Hyperbolic Polynomials and Convex Analysis

Heinz H. Bauschke, Osman Güler, Adrian S. Lewis and Hristo S. Sendov

Abstract. A homogeneous real polynomial p is *hyperbolic* with respect to a given vector d if the univariate polynomial $t \mapsto p(x - td)$ has all real roots for all vectors x. Motivated by partial differential equations, Gårding proved in 1951 that the largest such root is a convex function of x, and showed various ways of constructing new hyperbolic polynomials. We present a powerful new such construction, and use it to generalize Gårding's result to arbitrary symmetric functions of the roots. Many classical and recent inequalities follow easily. We develop various convex-analytic tools for such symmetric functions, of interest in interior-point methods for optimization problems over related cones.

1 Introduction

A homogeneous polynomial p is *hyperbolic* with respect to some fixed direction d if the univariate polynomial

$$(1) \qquad\qquad t \mapsto p(x - td)$$

has only real roots for all points x. A canonical example is the determinant, considered as a polynomial on the space of symmetric matrices: the determinant is hyperbolic with respect to the identity matrix (the roots then being simply the eigenvalues).

Interest in hyperbolic polynomials was originally motivated by the partial differential equations literature [8], [14]. An important paper of Gårding [9] showed various ways of constructing new hyperbolic polynomials from old ones; it is known they form a rich class. Our key result is a new such construction. Specifically, we show that if a symmetric polynomial q is hyperbolic with respect to the direction $(1, 1, \ldots, 1)$ then the composition $q \circ \lambda$ is hyperbolic with respect to d, where the components of λ are the roots of the original hyperbolic polynomial (1).

We apply this simple and powerful result to show that for any symmetric convex function f, the function $f \circ \lambda$ is convex. This generalizes a fundamental tool of Gårding: the largest root is a convex function. In the determinant example above, our result shows that any symmetric convex function of the eigenvalues of a symmetric

matrix is a convex function of that matrix, a beautiful theorem of Davis [6]. Another choice of hyperbolic polynomial gives von Neumann's famous characterization of unitarily invariant matrix norms as symmetric gauge functions of the singular values [33]. By varying the choice of polynomial we derive many elegant, often classical, inequalities in a remarkably unified way. Examples include diverse properties of the elementary symmetric functions, and in particular some sophisticated recent results of Krylov [19].

The second half of this paper is convex-analytic in character. Symmetric functions of eigenvalues are fundamental in eigenvalue optimization and semidefinite programming [24]. They have an attractive duality theory: the Fenchel conjugate is described elegantly by the formula $(f \circ \lambda)^* = f^* \circ \lambda$ [21]. Von Neumann proved a similar result for unitarily invariant norms, useful in matrix approximation problems [15].

Hyperbolic polynomials offer a a unifying framework in which to study such convexity and duality results. They also have potential application in modern interior point methodology. Following Gårding, we can associate with any hyperbolic polynomial the open convex cone

$$\{x : p(x - td) \neq 0, \forall t \leq 0\}.$$

In the determinant example this is simply the cone of positive definite matrices. Optimization problems over such cones are good candidates for interior point algorithms analogous to the dramatically successful techniques current in semidefinite programming [11]. With these aims in mind, we outline an attractive duality theory, developing convex-analytic tools for symmetric convex functions of the roots associated with general hyperbolic polynomials. For clarity, we omit some of the details: they may be found in [1], which is accessible at http://orion.uwaterloo.ca/~aslewis.

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 cl S (resp. bd S, conv S, span S). A *cone* is a nonempty set that contains every nonnegative multiple of all its members. 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. If X is a Euclidean space with inner product $\langle \cdot, \cdot \rangle$ and $h: X \to [-\infty, +\infty]$ is convex, then h^* (resp. ∂h, ∇h, dom h) stands for the *Fenchel conjugate* (resp. *subdifferential map, gradient map, domain*) of h. (Rockafellar's [32] 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 ran λ. 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], [9], and a deep inequality on elementary symmetric functions.

Hyperbolic Polynomials and Characteristic Roots

If p is a nonconstant polynomial on X and m is a positive integer then we say 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 2.1 (**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 2.2 (**"Characteristic Roots and Trace"**)   Suppose p is hyperbolic with respect to $d \in X$ of degree m. Then for every $x \in X$, we can write

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

and assume without loss of generality that $\lambda_1(x) \geq \lambda_2(x) \geq \cdots \geq \lambda_m(x)$. The corresponding map $X \to \mathbb{R}_{\downarrow}^m \colon x \mapsto \bigl(\lambda_1(x), \ldots, \lambda_m(x)\bigr)$ is denoted by λ and called the *characteristic map (with respect to p and d)*. We say that $\lambda_i(x)$ is the *i-th largest characteristic root of x (with respect to p and d)* and define *the sum of the k largest characteristic roots* 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 characteristic roots $\{\lambda_i(x)\}$ are thus the roots of the polynomial $t \mapsto p(x - td)$. It follows that the trace σ_m is linear (see also the paragraph following Fact 2.10).

Unless stated otherwise, we assume throughout the paper that

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

for every $1 \leq k \leq m$. The notions "characteristic root" 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 $1 \le k \le m$, 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.* Further examples are provided in Section 6.

A simple way to generate new hyperbolic polynomials is differentiation:

Proposition 2.3 *If $m > 1$, then $q(x) := \frac{d}{dt} p(x+td)|_{t=0} = \left(\nabla p(x) \right)(d)$ is hyperbolic with respect to d.*

Proof This is essentially a consequence of Rolle's theorem; see also [9, Lemma 1]. ∎

The following property of the characteristic roots is well-known [9, Equation (2)].

Fact 2.4 For all $r, s \in \mathbb{R}$ and every $1 \le i \le m$:

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

Hence the characteristic map λ is *positively homogeneous* ($\lambda(tx) = t\lambda(x)$, for all $t \ge 0$ and every $x \in X$) and continuous (use, for instance, [31, Appendix A]).

Gårding showed that the largest characteristic root is a *sublinear* function, that is, positively homogeneous and convex.

Theorem 2.5 (Gårding) *The largest characteristic root $\lambda_1(\cdot)$ is a sublinear function.*

Proof Positive homogeneity follows from Fact 2.4. Now Gårding showed that λ_m is concave (see [9, Theorem 2] and [8]), which is equivalent to the convexity of λ_1, since $\lambda_1(-x) = -\lambda_m(x)$, for every $x \in X$. ∎

Gårding's proof of his result is complex-variable-based. Other approaches include Lax's PDE-based argument for an analogous result on matrices [20], and an argument using Bezout's Theorem proposed (without proof) in [18].

The Hermitian Matrices (continued) It is well-known that the largest eigenvalue map is convex in this case; see, for instance, [13].

Hyperbolicity Cone

Definition 2.6 (Hyperbolicity Cone) The *hyperbolicity cone of p with respect to d,* written $C(d)$ or $C(p, d)$, is the set $\{x \in X : p(x - td) \ne 0, \forall t \le 0\}$.

Fact 2.7 $C(d) = \{x \in X : \lambda_m(x) > 0\}$. Hence $C(d)$ is an open convex cone that contains d with closure $\mathrm{cl}\, C(d) = \{x \in X : \lambda_m(x) \ge 0\}$. If $c \in C(d)$, then p is hyperbolic with respect to c and $C(c) = C(d)$.

474   H. H. Bauschke, O. Güler, A. S. Lewis and H. S. Sendov

Proof See Gårding's [8] and [9, Section 2]. ∎

Definition 2.8 (**Complete Hyperbolic Polynomial**) p is *complete* if

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

Fact 2.9   Suppose p is hyperbolic with respect to d, with corresponding characteristic map λ and hyperbolicity cone $C(d)$. Then

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

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

Proof See Gårding's [8] and [9, Section 3]. ∎

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

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 $u \in \mathbb{R}^m$. For any given integer $k = 1, 2, \ldots, m$, the map

$$E_k \colon \mathbb{R}^m \to \mathbb{R} \colon u \mapsto \sum_{i_1 < \cdots < i_k} \prod_{l=1}^{k} u_{i_l}$$

is called the k-th *elementary symmetric function* on $\mathbb{R}^m$. We also set $E_0 := 1$.
    Repeatedly applying Proposition 2.3 gives the following result.

Fact 2.10   For every $x \in X$ and all $t \in \mathbb{R}$,

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

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

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

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

    In particular, the trace $\sigma_m = E_1 \circ \lambda$ is a homogeneous (hyperbolic) polynomial of degree 1 (and hence linear).
    Notice the elementary symmetric functions themselves are hyperbolic:

Example 2.11 Let $X = \mathbb{R}^m$ and $d = (1, 1, \ldots, 1) \in \mathbb{R}^m$. Then for every $1 \le k \le m$, 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 characteristic map $\lambda(x) = x_\downarrow$. Since each E_k is symmetric, the result now follows from Fact 2.10. ∎

An Inequality in Elementary Symmetric Functions

The following inequality was discovered independently by McLeod [28] and by Bullen and Marcus [4, Theorem 3].

Fact 2.12 (**McLeod, 1959; Bullen and Marcus, 1961**)   Suppose $1 \le k \le l \le m$ 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 2.12. (Proofs can also be found in [2, Theorem 1.16], [5, Section V.4], and [30, Section VI.5].) We record two interesting consequences of Fact 2.12.

Corollary 2.13 (**Marcus and Lopes's [25, Theorem 2]**)   *The function* $-E_m^{1/m}$ *is sublinear on* $\mathbb{R}_+^m$, *and it vanishes on* $\operatorname{bd} \mathbb{R}_+^m$.

Recall that a function h is called *logarithmically convex*, if $\log \circ h$ is convex. The function q in Fact 2.12 is concave ("strictly modulo rays"), which easily yields logarithmic and strict convexity of $1/q$ (see [1]):

Proposition 2.14   *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) $q(u + v) \ge q(u) + q(v)$, *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.

Applying this to Fact 2.12 gives the following result.

Corollary 2.15   *Suppose* $1 \le k \le l \le m$. *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 $m \ge 2$.*

3   Composition and Convexity

We begin this section with our key result, a powerful new construction for hyperbolic polynomials.

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 characteristic map μ. Then $q \circ \lambda$ is a hyperbolic polynomial of degree n with respect to d and its characteristic 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, for example, [17, Proposition V.2.20.(ii)]) a polynomial in $E_1(y), \ldots, E_m(y)$. On the other hand, by Fact 2.10, $E_i \circ \lambda$ is a hyperbolic polynomial with respect to d of degree i, for $1 \leq i \leq m$. Altogether, $\tilde{p}(x) = q\big(\lambda(x)\big)$ 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 2.4, we obtain $\tilde{p}(tx) = q\big(\lambda(tx)\big) = t^n \tilde{p}(x)$, for all $t \in \mathbb{R}$ and every $x \in X$.

Step 3 $\tilde{p}(d) \neq 0$.
    Again using Fact 2.4, we have $\tilde{p}(d) = q\big(\lambda(d)\big) = q(e) \neq 0$.

Step 4 $\tilde{p}$ is hyperbolic with respect to d.
    Using once more Fact 2.4, we write for every $x \in X$ and all $t \in \mathbb{R}$:

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

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 characteristic roots are

$$\left\{ \frac{1}{k} \sum_{l=1}^{k} u_{i_l} : 1 \leq i_1 < i_2 < \cdots < i_k \leq m \right\}.$$

In particular, the largest characteristic root of q is the arithmetic mean of the k largest components of u.

Sublinearity of Sums of the Largest Characteristic Roots

We now present our generalization of Gårding's theorem (2.5): the sum of the k largest characteristic roots is sublinear.

Corollary 3.3   *For every* $1 \leq k \leq m$, *the function* σ_k *is sublinear.*

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 characteristic root of $\tilde{p}$ is equal to $\frac{1}{k}\sigma_k(x)$. Now Theorem 2.5 yields the sublinearity of σ_k.   ∎

This readily implies Lipschitzness of the characteristic map (with respect to any norm on X).

Corollary 3.4   *Each characteristic root* λ_k *is (globally) Lipschitz.*

Proof Recall that every sublinear finite function is globally Lipschitz (this follows from [32, Theorem 13.2, Corollary 13.2.2, and Corollary 13.3.3]); in particular, so is each σ_i. Thus λ_1 is Lipschitz. If $k \geq 2$, then $\lambda_k = \sigma_k - \sigma_{k-1}$ is—as the difference of two Lipschitz functions— Lipschitz as well.   ∎

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 Lipschitz; see, for instance, [13].

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.   ∎

We can rewrite Corollary 3.5 as

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

This leads to the following question:

Open Problem 3.6 (**The Lidskii Property**)   Decide whether or not

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

The condition means that the vector $\lambda(y)$ "majorizes" the vector $\lambda(x+y) - \lambda(x)$, for all $x, y \in X$; see [27, Proposition 4.B.8]. (The interested reader is referred to [27] for further information on majorization.) If this condition is satisfied, then we will simply say that "the Lidskii property holds".

The Hermitian Matrices (continued)   The Lidskii property does hold for the Hermitians, and indeed is a crucial tool in perturbation theory. A recent and very complete reference is Bhatia's [3]; see also [23] for a new proof rooted in nonsmooth analysis.

In Section 6, we point out further examples where the Lidskii property holds.

Convexity of Composition

We next aim to generalize Gårding's theorem still further. We begin with a useful tool.

Fact 3.7   Suppose $f\colon \mathbb{R}^m \to [-\infty, +\infty]$ 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 $\mathrm{conv}\big\{ u_{\pi(i)} : \pi$ is a permutation of $\{1, \ldots, m\}\big\}$ and $u \neq v$, then $f(u) > f(v)$.

Proof Imitate the proof of [22, Theorem 3.3] and consider [22, Example 7.1]. See also [27, 3.C.2.c on p. 68]. ∎

Lemma 3.8   Suppose $x, y \in X$, $\alpha \in (0, 1)$, and $f\colon \mathbb{R}^m \to [-\infty, +\infty]$ is convex and symmetric. Then

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

If f is strictly convex and $\alpha\lambda(x) + (1-\alpha)\lambda(y) \neq \lambda\big(\alpha x + (1-\alpha)y\big)$, then the inequality holds strictly.

Proof (See also [22, Proof of Theorem 4.3])   Fix an arbitrary $w \in \mathbb{R}^m_\downarrow$. Set $u := \alpha\lambda(x) + (1 - \alpha)\lambda(y)$ and $v := \lambda\big(\alpha x + (1 - \alpha)y\big)$. Then both u and v belong to $\mathbb{R}^m_\downarrow$. By Corollary 3.5, $w^T\lambda$ is convex on X. Therefore, $w^T\lambda\big(\alpha x + (1 - \alpha)y\big) \leq \alpha w^T\lambda(x) + (1-\alpha)w^T\lambda(y)$; equivalently, $w^T(u-v) \geq 0$. It follows that $u - v \in (\mathbb{R}^m_\downarrow)^\oplus$. By Fact 3.7, $f(u) \geq f(v)$, as required. The last part now follows from Fact 3.7. ∎

We hence obtain the following very pleasing generalization of Gårding's theorem (which is the case $f(u) = \max u_i$).

Theorem 3.9 (Convexity)   If $f\colon \mathbb{R}^m \to [-\infty, +\infty]$ is convex and symmetric then $f \circ \lambda$ is convex.

The Hermitian Matrices (continued)   In this case, the convexity of the composition is attributed to Davis [6]; see also [21, Corollary 2.7].

Another consequence is Gårding's inequality; see [11, Lemma 3.1].

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

Proof By Corollary 2.13, the function $-E_m^{1/m}$ is sublinear and symmetric on $\mathbb{R}^m_+$. Hence, by Theorem 3.9, the function $x \mapsto -E_m\big(\lambda(x)\big)^{1/m}$ is sublinear on $\{x \in X : \lambda(x) \geq 0\} = \mathrm{cl}\, C(d)$. The result follows, since $p(x) = p(d)E_m\big(\lambda(x)\big)$, for every $x \in X$. ∎

For a further discussion of this and other inequalities involving hyperbolic polynomials, see [18].

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 are positive semidefinite.

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

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

$$\|\lambda(x + y)\|^2 = w^T \lambda(x + y) \leq w^T\big(\lambda(x) + \lambda(y)\big)$$
$$\leq \|w\| \, \|\lambda(x) + \lambda(y)\| = \|\lambda(x + y)\| \, \|\lambda(x) + \lambda(y)\|.$$

The inequality follows. The condition for equality follows from the condition for equality in the Cauchy-Schwarz inequality. ∎

4   Euclidean Structure

The characteristic map λ induces a natural norm as follows.

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

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

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 $\|x\|^2 = \|\lambda(x)\|^2 = \sum_{i=1}^m \lambda_i(x)^2 = \big(E_1\big(\lambda(x)\big)\big)^2 - 2E_2\big(\lambda(x)\big)$. Facts 2.4 and 2.10 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 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 $F(x) := -\ln\big(p(x)\big)$. The Hessian norm at x is then given by

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

For small positive t we have $p(tx + d) = p(d) \prod_{i=1}^{m} \left(1 + t\lambda_i(x) \right)$ and hence

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

Expanding each side and comparing coefficients of t^2 gives the result. Further information can be found in [11].

Proposition 4.4 (Sharpened Cauchy-Schwarz) *Suppose p is complete. Then $\langle x, y \rangle \le \langle \lambda(x), \lambda(y) \rangle \le \|x\| \, \|y\|$, for all $x, y \in X$.*

Proof The Cauchy-Schwarz inequality in $\mathbb{R}^m$ and Corollary 3.11 imply

$$2\langle \lambda(x), \lambda(y) \rangle \ge \|\lambda(x + y)\|^2 - \|\lambda(x)\|^2 - \|\lambda(y)\|^2$$
$$= \|x + y\|^2 - \|x\|^2 - \|y\|^2 = 2\langle x, y \rangle,$$

as required. ∎

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 essentially due to von Neumann; see [21, Theorem 2.2] and the discussion therein.

We can now refine Theorem 3.9.

Theorem 4.5 (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

$$(f \circ \lambda)(\alpha x + \beta y) = \alpha(f \circ \lambda)(x) + \beta(f \circ \lambda)(y).$$

We have to show $x = y$. By Lemma 3.8 and convexity of f, we have

$$\alpha(f \circ \lambda)(x) + \beta(f \circ \lambda)(y) = (f \circ \lambda)(\alpha x + \beta y)$$
$$\le f\left(\alpha\lambda(x) + \beta\lambda(y) \right)$$
$$\le \alpha(f \circ \lambda)(x) + \beta(f \circ \lambda)(y);$$

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, Lemma 3.8 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. ∎

Theorem 4.5 gives an immediate, transparent proof of a recent result of Krylov: see [19, Thm 6.4.(ii)] (whose proof is considerably more involved).

Corollary 4.6 *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 4.5). 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 2.15. By Theorem 3.9 (or Theorem 4.5), F is convex (or 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 relates closely to Güler's recent work on hyperbolic barrier functions. With this new approach we have a simple proof of Güler's [11, Theorem 6.1]. The functions F and g below are crucial in interior-point theory, as they allow the construction of 'long-step' methods.

Corollary 4.7 *Suppose $p(d) > 0$ and c belongs to the hyperbolicity cone $C := C(d)$. Define*

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

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 4.6. Now let μ be the characteristic map corresponding to c. Then, by Fact 2.10, $p(x) = p(c)E_m(\mu(x))$ and $\big(\nabla p(x)\big)(c) = p(c)E_{m-1}(\mu(x))$. 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 4.6. ∎

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 semi-definite Hermitian matrices; this result is due to Fan [7].

We end this section with an easy exercise emphasizing the naturalness of the inner product.

Proposition 4.8 (**Trace**) $\sigma_m(x) = \langle d, x \rangle$, *for every $x \in X$.*

5   Duality

The last part of our development concerns duality. For this we need a little more structure.

Definition 5.1 **(Isometric Hyperbolic Polynomial)**   We say p is *isometric*, if for all $y, z \in X$, there exists $x \in X$ such that

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

Clearly, if p is isometric then ran λ is a closed convex cone contained in $\mathbb{R}^m_{\downarrow}$. However, the next example shows the range of λ need not be convex.

Example 5.2 **(A Hyperbolic Polynomial That Is Not Isometric)**   If the polynomial $p(x) = x_1 x_2 x_3$ is defined on $X = \text{span}\{(1, 1, 1), (3, 1, 0)\}$, 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 $\lambda(3, 1, 0) + \lambda(-3, -1, 0) = (3, 0, -3) \notin$ ran λ, the set ran λ 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

> p is complete, with corresponding inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$.

We call p "isometric" because of the equivalent condition (iii) in the following proposition [1, Proposition 5.4].

Proposition 5.3   *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 $u \in$ ran λ and every $y \in X$.
(iii) $d\big(u, \lambda(y)\big) = d\big(\lambda^{-1}(u), y\big)$, for all $u \in$ ran λ and every $y \in X$.

The Hermitian Matrices (continued)   Clearly, ran $\lambda = \mathbb{R}^m_{\downarrow}$ in this case, and an easy exercise shows $p = \det$ is indeed isometric [1].

The isometric property leads to a very concise duality result.

Theorem 5.4 **(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 ran $\lambda = \mathbb{R}^m_{\downarrow}$ then $(f \circ \lambda)^* = f^* \circ \lambda$.*

Proof Fix an arbitrary $y \in X$. Then, using Proposition 4.4, symmetry of f, and the Hardy-Littlewood-Pólya inequality (see [12, Section 10.2]), the inequality follows from

$$f^*\big(\lambda(y)\big) = \sup_{u \in \mathbb{R}^m} \{\langle u, \lambda(y)\rangle - f(u)\} = \sup_{u \in \mathbb{R}^m_\downarrow} \{\langle u, \lambda(y)\rangle - f(u)\}$$

$$\geq \sup_{u \in \text{ran } \lambda} \max_{x:\lambda(x)=u} \big\{\langle x, y\rangle - f\big(\lambda(x)\big)\big\} = \sup_{x \in X}\{\langle x, y\rangle - (f \circ \lambda)(x)\}$$

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

Now assume that p is isometric and $\text{ran } \lambda = \mathbb{R}^m_\downarrow$. Then Proposition 5.3 shows

$$f^*\big(\lambda(y)\big) = \sup_{u \in \mathbb{R}^m_\downarrow} \{\langle \lambda(y), u\rangle - f(u)\} \leq \sup_{u' \in \text{ran } \lambda} \{\langle \lambda(y), u'\rangle - f(u')\}$$

$$= \sup_{x \in X}\big\{\langle y, x\rangle - f\big(\lambda(x)\big)\big\} = (f \circ \lambda)^*(y). \qquad \blacksquare$$

Without the assumption $\text{ran } \lambda = \mathbb{R}^m_\downarrow$ this result may fail (see Example 6.1).

As a consequence of our Fenchel conjugacy formula we are easily able to identify subgradients of such convex functions.

Theorem 5.5 (Subgradients) *Suppose p is isometric, $\text{ran } \lambda = \mathbb{R}^m_\downarrow$, and $f \colon \mathbb{R}^m \to (-\infty, +\infty]$ is convex and symmetric. Let $x, y \in X$. Then*

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

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

Proof We know $(f \circ \lambda)^* = f^* \circ \lambda$. In view of Proposition 4.4, the following equivalences hold true: $y \in \partial(f \circ \lambda)(x) \Leftrightarrow (f \circ \lambda)(x) + (f \circ \lambda)^*(y) = \langle x, y\rangle \Leftrightarrow f\big(\lambda(x)\big) + f^*\big(\lambda(y)\big) = \langle \lambda(x), \lambda(y)\rangle$ and $\langle x, y\rangle = \langle \lambda(x), \lambda(y)\rangle \Leftrightarrow \lambda(y) \in \partial f\big(\lambda(x)\big)$ and $\langle x, y\rangle = \langle \lambda(x), \lambda(y)\rangle$. "Consequently": Clearly, by the above, $\lambda[\partial(f \circ \lambda)(x)] \subseteq \partial f\big(\lambda(x)\big)$. Conversely, pick $v \in \partial f\big(\lambda(x)\big)$. Then $f\big(\lambda(x)\big) + f^*(v) = \langle v, \lambda(x)\rangle$. By the assumption that $\text{ran } \lambda = \mathbb{R}^m$ and Proposition 5.3.(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 $v = \lambda(y) \in \lambda[\partial(f \circ \lambda)(x)]$. $\blacksquare$

The Hermitian Matrices (continued) Theorem 5.5 corresponds to [21, Theorem 3.2].

We end with another nice generalization of a result of Fan.

Corollary 5.6 (Variational Description of σ_k) *Let p be isometric, and suppose $\text{ran } \lambda = \mathbb{R}^m_\downarrow$. Let $1 \leq k \leq m$. Then for every $x \in X$,*

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

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

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 the set

$$\left\{ u \in \mathbb{R}^m : \sum u_i = k \text{ and each } 0 \leq u_i \leq 1 \right\}.$$

Now $\sigma_k = f \circ \lambda$ and so Theorem 5.4 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 (conclusion) Corollary 5.6 is a generalization of the variational formulations due to Rayleigh and Ky Fan; see also [13, Section 2].

6 Further Examples

This section contains some particularly natural examples (see [1]). Our aim is to get across the idea of the unity of the method and its breadth.

Reordering on $\mathbb{R}^n$

Consider the vector space $X = \mathbb{R}^n$, the polynomial

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

and the direction $d = (1, 1, \ldots, 1)$. Then p is hyperbolic and complete with characteristic map $\lambda(x) = x_{\downarrow}$. The induced norm and inner product in X are just the standard Euclidean ones in $\mathbb{R}^n$. We have $\text{ran}\lambda = \mathbb{R}^n_{\downarrow}$. It can be seen easily that p is isometric. In this case the sharpened Cauchy-Schwarz inequality (Proposition 4.4) reduces to the well-known Hardy-Littlewood-Pólya inequality (see [12, Chapter X])

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

Equality holds if and only if the vectors x and y can be simultaneously ordered with the same permutation. Since $\text{ran}\lambda = \mathbb{R}^n_{\downarrow}$, Theorem 5.4 shows that for every symmetric function $f \colon \mathbb{R}^n \to (-\infty, +\infty]$ we have $(f \circ \lambda)^* = f^* \circ \lambda$. Also the Lidskii property holds, because $\lambda(x)$ is the ordered set of eigenvalues of the symmetric matrix $\text{Diag}(x)$ (see [3, p. 69]).

Singular Values

Consider the vector space $M_{n,m}$ of n by m real matrices. We assume $m \leq n$ and denote the singular values of a matrix x in $M_{n,m}$ by $\sigma_1(x) \geq \sigma_2(x) \geq \cdots \geq \sigma_m(x)$. The Frobenius (or Hilbert-Schmidt) norm [15, pp. 291, 421] is defined by $\|x\|_F = \|\sigma(x)\|$, where the latter norm is the standard Euclidean norm in $\mathbb{R}^n$, and $\sigma(x) = \left(\sigma_1(x), \sigma_2(x), \ldots, \sigma_m(x)\right)$. 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 characteristic map

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

The induced norm and inner product are given by

$$\|(x, \alpha)\|^2 = 2m\alpha^2 + 2\|x\|_F^2,$$

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

for (x, α) and (y, β) in X. With the help of the singular value decomposition one can see that p is isometric. Notice that in this case the sharpened Cauchy-Schwarz inequality (Proposition 4.4) reduces to

$$\operatorname{tr} x^T y \le \sigma(x)^T \sigma(y).$$

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 $x = u\big(\operatorname{Diag} \sigma(x)\big) v$ and $y = u\big(\operatorname{Diag} \sigma(y)\big) v$). This is the classical result known as 'von Neumann's Lemma' (see for example [16, p. 182]). For a proof using results from this paper see [1]. We note that the Lidskii property holds too.

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 $X = \mathbb{C}_{n,m} \times \mathbb{R}$.

Interestingly, in this case we may have $(f \circ \lambda)^* \ne f^* \circ \lambda$.

Example 6.1 Consider the symmetric function $f(u) = \max u_i$. If $m = 2$ and $y \in X$ is such that $\lambda(y) = \frac{1}{4}(3, 1, 1, -1)$, a short calculation [1] shows that $0 = (f \circ \lambda)^*(y) \ne (f^* \circ \lambda)(y) = +\infty$.

Remark 6.2 Von Neumann's famous characterization of unitarily invariant matrix norms as symmetric gauge functions of the singular values [33] also follows naturally in this framework [1].

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 characteristic map

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

where $|x| = (|x_1|, |x_2|, \dots, |x_n|)$, and $e = (1, 1, \dots, 1) \in \mathbb{R}^{2n}$. Direct verification of the definition shows that p is isometric and furthermore that the Lidskii property holds. Note that the similarities with the previous example are not accidental. It corresponds to the subspace $(\operatorname{Diag} \mathbb{R}^n) \times \mathbb{R}$ of $M_{n,m} \times \mathbb{R}$.

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 $A = \mathrm{Diag}(1, -1, -1, \ldots, -1) \in M_n$ ($n \times n$ real matrices). Let the direction be $d = (d_1, d_2, \ldots, d_n) \in X$ such that $d_1^2 > d_2^2 + \cdots + d_n^2$. Then p is hyperbolic and complete with characteristic 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 $D(x) = (x^T A d)^2 - p(x)p(d)$ 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 known as Aczel's inequality, see [29, p. 57].) The induced norm and inner product are given by

$$\|x\|^2 = 2\frac{2(x^T A d)^2 - p(x)p(d)}{p(d)^2}, \quad \text{and}$$

$$\langle x, y \rangle = \frac{4(x^T A d)(y^T A d) - 2(x^T A y)p(d)}{p(d)^2},$$

for x and y in X. It is a bit trickier to see that p is isometric. Notice that in this case the sharpened Cauchy-Schwarz inequality (Proposition 4.4) becomes

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

and [1] gives the necessary and sufficient condition for equality. The Lidskii property holds as well.

The Degree 2 Case

Let the vector space be $X = \mathbb{R}^n$. We assume that $p(x)$ is homogeneous polynomial of degree two. Without loss of generality, we write

$$p(x) = x^T A x,$$

where $A \in S^n$. Fix a direction d in X with $p(d) \neq 0$. Then $p(x)$ is hyperbolic with respect to d if and only if the matrix $(d^T A d)^{-1} A$ has exactly one positive eigenvalue (see [9, p. 958]). Furthermore, p is complete if and only if A is nonsingular. Such a p is always isometric, and the Lidskii property holds.

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 (see Fact 2.10). We have

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

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 $\operatorname{tr}(\wedge^k x)$. For the first equality above, see [26], and justification for the use of the symbol $\operatorname{tr}(\wedge^k x)$ can be found in [10]. Now, from Corollary 3.10 (Gårding's inequality) and from the fact that $p_k(x) = \operatorname{tr}(\wedge^k x)$ is a homogeneous hyperbolic polynomial, it follows immediately that

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

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

References

[1] H. H. Bauschke, A. S. Lewis, O. Güler, and H. S. Sendov, *Hyperbolic Polynomials and Convex Analysis*. Research Report CORR 98-29, Department of Combinatorics and Optimization, University of Waterloo, June 1998.

[2] E. F. Beckenbach and R. Bellman, *Inequalities*. Springer-Verlag, 1961.

[3] R. Bhatia, *Matrix Analysis*. Springer-Verlag, 1997.

[4] P. Bullen and M. Marcus, *Symmetric means and matrix inequalities*. Proc. Amer. Math. Soc. **12**(1961), 285–290.

[5] P. S. Bullen, D. S. Mitrinović and P. M. Vasić, *Means and Their Inequalities*. Math. Appl. (East European Ser.), D. Reidel Publishing Company, Dordrecht, Holland, 1987.

[6] C. Davis, *All convex invariant functions of hermitian matrices*. Arch. Math. **8**(1957), 276–278.

[7] Ky Fan, *On a theorem of Weyl concerning eigenvalues of linear transformations II*. Proc. Nat. Acad. Sci. U.S.A. **36**(1950), 31–35.

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

[9] _____, *An inequality for hyperpolic polynomials*. J. Math. Mech. (6) **8**(1959), 957–965.

[10] W. H. Greub, *Multilinear Algebra*. Springer-Verlag, 1967.

[11] O. Güler, *Hyperbolic polynomials and interior point methods for convex programming*. Math. Oper. Res. (2) **22**(1997), 350–377.

[12] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*. Cambridge University Press, 2nd edition, 1952.

[13] J.-B. Hiriart-Urruty and D. Ye, *Sensitivity analysis of all eigenvalues of a symmetric matrix*. Numer. Math. **70**(1992), 45–72.

[14] L. Hörmander, *Notions of Convexity*. Progress in Mathematics **127**, Birkhäuser, Boston, 1994.

[15] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 2nd edition, 1985.

[16] _____, *Topics in Matrix Analysis*. Cambridge University Press, 1st edition, 1991; paperback edition with corrections, 1994.

[17] T. W. Hungerford, *Algebra*. Graduate Texts in Math. **73**, Springer-Verlag, New York, 1974.

[18] A. G. Khovanskii, *Analogues of the Aleksandrov-Fenchel inequalities for hyperbolic forms.* Soviet Math. Dokl. **29**(1984), 710–713.

[19] N. V. Krylov, *On the general notion of fully nonlinear second-order elliptic equations.* Trans. Amer. Math. Soc. (3) **347**(1995), 857–895.

[20] P. D. Lax, *Differential equations, difference equations and matrix theory.* Comm. Pure Applied Math. **11**(1958), 175–194.

[21] A. S. Lewis, *Convex analysis on the Hermitian matrices.* SIAM J. Optim. (1) **6**(1996), 164–177.

[22] ———, *Group invariance and convex matrix analysis.* SIAM J. Matrix Anal. Appl. (4) **17**(1996), 927–949.

[23] ———, *Lidskii's theorem via nonsmooth analysis.* SIAM J. Matrix Anal. Appl. **21**(1999), 379–381.

[24] A. S. Lewis and M. L. Overton, *Eigenvalue optimization.* Acta Numerica **5**(1996), 149–190.

[25] M. Marcus and L. Lopes, *Inequalities for symmetric functions and Hermitian matrices.* Canad. J. Math. **9**(1957), 305–312.

[26] M. Marcus and H. Minc, *The invariance of symmetric functions of singular values.* Pacific J. Math. **12**(1962), 327–332.

[27] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications.* Mathematics in Science and Engineering **143**, Academic Press, New York, 1979.

[28] J. B. McLeod, *On four inequalities in symmetric functions.* Proc. Edinburgh Math. Soc. **11**(1958/1959), 211–219.

[29] D. S. Mitrinović, *Analytic Inequalities.* Die Grundlehren der mathematischen Wisenschaften, Band 1965; Springer-Verlag, New York, 1970.

[30] D. S. Mitrinović, J. E. Pečarić and A. M. Fink, *Classical and New Inequalities in Analysis.* Math. Appl. (East European Ser.) **61**, Kluwer, 1993.

[31] A. M. Ostrowski, *Solution of Equations in Euclidean and Banach Spaces.* Academic Press, New York, 3rd edition, 1973.

[32] R. T. Rockafellar, *Convex Analysis.* Princeton University Press, Princeton, NJ, 1970.

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

Department of Mathematics and Statistics
Okanagan University College
Kelowna, British Columbia
V1V 1V7
email: bauschke@cecm.sfu.ca

Department of Mathematics and Statistics
University of Maryland Baltimore County
Baltimore, Maryland 21228,
USA
email: guler@math.umbc.edu

Department of Combinatorics & Optimization
University of Waterloo
Waterloo, Ontario
N2L 3G1
email: aslewis@math.uwaterloo.ca

Department of Combinatorics & Optimization
University of Waterloo
Waterloo, Ontario
N2L 3G1
email: hssendov@barrow.uwaterloo.ca