namely the eigenvalues $\lambda_i(Z)$. The properties of such polynomials play a significant role in the partial differential equations literature (see for example [13])), but we use just one, central result, due to Gårding [8]: the largest root $\lambda_1(\cdot)$ is always a convex function.
Working from Gårding's result, we show, just like Davis's theorem, that *any* symmetric convex function of the roots $\lambda_i(\cdot)$ is convex. The richness of the class of hyperbolic polynomials then allows us to derive many elegant (and often classical) inequalities in a unified fashion. Examples include beautiful properties of the elementary symmetric functions. One particular hyperbolic polynomial leads us back to the singular value example.
Associated with any hyperbolic polynomial comes a closed convex *hyperbolicity cone* which, with the above notation, we can write

$$\{Z : \lambda_i(Z) \geq 0 \ \forall i\}.$$

For example, in the symmetric matrix case this is simply the cone of positive semidefinite matrices. Güler has shown how optimization problems over such cones are good candidates for interior point algorithms analogous to the dramatically successful techniques current in semidefinite programming [10]. In part with that aim in mind, we develop an attractive duality theory and convex-analytic tools for symmetric convex functions of the roots associated with general hyperbolic polynomials.

Notation

We write $\mathbb{R}^m_{++}$ (resp. $\mathbb{R}^m_+$) for the set $\{u \in \mathbb{R}^m : u_i > 0, \forall i\}$ (resp. $\{u \in \mathbb{R}^m : u_i \geq 0, \forall i\}$. The *closure* (resp. *boundary, convex hull, linear span*) of a set S is denoted $\operatorname{cl} S$ (resp. $\operatorname{bd} S$, $\operatorname{conv} S$, $\operatorname{span} S$). A *cone* is a nonempty set that contains every nonnegative multiple of all its members; it thus always contains 0. If $u \in \mathbb{R}^m$, then $u_\downarrow$ is the vector u with its coordinates arranged decreasingly; also, $U_\downarrow := \{u_\downarrow : u \in U\}$, for every subset U of $\mathbb{R}^m$. The *transpose* of a matrix (or vector) A is denoted A^T. The *identity* matrix or map is written I. Suppose Y is an arbitrary Euclidean space with inner product $\langle\cdot,\cdot\rangle$ and $h : Y \to [-\infty,+\infty]$ is convex, then h^* (resp. ∂h, ∇h, $\operatorname{dom} h$) stands for the *Fenchel conjugate* (resp. *subdifferential map*, *gradient map*, *domain*) of h. (Rockafellar's monograph [33] is the standard reference for these notions from convex analysis.) Higher order derivatives are denoted by $\nabla^k h$. If $U \subseteq X$, then the *positive polar cone* is $U^{\oplus} := \{x \in X : \langle x, U\rangle \geq 0\}$. If A is a linear operator between Euclidean spaces, then its *conjugate* is written A^*. The *range* of a map λ is denoted by $\operatorname{ran}\lambda$. Finally, if A, B are two subsets of X, then $d(A,B) := \inf \|A-B\|$ is the *distance* between A and B.

2 Tools

We assume throughout the paper that

$$\boxed{X \text{ is a finite-dimensional real vector space.}}$$

This section contains a selection of important facts on hyperbolic polynomials from Gårding's fundamental work [8], and a deep inequality on elementary symmetric functions.

Hyperbolic polynomials and eigenvalues

Definition 2.1 (homogeneous polynomial) Suppose p is a nonconstant polynomial on X and m is a positive integer. Then p is *homogeneous of degree* m, if $p(tx) = t^m p(x)$, for all $t \in \mathbb{R}$ and every $x \in X$.

Definition 2.2 (hyperbolic polynomial) Suppose that p is a homogeneous polynomial of degree m on X and $d \in X$ with $p(d) \neq 0$. Then p *is hyperbolic with respect to* d, if the polynomial $t \mapsto p(x+td)$ (where t is a scalar) has only real zeros, for every $x \in X$.

Definition 2.3 ("eigenvalues and trace") Suppose p is hyperbolic with respect to $d \in X$ of degree m. Then for every $x \in X$, we can write

$$p(x+td) = p(d)\prod_{i=1}^{m}(t+\lambda_i(x))$$

and assume without loss of generality that $\lambda_1(x) \geq \lambda_2(x) \geq \cdots \geq \lambda_m(x)$. The corresponding map $X \to \mathbb{R}^m_\downarrow : x \mapsto (\lambda_1(x), \ldots, \lambda_m(x))$ is denoted by λ and called the *eigenvalue map (with respect to p and d)*. We say that $\lambda_i(x)$ is the i^{th} *largest eigenvalue of x (with respect to p and d)* and define *the sum of the k largest eigenvalues* by $\sigma_k := \sum_{i=1}^k \lambda_i$, for every $1 \leq k \leq m$. The function σ_m is called the *trace*.

The eigenvalues $\{\lambda_i(x)\}$ are thus the roots of the polynomial $t \mapsto p(x - td)$. It follows readily that the trace σ_m is linear (see also the paragraph following Fact 2.19).
Unless stated otherwise, we assume throughout the paper that

$$\boxed{\begin{array}{c} p \text{ is a hyperbolic polynomial of degree } m \text{ with respect to } d, \\ \text{with eigenvalue map } \lambda \text{ and } \sigma_k := \sum_{i=1}^k \lambda_k, \end{array}}$$

for every $1 \leq k \leq m$. The notions "eigenvalues" and "trace" are well-motivated by the the following example.

The Hermitian matrices. Let X be the real vector space of the $m \times m$ Hermitian matrices and $p := \det$. Then p is hyperbolic of degree m with respect to $d := I$ and λ maps $x \in X$ to its eigenvalues, arranged decreasingly. Thus for every $1 \leq k \leq m$, the function σ_k is indeed the sum of the k largest eigenvalues and σ_m is the (ordinary) trace.

As we go, we will point out what some of the results become in the important case of the *Hermitian matrices*. Details and further examples are provided in Section 7.
We now introduce the notion of isomorphic triples, which will simplify the analysis of homogeneous polynomials in Section 7 considerably.

Definition 2.4 Suppose p (resp. q) is a homogeneous polynomial on X (resp. Y) and $d \in X$ (resp. $e \in Y$). If there exists a linear one-to-one map Φ from X onto Y with $p = q \circ \Phi$ and $\Phi(d) = e$, then we say that the triple (X, p, d) is *isomorphic* to (Y, q, e) (by Φ), and we write $(X, p, d) \simeq (Y, q, d)$.

It is clear that the binary operation $\simeq$ defines an equivalence relation on all triples. The following basic properties are easy to verify.

Proposition 2.5 Suppose (X, p, d) is isomorphic to (Y, q, e) by Φ. Then:

(i) The degrees of p and q coincide.

(ii) p is hyperbolic with respect to d if and only if q is hyperbolic with respect to e.

(iii) If p (resp. q) is hyperbolic with respect to d (resp. e) with corresponding eigenvalue map λ (resp. μ), then $\lambda = \mu \circ \Phi$.

Many examples of hyperbolic polynomials can be obtained as described below.

Proposition 2.6

(i) If q is hyperbolic with respect to the same d, then so is pq.

(ii) If $m > 1$, then $q(x) := \frac{d}{dt} p(x+td)\big|_{t=0} = \big(\nabla p(x)\big)(d)$ is hyperbolic with respect to d.

(iii) If Y is a subspace of X and $d \in Y$, then the restriction $p|_Y$ is hyperbolic with respect to d.

Proof. (i) is elementary. (ii) essentially follows from Rolle's Theorem; see also [8, Lemma 1]. (iii) is obvious. ■

The technique of Proposition 2.6.(ii) has a higher order analog; see Fact 2.19 below.
Given a hyperbolic polynomial on $\mathbb{R}^n$, we can construct a related one on $\mathbb{R}^{n-1}$ as follows.

Proposition 2.7 Suppose p is hyperbolic with respect to $d \in \mathbb{R}^n$ with eigenvalue map λ. Assume that $d_i \neq 0$ and define q on $\mathbb{R}^{n-1}$ by

$$q(y_1, \dots, y_{n-1}) = p(y_1, \dots, y_{n-1}, \tfrac{y_i}{d_i} d_n).$$

Then q is hyperbolic with respect to $e := (d_1, \dots, d_{n-1})$ and its eigenvalue map μ satisfies $\mu(y_1, \dots, y_{n-1}) = \lambda(y_1, \dots, y_{n-1}, \frac{y_i}{d_i} d_n)$.

Proof. Straightforward. ■

The following property of the eigenvalues is well-known ([8, Equation (2)]) and easily verified.

Fact 2.8 For all $r, s \in \mathbb{R}$ and every $1 \le i \le m$:

$$\lambda_i(rx + sd) = \begin{cases} r\lambda_i(x) + s, & \text{if } r \ge 0; \\ r\lambda_{m+1-i}(x) + s, & \text{otherwise.} \end{cases}$$

It follows that the eigenvalue map λ is *positively homogeneous* ($\lambda(tx) = t\lambda(x)$, for all $t \ge 0$ and every $x \in X$) and continuous (the zeros of a polynomial are continuous with respect to the coefficients; see, for instance, [32, Appendix A]).
Gårding showed that the largest eigenvalue map is *sublinear*, that is, positively homogeneous and convex.

Theorem 2.9 (Gårding) The largest eigenvalue map λ_1 is sublinear.

Proof. Positive homogeneity follows from Fact 2.8. Now Gårding showed that λ_m is concave [8, Theorem 2], which is equivalent to the convexity of λ_1, since $\lambda_1(-x) = -\lambda_m(x)$, for every $x \in X$. ■

The Hermitian matrices (continued). It is well-known that the largest eigenvalue map is convex in this case; see, for instance, [12].

Hyperbolicity cone

Definition 2.10 (hyperbolicity cone) The *hyperbolicity cone of p with respect to d*, written $C(d)$ or $C(p, d)$, is the set $\{x \in X : p(x+td) \neq 0, \forall t \ge 0\}$.

Fact 2.11 $C(d) = \{x \in X : \lambda_m(x) > 0\}$. Hence $C(d)$ is an open convex cone that contains d with closure $\operatorname{cl} C(d) = \{x \in X : \lambda_m(x) \ge 0\}$. If $c \in C(d)$, then p is hyperbolic with respect to c and $C(c) = C(d)$.

Proof. See Gårding's [8, Section 2]. ■

Remark 2.12 Note that $\lambda_m(x) > 0$ if and only if $\lambda_1(-x) < 0$ by Fact 2.8. Hence Gårding's result (Theorem 2.9) implies the convexity of $C(d)$. In fact,

$C(d)$ is a convex cone precisely because λ_1 is a convex function.

To see why convexity of $C(d)$ yields convexity of λ_1, fix x and y in X and observe that $x - \lambda_m(x)d$ and $y - \lambda_m(y)d$ both belong to $\operatorname{cl} C(d)$. By assumption, $(x + y) - (\lambda_m(x) + \lambda_m(y)) \in \operatorname{cl} C(d)$. On the other hand, the smallest t such that $(x + y) + td$ belongs to $\operatorname{cl} C(d)$ is $-\lambda_m(x + y)$. Altogether, $\lambda_m(x) + \lambda_m(y) \le \lambda_m(x + y)$ and the concavity of λ_m (or convexity of λ_1) follows.

Definition 2.13 (complete hyperbolic polynomial) p is *complete* if

$$\{x \in X : \lambda(x) = 0\} = \{0\}.$$

The following result, which follows easily from Proposition 2.5.(iii), considers the concepts just introduced for isomorphic triples.

Proposition 2.14 Suppose (X, p, d) is isomorphic to (Y, q, e) by Φ. Then:

(i) $C(q, e) = \Phi(C(p, d))$.

(ii) p is complete if and only if q is.

Fact 2.15 Suppose p is hyperbolic with respect to d, with corresponding eigenvalue map λ and hyperbolicity cone $C(d)$. Then

$$\begin{aligned}\{x \in X : \lambda(x) = 0\} &= \{x \in X : x + C(d) = C(d)\} \\ &= \{x \in X : p(tx + y) = p(y), \forall y \in X, \forall t \in \mathbb{R}\}.\end{aligned}$$

Consequently, $\{x \in X : \lambda(x) = 0\} = \operatorname{cl} C(d) \cap (-\operatorname{cl} C(d))$.

Proof. See Gårding's [8, Section 3]. The "Consequently" part follows readily from the displayed equation and the openness of $C(d)$. ∎

It is always possible to find a restriction of p that is complete: indeed, $d \notin \{x \in X : \lambda(x) = 0\}$; consequently, if Y is any subspace of X which contains d and is algebraically complemented to $\{x \in X : \lambda(x) = 0\}$, then $p|_Y$ is hyperbolic with respect to d (Proposition 2.6.(iii)) and complete.

Example 2.16 We let $X = \mathbb{R}^n$, $p(x) = \sum_j x_j$ and $d = (1, 1, \ldots, 1)$ in X. Then p is hyperbolic with respect to d of degree $m = 1$ and $\lambda(x) = \frac{1}{n} \sum_{j=1}^n x_j$. It follows that p is complete only when $n = 1$.

The Hermitian matrices (continued). The hyperbolicity cone of $p = \det$ with respect to $d = I$ is the set of all positive definite matrices. The polynomial $p = \det$ is complete, since every nonzero Hermitian matrix has at least one nonzero eigenvalue.

Elementary symmetric functions

Definition 2.17 (symmetric function) A function f on $\mathbb{R}^m$ is *symmetric*, if $f(u) = f(u_{\pi(i)})$, for all permutations π of $\{1, \ldots, m\}$ and every $u \in \mathbb{R}^m$.

Definition 2.18 (elementary symmetric functions) For any given integer $k = 1, 2, \ldots, m$, the map $E_k : \mathbb{R}^m \to \mathbb{R} : u \mapsto \sum_{i_1 < \cdots < i_k} \prod_{l=1}^k u_{i_l}$ is called the k^{th} *elementary symmetric function* on $\mathbb{R}^m$. We also set $E_0 := 1$.

Fact 2.19 For every $x \in X$ and all $t \in \mathbb{R}$,

$$p(x + td) = p(d) \prod_{i=1}^{m} (t + \lambda_i(x)) = p(d) \sum_{i=0}^{m} E_i(\lambda(x)) t^{m-i}$$

and for every $0 \leq i \leq m$,

$$p(d) E_i(\lambda(x)) = \frac{1}{(m-i)!} \nabla^{m-i} p(x) [\underbrace{d, d, \ldots, d}_{m-i \text{ times}}].$$

If $1 \leq i \leq m$, then $E_i \circ \lambda$ is hyperbolic with respect to d of degree i.

Proof. The first displayed equation is elementary while the second displayed equation is a consequence of Taylor's Theorem. The "If" part follows by employing Proposition 2.6.(ii) repeatedly. ■

Fact 2.19 gives a very transparent proof of the linearity of trace: indeed, $\sigma_m = E_1 \circ \lambda$ is a homogeneous (hyperbolic) polynomial of degree 1 and hence linear.

We also note that the elementary symmetric functions themselves are hyperbolic:

Example 2.20 Let $X = \mathbb{R}^m$ and $d = (1, 1, \ldots, 1) \in \mathbb{R}^m$. Then for every $1 \leq k \leq m$, the k^{th} elementary symmetric function E_k is hyperbolic of degree k with respect to d.

Proof. Let $p := E_m$. It is straightforward to check that E_m is hyperbolic of degree m with respect to d with corresponding eigenvalue map $\lambda(x) = x_{\downarrow}$. Since each E_k is symmetric, the result now follows from Fact 2.19. ■

An inequality in elementary symmetric functions

The following inequality was discovered independently by McLeod [29] and by Bullen and Marcus [3, Theorem 3].

Fact 2.21 (McLeod, 1959; Bullen and Marcus, 1961) Suppose $1 \leq k \leq l \leq m$ and $u, v \in \mathbb{R}^m_{++}$. Set $q := (E_l/E_{l-k})^{1/k}$. Then

$$q(u+v) > q(u) + q(v),$$

unless u and v are proportional or $k = l = 1$, in which case we have equality.

Bullen and Marcus's proof relies on an inequality by Marcus and Lopes ([25, Theorem 1], which is the case $k = 1$ in Fact 2.21. (Proofs can also be found in [1, Theorem 1.16], [4, Section V.4], and [31, Section VI.5].)
We record two interesting consequences of Fact 2.21.

Corollary 2.22 (Marcus and Lopes's [25, Theorem 2]) The function $-E_m^{1/m}$ is sublinear on $\mathbb{R}^m_+$, and it vanishes on $\operatorname{bd} \mathbb{R}^m_+$.

Proof. Set $k = l = m$ in Fact 2.21 and use continuity. ■

Recall that a function h is called *logarithmically convex*, if $\log \circ h$ is convex. The function q in Fact 2.21 is concave ("strictly modulo rays"), which yields logarithmic and strict convexity of $1/q$:

Proposition 2.23 Suppose q is a function defined on $\mathbb{R}^m_{++}$. Consider the following properties:
(i) the range of q is contained in $(0, +\infty)$;
(ii) $q(ru) = rq(u)$, for all $r > 0$ and every $u \in \mathbb{R}^m_{++}$;
(iii) $q(u + v) \geq q(u) + q(v)$, for all $u, v \in \mathbb{R}^m_{++}$;
(iv) if $u, v \in \mathbb{R}^m_{++}$ with $q(u + v) = q(u) + q(v)$, then $v = \rho u$, for some $\rho > 0$.
Suppose q satisfies (i)–(iii). Then $1/q$ is logarithmically convex. If furthermore (iv) holds, then $1/q$ is strictly convex.

Proof. (i)–(iii) implies that q is a concave function with range in $(0, +\infty)$. It follows that $\ln \circ q$ is concave (since ln is increasing and concave). Hence $-\ln \circ q$ is convex; equivalently, $1/q$ is logarithmically convex. Now assume that (iv) holds as well and fix $u, v \in \mathbb{R}^m_{++}$ with $1/q(\frac{1}{2}u + \frac{1}{2}v) = \frac{1}{2}\frac{1}{q(u)} + \frac{1}{2}\frac{1}{q(v)}$.

It suffices to show that $u = v$. Because q is concave on $\mathbb{R}^m_{++}$ and $r \mapsto 1/r$ is convex on $(0, +\infty)$, we estimate

$$\frac{1}{\frac{1}{2}q(u) + \frac{1}{2}q(v)} \geq \frac{1}{q(\frac{1}{2}u + \frac{1}{2}v)} = \frac{1}{2}\frac{1}{q(u)} + \frac{1}{2}\frac{1}{q(v)} \geq \frac{1}{\frac{1}{2}q(u) + \frac{1}{2}q(v)}.$$

Hence equality holds throughout and so $v = \rho u$, for some $\rho > 0$. The equalites imply $\frac{1}{1+\rho} = \frac{1}{4}(1 + \frac{1}{\rho})$, which yields $\rho = 1$. Thus $u = v$ and therefore $1/q$ is strictly convex. ■

Corollary 2.24 Suppose $1 \leq k \leq l \leq m$. Then the function $(E_{l-k}/E_l)^{1/k}$ is symmetric, positively homogeneous, and logarithmically convex. Moreover, the function is strictly convex on $\mathbb{R}^m_{++}$ unless $l = 1$ and $m \geq 2$.

Proof. Positive homogeneity and symmetry are clear. Log convexity follows by combining Proposition 2.23 and Fact 2.21; this even yields strict convexity unless $k = l = 1$. But if $k = l = 1$, then the function becomes $1/\sum_{i=1}^m u_i$, which is strictly convex exactly when $m = 1$. ■

3 Convexity

Sublinearity of the sum of the largest eigenvalues

Theorem 3.1 Suppose q is a homogeneous symmetric polynomial of degree n on $\mathbb{R}^m$, hyperbolic with respect to $e := (1, 1, \ldots, 1) \in \mathbb{R}^m$, with eigenvalue map μ. Then

$$q \circ \lambda$$

is a hyperbolic polynomial of degree n with respect to d and its eigenvalue map is $\mu \circ \lambda$.

Proof. For simplicity, write $\tilde{p}$ for $q \circ \lambda$.
Step 1: $\tilde{p}$ is a polynomial on X.
Since $q(y)$ is a symmetric polynomial on $\mathbb{R}^m$, it is (by, e.g., [16, Proposition V.2.20.(ii)]) a polynomial in $E_1(y), \ldots, E_m(y)$. On the other hand, by Fact 2.19, $E_i \circ \lambda$ is hyperbolic with respect to d of degree i, for $1 \leq i \leq m$. Altogether, $\tilde{p}(x) = q(\lambda(x))$ is a polynomial on X.

Step 2: $\tilde{p}$ is homogeneous of degree n.
Since q is symmetric and homogeneous, and in view of Fact 2.8, we obtain $\tilde{p}(tx) = q(\lambda(tx)) = t^n \tilde{p}(x)$, for all $t \in \mathbb{R}$ and every $x \in X$.
Step 3: $\tilde{p}(d) \neq 0$.
Again using Fact 2.8, we have $\tilde{p}(d) = q(\lambda(d)) = q(e) \neq 0$.
Step 4: $\tilde{p}$ is hyperbolic with respect to d.
Using once more Fact 2.8, we write for every $x \in X$ and all $t \in \mathbb{R}$:

$$\tilde{p}(x + td) = q(\lambda(x + td)) = q(\lambda(x) + te) = q(e) \prod_{k=1}^{n} (t + \mu_k(\lambda(x))). \quad \blacksquare$$

The next example is easy to check.

Example 3.2 Fix $1 \leq k \leq m$, set $e := (1, 1, \ldots, 1) \in \mathbb{R}^m$, and let

$$q(u) := \prod_{1 \leq i_1 < i_2 < \cdots < i_k \leq m} \sum_{l=1}^{k} u_{i_l}.$$

Then q is a homogeneous symmetric polynomial on $\mathbb{R}^m$ of degree $\binom{m}{k}$, hyperbolic with respect to e, and its eigenvalues are $\{\frac{1}{k} \sum_{l=1}^{k} u_{i_l} : 1 \leq i_1 < i_2 < \cdots < i_k \leq m\}$. In particular, the largest eigenvalue of q is the sum of the k largest components of u.

We now present our main result, the generalization of Theorem 2.9: the sum of the largest eigenvalues is sublinear. This readily implies local Lipschitzness of each eigenvalue map (see also [36]).

Corollary 3.3 For every $1 \leq k \leq m$, the function σ_k is sublinear and λ_k is locally Lipschitz.

Proof. Fix $1 \leq k \leq m$, define q as in Example 3.2, and consider $\tilde{p} := q \circ \lambda$. By Theorem 3.1 and Example 3.2, the largest eigenvalue of $\tilde{p}$ is equal to $\frac{1}{k}\sigma_k(x)$. Now Theorem 2.9 yields the sublinearity of σ_k. Finally, recall that every convex function is locally Lipschitz ([33, Theorem 10.4]), hence so is each σ_i. So λ_1 is locally Lipschitz. If $k \geq 2$, then $\lambda_k = \sigma_k - \sigma_{k-1}$ is — as the difference of two locally Lipschitz functions — locally Lipschitz, too. $\blacksquare$

The Hermitian matrices (continued). Here it is well known that the sum of the k largest eigenvalues is a convex function and that the k^{th} largest eigenvalue map is locally Lipschitz; see, for instance, [12].

Remark 3.4 Consider the polynomial $\tilde{p}$ constructed in the proof of Corollary 3.3 in the context of the Hermitian matrices. Then

$$(-1)^{\binom{m}{k}}\tilde{p}(x - \tfrac{t}{k}I) = \det(tI - \Delta_k(x)),$$

where $\Delta_k(x)$ denotes the k^{th} additive compound of x. (See [28, Section 19.F] for more on compound matrices.)

Corollary 3.5 The function $w^T\lambda(\cdot)$ is sublinear, for every $w \in \mathbb{R}^m_{\downarrow}$.

Proof. Write $w^T\lambda = \sum_{i=1}^m w_i\lambda_i = w_m\sigma_m + \sum_{i=1}^{m-1}(w_i - w_{i+1})\sigma_i$ and then apply Corollary 3.3. ■

Note that we can rewrite Corollary 3.5 quite artificially as $w^T(\lambda(x+y) - \lambda(x)) \leq w_{\downarrow}^T\lambda(y)$, for all $x, y \in X$ and $w \in \mathbb{R}^m_{\downarrow}$.
It would be interesting to find out about the following generalization:

Open Problem 3.6 (Lidskii's theorem) Decide whether or not

$$w^T(\lambda(x+y) - \lambda(x)) \leq w_{\downarrow}^T\lambda(y), \quad \text{for all } x, y \in X \text{ and } w \in \mathbb{R}^m .$$

If this condition is satisfied, then we say that *Lidskii's theorem holds* for the triple (X, p, d).
The condition means that the vector $\lambda(y)$ "majorizes" the vector $\lambda(x+y) - \lambda(x)$, for all $x, y \in X$; see [28, Proposition 4.B.8]. (The interested reader is referred to [28] for further information on majorization.)

The Hermitian matrices (continued). Lidskii's theorem does hold for the Hermitians. A recent and very complete reference is Bhatia's [2]; see also [22] for a new proof rooted in nonsmooth analysis.

In Section 7, we point out that Lidskii's theorem holds valid for all our examples. It will be convenient to have the following simple result ready:

Proposition 3.7 Suppose (X, p, d) is isomorphic to (Y, q, e). Then Lidskii's theorem holds for (X, p, d) if and only if it does for (Y, q, e).

Proof. Immediate from Proposition 2.5.(iii). ■

Convexity of composition

Fact 3.8 Suppose $f : \mathbb{R}^m \to [-\infty, +\infty]$ is convex and symmetric. Suppose further $u, v \in \mathbb{R}^m_{\downarrow}$ and $u - v \in (\mathbb{R}^m_{\downarrow})^{\oplus}$. Then $f(u) \geq f(v)$. Moreover: if f is strictly convex on $\operatorname{conv}\{u_{\pi(i)} : \pi \text{ is a permutation of } \{1, \ldots, m\}\}$ and $u \neq v$, then $f(u) > f(v)$.

Proof. Imitate the proof of [20, Theorem 3.3] and consider [20, Example 7.1]. See also [28, 3.C.2.c on page 68]. ■

Theorem 3.9 (convexity) Suppose $x, y \in X$, $\alpha \in (0, 1)$, and $f : \mathbb{R}^m \to [-\infty, +\infty]$ is convex and symmetric. Then

$$f(\lambda(\alpha x + (1-\alpha)y)) \leq f(\alpha\lambda(x) + (1-\alpha)\lambda(y))$$

and hence the composition $f \circ \lambda$ is convex. If f is strictly convex and $\alpha\lambda(x) + (1-\alpha)\lambda(y) \neq \lambda(\alpha x + (1-\alpha)y)$, then $f(\lambda(\alpha x + (1-\alpha)y)) < f(\alpha\lambda(x) + (1-\alpha)\lambda(y))$.

Proof. (See also [20, Proof of Theorem 4.3].) Fix an arbitrary $w \in \mathbb{R}^m_{\downarrow}$. Set $u := \alpha\lambda(x) + (1-\alpha)\lambda(y)$ and $v := \lambda(\alpha x + (1-\alpha)y)$. Then both u and v belong to $\mathbb{R}^m_{\downarrow}$. By Corollary 3.5, $w^T\lambda$ is convex on X. Therefore, $w^T\lambda(\alpha x + (1-\alpha)y) \leq \alpha w^T\lambda(x) + (1-\alpha)w^T\lambda(y)$; equivalently, $w^T(u - v) \geq 0$. It follows that $u - v \in (\mathbb{R}^m_{\downarrow})^{\oplus}$. By Fact 3.8, $f(u) \geq f(v)$, which is the second displayed statement. The convexity of $f \circ \lambda$ follows. Finally, the "If" part is implied by the above and the "Moreover" part of Fact 3.8. ■

The Hermitian matrices (continued). In this case, the convexity of the composition is attributed to *Davis* [5]; see also [19, Corollary 2.7].

Another consequence is Gårding's inequality; see [10, Lemma 3.1].

Corollary 3.10 (Gårding's inequality) Suppose $p(d) > 0$. Then function $x \mapsto -(p(x))^{1/m}$ is sublinear on the hyperbolicity cone $C(d)$, and it vanishes on its boundary.

Proof. By Corollary 2.22, the function $-E_m^{1/m}$ is sublinear and symmetric on $\mathbb{R}^m_+$. Hence, by Theorem 3.9, the function $x \mapsto -(E_m(\lambda(x)))^{1/m}$ is sublinear on $\{x \in X : \lambda(x) \geq 0\} = \operatorname{cl} C(d)$. The result follows, since $p(x) = p(d)E_m(\lambda(x))$, for every $x \in X$. ■

The Hermitian matrices (continued). Corollary 3.10 implies the *Minkowski Determinant Theorem*: $\sqrt[m]{\det(x+y)} \geq \sqrt[m]{\det x} + \sqrt[m]{\det y}$, whenever $x, y \in X$ are positive semi-definite.

Corollary 3.11 Suppose $x, y \in X$. Then:

(i) $\|\lambda(x+y)\| \leq \|\lambda(x)+\lambda(y)\|$.

(ii) $\|\lambda(x+y)\|^2 - \|\lambda(x)\|^2 - \|\lambda(y)\|^2 \leq 2\langle\lambda(x), \lambda(y)\rangle$.

Moreover, equality holds in (i) or (ii) if and only if $\lambda(x+y) = \lambda(x)+\lambda(y)$.

Proof. (i): Let $w := \lambda(x+y) \in \mathbb{R}^m_{\downarrow}$. Then, using Corollary 3.5 and the Cauchy-Schwarz inequality in $\mathbb{R}^m$, we estimate

$$\begin{aligned}\|\lambda(x+y)\|^2 &= w^T\lambda(x+y) \leq w^T(\lambda(x)+\lambda(y)) \\ &\leq \|w\|\|\lambda(x)+\lambda(y)\| = \|\lambda(x+y)\|\|\lambda(x)+\lambda(y)\|.\end{aligned}$$

The inequality follows. The condition for equality follows from the condition for equality in the Cauchy-Schwarz inequality.
(ii): The condition is equivalent to (i). ■

4 Making X Euclidean

Definition 4.1 Define $\|\cdot\| : X \to [0, +\infty) : x \mapsto \|\lambda(x)\|$ and

$$\langle\cdot,\cdot\rangle : X \times X \to \mathbb{R} : (x, y) \mapsto \tfrac{1}{4}\|x+y\|^2 - \tfrac{1}{4}\|x-y\|^2.$$

Theorem 4.2 Suppose p is complete. Then X equipped with $\langle\cdot,\cdot\rangle$ is a Euclidean space with induced norm $\|\cdot\|$.

Proof. We have $\|x\|^2 = \|\lambda(x)\|^2 = \sum_{i=1}^m \lambda_i(x)^2 = (E_1(\lambda(x)))^2 - 2E_2(\lambda(x))$. Facts 2.8 and 2.19 imply that $\|\cdot\|^2$ is a homogeneous polynomial of degree 2 on X. Since $\|\cdot\| \geq 0$ and p is complete, the result now follows from the Polarization Identity. ■

Remark 4.3 The Euclidean norm $\|\cdot\|$ defined in Definition 4.1 is precisely the *Hessian norm* used in interior point methods and thus well-motivated. To see this, assume that p is complete and recall that the *hyperbolic barrier function* is defined by $F(x) := -\ln(p(x))$. The Hessian norm at x is then given by

$$\|x\|_d^2 := \nabla^2 F(d)[x, x].$$

For t positive and sufficiently small, we have $p(tx+d) = p(d)\prod_{i=1}^{m}(1+t\lambda_i(x))$ and hence (after taking logarithms)

$$F(d+tx) = F(d) - \sum_{i=1}^{m} \ln(1+t\lambda_i(x)).$$

Expand the left (resp. right) side of this equation into a Taylor (resp. log) series. Then compare coefficients of t^2 to conclude $\nabla^2 F(d)[x,x]/2! = \|\lambda(x)\|^2/2$. Thus $\|\cdot\|_d = \|\cdot\|$. Further information can be found in [10]; see, in particular, [10, equation 16].

The norm constructed above has the pleasant property that any isomorphism to another triple is actually an isometry:

Proposition 4.4 Suppose p is complete and the triple (X, p, d) is isomorphic to the triple (Y, q, e) by Φ. Then Φ is an isometry from X onto Y.

Proof. Denote the eigenvalue map of p (resp. q) by λ (resp. μ). Using Proposition 2.5.(iii) and the definitions of the norms in X and Y, we have for every $x \in X$: $\|x\| = \|\lambda(x)\| = \|\mu(\Phi(x))\| = \|\Phi(x)\|$. ■

Proposition 4.5 (sharpened Cauchy-Schwarz) Suppose p is complete. Then

$$\langle x, y\rangle \leq \langle \lambda(x), \lambda(y)\rangle \leq \|x\|\|y\|, \qquad \text{for all } x, y \in X.$$

Proof. By the Cauchy-Schwarz inequality in $\mathbb{R}^m$ and Corollary 3.11.(ii), $2\langle\lambda(x), \lambda(y)\rangle \geq \|\lambda(x+y)\|^2 - \|\lambda(x)\|^2 - \|\lambda(y)\|^2 = \|x+y\|^2 - \|x\|^2 - \|y\|^2 = 2\langle x, y\rangle$. ■

The Hermitian matrices (continued). The inner product on the Hermitian matrices is precisely what one would expect: $\langle x, y\rangle = \text{trace}\,(xy)$. The sharpening of the Cauchy-Schwarz inequality is due to *von Neumann*; see [19, Theorem 2.2] and the discussion therein.

We can now refine Theorem 3.9.

Theorem 4.6 (strict convexity) Suppose p is complete and $f : \mathbb{R}^m \to [-\infty, +\infty]$ is strictly convex and symmetric. Then the composition $f \circ \lambda$ is strictly convex on X.

Proof. Fix $\alpha \in (0,1)$, $x, y \in X$ and set $\beta := 1 - \alpha$. Suppose that $(f \circ \lambda)(\alpha x + \beta y) = \alpha(f \circ \lambda)(x) + \beta(f \circ \lambda)(y)$. We have to show that $x = y$. Using Theorem 3.9 and convexity of f, we estimate

$$\begin{aligned}\alpha(f \circ \lambda)(x) + \beta(f \circ \lambda)(y) &= (f \circ \lambda)(\alpha x + \beta y) \\ &\leq f(\alpha\lambda(x) + \beta\lambda(y)) \\ &\leq \alpha(f \circ \lambda)(x) + \beta(f \circ \lambda)(y);\end{aligned}$$

hence equality must hold throughout. By strict convexity of f, we conclude that $\lambda(x) = \lambda(y)$. We also know that $\alpha\lambda(x) + \beta\lambda(y) = \lambda(\alpha x + \beta y)$ (otherwise, Theorem 3.9 would imply that the first displayed inequality is strict, which is a contradiction). Thus $\lambda(x) = \lambda(y) = \alpha\lambda(x) + \beta\lambda(y) = \lambda(\alpha x + \beta y)$. Since λ is norm preserving, we obtain $\|x\| = \|y\| = \|\alpha x + \beta y\|$. But $\|\cdot\|$ is induced by an inner product, whence $\|\cdot\|^2$ is strictly convex. Therefore, $x = y$ and the proof is complete. ■

We now demonstrate how Theorem 4.6 can be used to recover a recent result by Krylov (see [17, Theorem 6.4.(ii)]). Our proof appears to be more transparent than Krylov's.

Corollary 4.7 Suppose $p(d) > 0$. Then each of the following functions is convex on the hyperbolicity cone $C(d)$:

$$-\ln p, \quad \ln \frac{E_{m-1} \circ \lambda}{E_m \circ \lambda}, \quad \frac{E_{m-1} \circ \lambda}{E_m \circ \lambda}.$$

If p is complete, then each of these functions is strictly convex.

Proof. Define first $f(u) := -\ln p(d) - \sum_{i=1}^m \ln u_i$ on $\mathbb{R}^m_{++}$ and $F(x) := -\ln p(x)$ on $C(d)$. Then f is strictly convex and symmetric. Since $p(x) = p(d)E_m(\lambda(x))$, we have $F = f \circ \lambda$. It follows that F is convex (by Theorem 3.9), even strictly if p is complete (by Theorem 4.6). This proves the result for the first function. Now let $f := \ln(E_{m-1}/E_m)$ on $\mathbb{R}^m_{++}$ and $F := \ln \frac{E_{m-1} \circ \lambda}{E_m \circ \lambda}$ on $C(d)$. Then f is strictly convex by Corollary 2.24. By Theorem 3.9 (resp. Theorem 4.6), F is convex (resp. strictly convex, if p is complete). This yields the statement for the second function. Finally observe that the third function is obtained by taking the exponential of the second function. But this operation preserves (strict) convexity. ■

Krylov's result is closely related to parts of Güler's very recent work on hyperbolic barrier functions. We now give a simple proof of Güler's [10, Theorem 6.1]. The functions F and g below play a crucial role in interior point method, as they allow the construction of long-step interior-point methods using the hyperbolic barrier function F.

Corollary 4.8 Suppose $p(d) > 0$ and c belongs to the hyperbolicity cone $C := C(d)$. Define

$$F : C \to \mathbb{R} : x \mapsto -\ln(p(x)) \quad \text{and} \quad g : C \to \mathbb{R} : x \mapsto -(\nabla F(x))(c).$$

Then F and g are convex on C. If p is complete, then both F and g are strictly convex.

Proof. The statement on F is already contained in Corollary 4.7. Now let μ be the eigenvalue map corresponding to c. Then, by Fact 2.19, $p(x) = p(c)E_m(\mu(x))$ and $(\nabla p(x))(c) = p(c)E_{m-1}(\mu(x))$. Thus

$$g(x) = \frac{1}{p(x)}\big(\nabla p(x)\big)(c) = \frac{E_{m-1}(\mu(x))}{E_m(\mu(x))}.$$

Now argue as for the second function in the proof of Corollary 4.7. ■

The Hermitian matrices (continued). The statement on F corresponds to strict convexity of the function $x \mapsto -\ln\det(x)$ on the cone of positive semidefinite Hermitian matrices; this result is due to *Fan* [6].

Remark 4.9 It is worthwhile to point out that Krylov [17] and Güler derive their results from hyperbolic function theory whereas we here "piggyback" on inequalities in elementary symmetric functions. The latter approach is far more elementary.

We already pointed out that the trace σ_m is linear. With the notation introduced in Definition 4.1, we can express this more elegantly.

Proposition 4.10 (trace) $\sigma_m(x) = \langle d, x\rangle$, for every $x \in X$.

Proof. Fix $x \in X$. By Fact 2.8, $\|x\pm d\|^2 = \sum_{i=1}^m(\lambda_i(x\pm d))^2 = \sum_{i=1}^m(\lambda_i(x)\pm 1)^2 = \|x\|^2 \pm 2\sigma_m(x) + m$. So $4\langle x, d\rangle = \|x+d\|^2 - \|x-d\|^2 = 4\sigma_m(x)$. ■

5 Convex calculus

Definition 5.1 (isometric hyperbolic polynomial) We say p is *isometric (with respect to d)*, if for all $y, z \in X$, there exists $x \in X$ such that

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

Isometricity depends only on equivalence classes of triples:

Proposition 5.2 Suppose (X, p, d) is isomorphic to (Y, q, e). Then p is isometric if and only if q is.

Proof. Immediate from Proposition 2.5.(iii). ■

It is clear that if p is isometric, then $\operatorname{ran}\lambda$ is a closed convex cone contained in $\mathbb{R}^m_{\downarrow}$. The next example shows that the range of λ may be nonconvex in general.

Example 5.3 (a hyperbolic polynomial that is not isometric) If the polynomial $p(x) = x_1x_2x_3$ is defined on $X = \operatorname{span}\big\{(1,1,1),(3,1,0)\big\}$, then p is hyperbolic of degree $m = 3$ with respect to $d = (1,1,1)$. Hence $\lambda(x) = x_{\downarrow}$ and p is complete. It follows that for all $\alpha, \beta \in \mathbb{R}$,

$$\lambda\big(\alpha(1,1,1) + \beta(3,1,0)\big) = \begin{cases} \alpha(1,1,1) + \beta(3,1,0), & \text{if } \beta \geq 0; \\ \alpha(1,1,1) + \beta(0,1,3), & \text{otherwise.} \end{cases}$$

Since $\lambda(3,1,0) + \lambda(-3,-1,0) = (3,0,-3) \notin \operatorname{ran}\lambda$, the set $\operatorname{ran}\lambda$ is a closed *nonconvex* cone in $\mathbb{R}^3_{\downarrow}$. In particular, p is not isometric.

Unless stated otherwise, we assume from now on that

$$\boxed{p \text{ is complete, with corresponding inner product } \langle\cdot,\cdot\rangle \text{ and norm } \|\cdot\|.}$$

We chose the name "isometric" because of the equivalent condition (iii) in the following proposition.

Proposition 5.4 The following are equivalent:

(i) p is isometric.

(ii) $\max_{x:\lambda(x)=u}\langle x, y\rangle = \langle u, \lambda(y)\rangle$, for all $u \in \operatorname{ran}\lambda$ and every $y \in X$.

(iii) $d(u, \lambda(y)) = d(\lambda^{-1}(u), y)$, for all $u \in \operatorname{ran}\lambda$ and every $y \in X$.

Proof. "(i)⇒(ii)": fix $u \in \operatorname{ran}\lambda$ and $y \in X$. If $x \in X$ with $\lambda(x) = u$, then $\langle x, y\rangle \leq \langle u, \lambda(y)\rangle$ by Proposition 4.5. Since $u \in \operatorname{ran}\lambda$, there exists $z \in X$ such that $\lambda(z) = u$. By isometricity of p, there exists $x \in X$ such that $\lambda(x) = u$ and $\lambda(x + y) = \lambda(x) + \lambda(y)$. Now Corollary 3.11.(ii) and its condition for equality implies that $\langle x, y\rangle = \langle u, \lambda(y)\rangle$ and (ii) follows.
"(ii)⇒(iii)": fix $u \in \operatorname{ran}\lambda$ and $y \in X$. If $x \in X$ with $\lambda(x) = u$, then (by Proposition 4.5) $\|u - \lambda(y)\|^2 = \|x\|^2 + \|y\|^2 - 2\langle\lambda(x), \lambda(y)\rangle \leq \|x\|^2 + \|y\|^2 - 2\langle x, y\rangle = \|x - y\|^2$ and hence $d(u, \lambda(y)) \leq d(\lambda^{-1}(u), y)$. Equality follows if we pick $\bar{x} \in X$ such that $\lambda(\bar{x}) = u$ and $\langle \bar{x}, y\rangle = \langle u, \lambda(y)\rangle$.
"(iii)⇒(i)": fix $y, z \in X$. Set $u := \lambda(z)$. Note that $d(\lambda^{-1}(u), y)$ is attained, since the closed set $\{x \in X : \lambda(x) = u\}$ is contained in $\{x \in X : \|x\| = \|u\|\}$, which is compact. So pick $x \in X$ with $\lambda(x) = u = \lambda(z)$ and $\|u - \lambda(y)\| = \|x - y\|$. Squaring and simplifying yields $\langle\lambda(x), \lambda(y)\rangle = \langle x, y\rangle$. Now Corollary 3.11.(ii) and its condition for equality yields $\lambda(x+y) = \lambda(x)+\lambda(y)$. Hence p is isometric. ■

The Hermitian matrices (continued). Here $\operatorname{ran}\lambda = \mathbb{R}^m_{\downarrow}$ and $p = \det$ is isometric (we will discuss this in Section 6).

Theorem 5.5 (Fenchel conjugacy) Suppose that $f : \mathbb{R}^m \to (-\infty, +\infty]$ is symmetric. Then $(f \circ \lambda)^* \leq f^* \circ \lambda$. If p is isometric and $f(P_{\operatorname{ran}\lambda} u) \leq f(u)$, for every $u \in (\operatorname{dom} f)_{\downarrow}$, then $(f \circ \lambda)^* = f^* \circ \lambda$.

Proof. Fix an arbitrary $y \in X$. Then, using Proposition 4.5, symmetry of f, and the Hardy-Littlewood-Pólya inequality (see [11, Section 10.2]), the inequality follows from

$$\begin{aligned} f^*(\lambda(y)) &= \sup_{u \in \mathbb{R}^m}\big\{\langle u, \lambda(y)\rangle - f(u)\big\} = \sup_{u \in \mathbb{R}^m_{\downarrow}}\big\{\langle u, \lambda(y)\rangle - f(u)\big\} \\ &\geq \sup_{u \in \operatorname{ran}\lambda}\ \max_{x:\lambda(x)=u}\big\{\langle x, y\rangle - f(\lambda(x))\big\} = \sup_{x \in X}\big\{\langle x, y\rangle - (f \circ \lambda)(x)\big\} \\ &= (f \circ \lambda)^*(y). \end{aligned}$$

Now assume that p is isometric and $f(P_{\operatorname{ran}\lambda} u) \leq f(u)$, for every $u \in (\operatorname{dom} f)_{\downarrow}$. Fix momentarily an arbitrary $u \in \mathbb{R}^m_{\downarrow}$. Then, on the one hand, $f(P_{\operatorname{ran}\lambda} u) \leq f(u)$ (if $u \in (\operatorname{dom} f)_{\downarrow}$, then the inequality follows by assumptions; otherwise, the inequality is trivial). Since $\operatorname{ran}\lambda$ is a closed convex cone that contains $\lambda(y)+P_{\operatorname{ran}\lambda}u$, a well-known property of projections yields on the other hand $\langle u-P_{\operatorname{ran}\lambda}u,\lambda(y)\rangle \leq 0$. Altogether, $\langle u,\lambda(y)\rangle - f(u) \leq \langle P_{\operatorname{ran}\lambda}u,\lambda(y)\rangle - f(P_{\operatorname{ran}\lambda}u)$. Therefore, using Proposition 5.4,

$$\begin{aligned} f^*(\lambda(y)) &= \sup_{u\in\mathbb{R}^m_\downarrow}\big\{\langle\lambda(y),u\rangle - f(u)\big\} \leq \sup_{u'\in\operatorname{ran}\lambda}\big\{\langle\lambda(y),u'\rangle - f(u')\big\} \\ &= \sup_{x\in X}\big\{\langle y,x\rangle - f(\lambda(x))\big\} = (f\circ\lambda)^*(y). \quad \blacksquare \end{aligned}$$

The assumption that $f(P_{\operatorname{ran}\lambda}u)\leq f(u)$, for every $u\in(\operatorname{dom} f)_\downarrow$ is important: in Section 7, we present an isometric hyperbolic polynomial and a convex symmetric function f with $(f\circ\lambda)^*\neq f^*\circ\lambda$.

Corollary 5.6 Suppose p is isometric and $f:\mathbb{R}^m\to(-\infty,+\infty]$ is symmetric. Suppose one of the following conditions holds:

(i) $(\operatorname{dom} f)_\downarrow \subseteq \operatorname{ran}\lambda$.

(ii) $\operatorname{ran}\lambda = \mathbb{R}^m_\downarrow$.

(iii) f is convex and $P_{\operatorname{ran}\lambda}u \in \operatorname{conv}\{u_{\pi(i)} : \pi \text{ permutes } \{1,\dots,m\}\}$, for every $u\in(\operatorname{dom} f)_\downarrow$.

Then $(f\circ\lambda)^* = f^*\circ\lambda$.

Proof. (i) is clear from Theorem 5.5. (ii) is implied by (i). (iii): fix $u\in(\operatorname{dom} f)_\downarrow$ and write $P_{\operatorname{ran}\lambda}u = \sum_i \rho_i u^i$, where each ρ_i is nonnegative, $\sum_i\rho_i = 1$, and each u^i is some permutation of u. By convexity and symmetry of f, we conclude $f(P_{\operatorname{ran}\lambda}u)\leq f(u)$. Apply again Theorem 5.5. $\blacksquare$

Theorem 5.7 (subgradients) Suppose p is isometric, $\operatorname{ran}\lambda = \mathbb{R}^m_\downarrow$, and $f:\mathbb{R}^m\to(-\infty,+\infty]$ is convex and symmetric. Let $x,y\in X$. Then

$$y\in\partial(f\circ\lambda)(x) \text{ if and only if } \lambda(y)\in\partial f(\lambda(x)) \text{ and } \langle x,y\rangle = \langle\lambda(x),\lambda(y)\rangle.$$

Consequently, $\lambda\big[\partial(f\circ\lambda)(x)\big] = \partial f(\lambda(x))$.

Proof. Since $\operatorname{ran}\lambda=\mathbb{R}^m_\downarrow$, we have (Corollary 5.6.(ii)) $(f\circ\lambda)^* = f^*\circ\lambda$. In view of Proposition 4.5, the following equivalences hold true: $y\in\partial(f\circ\lambda)(x)$ $\Leftrightarrow$ $(f\circ\lambda)(x)+(f\circ\lambda)^*(y) = \langle x,y\rangle$ $\Leftrightarrow$ $f(\lambda(x))+f^*(\lambda(y)) = \langle\lambda(x),\lambda(y)\rangle$ and

$\langle x,y\rangle = \langle\lambda(x),\lambda(y)\rangle \Leftrightarrow \lambda(y)\in\partial f(\lambda(x))$ and $\langle x,y\rangle = \langle\lambda(x),\lambda(y)\rangle$. "Consequently": Clearly, by the above, $\lambda\big[\partial(f\circ\lambda)(x)\big]\subseteq\partial f(\lambda(x))$. Conversely, pick $v\in\partial f(\lambda(x))$. Then $f(\lambda(x))+f^*(v)=\langle v,\lambda(x)\rangle$. By the assumption that $\operatorname{ran}\lambda=\mathbb{R}^m$ and Proposition 5.4.(ii), $\langle v,\lambda(x)\rangle=\langle y,x\rangle$, for some y with $\lambda(y)=v$. Hence $(f\circ\lambda)(x)+(f\circ\lambda)^*(y)=\langle y,x\rangle$ and so $y\in\partial(f\circ\lambda)(x)$, which implies $v=\lambda(y)\in\lambda\big[\partial(f\circ\lambda)(x)\big]$. ■

The Hermitian matrices (continued). Theorem 5.7 corresponds to [19, Theorem 3.2].

Corollary 5.8 (differentiability) Suppose p is isometric, $\operatorname{ran}\lambda=\mathbb{R}^m_\downarrow$, and $f:\mathbb{R}^m\to(-\infty,+\infty]$ is convex and symmetric. Let $x,y\in X$. Then $f\circ\lambda$ is differentiable at x and $y=\nabla(f\circ\lambda)(x)$ if and only if f is differentiable at $\lambda(x)$ and $\{y'\in X:\lambda(y')=\nabla f(\lambda(x)),\langle x,y'\rangle=\langle\lambda(x),\lambda(y')\rangle\}=\{y\}$.

Proof. Clear from Theorem 5.7. ■

Corollary 5.9 (variational description of σ_k) Let p be isometric, and suppose $\operatorname{ran}\lambda=\mathbb{R}^m_\downarrow$. Let $1\leq k\leq m$. Then for every $x\in X$,

$$\sigma_k(x)=\max_{y:\lambda(y)\geq 0,\sigma_m(y)=k,\lambda_1(y)\leq 1}\langle x,y\rangle$$

and $\partial\sigma_k(x)=\{y\in X:\langle x,y\rangle=\sigma_k(x),\lambda(y)\geq 0,\sigma_m(y)=k,\lambda_1(y)\leq 1\}$.

Proof. Define $f(u):=\max_{i_1<\cdots<i_k}\sum_{l=1}^k u_{i_l}$. Then f is symmetric and convex on $\mathbb{R}^m$ and f^* is the indicator function of $\{u\in\mathbb{R}^m:\sum_i u_i=k$ and each $0\leq u_i\leq 1\}$. Now $\sigma_k=f\circ\lambda$ and so Corollary 5.6 yields $\sigma_k^*=f^*\circ\lambda$. Thus $y\in\partial\sigma_k(x)\Leftrightarrow x\in\partial\sigma_k^*(y)\Leftrightarrow\langle x,y\rangle=\sigma_k(x)$, $\lambda(y)\geq 0$, $\sigma_m(y)=k$, and $\lambda_1(y)\leq 1$. ■

The Hermitian matrices (continued). Corollary 5.9 is a direct generalization of the variational formulations due to *Rayleigh* and *Ky Fan*; see [12, Section 2] for more details.

6 Diagonalization

We uphold the assumption that

$$\boxed{p \text{ is complete, with corresponding inner product } \langle\cdot,\cdot\rangle \text{ and norm } \|\cdot\|.}$$

Definition 6.1 (invariance group) Let $\mathcal{L} = \mathcal{L}(X)$ be the *(general) linear group on* X (with composition), endowed with the natural topology of pointwise convergence, and let $\mathcal{O} = \mathcal{O}(X) := \{A \in \mathcal{L}(X) : A^* = A^{-1}\}$ be the *orthogonal group.* We define the *invariance group* by

$$\mathcal{G} = \mathcal{G}(p) := \{A \in \mathcal{L} : \lambda \circ A = \lambda\}.$$

Since $\lambda(x) = \lambda(y)$ precisely when $p(x+td) = p(y+td)$ (equality as polynomials in t), it follows that $A \in \mathcal{G}$ if and only if $A \in \mathcal{L}$ and $p(x+td) = p(Ax+td)$, for all $t \in \mathbb{R}$ and every $x \in X$.

Proposition 6.2 $\mathcal{G}$ is a closed subgroup of $\mathcal{O}$.

Proof. Elementary. ■

The Hermitian matrices (continued). The invariance group $\mathcal{G}$ contains all unitary similarity transformations $x \mapsto u^*xu$ for Hermitian x and unitary u. (This actually describes the entire invariance group; see Section 7 below.)

Definition 6.3 (diagonalizability) We say that p *allows diagonalization (with respect to* d*)*, if there is some linear isometry Δ from $\operatorname{span}\operatorname{ran}\lambda$ to X such that

$$\text{for every } x \in X, \text{ there exists } A \in \mathcal{G} \text{ with } x = A\Delta\lambda(x).$$

We refer to Δ as a *diagonalizing map (of* p *with respect to* d*).*

It follows readily that $\lambda \circ \Delta \circ \lambda = \lambda$.

The Hermitian matrices (continued). The polynomial $p = \det$ allows diagonalization and a diagonalizing map is $\Delta := \operatorname{Diag}$, which sends a vector in $\mathbb{R}^m$ to the corresponding diagonal matrix.

Proposition 6.4 Suppose (X, p, d) is isomorphic to (Y, q, e) by Φ and let μ be the eigenvalue map corresponding to q. Then:

(i) $\mathcal{L}(Y)\Phi = \Phi\mathcal{L}(X)$ and $\mathcal{G}(q)\Phi = \Phi\mathcal{G}(p)$.

(ii) Suppose Δ is a diagonalizing map of p, $x \in X$, $y = \Phi x$, and $A \in \mathcal{G}(p)$ with $x = A\Delta\lambda(x)$. Then $\Phi\Delta$ is a diagonalizing map of q, $\Phi A\Phi^* \in \mathcal{G}(q)$, and $y = (\Phi A\Phi^*)(\Phi\Delta)\mu(y)$.

(iii) p allows diagonalization if and only if q does.

Proof. Denote the eigenvalue map of q by μ and recall that $\lambda = \mu \circ \Phi$ (Proposition 2.5.(iii)) and that Φ is a surjective isometry from X to Y (Proposition 4.4). Hence $\Phi^* = \Phi^{-1}$ and (i) follows. (ii): Apply Φ to both sides of $x = A\Delta\lambda(x)$ and use (i). (iii) follows from (ii) by symmetry. ■

Theorem 6.5 If p allows diagonalization, then it is isometric.

Proof. Let Δ be a diagonalizing map, fix $y \in X$ and $u \in \operatorname{ran}\lambda$, say $u = \lambda(z)$, for some $z \in X$. Obtain $A \in \mathcal{G}$ such that $y = A\Delta\lambda(y)$. Set $x := A\Delta\lambda(z)$. Then $\lambda(x) = u$ and $\langle x, y\rangle = \langle A\Delta\lambda(z), A\Delta\lambda(y)\rangle = \langle u, \lambda(y)\rangle$. Therefore, by Proposition 5.4, p is isometric. ■

Theorem 6.6 Suppose p allows diagonalization and let Δ be a diagonalizing map. Let $x, y \in X$. Then $\langle x, y\rangle = \langle \lambda(x), \lambda(y)\rangle$ if and only if there exists $A \in \mathcal{G}$ such that simultaneously $x = A\Delta\lambda(x)$ and $y = A\Delta\lambda(y)$.

Proof. "$\Leftarrow$": $\langle \lambda(x), \lambda(y)\rangle = \langle A\Delta\lambda(x), A\Delta\lambda(y)\rangle = \langle x, y\rangle$.
"$\Rightarrow$": Pick $A \in \mathcal{G}$ such that $x + y = A\Delta\lambda(x+y)$. Then $\langle \lambda(x), \lambda(x+y)\rangle = \langle A\Delta\lambda(x), A\Delta\lambda(x+y)\rangle = \langle A\Delta\lambda(x), x+y\rangle$. Using this and Proposition 4.5, we estimate

$$\begin{aligned}
\|x - A\Delta\lambda(x)\|^2 &= \|x\|^2 + \|A\Delta\lambda(x)\|^2 - 2\langle x, A\Delta\lambda(x)\rangle \\
&= 2\|x\|^2 - 2\langle x+y, A\Delta\lambda(x)\rangle + 2\langle y, A\Delta\lambda(x)\rangle \\
&= 2\|x\|^2 - 2\langle \lambda(x), \lambda(x+y)\rangle + 2\langle y, A\Delta\lambda(x)\rangle \\
&\le 2\|x\|^2 - 2\langle x, x+y\rangle + 2\langle \lambda(y), \lambda(x)\rangle \\
&= 2\|x\|^2 - 2\langle x, x+y\rangle + 2\langle y, x\rangle \\
&= 0.
\end{aligned}$$

Hence $x = A\Delta\lambda(x)$. By symmetry, $y = A\Delta\lambda(y)$. ■

The Hermitian matrices (continued). Theorem 6.6 becomes a classic criterion for simultaneous ordered spectral decomposition due to *Theobald* [34].

Corollary 6.7 Suppose p allows diagonalization and let Δ be a diagonalizing map. Suppose further $\operatorname{ran}\lambda = \mathbb{R}^m_{\downarrow}$, and $f : \mathbb{R}^m \to (-\infty, +\infty]$ is convex and symmetric. Let $x, y \in X$. Then $y \in \partial(f \circ \lambda)(x)$ if and only if $\lambda(y) \in \partial f(\lambda(x))$ and there exists $A \in \mathcal{G}$ such that $x = A\Delta\lambda(x)$ and $y = A\Delta\lambda(y)$. Hence $\partial(f \circ \lambda)(x) = \{A\Delta\partial f(\lambda(x)) : x = A\Delta\lambda(x)\}$. Consequently, $f \circ \lambda$ is differentiable at x if and only if f is differentiable at $\lambda(x)$.

Proof. Combine Theorem 5.7 with Theorem 6.5 and Theorem 6.6. "Consequently": if $f \circ \lambda$ is differentiable at x and $x = A\Delta\lambda(x)$, for some $A \in \mathcal{G}$, then $A\Delta\partial f(\lambda(x))$ is singleton. Hence f is differentiable at $\lambda(x)$. Conversely, assume that f is differentiable at $\lambda(x)$. Then each element in the convex set $\partial(f \circ \lambda)(x)$ has the same norm, namely $\|\nabla f(\lambda(x))\|$. This can only happen when the set is a singleton and hence $f \circ \lambda$ must be differentiable at x. ■

The Hermitian matrices (conclusion). Corollary 6.7 recovers results recently established by Lewis, see [19, Section 3].

We conclude by connecting the present framework to Lewis's framework of normal decomposition systems [20]:

Theorem 6.8 (normal decomposition system) Suppose p allows diagonalization, and let Δ be a diagonalizing map. Set $\gamma := \Delta \circ \lambda$. Then X is a Euclidean space, $\mathcal{G}$ is a closed subgroup of $\mathcal{O}$, and γ is a selfmap of X. Also:

(i) $\gamma(Ax) = \gamma(x)$, for every $x \in X$ and all $A \in \mathcal{G}$.

(ii) For every point $x \in X$, there exists an operator $A \in \mathcal{G}$ with $x = A\gamma(x)$.

(iii) For all $x, y \in X$, the inequality $\langle x, y\rangle \leq \langle \gamma(x), \gamma(y)\rangle$ holds.

In other words, $(X, \mathcal{G}, \gamma)$ is a *normal decomposition system.*

Proof. Only (iii) is not immediately clear. Suppose $x, y \in X$. Then, using Proposition 4.5 and the fact that Δ is an isometry, we estimate $\langle x, y\rangle \leq \langle \lambda(x), \lambda(y)\rangle = \langle \Delta\lambda(x), \Delta\lambda(y)\rangle = \langle \gamma(x), \gamma(y)\rangle$. ■

7 Examples

7.1 $\mathbb{R}^n$

Consider the vector space

$$X = \mathbb{R}^n,$$

the polynomial

$$p(x) = \prod_{i=1}^{n} x_i,$$

and the direction

$$d = (1, 1, ..., 1).$$

Then p is hyperbolic and complete with eigenvalue map

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

The induced norm and inner product in X are just the standard Euclidean ones in $\mathbb{R}^n$. The invariance group $\mathcal{G}$ is the set of all linear transformations of the form

$$\mathcal{G} = \{(x_1, x_2, ..., x_n) \mapsto (x_{\pi(1)}, x_{\pi(2)}, ..., x_{\pi(n)}) \,|\, \pi \text{ is a permutation}\}.$$

Clearly $\mathcal{G}$ is isomorphic to the symmetric group S_n. We have $\mathrm{ran}\lambda = \mathbb{R}^n_{\downarrow}$ and so we can choose a diagonalizing map $\Delta :$ span ran $\lambda \mapsto X$ to be the identity map. Hence, by Theorem 6.5, p is isometric. In this case the sharpened Cauchy-Schwarz inequality (Proposition 4.5) reduces to the well-known Hardy-Littlewood-Pólya inequality (see [11, Chapter X]).

$$x^T y \leq x_{\downarrow}^T y_{\downarrow}$$

and Theorem 6.6 shows equality holds if and only if the vectors x and y can be simultaneously ordered with the same permutation. Since $\mathrm{ran}\lambda = \mathbb{R}^n_{\downarrow}$, Corollary 5.6 shows that for every symmetric function $f : \mathbb{R}^n \to (-\infty, +\infty]$ we have

$$(f \circ \lambda)^* = f^* \circ \lambda.$$

Also Lidskii's Theorem holds, because $\lambda(x)$ is the ordered set of eigenvalues of the symmetric matrix $\mathrm{Diag}(x)$ (see [2, page 69]).

7.2 Hermitian matrices

In this section we summarize the example we have followed throughout the paper. Consider the vector space H^n (of $n \times n$ Hermitians matrices), and denote the ordered eigenvalues of a matrix $x \in H^n$ by $\tilde{\lambda}_1(x) \geq \tilde{\lambda}_2(x) \geq ... \geq \tilde{\lambda}_n(x)$. In the case of Hermitian matrices, the Frobenius [14, page 291] norm can be defined by $\|x\|_F = \|\tilde{\lambda}(x)\|$, where the last norm is the standard Euclidean norm in $\mathbb{R}^n$. Let

$$X = H^n,$$

the polynomial be

$$p(x) = \det x,$$

and the direction be

$$d = I.$$

Then p is hyperbolic and complete with eigenvalue map

$$\lambda(x) = \tilde{\lambda}(x).$$

The induced norm and inner product in X are given by

$$\|x\|^2 = \|x\|_F^2,$$
$$\langle x, y\rangle = \operatorname{tr} xy.$$

The invariance group $\mathcal{G}$ consists of all linear transformations on X that preserve the eigenvalues of every $x \in X$. Then from [27, Theorem 4] it follows that

$$\mathcal{G} = \{x \mapsto u^* x u,\ x \mapsto u^* x^T u \,|\, u \,\text{unitary}\}.$$

Clearly we have $\operatorname{ran}\lambda = \mathbb{R}^n_{\downarrow}$. We can choose a diagonalizing map Δ : span ran $\lambda \mapsto X$ to be

$$\Delta((x_1, x_2, ..., x_n)) = \text{Diag }(x_1, x_2, ..., x_n).$$

Hence, by Theorem 6.5, p is isometric. In this case the sharpened Cauchy-Schwarz inequality (Proposition 4.5) reduces to Fan's inequality:

$$\operatorname{tr} x^T y \le \tilde{\lambda}(x)^T \tilde{\lambda}(y)$$

and Theorem 6.6 shows equality holds if and only if the matrices x and y can be simultaneously unitarily diagonalized (with eigenvalues in decreasing order), which is due to Theobald. Since $\operatorname{ran}\lambda = \mathbb{R}^n_{\downarrow}$, Corollary 5.6 implies that for every symmetric function $f : \mathbb{R}^n \to (-\infty, +\infty]$ we have

$$(f \circ \lambda)^* = f^* \circ \lambda.$$

It is well known that Lidskii's theorem holds in this case (see[2, Section III.4]).
Note that there is an entirely analogous example on the space of n by n real symmetric matrices.

7.3 Singular values

Consider the vector space $M_{n,m}$ (of n by m real matrices). We assume $m \le n$ and denote the singular values of a matrix x in $M_{n,m}$ by $\sigma_1(x) \ge \sigma_2(x) \ge ... \ge \sigma_m(x)$. The Frobenius norm [14, page 291 & page 421] is defined by

$\|x\|_F = \|\sigma(x)\|$, where the last norm is the standard Euclidean norm in $\mathbb{R}^n$, and $\sigma(x) = (\sigma_1(x), \sigma_2(x), ..., \sigma_m(x))$. Now consider the vector space

$$X = M_{n,m} \times \mathbb{R},$$

the polynomial

$$p(x, \alpha) = \det(\alpha^2 I_m - x^T x) \quad (x \in M_{n,m}, \alpha \in \mathbb{R}),$$

and the direction

$$d = (0, 1).$$

Then p is hyperbolic and complete with eigenvalue map

$$\lambda(x, \alpha) = (\alpha + \sigma_1(x), \alpha + \sigma_2(x), ..., \alpha - \sigma_2(x), \alpha - \sigma_1(x)).$$

The induced norm and inner product are given by

$$\|(x, \alpha)\|^2 = 2m\alpha^2 + 2\|x\|_F^2,$$
$$\langle (x, \alpha), (x, \beta) \rangle = 2m\alpha\beta + 2\mathrm{tr}\, x^T y,$$

for (x, α) and (y, β) in X. It is not difficult to see that if $T \in \mathcal{G}$ then $T(x, \alpha) = (\tilde{T}x, \alpha)$, where $\tilde{T}$ is a linear operator on $M_{n,m}$ preserving all singular values. Then from the main theorem in [26] or in [24], it follows that the invariance group $\mathcal{G}$ is:

$$\mathcal{G} = \begin{cases} \{(x, \alpha) \mapsto (uxv, \alpha) \,|\, u, v \text{ orthogonal}\}, & \text{(m<n)} \\ \{(x, \alpha) \mapsto (uxv, \alpha),\ (x, \alpha) \mapsto (ux^T v, \alpha) \,|\, u, v \text{ orthogonal}\}, & \text{(m=n)} \end{cases}$$

The span of ranλ decomposes as a direct sum:

$$\text{span ran } \lambda = \mathbb{R}e + \{s \in \mathbb{R}^{2m} \,|\, s_{2m-i+1} = -s_i,\ \forall i\},$$

where the vector $e \in \mathbb{R}^{2m}$ has all components 1, and then we can choose a diagonalizing map Δ: span ran $\lambda \mapsto X$ to be

$$\Delta(\alpha e + s) = (\mathrm{Diag}(s_1, s_2, ..., s_m), \alpha),$$

where $(\mathrm{Diag}(s_1, s_2, ..., s_m))_{i,j} = s_i$ if $i = j$ and 0 otherwise. Hence, by Theorem 6.5, p is isometric. Notice that in this case the sharpened Cauchy-Schwarz inequality (Proposition 4.5) reduces to

$$\mathrm{tr}\, x^T y \leq \sigma(x)^T \sigma(y),$$

and Theorem 6.6 shows equality holds if and only if x and y have a simultaneous 'ordered' singular value decomposition (that is, there are unitary matrices u and v such that $x = u(\mathrm{Diag}\,\sigma(x))v$ and $y = u(\mathrm{Diag}\,\sigma(y))v$). This is the classical result known as 'von Neumann's Lemma' (see for example [15, page 182]).
Note that when $m = 1$ we get the Lorentz Cone example which is discussed below. An analogous example can be obtained by considering the vector space $X = \mathbb{C}_{n,m} \times \mathbb{R}$.

***

We now show that for some functions in the singular value case we have $(f \circ \lambda)^* \neq f^* \circ \lambda$. Consider the symmetric function

$$f(u) = \max_{1 \leq i \leq m} u_i.$$

Then

$$f^*(v) = \begin{cases} 0, & \sum_{i=1}^m v_i = 1,\ v_i \geq 0 \\ +\infty, & \text{else} \end{cases}$$

Now let $m = 2$. Then $\mathrm{ran}\lambda = \{\alpha e + (\beta, \gamma, -\gamma, -\beta) \,|\, \beta \geq \gamma \geq 0\}$. Let $v = \frac{1}{4}(3, 1, 1, -1) \in \mathrm{ran}\,\lambda$. Let $y \in X$ be such that $\lambda(y) = v$. It is straightforward to check that $\langle \lambda(z), \lambda(y) \rangle = \lambda_1(z)\ \forall\, z \in$ X. It follows from the sharpened Cauchy-Schwarz inequality (Proposition 4.5) that $\langle z, y \rangle \leq \lambda_1(z)\ \forall\, z \in X$. Then

$$(f \circ \lambda)^*(y) = \lambda_1^*(y) = \sup_{z \in X} \{\langle z, y \rangle - \lambda_1(z)\} = 0.$$

On the other hand clearly

$$(f^* \circ \lambda)(y) = f^*(v) = +\infty.$$

***

In this subsection we show Lidskii's theorem holds for this example. So we want to show that for all $(x, \alpha), (y, \beta) \in X$

$$w^T \left(\lambda(x + y, \alpha + \beta) - \lambda(x, \alpha)\right) \leq w_\downarrow^T \lambda(y, \beta)\ \ \forall w \in \mathbb{R}^{2m}.$$

This is equivalent to

$$w^T \left((\sigma(x + y), (-\sigma(x + y))_\downarrow) - (\sigma(x), (-\sigma(x))_\downarrow)\right) \leq w_\downarrow^T (\sigma(y), (-\sigma(y))_\downarrow),$$

for all $w \in \mathbb{R}^{2m}$. Now assume $n = m$ and let Λ be the eigenvalue map (ordered decreasingly) in H^{2m}. But we have (see [14, Theorem 7.3.7])

$$\Lambda \begin{pmatrix} 0 & x \\ x^T & 0 \end{pmatrix} = \begin{pmatrix} \sigma(x) \\ (-\sigma(x))_\downarrow \end{pmatrix},$$

so the above inequality is equivalent to

$$w^T \left(\Lambda \begin{pmatrix} 0 & x+y \\ (x+y)^T & 0 \end{pmatrix} - \Lambda \begin{pmatrix} 0 & x \\ x^T & 0 \end{pmatrix} \right) \leq w_\downarrow^T \Lambda \begin{pmatrix} 0 & y \\ y^T & 0 \end{pmatrix},$$

for all w in $\mathbb{R}^{2m}$, which is true by Lidskii's Theorem in H^{2m}. Hence Lidskii's theorem holds when $n = m$.

7.4 Absolute reordering

Consider the vector space

$$X = \mathbb{R}^n \times \mathbb{R}$$

Let the polynomial be

$$p(x, \alpha) = \prod_{i=1}^{n} (\alpha^2 - x_i^2),$$

and the direction be

$$d = (0, 1).$$

Then p is hyperbolic and complete with eigenvalue map

$$\lambda(x, \alpha) = (|x|_\downarrow, (-|x|)_\downarrow) + \alpha e,$$

where $|x| = (|x_1|, |x_2|, \ldots, |x_n|)$, and $e = (1, 1, \ldots, 1) \in \mathbb{R}^{2n}$. If $\|x\|_2$ denotes the standard Euclidean norm in $\mathbb{R}^n$, then the induced norm and inner product in X are given by

$$\|(x, \alpha)\|^2 = 2\|x\|_2^2 + 2n\alpha^2,$$

$$\langle (x, \alpha), (y, \beta) \rangle = 2 \sum_{i=1}^{n} x_i y_i + 2n\alpha\beta.$$

The invariance group $\mathcal{G}$ is

$$\mathcal{G} = \{(x, \alpha) \mapsto (P_{(-)} x, \alpha) \mid P_{(-)} \text{ is a signed permutation matrix}\},$$

where a signed permutation matrix has only one nonzero entry in each row and column which is either $+1$ or -1. The span of ranλ decomposes as a direct sum:

$$\text{span ran } \lambda \;=\; \mathbb{R}e + \{s \in \mathbb{R}^{2n}_{\downarrow} \,|\, s_{2n-i+1} = -s_i, \;\forall i\}.$$

We can choose a diagonalizing map Δ: span ran $\lambda \mapsto X$ to be

$$\Delta(\alpha e + s) \;=\; ((s_1, s_2, \ldots, s_n), \alpha)\,.$$

Hence, by Theorem 6.5, p is isometric. In this case the sharpened Cauchy-Schwarz inequality (Proposition 4.5) reduces to the well-known inequality (see [20, section 7])

$$x^T y \;\leq\; |x|_{\downarrow}^T |y|_{\downarrow}$$

and Theorem 6.6 shows equality holds if and only if the vectors x and y can be simultaneously ordered with the same signed permutation.
Note that the similarities with the previous example are not accidental. It corresponds to the subspace (Diag $\mathbb{R}^n$) $\times \mathbb{R}$ of $M_{n,m} \times \mathbb{R}$. So we can immediately see that for some functions f we have $(f \circ \lambda)^* \neq f^* \circ \lambda$. Also because $|x|_{\downarrow} \;=\; \sigma(\text{Diag}(x))$, one sees, from the corresponding part in the previous example, that Lidskii's Theorem holds.

7.5 Lorentz cone

Let the vector space be

$$X = \mathbb{R}^n,$$

and the polynomial be

$$p(x) = x^T A x = x_1^2 - x_2^2 - \cdots - x_n^2,$$

where $A \;=\; \text{Diag}(1, -1, -1, ..., -1) \;\in\; M_n$ ($n \times n$ real matrices). Let the direction be

$$d = (d_1, d_2, ..., d_n) \;\in\; X \;\text{ such that } d_1^2 > d_2^2 + \cdots + d_n^2.$$

Then p is hyperbolic and complete with eigenvalue map

$$\lambda(x) = \left(\frac{x^T A d + \sqrt{D(x)}}{p(d)}, \frac{x^T A d - \sqrt{D(x)}}{p(d)} \right),$$

where $D(x) = (x^TAd)^2 - p(x)p(d)$ is the discriminant of $p(x+td)$ considered as a quadratic polynomial in t. (The fact that $D(x) \geq 0$ for each x, and so that $p(x)$ is hyperbolic, is the well-known Aczel inequality, see [30, p.57].) The induced norm and inner product are given by

$$\|x\|^2 = 2\frac{2(x^TAd)^2 - p(x)p(d)}{p(d)^2}, \quad \text{and}$$
$$\langle x, y\rangle = \frac{4(x^TAd)(y^TAd) - 2(x^TAy)p(d)}{p(d)^2},$$

for x and y in X. Immediately from the definition the invariance group $\mathcal{G}$ is

$$\mathcal{G} = \{B \in M_n \mid B^TAB = A \text{ and } B^TAd = Ad\}.$$

We now show that the mapping $\lambda : X \to \mathbb{R}^2_{\downarrow}$ is onto. Indeed, fix $(t_1, t_2) \in \mathbb{R}^2_{\downarrow}$, and let l be an arbitrary, fixed nonzero vector from $\{d\}^{\perp} \subset X$. (The reader can easily verify that $l \in \{d\}^{\perp}$ if and only if $l^TAd = 0$.) Set $\alpha := \frac{1}{2}(t_1 + t_2)$, and $v := \sqrt{-\frac{p(d)}{p(l)}}\left(\frac{t_1 - t_2}{2}\right) l$. Then we have $\lambda(\alpha d + v) = (t_1, t_2)$. Clearly then

$$\text{span ran } \lambda = \mathbb{R}^2.$$

Above we have to make sure that $p(l) < 0$. Indeed, because the discriminant of $p(x)$ is always nonnegative we get that $p(l) \leq 0$. Suppose that $p(l) = 0$, then this together with $l^TAd = 0$, and $d^tAd > 0$ gives us the three relations: $l_1^2 = \bar{l}^T\bar{l}$; $d_1l_1 = \bar{d}^T\bar{l}$; $d_1^2 > \bar{d}^T\bar{d}$, where we have used the notation $\bar{x} = (x_2, ..., x_n)$, and the dot product in the relations is the usual one in $\mathbb{R}^{n-1}$. Notice that $\bar{l} \neq 0$ or otherwise $l = 0$. Then from the Cauchy-Schwarz inequality we get: $|d_1l_1|^2 = |\bar{d}^T\bar{l}|^2 \leq |\bar{d}^T\bar{d}||\bar{l}^T\bar{l}| < d_1^2l_1^2$, contradiction.
We can choose a diagonalizing map Δ: span ran $\lambda \mapsto X$ to be

$$\Delta(u_1, u_2) = \frac{u_1 + u_2}{2}\, d + \sqrt{-\frac{p(d)}{p(l)}}\left(\frac{u_1 - u_2}{2}\right) l,$$

where again, l is an arbitrary, fixed, nonzero vector from $\{d\}^{\perp} \subset X$. (The reader can easily verify that Definition 6.3 of the map Δ is satisfied.) Hence, by Theorem 6.5, p is isometric. Notice that in this case the sharpened Cauchy-Schwarz inequality (Proposition 4.5) becomes

$$(x^TAd)(y^TAd) - (x^TAy)p(d) \leq \sqrt{D(x)D(y)},$$

and Theorem 6.6 gives the necessary and sufficient condition for equality. Let $f = (1, 0, ..., 0) \in \mathbb{R}^n$. The fact that Lidskii's Theorem holds for the polynomial $p(x)$ in the direction f is clear from the corresponding discussion in section 7.3. For arbitrary direction, note that, by [7, pages 7-8], we have that the hyperbolicity cone $C(p, d) = \{x \in \mathbb{R}^n | x_1^2 > x_2^2 + \cdots + x_n^2\}$ is *homogeneous*: that is, there is an orthogonal linear map $U : \mathbb{R}^n \to \mathbb{R}^n$ such that
1) $p(Ux) = p(x)$
2) $Ud = f$.
Hence *the triples* (X, p, d) *and* (X, p, f) *are isomorphic.* So from Proposition 3.7 we see that Lidskii's Theorem holds again.

***

We note that if Y is a subspace of H^s (for some positive integer s), $d \in Y$ and $d \succ 0$, then $q(y) = \det y$ is a hyperbolic polynomial over Y with respect to the direction d. Indeed $q(y + td) = \det(d) \det(d^{-\frac{1}{2}} y d^{-\frac{1}{2}} + tI)$ and all the eigenvalues of $d^{-\frac{1}{2}} y d^{-\frac{1}{2}}$ are real numbers because it is a hermitian matrix. Triples of this type will be called **standard hyperbolic triples**.

Many of our examples are isomorphic to a standard hyperbolic triple. For the example in section 7.1, consider the map $\phi(x) = \mathrm{Diag}(x)$. Then clearly $p(x) = \det \phi(x)$. For the example in section 7.2 it is clear. The example in section 7.3 is 'almost' isomorphic to a standard hyperbolic triple as well. Indeed, consider the mapping $\phi : M_{n,m} \times \mathbb{R} \to H^{n+m}$ defined by:

$$(x, \alpha) \longmapsto \begin{pmatrix} \alpha I_n & x \\ x^T & \alpha I_m \end{pmatrix},$$

then $\alpha^{n-m} p(x, \alpha) = \det \phi(x, \alpha)$.

If we consider (a slight variation) the hyperbolic polynomial

$$p(x, \alpha) = \det(\alpha^2 I - x x^T)$$

with respect to $d = (0, 1)$, where again $x \in X = M_{n,m} \times \mathbb{R}$. Then the mapping $\phi : M_{n,m} \times \mathbb{R} \to H^{2n}$ defined by:

$$(x, \alpha) \longrightarrow \begin{pmatrix} \alpha I_{n-m} & 0 & 0 \\ 0 & \alpha I_m & x^T \\ 0 & x & \alpha I_n \end{pmatrix}$$

gives an isomorphism between (X, p, d) and a standard hyperbolic triple. The fact that $p(x, \alpha) = \det \Phi(x, \alpha)$ follows from the identity:

$$\det \begin{pmatrix} \alpha I_{n-m} & 0 & 0 \\ 0 & \alpha I_m & x^T \\ 0 & x & \alpha I_n \end{pmatrix} =$$

$$\det \begin{pmatrix} I_{n-m} & 0 & 0 \\ 0 & I_m & 0 \\ 0 & -\frac{1}{\alpha}x & I_n \end{pmatrix} \begin{pmatrix} \alpha I_{n-m} & 0 & 0 \\ 0 & \alpha I_m & x^T \\ 0 & x & \alpha I_n \end{pmatrix} \begin{pmatrix} I_{n-m} & 0 & 0 \\ 0 & I_m & -\frac{1}{\alpha}x^T \\ 0 & 0 & I_n \end{pmatrix}$$

$$= \det \begin{pmatrix} \alpha I_{n-m} & 0 & 0 \\ 0 & \alpha I_m & 0 \\ 0 & 0 & \alpha I_n - \frac{1}{\alpha}xx^T \end{pmatrix} = \det(\alpha^2 I_n - xx^T).$$

When $\alpha = 0$ the conclusion of the above identity still holds, one just needs to consider the two cases $n = m$ and $n < m$ separately.

In general though it is not true that every hyperbolic triple is isomorphic to a standard hyperbolic triple: consider for example $X = \mathbb{R}^5$,

$$p(x) = x_0^2 - x_1^2 - x_2^2 - x_3^2 - x_4^2, \quad d = (1, 0, 0, 0, 0).$$

Suppose there is a linear isomorphism $\phi : X \to Y \subset H^s$, such that $p(x) = \det \phi(x)$, and $\phi(d) \succ 0$.
Because p is homogeneous of degree 2 we have $t^2 p(x) = p(tx) = \det \phi(tx) = \det t\phi(x) = t^s \det \phi(x)$. Hence we see that $s = 2$.
Because ϕ is linear, there are vectors $a, b, c, f \in \mathbb{R}^5$ such that for every $x \in \mathbb{R}^5$ we have

$$p(x) = \det \begin{pmatrix} a^T x & b^T x + ic^T x \\ b^T x - ic^T x & f^T x \end{pmatrix}.$$

There is a nonzero vector $x \in \mathbb{R}^5$ such that $x_0 = 0$, and $x \perp \operatorname{span}\{a, b, c\}$. So $0 \neq -\|x\|^2 = p(x) = \det \phi(x) = 0$, a contradiction. Of course this example doesn't disprove the conjecture made in [18], which concerns polynomials in only two variables.

7.6 The degree 2 case

In this section we show that every complete hyperbolic polynomial of degree two is isometric. Let the vector space be

$$X = \mathbb{R}^n.$$

We will assume that $p(x)$ is hyperbolic polynomial of degree two with respect to a vector d. Without loss of generality, we write

$$p(x) \;=\; x^T Ax,$$

where $A \in H^n$. Proposition 2.5 implies that if $S : X \to X$ is a nonsingular linear transformation, then $q(y) := p(Sy)$ is hyperbolic with respect to $l = S^{-1}d$.

Lemma 7.1 If $p(x) = x^t Ax$ is hyperbolic, then p is complete if and only if A is nonsingular.

Proof. Because of Fact 2.15, the linearity space of $p(x)$ in our case is $\{x \in X : (tx + y)^T A(tx + y) = y^T Ay, \forall y \in X, \forall t \in \mathbb{R}\} = \{x \in X : x^T Ax\, t^2 + 2x^T Ay\, t = 0 \ \forall y \in X, \forall t \in \mathbb{R}\} = \{x \in X : x^T Ax = 0 \text{ and } x^T Ay = 0 \ \forall y \in X\} = \{x \in X : Ax = 0\} = \{0\}$ iff A is nonsingular. ∎

Proposition 2.14 now says that if $p(x)$ is a complete hyperbolic polynomial with respect to d, and $S : X \to X$ is a nonsingular linear transformation, then $q(y) := p(Sy)$ is also a complete hyperbolic polynomial with respect to $l = S^{-1}d$.

Lemma 7.2 Let $p(x) = x^T Ax$ be a complete, hyperbolic polynomial, with respect to d of degree two. Then the symmetric matrix A is nonsingular and has exactly $(n-1)$ eigenvalues of one sign, and 1 eigenvalue with the opposite sign.

Proof. The nonsingularity of A follows from the previous lemma. Now, because $p(x)$ is hyperbolic with respect to d, we have that the discriminant of the quadratic function

$$t \mapsto (x + td)^T A(x + td),$$

$(d^T Ax)^2 - (d^T Ad)(x^T Ax)$ is nonnegative $\forall x \in X$. This inequality implies two things. First A cannot be positive definite because then the Cauchy-Schwarz inequality for the scalar product defined by A contradicts the nonnegativity of the discriminant. Similarly, A cannot be negative definite. Without loss of generality we can assume that that $d^T Ad > 0$, so for every x in the $(n-1)$-dimensional orthogonal complement (with respect to the usual inner product) of the vector Ad we have $0 \geq x^T Ax$. This implies that A has at least $(n-1)$ nonpositive eigenvalues, but none of them can be zero, so A has $(n-1)$ strictly negative eigenvalues. The last eigenvalue must be strictly positive, because A cannot be negative semidefinite. The case $d^T Ad < 0$ is handled analogously. ■

Now, Proposition 5.2 says that if $p(x)$ is an isometric, complete hyperbolic polynomial with respect to d, and $S : X \to X$ is a nonsingular linear transformation, then $q(y) := p(Sy)$ is also an isometric, complete, hyperbolic polynomial with respect to $l = S^{-1}d$.

Let $p(x) = x^T Ax$ be isometric with respect to d. Without loss of generality we can assume that $p(d) > 0$. By Sylvester's theorem in the linear algebra (see for example [14], Theorem 4.5.8), there exists a nonsingular transformation $x = Sy$ of the variable x such that $q(y) := p(Sy)$ has the form: $q(y) = y_1^2 - y_2^2 - \cdots y_n^2$. Moreover, from the above, $q(y)$ is hyperbolic with respect to $S^{-1}d$. Because the subsection about the Lorentz cone showed that $q(y) = y_1^2 - y_2^2 - \cdots - y_n^2$ is isometric with respect to any d in the hyperbolicity cone of q, and $C(q, l) = S^{-1}(C(p, d))$ we answered the question about isometricity for the whole class of hyperbolic polynomials of degree two.

7.7 Antisymmetric tensor powers

Consider the function $p(x) = \det x$ on the vector space of $n \times n$ real symmetric (or Hermitian) matrices, and let $q = E_k$ be the elementary symmetric function of order k and $p_k(x) = E_k \circ \lambda(x)$. We saw earlier that p_k is a hyperbolic polynomial with respect to the identity matrix I. We have

$$p_k(x) = \sum_{\alpha=(i_1<i_2<\cdots<i_k)} \det x[\alpha|\alpha] = \operatorname{tr}\left(\wedge^k x\right),$$

where $x[\alpha|\alpha]$ is the principal submatrix obtained from x by keeping its rows and columns $i_1, \ldots, i_k$, and the second equality above can be regarded as the definition of the symbol $\operatorname{tr}\left(\wedge^k x\right)$. For the first equality above one can see [26], and justification for the use the the symbol $\operatorname{tr}\left(\wedge^k x\right)$ can be found in the explanations below. Now, from Corollary 3.10, Definition 2.10 and from the fact that $p_k(x) = \operatorname{tr}\left(\wedge^k x\right)$ is a homogeneous hyperbolic polynomial, it follows immediately that

$$\mathrm{tr}\ \left(\wedge^k(x+y)\right)^{1/k} \geq \mathrm{tr}\ \left(\wedge^k x\right)^{1/k} + \mathrm{tr}\ \left(\wedge^k y\right)^{1/k},$$

when x, y are symmetric and positive definite. This is one of the main results in [25].

***

We give some preliminary definitions and facts about tensor spaces, they can be found in [9] and [2]. Let $E_1, \ldots, E_k$ be k copies of a given vector space E. Consider the free vector space $C = C(E_1 \times \cdots \times E_k)$ over the set $E_1 \times \cdots \times E_k$. Let $N = N(E_1, \ldots, E_k)$ be the subspace of C generated by the elements

$$(x_1, \ldots, \alpha y_1 + \beta y_2, \ldots, x_k) - \alpha(x_1, \ldots, y_1, \ldots, x_k) - \beta(x_1, \ldots, y_2, \ldots, x_k),$$

for all indexes i and for all $y_1, y_2 \in E_i$ and $x_j \in E_j \,\forall\, j \neq i$. Denote by π the canonical projection of C onto the space $G = C/N$ and define a mapping

$$\varphi : E_1 \times \cdots \times E_k \to C/N$$

by setting

$$\varphi(x_1, \ldots, x_k) = \pi(x_1, \ldots, x_k).$$

We call the pair (G, φ) a **tensor product** of $E_1, \ldots, E_k$ and we will denote $\varphi(x_1, \ldots, x_k)$ by $x_1 \otimes \cdots \otimes x_k$ - the tensor product of the vectors $x_1, \ldots, x_k$. The vector space G is sometimes denoted by $\otimes^k E$. The **antisymmetric tensor product** of vectors $x_1, \ldots, x_k \in E$ is defined and denoted by

$$x_1 \wedge \cdots \wedge x_k = (k!)^{\frac{1}{2}} \sum_\sigma \epsilon_\sigma x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(k)},$$

where σ runs over all permutations of the k indices and ϵ_σ is ± 1, depending on whether σ is an even or odd permutation. Clearly $x_1 \wedge \cdots \wedge x_k \in \otimes^k E$. The span of all antisymmetric tensors $x_1 \wedge \cdots \wedge x_k$ in $\otimes^k E$ is denoted by $\wedge^k E$ and is called the kth **antisymmetric tensor product** of E. If the vector space E is an Euclidean space then $\wedge^k E$ can also be made Euclidean by defining inner product as follows

$$\langle x_1 \wedge \cdots \wedge x_k, y_1 \wedge \cdots \wedge y_k \rangle = \det(\langle x_i, y_j \rangle)_{i,j=1}^{k,k},$$

and extending it linearly to the whole space $\wedge^k E$ (it is well defined!). If the vector space E has dimension n and $e_1, \dots, e_n$ is an orthonormal basis, then $e_{\mathcal{I}} = e_{i_1} \wedge \dots \wedge e_{i_k}$ is an orthonormal basis of $\wedge^k E$, where $\mathcal{I} \in \{(i_1, i_2, \dots, i_k) \,|\, 1 \le i_1 < i_2 < \dots < i_k \le n\}$. Moreover if we are given linear operator A on E it can be extended in a unique way to a linear operation on $\wedge^k E$ by

$$A(x_1 \wedge \dots \wedge x_k) = A(x_1) \wedge \dots \wedge A(x_k),$$

and extending it linearly to the whole space $\wedge^k E$ (it is again well defined!). The linear operator on $\wedge^k E$ induced by A will be denoted by $\wedge^k A$.

7.8 Unitary invariant norms

In this section we derive a well known von Neumann's theorem about unitary invariant norms as a consequence of the convexity results in this paper.

In 1937, von Neumann [35] gave a famous characterization of unitary invariant matrix norms (that is, norms f on $\mathbb{C}^{m\times n}$ satisfying $f(uxv) = f(x)$ for all unitary matrices u and v and matrices x in $\mathbb{C}^{m\times n}$). His result states that such norms are those functions of the form $g \circ \sigma$, where the map

$$x \in \mathbb{C}^{m\times n} \mapsto \sigma(x) \in \mathbb{R}^m$$

has components the singular values $\sigma_1(x) \ge \sigma_2(x) \ge \dots \ge \sigma_m(x)$ of x (assuming $m \le n$) and g is a norm on $\mathbb{R}^m$, invariant under sign changes and permutations of components. Proof of this can be found also in [14, Theorem 7.4.24].

For $x \in \mathbb{R}^m$, let $|x|_\downarrow$ have components $|x_i|$ arranged in decreasing order.

Lemma 7.3 For $x, y, \omega \in \mathbb{R}^m$, such that $\omega_1 \ge \omega_2 \ge \dots \ge \omega_m \ge 0$, and $\lambda \in [0,1]$, we have

$$\langle \omega, |\lambda x + (1-\lambda) y|_\downarrow \rangle \le \langle \omega, \lambda |x|_\downarrow + (1-\lambda)|y|_\downarrow \rangle.$$

Proof. Apply Theorem 2.4 and Example 7.2 from [20], with $X = \mathbb{R}^m$, G =signed permutation matrices, $\gamma(x) = |x|_\downarrow$. ■

Now define $H : \mathbb{R}^{2n} \to \mathbb{R}^n$ by

$$H(u) = \tfrac{1}{2}(v_1 + v_2, v_3 + v_4, \dots, v_{2n-1} + v_{2n}),$$

where $v = |u|_\downarrow$.

Lemma 7.4 For $u, v \in \mathbb{R}^{2n}$, $z \in \mathbb{R}^n$ such that $z_1 \geq z_2 \geq ... \geq z_n \geq 0$, and $\lambda \in [0, 1]$ we have

$$\langle z, H(\lambda u + (1-\lambda)v)\rangle \leq \langle z, \lambda H(u) + (1-\lambda)H(v)\rangle.$$

Proof. Apply Lemma 7.3 with $m = 2n$ and $\omega_{2i-1} = \omega_{2i} = z_i$. ■

Now suppose $g : \mathbb{R}^n \mapsto (-\infty, +\infty]$ is convex and absolutely symmetric (that is, $g(x) = g(|x|_\downarrow)$, $\forall x$).

Lemma 7.5 $g(H(\lambda u + (1-\lambda)v)) \leq \lambda g(H(u)) + (1-\lambda)g(H(v))$.

Proof. Apply Theorem 3.3 from [20] to Lemma 7.4. ■

Now define $f : \mathbb{R}^{2n} \mapsto (-\infty, +\infty]$ by $f(u) = g(H(u))$.

Lemma 7.6 The function f is absolutely symmetric and convex.

Proof. Notice that $H(|u|_\downarrow) = H(u)$. Consequently, $f(|u|_\downarrow) = g(H(|u|_\downarrow)) = g(H(u)) = f(u)$, $\forall u$. So f is absolutely symmetric. The convexity follows from Lemma 7.5. ■

Theorem 7.7 (von Neumann) The function $g \circ \sigma$ is convex.

Proof. Using Section 7.3 where $X = M_{n,m} \times \mathbb{R}$, $p(x, \alpha) = \det(\alpha^2 I - x^T x)$, and $d = (0, 1)$, we have that $\lambda(x, 0) = (\sigma_1(x), ..., \sigma_m(x), -\sigma_m(x), ..., -\sigma_1(x))$. So $H(\lambda(x, 0)) = \sigma(x)$. Then finally $g(\sigma(x)) = f(\lambda(x, 0))$, which, because of Theorem 3.9, is convex in x. ■

References

[1] E.F. BECKENBACH and R. BELLMAN. *Inequalities.* Springer-Verlag, 1961.

[2] R. BHATIA. *Matrix Analysis.* Springer-Verlag, 1997.

[3] P. BULLEN and M. MARCUS. Symmetric means and matrix inequalities. *Proceedings of the American Mathematical Society*, 12:285–290, 1961.

[4] P.S. BULLEN, D.S. MITRINOVIĆ, and P.M. VASIĆ. *Means and their inequalities.* Mathematics and its applications. East European Series. D. Reidel Publishing Company, Dordrecht, Holland, 1987.

[5] C. DAVIS. All convex invariant functions of hermitian matrices. *Archiv der Mathematik*, 8:276–278, 1957.

[6] K. FAN. On a theorem of Weyl concerning eigenvalues of linear transformations II. *Proceedings of the National Academy of Sciences of the United States of America*, 36:31–35, 1950.

[7] J. FARAUT and A. KORANYI. *Analysis on symmetric cones.* Oxford University Press, 1994.

[8] L. GÅRDING. An inequality for hyperpolic polynomials. *Journal of Mathematics and Mechanics*, 8(6):957–965, 1959.

[9] W. H. GREUB. *Multilinear Algebra.* Springer-Verlag, 1967.

[10] O. GÜLER. Hyperbolic polynomials and interior point methods for convex programming. *Mathematics of Operations Research*, 22(2):350–377, 1997.

[11] G. HARDY, J.E.LITTLEWOOD, and G. PÓLYA. *Inequalities.* Cambridge University Press, second edition, 1952.

[12] J.-B. HIRIART-URRUTY and D. YE. Sensitivity analysis of all eigenvalues of a symmetric matrix. *Numerische Mathematik*, 70:45–72, 1992.

[13] L. HÖRMANDER. *Notions of convexity*, volume 127 of *Progress in mathematics.* Birkhäuser, Boston, 1994.

[14] R.A. HORN and C.R. JOHNSON. *Matrix analysis.* Cambridge University Press, second edition, 1985.

[15] R.A. HORN and C.R. JOHNSON. *Topics in matrix analysis.* Cambridge University Press, first edition, 1991. Paperback edition with corrections, 1994.

[16] T.W. HUNGERFORD. *Algebra*, volume 73 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1974.

[17] N.V. KRYLOV. On the general notion of fully nonlinear second-order elliptic equations. *Transactions of the American Mathematical Society*, 347(3):857–895, March 1995.

[18] P.D. LAX. Differential Equations, Difference Equations and Matrix Theory. *Communications on pure and applied mathematics*, 11:175–194, 1958.

[19] A.S. LEWIS. Convex analysis on the Hermitian matrices. *SIAM Journal on Optimization*, 6(1):164–177, February 1996.

[20] A.S. LEWIS. Group invariance and convex matrix analysis. *SIAM Journal on Matrix Analysis*, 17(4):927–949, 1996.

[21] A.S. LEWIS. Convex analysis on Cartan subspaces, April 1997. Preprint.

[22] A.S. LEWIS. Lidskii's theorem via nonsmooth analysis. Research Report CORR 98-11, Department of Combinatorics and Optimization, University of Waterloo, April 1998.

[23] A.S. LEWIS and M.L. OVERTON. Eigenvalue optimization. *Acta Numerica*, 5:149–190, 1996.

[24] C.-K. LI and N.-K. TSING. Linear operators preserving certain functions on singular values of matrices. *Linear and Multilinear Algebra*, 26(1-2):133–143, 1990.

[25] M. MARCUS and L. LOPES. Inequalities for symmetric functions and Hermitian matrices. *Canadian Journal of Mathematics*, 9:305–312, 1957.

[26] M. MARCUS and H. MINC. The invariance of symmetric functions of singular values. *Pacific Journal of Mathematics*, 12:327–332, 1962.

[27] M. MARCUS and B. MOYLS. Linear transformations on algebras of matrices. *Canadian Journal of Mathematics*, 11:61–66, 1959.

[28] A.W. MARSHALL and I. OLKIN. *Inequalities: Theory of Majorization and Its Applications*, volume 143 of *Mathematics in Science and Engineering*. Academic Press, New York, 1979.

[29] J.B. McLEOD. On four inequalities in symmetric functions. *Proceedings of the Edinburgh Mathematical Society (Series II)*, 11:211–219, 1958-1959.

[30] DRAGOSLAV S. MITRINOVIĆ. *Analytic inequalities.* Springer-Verlag, New York, 1970. In cooperation with P. M. Vasić. Die Grundlehren der mathematischen Wisenschaften, Band 1965.

[31] D.S. MITRINOVIĆ, J.E. PEČARIĆ, and A.M. FINK. *Classical and New Inequalities in Analysis*, volume 61 of *Mathematics and Its Applications (East European Series).* Kluwer, 1993.

[32] A.M. OSTROWSKI. *Solution of equations in Euclidean and Banach spaces.* Academic Press, New York, third edition, 1973.

[33] R.T. ROCKAFELLAR. *Convex Analysis.* Princeton University Press, Princeton, NJ, 1970.

[34] C.M. THEOBALD. An inequality for the trace of the product of two symmetric matrices. *Mathematical Proceedings of the Cambridge Philosophical Society*, 77:265–267, 1975.

[35] J. von NEUMANN. Some matrix inequalities and metrization of matric-space. *Tomsk University Review*, 1:286–300, 1937. In: *Collected Works*, Pergamon, Oxford, 1962, Volume IV, 205-218.

[36] S. WAKABAYASHI. Remarks on hyperbolic polynomials. *Tsukuba Journal of Mathematics*, 10(1):17–28, 1986.