Geometry of convex sets arising from hyperbolic polynomials

by

Tor Gunnar Josefsson Myklebust

A thesis

presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Combinatorics & Optimization

Waterloo, Ontario, Canada, 2008

©Tor Gunnar Josefsson Myklebust 2008

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This thesis focuses on convex sets and convex cones defined using hyperbolic polynomials.

We first review some of the theory of convex sets in $\mathbb{R}^d$ in general. We then review some classical algebraic theorems concerning polynomials in a single variable, as well as presenting a few more modern results about them. We then discuss the theory of hyperbolic polynomials in several variables and their associated hyperbolicity cones. We survey various ways to build and decompose hyperbolic cones and we prove that every nontrivial hyperbolic cone is the intersection of its derivative cones. We conclude with a brief discussion of the set of extreme rays of a hyperbolic cone.

Acknowledgments

I thank my friends and family for their patience and generous help during my time at Waterloo. I particularly thank my advisor, Dr. Levent Tunçel, for sharing his wealth of good advice over the course of my Master's, and I further thank him and my readers, Drs. Stephen Vavasis and David Wagner, for their many useful comments on early versions of this thesis.

Contents

Chapter 1

Introduction and background

1.1 Notational conventions

Unless otherwise specified, G is a convex set, K is a convex cone, $\bar{e}$ is the vector of all ones of the appropriate dimension, $\mathbb{R}^d$ is the set of d-tuples of real numbers, Σ^d is the set of $d \times d$ symmetric matrices, Σ^d_+ is the set of $d \times d$ symmetric and positive semidefinite matrices, and Σ^d_{++} is the set of $d \times d$ symmetric and positive definite matrices.

We define arithmetic on $(-\infty, \infty] = \mathbb{R} \cup \{\infty\}$ by setting $\infty + x = \infty$ for all $x \in (-\infty, \infty]$, $\lambda\infty = \infty$ for $\lambda \in (0, \infty]$, and $0 \cdot \infty = \infty \cdot 0 = 0$. We do not permit ∞ to be subtracted from ∞ or ∞ to be multiplied by a negative number.

A function $f : \mathbb{R}^d \to \mathbb{R}^{d'}$ is said to be $\mathcal{C}^k$ if it is k times continuously differentiable. We denote the kth derivative of a $\mathcal{C}^k$ function f by $D^k f$; evaluated at a point x it is $(D^k f)(x)$. $(D^k f)(x)$ is a function mapping k-tuples of vectors in $\mathbb{R}^d$ to $\mathbb{R}^{d'}$; its evaluation at the vectors $v_1, \ldots, v_k$ is written $(D^k f)(x)[v_1, \ldots, v_k]$. Out of habit, we have special notations for the first derivative of a $\mathcal{C}^1$ function and the second derivative of a $\mathcal{C}^2$ function — $(\nabla f)(x)$ is the first derivative of f evaluated at x and $(Hf)(x)$ is the second derivative of f evaluated at x. If $d' = 1$, then $(\nabla f)(x)$ is a linear functional on $\mathbb{R}^d$ and so we often omit the square brackets around its argument.

If $G \subseteq \mathbb{R}^d$, then $\operatorname{cl} G$ is the closure of G, $\operatorname{int} G$ is the interior of G, and $\operatorname{bd} G$ is the boundary of G.

1.2 Introduction

A wide variety of real-world problems can be modelled effectively as continuous optimisation problems. Continuous optimisation problems that are *convex* tend to be better-behaved and tend to admit solution algorithms that are faster and more robust than their nonconvex relatives. (A convex optimisation problem, for our purposes, is a problem of minimising some fixed convex function (defined in Section 1.4), called the *objective function*, over a fixed convex set (defined in Section 1.3), called the *feasible set*.)

Some effective approaches to nonconvex continuous optimisation problems rely heavily on the ability to solve a sequence of convex optimisation problems efficiently.

Convex optimisation in general is also a very powerful tool for proving that a given optimisation problem may be approximated efficiently. This power stems from the ellipsoid method of D. Yudin and A. Nemirovskii [43] and N. Shor [36]; subject to fairly weak polynomial-time computability conditions on the objective function and the feasible set, the ellipsoid method will take a parameter $\epsilon > 0$ and either produce a feasible point whose objective value is within ϵ of optimal or conclude that no such point exists in polynomial time.

Certain classes of convex optimisation problems are especially well-behaved and admit especially fast and robust solution algorithms. Most notably, *linear programs*, where the objective function is linear and the feasible region is polyhedral, have a simple and well-understood duality theory and especially efficient algorithms for their solution. Nontrivial linear programs defined over spaces with millions of dimensions are routinely solved to optimality using modern linear program solvers. Another notable class are the *semidefinite programs*, where the objective function is linear and the feasible region is the intersection of an affine space with the set of positive semidefinite matrices of a given size. Software packages exist that take advantage of the underlying structure of the set of positive semidefinite matrices to get substantially better performance than could be expected from a package for general convex programming.

Another class of convex optimisation problems, the *hyperbolic programs*, has recently gained attention. In a hyperbolic program, the objective is again linear, but the feasible set is defined using a polynomial that satisfies a certain stability property known as *hyperbolicity*. Hyperbolic programs appear to be more general than semidefinite programs — every semidefinite program can be formulated as a hyperbolic program, but the converse is not known.

This thesis is concerned with hyperbolic polynomials and, in particular, the structure of the feasible set of hyperbolic programming problems.

This notion of hyperbolicity first arose in partial differential equations in the work of I. Petrovsky. The theory of hyperbolic partial differential equations was notably furthered by M. Atiyah, R. Bott, and L. Gårding ([1], [2]). It is discussed in L. Hörmander's four-volume set [22].

There are also deep connections to discrete mathematics. Y. Choe, J. Oxley, A. Sokal, and D. Wagner [10], P. Brändén [7] and D. Wagner and Y. Wei [42] used hyperbolic polynomials to prove theorems in matroid theory. L. Gurvits [18] used hyperbolic polynomials to give simple proofs of two important theorems in combinatorics.

A connection to matrix theory exists; J. Borcea, P. Brändén, and B. Shapiro [6] used hyperbolic polynomials to prove three conjectures of C. Johnson in the theory of matrices. Further, P. Brändén, J. Borcea, and T. Liggett [8] established results in probability theory using hyperbolic polynomials.

1.3 Convex sets

Let $G \subseteq \mathbb{R}^d$. We say that G is convex if, for each x and y in G, the line segment $[x, y]$ between x and y is contained in G. The class of convex subsets of $\mathbb{R}^d$ is closed under intersection, so one can speak of the "smallest convex set" containing a given set $G \subseteq \mathbb{R}^d$, meaning simply the intersection of all convex sets containing G. This "smallest convex set" is known as the convex hull; the convex hull of a set X is denoted $\operatorname{conv} X$.

A classical theorem asserts that the convex hull of a set X is simply the set of all affine combinations of elements of X in which the coefficients of elements of X are nonnegative and sum to 1. (Such affine combinations are known as *convex combinations*.) A theorem of C. Carathéodory asserts that, in fact, each point of $\operatorname{conv} X$ can be written as a convex combination of at most $d+1$ points of X.

1.4 Convex functions

Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$. We say that f is *convex* if $f(\lambda a + (1-\lambda)b) \leq \lambda f(a) + (1-\lambda)f(b)$ for all $a \in \mathbb{R}^d$, $b \in \mathbb{R}^d$, and $\lambda \in [0, 1]$.

There is a correspondence between convex functions and convex sets. If f is convex, define its *epigraph* $\operatorname{epi} f$ by

$$\operatorname{epi} f = \{(x, y) \in \mathbb{R}^d \times \mathbb{R} : y \geq f(x)\}.$$

Then $\operatorname{epi} f$ is a convex set. Further, if f is differentiable at x, then no point $(y, z) \in \operatorname{epi} f$ satisfies $(\nabla f)(x)[y] > z$.

Conversely, given a convex set G, one can define the *indicator function* of G, $1/\chi_G$, by

$$1/\chi_G(x) := \begin{cases} \infty & \text{if } x \in G \\ 1 & \text{otherwise} \end{cases}$$

1.5 Convex optimisation

A general convex optimisation problem can be stated in the form

$$\begin{array}{rrcl} \inf & f(x) & & \\ \text{subject to} & g_1(x) & \leq & 0 \\ & \vdots & \vdots & \vdots \\ & g_m(x) & \leq & 0 \end{array} \tag{1.1}$$

for some convex functions f and $g_1, \ldots, g_m$ mapping $\mathbb{R}^d$ to $\mathbb{R}$. No generality is lost by insisting that the objective function is linear; the problem above is equivalent to the problem (in one more variable, λ)

$$\begin{array}{rrcl} \inf & \lambda & & \\ \text{subject to} & g_1(x) & \leq & 0 \\ & \vdots & \vdots & \vdots \\ & g_m(x) & \leq & 0 \\ & f(x) - \lambda & \leq & 0 \end{array}$$

The transformed problem is still a convex optimisation problem. The optimal objective values, if such exist, are the same. If one problem has no feasible solution, neither does the other, and if there are feasible points with arbitrarily negative objective value in one problem, the same is true for the other.

Another form, more convenient for our purposes, is the *conic form.* A convex optimisation problem in conic form is phrased as

$$\begin{array}{rrcl} \inf & c^T x & & \\ \text{subject to} & x & \in & K \cap (V + b) \end{array}$$

where c is a vector, K is a convex cone, V is a linear space, and b is a vector. We can transform a problem in the form of 1.1 whose objective function is linear to an equivalent problem in conic form by taking

$$K = \left\{ (x, y) \in \mathbb{R}^d \times \mathbb{R}_{++} : g_i\left(\frac{1}{y} x\right) \leq 0 \text{ for } 1 \leq i \leq n \} \cup \{0 \right\}$$

and letting $V = \{(x, y) \in \mathbb{R}^d \times \mathbb{R}_{++} : y = 1\}$. Optimisation problems in conic form are a convenient setting both for practical implementations and for our theoretical development.

When the set $V \cap K$ contains an interior point of K, a remarkable duality theorem (see, for instance, [38]) holds: the optimal objective value of the "primal" problem

$$\begin{array}{rrcl} \inf & c^T x & & \\ \text{subject to} & x & \in & K \cap (V + b) \end{array}$$

is equal to the optimal objective value of the "dual" problem

$$\begin{array}{rrcl} \sup & b^T y & & \\ \text{subject to} & y & \in & K^* \cap (V^\perp + c) \end{array}$$

where K^* is the dual cone to K (defined in Section 2.8).

Interior-point methods are a class of particularly effective methods for solving convex optimisation problems in conic form. One constructs a convex *barrier function* for the cone K — a function $f : \operatorname{int} K \to \mathbb{R}$ such that if $\{x_n\}_{n=1}^\infty$ tends to $x \in \operatorname{bd} K$, then $f(x_n)$ tends to ∞. If one further requires that f be three-times differentiable and satisfy the *self-concordance* conditions

$$|(D^3 f)[h, h, h]| \le 2\sqrt{|(Hf)[h, h]|^3} \quad \text{for all } h \in \mathbb{R}^d$$

and

$$|(\nabla f)[h]| \le \sqrt{\theta (Hf)[h, h]} \quad \text{for all } h \in \mathbb{R}^d$$

for some θ, one can apply the rich theory of Y. Nesterov and A. Nemirovskii [28]. In particular, one can obtain an algorithm that finds a solution within an error of $\frac{1}{\epsilon}$ from optimal after doing $O(\sqrt{\theta} \log \frac{1}{\epsilon})$ iterations.

1.6 Hyperbolic polynomials and hyperbolic cones

1.6.1 Polynomials in general

The following fundamental theorem shows that the function mapping the coefficients of a polynomial to its roots is continuous.

Theorem 1.6.1. *Let P_n be the set of univariate monic polynomials of degree n with complex coefficients. The map $Z : P_n \to 2^{\mathbb{C}}$ sending a polynomial to its set of roots is continuous (in the Hausdorff metric).*

PROOF : Let $p \in P_n$ and fix $\epsilon > 0$. We shall assume that ϵ is so small that no two distinct roots of p are within 2ϵ of one another. Let $\delta > 0$ be such that $|p(x)| < \delta/2$ whenever x is at most ϵ away from a root of p.

Let $q \in P_n$ be such that $|q - p| < \delta/2$ on every circle of radius ϵ centred at a root of p. (The set of such polynomials contains p as an interior point.) Let r be a root of p, and let γ be the contour running once clockwise around the circle of radius ϵ centred at r. Since $|q - p| < \delta < |p|$ on γ, Rouché's theorem informs us that $p + (q - p) = q$ has the same number of roots as p within this contour. q has the same number of roots as p, so all of q's roots must be within ϵ of a root of p. The theorem follows. $\square$

Corollary 1.6.2. *Let P_n be as before. Let S be the set of nonzero polynomials of degree n with only real roots. Then, for $1 \leq k \leq n$, the function $\lambda_k(p)$ mapping a polynomial $p \in S$ to its kth smallest root is continuous.*

Definition 1.6.3. A multivariate polynomial $p \in \mathbb{R}[x_1, \ldots, x_d]$ is said to be *homogeneous* if there is an n for which, for all $x \in \mathbb{R}^d$ and real λ, $p(\lambda x) = \lambda^n p(x)$. Equivalently, p is homogeneous if all terms of p have the same total degree.

1.6.2 Hyperbolic polynomials and their cones

Definition 1.6.4. p is said to be *hyperbolic* in a direction $e \in \mathbb{R}^d$ if p is homogeneous, $p(e) > 0$, and the univariate polynomial $q(\lambda) = p(x + \lambda e)$ only has real roots for all $x \in \mathbb{R}^d$.

If p is hyperbolic in direction e, let $\Lambda_{++}(p, e)$ be the connected component of e in the set $\{x : p(x) \neq 0\}$. The *hyperbolicity cone* of p in direction e, denoted $\Lambda_+(p, e)$, is the closure of $\Lambda_{++}(p, e)$. If K is the hyperbolicity cone of some hyperbolic polynomial in some direction we say that K is a *hyperbolic cone.*

Important examples of hyperbolic polynomials are linear functionals (which are hyperbolic in the direction e if they are positive when evaluated at e), quadratics of the form $\sum_{i=2}^d x_i^2 - x_1^2$, which are hyperbolic in the direction $(1, 0, 0, \ldots, 0)^T$, and the determinant on the space of symmetric matrices of a fixed size, which is hyperbolic in the direction of the identity matrix I. The corresponding hyperbolic cones are halfplanes through the origin in the case of linear functionals, the *second-order cone* in the case of $x_2^2 + \ldots + x_d^2 - x_1^2$, and the cone of symmetric, positive semidefinite matrices in the case of det.

The hyperbolicity cone of a hyperbolic polynomial in a direction of hyperbolicity is clearly a cone, since hyperbolic polynomials are required to be homogeneous. It is a rather deeper theorem due to L. Gårding [15] that *every* hyperbolic cone is *convex*.

Motivated by the example of the positive semidefinite cone, we define the *eigenvalues* of a point $x \in \mathbb{R}^d$ with respect to the direction $e' \in \Lambda_{++}(p, e)$ to be the roots (with multiplicity) of the polynomial $\lambda \mapsto p(x + \lambda e')$. These are the usual eigenvalues if $p = \det$ and $e' = I$, and they are the entries of x if $p = x_1 \dots x_n$ and $e' = \bar{e}$.

1.7 Hyperbolic programs

One can define conic-form convex optimisation problems over hyperbolic cones. Let p be hyperbolic in direction e. A *hyperbolic program* is an optimisation problem of the form

$$\begin{array}{rrcl} \inf & c^T x & & \\ \text{subject to} & x & \in & \Lambda_+(p, e) \cap (V + b) \end{array}$$

for some linear space V and vector b.

It appears that hyperbolic programs can easily model a wider class of problems than semidefinite programs can.

We note that $-\log p(x)$ is a barrier function for $\Lambda_+(p, e)$; it is defined on $\Lambda_{++}(p, e)$ and tends to infinity near its boundary. It follows from results of O. Güler [17] that $-\log p$ is a self-concordant barrier for $\Lambda_+(p, e)$ with barrier parameter $\theta = \deg p$. Moreover, $-\log p$ shares many of the properties of $-\log\det$ that make interior-point methods for semidefinite programming so successful.

In Chapter 2, we discuss aspects of the theory of general convex sets in $\mathbb{R}^d$. In Chapter 3, we collect many useful results about polynomials in a single variable. In Chapter 4, the main body of this thesis, we discuss the theory of hyperbolic polynomials and hyperbolic cones.

It is prudent to remark that, while I have directed significant effort at giving proper attribution for results previously known, there are probably cases of lacking or improper attribution that have thus far escaped my notice. Almost everywhere, if attribution is missing, I implicitly attribute the result to the "folklore" of mathematics; the exceptions are Theorem 4.2.4 and Section 4.5.

Chapter 2

Convex sets

2.1 Convex bodies and cones

Let $G \subseteq \mathbb{R}^d$.

Definition 2.1.1. G is said to be *convex* if, for all x and y in G, every point on the line segment between x and y also lies in G. That is, G is convex if, for all x and y in G and $\lambda \in [0, 1]$, the point $\lambda x + (1 - \lambda)y$ lies in G.

The *convex hull* of G is the intersection of all convex sets containing G. The convex hull of G is denoted by $\operatorname{conv} G$.

$K \subseteq \mathbb{R}^d$ is said to be a *cone* if, for all $x \in K$ and $\lambda \geq 0$ in $\mathbb{R}$, one has $\lambda x \in K$.

G is said to be a *convex cone* if, for all x and y in G (not necessarily distinct) and nonnegative real λ and μ, the point $\lambda x + \mu y$ also lies in G. We remark that a set is a convex cone if and only if it is convex and a cone.

Proposition 2.1.2. *The convex hull of a set G is precisely the set*

$$\left\{ \sum_{g \in G} \lambda_g g : \sum_{g \in G} \lambda_g = 1, \lambda_g \geq 0 \ \ \textit{for all } g \in G \right\}$$

where all sums are understood to have finite support.

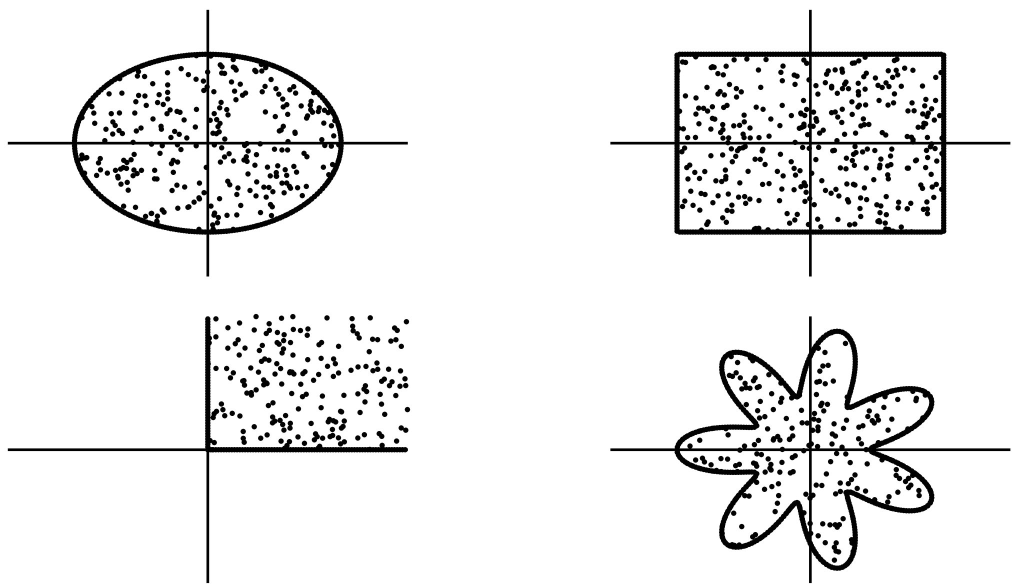

Figure 2.1: The first three pictures depict convex sets in the plane. The fourth depicts a nonconvex set. The third picture depicts a convex cone; the other three do not.

PROOF : Omitted. This is Theorem 1.1.2 of the book by R. Schneider [35]. □

Proposition 2.1.3. *Let P be a linear map from $\mathbb{R}^d$ to $\mathbb{R}^{d'}$ and let $G \subseteq \mathbb{R}^d$ Then:*

(a). If G is convex, so is PG.

(b). $P \operatorname{conv} G = \operatorname{conv} PG$.

(c). If G is a cone, so is PG.

(d). If G is a convex cone, so is PG.

(e). If $G \subseteq \mathbb{R}^d$ and $H \subseteq \mathbb{R}^e$ are two convex sets, then their Cartesian product $G \times H \subseteq \mathbb{R}^{d+e}$ is also a convex set.

(f). If $G \subseteq \mathbb{R}^d$ and $H \subseteq \mathbb{R}^d$ are two convex sets, then their Minkowski sum $G + H = \{g + h : g \in G, h \in H\}$ is also convex.

PROOF :

(a). Let x, y, and λ be such that $x \in PG$, $y \in PG$, and $\lambda \in [0, 1]$. Let x' and y' be members of G such that $Px' = x$ and $Py' = y$. Then $\lambda x + (1-\lambda)y = \lambda Px' + (1-\lambda)Py' = P\lambda x' + P(1-\lambda)y' = P(\lambda x' + (1-\lambda)y')$. Since G is convex, $\lambda x' + (1-\lambda)y'$ is in G, and hence $\lambda x + (1-\lambda)y$ is in PG.

(b). Let A be a convex set containing G. Then PA is convex (by the preceding part) and PA contains PG. Thus, $\operatorname{conv} PG \subseteq PA$. It follows that $\operatorname{conv} PG \subseteq P \operatorname{conv} G$. Now let B be a convex set containing PG. The preimage $P^{-1}B$ contains $P^{-1}PG \supseteq G$, establishing that $P \operatorname{conv} G \subseteq \operatorname{conv} PG$. The result is thus proved.

(c). Let $x \in PG$. Let g be such that $Pg = x$. Let $\lambda \geq 0$. Then $\lambda x = \lambda Pg = P(\lambda g) \in PG$. Thus PG is a cone.

(d). By (a), PG is convex. By (c), PG is a cone. Thus it follows that PG is a convex cone.

(e). Let (g_1, h_1) and $(g_2, h_2$ be in $G \times H$, and let $\lambda \in [0, 1]$. Then $\lambda(g_1, h_1) + (1 - \lambda)(g_2, h_2) = (\lambda g_1 + (1 - \lambda)g_2, \lambda h_1 + (1 - \lambda)h_2)$. Since G and H are convex, $\lambda g_1 + (1 - \lambda)g_2$ is in G and $\lambda h_1 + (1 - \lambda)h_2)$ is in H. Thus $\lambda(g_1, h_1) + (1 - \lambda)(g_2, h_2)$ is in $G \times H$. Thus $G \times H$ is convex.

(f). Let P be the linear map from $\mathbb{R}^d \times \mathbb{R}^d$ such that $P(x, y) = x + y$. Then $G + H = P(G \times H)$, so $G + H$ is a linear image of a Cartesian product of convex sets. By (a) and (d), therefore, $G + H$ is convex.

□

2.2 Combinatorial theorems about convexity

Though we shall not have occasion to use any of the following theorems except for Carathéodory's, we present the results and their proofs because they are elegant and fundamental to the broader study of convex sets in finite dimensions.

Theorem 2.2.1 (Carathéodory's Theorem (1911) [9])**.** *Let X be a subset of $\mathbb{R}^d$, and let $y \in \operatorname{conv} X$. Then there are $x_1, \ldots, x_{d+1} \in X$ such that $y \in \operatorname{conv}\{x_1, \ldots, x_{d+1}\}$.*

Proof : Let d be the least dimension in which Carathéodory's theorem fails, and let X be a d-dimensional set and y a point of $\operatorname{conv} X$ such that for any choice of $x_1, \ldots, x_{d+1}$ in X one has that $y \notin \operatorname{conv}\{x_1, \ldots, x_n\}$. Carathéodory's theorem is easily verified in one dimension, so we may assume that $d \geq 2$.

By Theorem 2.1.2, there is a finite subset $\{x_1, \ldots, x_n\} \subseteq X$ and real numbers $\lambda_1, \ldots, \lambda_n$ such that $\lambda_i > 0$ for each i, $\sum_{i=1}^n \lambda_i = 1$, and $\sum_{i=1}^n \lambda_i x_i = y$. Choose such a subset of minimum cardinality. Since $n > d + 1$, there is an affine dependence among $x_1, \ldots, x_n$ — there are $\mu_1, \ldots, \mu_n$, not all zero, such that $\sum_{i=1}^n \mu_i = 0$ and $\sum_{i=1}^n \mu_i x_i = 0$.

Choose a scalar ν such that $\lambda_i + \nu\mu_i \geq 0$ for all i and there exists a k for which $\lambda_k + \nu\mu_k = 0$. Then $\sum_{i=1}^n \lambda_i + \nu\mu_i = 0$ and $\sum_{i=1}^n (\lambda_i + \nu\mu_i)x_i = y$. Since

$\lambda_k + \nu\mu_k = 0$, however, one may omit any mention of x_k. This contradicts minimality of $\{x_1, \ldots, x_n\}$. $\square$

Theorem 2.2.2 (Radon's Theorem (1921) [32])**.** *Let* $S \subseteq \mathbb{R}^d$ *contain at least* $d+2$ *points. Then there is a subset* $T \subseteq S$ *such that* $(\operatorname{conv} T) \cap (\operatorname{conv}(S \setminus T))$ *is nonempty.*

Proof : Without loss of generality, assume that S contains exactly $d+2$ points. We again induct on d. If $d = 1$, Radon's theorem informs us that, given three points on a line, one is between the other two. This is obvious; we therefore assume that $d > 1$.

Suppose that not all points of S are vertices of $\operatorname{conv} S$. Then there is an $s \in S$ such that $s \in \operatorname{conv}(S \setminus \{s\})$; take $T = \{s\}$. We therefore assume that all points of S are vertices of $\operatorname{conv} S$. (One particular consequence of this assumption is that no three points of S are on the same straight line.)

Again without loss of generality, we shall assume that the origin is a member of S. (If not, perform an appropriate translation.) Since 0 is a vertex of $\operatorname{conv} S$, there is a vector $a \in \mathbb{R}^d$ and scalar $\beta > 0$ such that $a^T x < \beta$ for all $x \in S \setminus \{0\}$.

Let $P : \mathbb{R}^d \setminus \{0\} \to \{x \in \mathbb{R}^d : a^T x = \beta\}$ be defined by $P(x) = \frac{\beta x}{a^T x}$. (Then P is perspective projection onto the plane $a^T x = \beta$ with the camera placed at the origin.) One can show that P maps line segments to line segments and that P maps the line segment $[a, b]$ to a single point if and only if a, b, and 0 lie on the same line. This implies, in particular, that $P(\operatorname{conv} X) = \operatorname{conv} P(X)$ for any finite set X.

We apply Radon's theorem on $P(S \setminus \{0\})$, which is a set of $d+1$ points in the $(d-1)$-dimensional affine subspace $\{x \in \mathbb{R}^d : a^T x = \beta\}$. There is therefore a $T \subseteq S \setminus \{0\}$ such that $\operatorname{conv} P(T) \cap \operatorname{conv} P(S \setminus \{0\} \setminus T)$ is nonempty; let p be a point in that intersection. Then there are points $t \in \operatorname{conv} T$ and $s \in \operatorname{conv}(S \setminus \{0\} \setminus T)$ that lie on the line through 0 and p. One of these two is farther from the origin; without loss of generality, say s. Then $\operatorname{conv}(S \setminus T) \cap \operatorname{conv} T$ is nonempty, as desired. $\square$

Theorem 2.2.3 (Helly's Theorem (1923) [19])**.** *Let* $S_1, \ldots, S_n$ *be closed,* *convex subsets of* $\mathbb{R}^d$*. If, for every* $I \subseteq \{1, \ldots, n\}$ *with* $|I| \leq d+1$*, one has*

$$\cap_{i \in I} S_i \neq \emptyset$$

then

$$\cap_{i=1}^n S_i \neq \emptyset.$$

Proof : We induct on n. If $n \leq d+1$, the result clearly follows from the hypotheses. Thus we see that $n \geq d+2$. For each i with $1 \leq i \leq n$, let x_i be some point of $\cap_{j \neq i} S_j$. (This is nonempty by induction.) By Radon's theorem, one can partition the set $\{x_i : 1 \leq i \leq n\}$ into two subsets whose convex hulls intersect. Thus let I_1 and I_2 be such that $I_1 \cap I_2 = \emptyset$, $I_1 \cup I_2 = \{1, \ldots, n\}$, and $\operatorname{conv}\{x_i : i \in I_1\} \cap \operatorname{conv}\{x_i : i \in I_2\}$ intersect. Each point of $\{x_i : i \in I_2\}$ belongs to $\cap_{i \in I_1} S_i$; by convexity, so does each point of $\operatorname{conv}\{x_i : i \in I_2\}$. Similarly, each point of $\operatorname{conv}\{x_i : i \in I_1\}$ belongs to $\cap_{i \in I_2} S_i$. Let $x \in \operatorname{conv}\{x_i : i \in I_1\} \cap \operatorname{conv}\{x_i : i \in I_2\}$. Then $x \in S_i$ for all $i \in I_1$ and $x \in S_i$ for all $i \in I_2$; it follows that $x \in \cap_{i=1}^n S_i$. Thus $\cap_{i=1}^n S_i$ is nonempty, as desired. $\square$

The bounds in all three theorems are tight. For Carathéodory's theorem, $|T| \leq d$ does not suffice if one takes S to be the vertices of a d-simplex and x a point in its interior. For Radon's theorem, $|S| \geq d+1$ does not suffice since one can choose S to be the set of vertices of a d-simplex. For Helly's theorem, taking all $\mathcal{S}$ with $|\mathcal{S}| \leq d$ does not suffice since one can choose $\mathcal{F}$ to be the set of facets of a d-simplex.

Carathéodory's theorem has the following "interior" analogue due to E. Steinitz in the early 20th century. We do not prove it.

Theorem 2.2.4. *Let* $S \subseteq \mathbb{R}^d$ *be finite and let* $y \in \operatorname{int}\operatorname{conv} S$*. Then there exists a subset* $T \subseteq S$ *of size at most* $2d$ *such that* $y \in \operatorname{int}\operatorname{conv} T$*.*

2.3 Convex sets and topology

Proposition 2.3.1. *Let* $G \subseteq \mathbb{R}^d$ *be convex and let* $X \subseteq \mathbb{R}^d$*.*

(a). If $x \in \operatorname{int} G$ *and* $y \in \operatorname{cl} G$*, then* $[x, y) \subseteq \operatorname{int} G$*.*

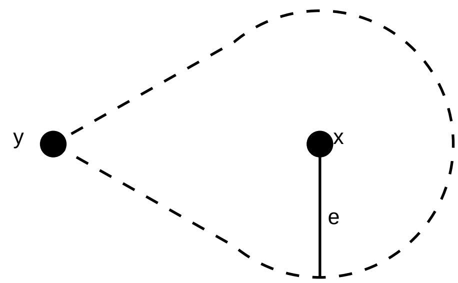

Figure 2.2: If $x \in \operatorname{int} G$, then G contains $B(x, \epsilon)$ for some $\epsilon > 0$. If $y \in \operatorname{cl} G$, then $\operatorname{int} G$ also contains the region enclosed by the dotted curve.

(b). $\operatorname{cl} G$ *and* $\operatorname{int} G$ *are convex.*

(c). If X *is open,* $\operatorname{conv} X$ *is open.*

(d). If X *is compact,* $\operatorname{conv} X$ *is compact.*

(e). $\operatorname{conv} X \subseteq \operatorname{conv} \operatorname{cl} X \subseteq \operatorname{cl} \operatorname{conv} X$.

(f). If X *is bounded,* $\operatorname{conv} \operatorname{cl} X = \operatorname{cl} \operatorname{conv} X$.

Proof : We omit the messy algebra needed to prove part (a), but the idea is contained in Figure 2.2.To see that $\operatorname{cl} G$ is convex, let x and y be in $\operatorname{cl} G$. Let $x_n \to x$ and $y_n \to y$; by continuity, $\lambda x_n + (1-\lambda) y_n \to \lambda x + (1 - \lambda y)$. Convexity of $\operatorname{int} G$ follows from part (a). If a convex set contains an open set, then so does its interior; (c) quickly follows. It is clear that $\operatorname{conv} X \subseteq \operatorname{conv} \operatorname{cl} X$.

Let $\Delta_d = \{\lambda \in \mathbb{R}^d_+ : \sum \lambda_i = 1\}$. By Carathéodory's theorem, $\operatorname{conv} X$ is a continuous image of the compact set $\Delta_d \times X^{d+1}$. If X is compact, it follows that $\operatorname{conv} X$ is compact as well. Let $x \in \operatorname{conv} \operatorname{cl} X$. Draw sequences $\{\lambda_n\}_{n=1}^{\infty}$ from Δ_d and $\{x_n\}_{n=1}^{\infty}$ from X^{d+1} such that $\sum \lambda_{n,i} x_{n,i} \to x$, λ_n converges, and x_n converges. It is then immediate that $x \in \operatorname{cl} \operatorname{conv} X$.

If X is bounded, then $\operatorname{cl} X$ is compact; hence $\operatorname{cl} \operatorname{conv} X \subseteq \operatorname{cl} \operatorname{conv} \operatorname{cl} X = \operatorname{conv} \operatorname{cl} X$. The reverse containment follows from part (e). □

Definition 2.3.2. The *affine hull* of G is the intersection of all affine spaces in $\mathbb{R}^d$ containing G. It is denoted by $\operatorname{aff} G$.

Proposition 2.3.3. $\operatorname{aff} G$ *is the set of all* $x \in \mathbb{R}^d$ *that can be written as* $\sum_{i=1}^{d+1} \lambda_i x_i$ *for some* λ *with* $\vec{e}^T \lambda = 1$ *and some* $x_1, \ldots, x_{d+1}$ *in* G.

Proof : Omitted. □

Definition 2.3.4. $x \in G$ is in the *relative interior* of G if x is in the interior of G in the usual topology on $\operatorname{aff} G$.

Proposition 2.3.5. *Let* $G \subseteq \mathbb{R}^d$ *be convex. If* G *is nonempty, then* G *has nonempty relative interior.*

Proof : Omitted. □

2.4 The metric projection

Given any closed, convex set and any point, there is a unique nearest point in the convex set to the point. The proof is elementary and is omitted; however, this fact is so important that it earns a name and a convenient notation:

Definition 2.4.1. The *metric projection* to a convex set G maps a point x to its nearest point in G. The metric projection of x onto G is denoted by $\operatorname{proj}(G, x)$.

It will be convenient to fix a closed, convex set $G \subseteq \mathbb{R}^d$ for the remainder of this section.

If $x \notin G$, then the angle from a point of G to $\operatorname{proj}(G, x)$ to x is always obtuse:

Proposition 2.4.2. *Let* $x \notin G$ *and let* $z \in G$. *Then*

$$(z - \operatorname{proj}(G, x))^T (x - \operatorname{proj}(G, x)) \leq 0.$$

Proof : Suppose otherwise. Let $z \in G$ be such that $(z - \operatorname{proj}(G, x))^T (x - \operatorname{proj}(G, x)) > 0$. Then, by convexity, $[\operatorname{proj}(G, x), z] \subseteq G$. We observe that the function $f(\lambda) = ||x - \lambda z + (1 - \lambda) \operatorname{proj}(G, x)||^2$ is differentiable. Since proj(G, x) is the nearest point of G to x, f must have nonnegative derivative at zero. Sadly, its derivative at zero is

$$f'(0) = 2(x - \text{proj}(G, x))^T (z - \text{proj}(G, x)),$$

from which the theorem follows. $\square$

In particular, the hyperplane

$$\{z \in \mathbb{R}^d : (x - \text{proj}(G, x))^T z = (x - \text{proj}(G, x))^T \text{proj}(G, x)\}$$

defines a hyperplane separating G from x.

Proposition 2.4.3. *The metric projection is Lipschitz with constant 1 and hence continuous.*

Proof : Let x and y be in $\mathbb{R}^d$. Let $a = \text{proj}(G, x)$ and $b = \text{proj}(G, y)$. Then $(b-a)^T(x-a) \leq 0$ and $(a-b)^T(y-b) \leq 0$; thus, $(b-a)^T(x-a-y+b) = (b-a)^T(x-y) + ||b-a||^2 \leq 0$. We note that $|(b-a)^T(x-y)| \leq ||b-a||||y-x||$; it quickly follows that $||y-x|| \geq ||b-a||$. $\square$

Proposition 2.4.4. *Let $x \in \mathbb{R}^d \setminus G$. Let $y = \text{proj}(G, x) + \lambda(x - \text{proj}(G, x))$ for some $\lambda \geq 0$. Then* proj$(G, y) = $ proj(G, x)*.*

Proof : Omitted. $\square$

Proposition 2.4.5. *Let G be compact and convex. Let R be so large that $G \subseteq B(0, R)$. The image of the sphere $S(0, R)$ under the metric projection* proj$(G, \cdot)$ *is* bd G*.*

Proof : $S(0, R)$ is compact, so its image under the continuous function proj$(G, \cdot)$ is compact as well. It is certainly contained in bd G. To prove the reverse inclusion, suppose $x \in \text{bd}\, G$ is not in proj$(G, S(0, R))$. By compactness, there is an $\epsilon > 0$ such that $B(x, \epsilon) \cap \text{proj}(G, S(0, R)) = \emptyset$; choose $y \in B(x, \epsilon) \setminus G$. Since proj$(G, \cdot)$ is Lipschitz of constant 1, proj$(G, x) = x$, $||x - y|| < \epsilon$, and there are no points of proj$(G, S(0, R))$ within ϵ of x, we see that proj$(G, y) \notin$ proj$(G, S(0, R))$. Sadly, there is some point of $S(0, R)$ on the ray from proj(G, y) through y, and all such points map to proj(G, y) under the metric projection; this is a contradiction. $\square$

Definition 2.4.6. A *supporting hyperplane* of G is a hyperplane $H = \{x \in \mathbb{R}^d : a^T x = b\}$ such that all points g of G satisfy $a^T g \leq b$.

Theorem 2.4.7. *Let $x \in \operatorname{bd} G$. There is a supporting hyperplane to G containing x.*

PROOF : Let $y \in \mathbb{R}^d \setminus G$ be such that $\operatorname{proj}(G, y) = x$. Then, by Proposition 2.4.2 $(z - x)^T (y - x) \leq 0$ for all $z \in G$; take the hyperplane $\{z \in \mathbb{R}^d : (y - x)^T z = (y - x)^T x\}$. $\square$

2.5 Faces

Definition 2.5.1. $F \subseteq G$ is a *face* of G if F is convex and, whenever x and y are in G and $(x + y)/2$ is in F, both x and y must also be in F.

A subset F of G is an *exposed face* if it is the intersection of G with a supporting hyperplane.

Example 2.5.2. As the nomenclature implies (and as can easily be proven), all exposed faces are faces. However, the converse is not true, as can be seen by examining the point $(1, 1)$ in the Minkowski sum $[(0, 0), (1, 0)] + B(0, 1)$ depicted in Figure 4.2.

Proposition 2.5.3. *Faces of faces of G are faces of G.*

PROOF : Let F be a face of G and let E be a face of F. Let $x \in E$, and suppose that $x = (y + z)/2$ where y and z are in G. Since $x \in F$ and F is a face, y and z must be in F. Sadly, E is a face of F; y and z must therefore lie in E. $\square$

Proposition 2.5.4. *Exposed faces of exposed faces of G are not necessarily exposed faces of G.*

PROOF : Let G be the set $[(0, 0)^T, (1, 0)^T] + B(0, 1)$. Then $F = [(0, 1)^T, (1, 1)^T]$ is an exposed face of G and $E = \{(0, 1)^T\}$ is an exposed face of F. Sadly, E is not an exposed face of G. $\square$

Proposition 2.5.5. *Every point of G lies in the relative interior of some face of G.*

PROOF : We induct on the dimension of G. The result is clear if the dimension is zero. Let $x \in G$. If $x \in \operatorname{int} G$, then x lies in the relative interior of G. Otherwise, $x \in \operatorname{bd} G$. Then x lies in some exposed face F of G. F has dimension strictly smaller than $\dim G$, so x lies in the relative interior of some face E of F. E is then a face of G, yielding the desired result. $\square$

Definition 2.5.6. Let $G \subseteq \mathbb{R}^d$ be a closed, convex set. An *extreme point* of G is a point $x \in G$ such that $\{x\}$ is a face of G. An *exposed extreme point* of G is a point $x \in G$ such that $\{x\}$ is an exposed face.

Let $K \subseteq \mathbb{R}^d$ be a convex cone. An *extreme ray* of K is a ray $\{\lambda x : \lambda \geq 0\}$ that is a face of G and is not $\{0\}$. An *exposed extreme ray* of G is a ray $\{\lambda x : \lambda \geq 0\}$ that is an exposed face of G and is not $\{0\}$.

Proposition 2.5.7. *The extreme rays of $\mathbb{R}^d_+$ are the rays defined by the vectors of the standard basis.*

PROOF : Let $x \in \mathbb{R}^d_+ \setminus \{0\}$. Suppose that $x_i \neq 0$ and $x_j \neq 0$. Fix $\epsilon < x_i$. Then $x \pm \epsilon e_i$ is not a scalar multiple of x. Further, $x = \frac{1}{2}(x + \epsilon e_i) + \frac{1}{2}(x - \epsilon e_i)$. It follows that x does not define an extreme ray if x has at least two nonzero components.

Now suppose that x has exactly one nonzero component — that is, $x = x_i e_i$ for some i. Further suppose that $x = \lambda v + (1 - \lambda) w$ for some v and w in $\mathbb{R}^d_+$ and $\lambda \in (0, 1)$. For all $j \neq i$, we have that $x_j = 0$, $v_j \geq 0$, and $w_j \geq 0$. It follows that $v_j = w_j = 0$. Thus v and w are scalar multiples of e_i. This proves that x defines an extreme ray. $\square$

Proposition 2.5.8. *Every nonzero boundary point of a second-order cone*

$$\{x \in \mathbb{R}^d : x_2^2 + \ldots + x_d^2 \leq x_1^2\}$$

defines an extreme ray.

PROOF : Let x be such that $x_2^2 + \ldots + x_d^2 = x_1^2$ and $x_1 > 0$ — that is, x is a nonzero boundary point of the d-dimensional second-order cone. Suppose that $x = \lambda v + (1-\lambda)w$ for some v and w in SOC^d and $\lambda \in (0,1)$. By the triangle inequality,

$$\begin{aligned}&\sqrt{x_2^2 + \ldots + x_d^2} \\ &\quad \leq \lambda\sqrt{v_2^2 + \ldots + v_d^2} + (1-\lambda)\sqrt{w_2^2 + \ldots + w_d^2} \leq \lambda v_1 + (1-\lambda)w_1 = x_1\end{aligned}$$

Equality holds in the first inequality if and only if $(v_2, \ldots, v_d)$ and $(w_2, \ldots, w_d)$ are scalar multiples of one another. Equality holds in the second inequality if and only if $v_1 = \sqrt{v_2^2 + \ldots + v_d^2}$ and $w_1 = \sqrt{w_2^2 + \ldots + w_d^2}$. Thus one has equality if and only if v and w are scalar multiples of one another. It follows that x defines an extreme ray. $\square$

Proposition 2.5.9. *The extreme rays of a positive semidefinite cone are defined by the positive semidefinite matrices of rank one.*

PROOF : Let M be a positive semidefinite matrix of rank at least two. By the spectral theorem, one may write

$$M = \sum_{i=1}^{\operatorname{rank} M} \lambda_i P_i$$

where, for each i, P_i is a rank-one positive semidefinite matrix. The spectral theorem further guarantees that the matrices P_i are linearly independent. One can thus write M, for instance, as $M = (\lambda_1 P_1) + (\lambda_2 P_2 + \ldots + \lambda_n P_n)$ and it is clear from this and linear independence that M defines no extreme ray.

Thus all extreme rays of a positive semidefinite are defined by matrices of rank one. We now check that every matrix of rank one defines an extreme ray. We recall that every rank-one positive semidefinite matrix is of the form vv^T for some vector v. Suppose that $vv^T = \lambda A + (1-\lambda)B$ for positive semidefinite A and B and $\lambda \in (0,1)$. If $w \perp v$, then $0 = w^T vv^T w = \lambda w^T A w + (1-\lambda) w^T B w$. By positive semidefiniteness, we must have $w^T A w = w^T B w = 0$.

We recall that, if M is positive semidefinite, then $w^T M w = 0$ iff $w \in \ker M$. Hence $v^\perp \subseteq \ker A$ and $v^\perp \subseteq \ker B$. Thus A and B are both of rank one; let $A = aa^T$ and $B = bb^T$. Since $v^\perp \subseteq \ker aa^T$, we see that $a \in (v^\perp)^\perp = \operatorname{span} v$.

Similarly, $b \in \operatorname{span} v$. Thus there are α and β such that $a = \alpha v$ and $b = \beta v$; it follows that aa^T and bb^T are scalar multiples of vv^T. Thus vv^T defines an extreme ray of the positive semidefinite cone. $\square$

2.6 Curvature and smoothness

Let G be a closed convex set with interior. We study the consequences of various differentiability assumptions on $\operatorname{bd} G$.

Definition 2.6.1. Let $x \in \operatorname{bd} G$. The *normal cone* to G at x is

$$N(G, x) = \{y \in \mathbb{R}^d : \operatorname{proj}(G, x + y) = x\}.$$

Let $x \in \operatorname{bd} G$. x is said to be a *regular* boundary point if $N(G, x)$ is a ray. The *outward unit normal* to G at a regular boundary point x is the unique element of $N(G, x)$ of norm one.

If x is a regular boundary point, the *tangent space* to G at x is the set

$$T_x G = \{y \in \mathbb{R}^d : z \in N(G, x) \implies \langle y, z\rangle = 0\}.$$

Example 2.6.2. Let $G = \{x \in \mathbb{R}^d : Ax \leq b\}$ be a polytope. Let $x \in \operatorname{bd} G$. Let $A^= x \leq b^=$ be the subsystem of the defining inequalities that are satisfied with equality at x. Then $N(G, x)$ is the convex cone generated by the rows of $A^=$. Thus x is a regular boundary point iff x lies in the relative interior of a facet of G. The tangent space to G at a regular boundary point x is the affine hull of the facet containing x in its relative interior.

Example 2.6.3. Let G be the unit disc in the plane. The normal cone to a point x on the boundary of G is the ray generated by x. It follows that every boundary point of G is regular. The tangent space to G at a boundary point x is the translate of the line tangent to the unit circle at x that passes through the origin.

The next result states that G near a boundary point is locally the epigraph of a convex function. We call this function the local representation of G at x.

Theorem 2.6.4. *Let $x \in \operatorname{bd} G$ be a regular boundary point, and let $-e$ be the outward unit normal to G at x. Then there exists an $\epsilon > 0$ and a convex real function f on $T_xG \cap B(0,\epsilon)$ such that*

- $f(0) = 0$.
- $f(y) \geq 0$ *for all* $y \in T_xG \cap B(0,\epsilon)$.
- $x + y + f(y)e \in \operatorname{bd} G$ *for all* $y \in T_xG \cap B(0,\epsilon)$.

Proof : The proof of this theorem lies outside the scope of our investigation; we therefore omit it. □

We say that G is "$\mathcal{C}^k$ at x" if its local representation at x is $\mathcal{C}^k$ at 0.

Theorem 2.6.5. *G is $\mathcal{C}^1$ if and only if every boundary point is regular.*

Definition 2.6.6. Let G be closed and convex and let $x \in \operatorname{bd} G$. Let f be the local representation of G at x. We define the *upper curvature* $\overline{\kappa}$ by

$$\overline{\kappa}(G,x,h) = \limsup_{t \to 0^+} \frac{2f(th)}{t^2}$$

and *lower curvature* $\underline{\kappa}$ by

$$\underline{\kappa}(G,x,h) = \liminf_{t \to 0^+} \frac{2f(th)}{t^2},$$

both for any $h \in T_xG$.

When the two agree, we simply call them the *curvature* and denote it by $\kappa(G,x,h)$. We shall show that the two agree precisely when G is $\mathcal{C}^2$ at x.

Proposition 2.6.7. *Suppose G is $\mathcal{C}^2$ at x. Then, for all $h \in T_xG$, $\overline{\kappa}(G,x,h) = \underline{\kappa}(G,x,h)$. Furthermore, κ defines a bilinear form on T_xG.*

Proof : Let f be the local representation of G at x. Then $f(0) = 0$ and $(\nabla f)(0) = 0$. Thus, $f(-v) = f(v) + o(||v||^2)$. We therefore compute

$$\begin{aligned}
\lim_{t \to 0^+} \frac{2f(th)}{t^2} = \lim_{t \to 0^+} & \frac{f(th) + f(-th) + o(t^2)}{t^2} \\
&= \lim_{t \to 0^+} \frac{f(th) - 2f(0) + f(-th)}{t^2} + o(1) = \frac{\partial^2}{\partial t^2} f(th).
\end{aligned}$$

Thus $\underline{\kappa} = \overline{\kappa}$ and κ is merely the Hessian of f at zero. □

2.7 Convex sets described by the zero set of a function

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a function such that $G := \{x \in \mathbb{R}^d : f(x) \geq 0\}$ is convex. Assume that $\operatorname{bd} G = \{x \in \mathbb{R}^d : f(x) = 0\}$ and that $\operatorname{int} G = \{x \in \mathbb{R}^d : f(x) > 0\}$.

Proposition 2.7.1. *Suppose that f is $\mathcal{C}^1$ and $x \in \operatorname{bd} G$. Then, if $(\nabla f)(x) \neq 0$, a supporting hyperplane for G at x is $\{y : (\nabla f)(x)y = 0\}$.*

Proof : If y is such that $(\nabla f)(x)y < 0$, then, for some point z on the line segment $[y, x)$, we have that $(\nabla f)(x)z < 0$. Thus having $y \in G$ would violate convexity of G. □

Proposition 2.7.2. *If f is $\mathcal{C}^2$, $x \in \operatorname{bd} G$, and $(\nabla f)(x) \neq 0$, then the curvature map κ at x is simply*

$$\frac{1}{2||(\nabla f)(x)||}(Hf)(x).$$

Proof : Let g be the local representation of G at x. Let $h \in T_x G$. We note that

$$f(x+th) = f(x)+t(\nabla f)(x)h+t^2(Hf)(x)[h,h]+o(t^2) = t^2(Hf)(x)[h,h]+o(t^2).$$

(The linear term vanishes because the gradient must be orthogonal to the tangent space.)

When t is near zero, we have

$$\begin{aligned} g(th) &= f(x+th)/\frac{\partial f(x+th+\lambda(\nabla f)(x+th))}{\partial \lambda} \\ &= f(x+th)/||(\nabla f)(x+th)|| = t^2(Hf)(x)[h,h]/||(\nabla f)(x+th)|| + o(t^2). \end{aligned}$$

Thus,

$$\begin{aligned} \kappa(x)[h,h] &= \lim_{t\to 0} \frac{g(th)}{t^2} = \lim_{t\to 0} \frac{t^2(Hg)(0)[h,h]+o(t^2)}{t^2} \\ &= \lim_{t\to 0}(Hg)(0)[h,h]+o(1) = (Hg)(0)[h,h] = \frac{(Hf)(x)[h,h]}{2||(\nabla f)(x)||}. \end{aligned}$$

□

2.8 Polarity and duality

We next consider three notions of duality in convex geometry. We roughly follow the presentation in the book by R. Schneider [35].

Definition 2.8.1. Let conv^d be the set of nonempty compact convex subsets of $\mathbb{R}^d$. We call sets in conv^d "*convex bodies.*" Let conv^d_0 be the set of convex bodies with interior. Let conv^d_{00} be the set of convex bodies containing the origin in their interior.

Let G be a convex subset of $\mathbb{R}^d$. The *polar body* of G, denoted G^*, is the set

$$G^* = \{y \in \mathbb{R}^d : \langle x, y\rangle \leq 1 \text{ for all } x \in G\}.$$

Thus, G^* can be considered the set of linear functionals on $\mathbb{R}^d$ mapping each point of G to a value at most 1. Polarity behaves particularly well on conv^d_{00}:

Proposition 2.8.2. *If $G \in \text{conv}^d$, then $G^* \in \text{conv}^d$ if and only if $G \in \text{conv}^d_{00}$. Furthermore, if $G \in \text{conv}^d$, then $G^{**} = G$.*

Proof : It is immediate from the definition that $G \subseteq G^{**}$. Thus we prove only that $G^{**} \subseteq G$. Let $x \in G^{**} \setminus G$. Then there is a vector $a \in \mathbb{R}^d$ for which $a^T x > 1$ and $a^T y \leq 1$ for all $y \in G$. Since $a^T y \leq 1$ for all $y \in G$, we see that $a \in G^*$. Unfortunately, $a^T x > 1$, so $x \notin G^{**}$. This is a contradiction from which we conclude that $G^{**} \subseteq G$. Thus equality is proven. □

Proposition 2.8.3. *Let A and B be members of conv^d_{00} such that $A \subseteq B$. Then $A^* \supseteq B^*$.*

Proof : Let $y \in B^*$. Then, for every $x \in B$ — and, in particular, every $x \in A$ — we have $x^T y \leq 1$. Thus $y \in A^*$ and the theorem is proved. □

Proposition 2.8.4. *Let A and B be in conv^d_{00}. Then $(A \cap B)^* = \text{conv}(A^* \cup B^*)$ and $\text{conv}(A \cup B)^* = A^* \cap B^*$.*

Proof : Let $x \in \text{conv}(A^* \cup B^*)$. Let $a \in A^*$, $b \in B^*$, and $\lambda \in [0, 1]$ be such that $x = \lambda a + (1 - \lambda)b$. Now let $y \in A \cap B$. Since $y \in A$, $a^T y \leq 1$; since $y \in B$, $b^T y \leq 1$. Thus $x^T y \leq 1$. It follows that $\text{conv}(A^* \cup B^*) \subseteq (A \cap B)^*$.

Applying this result to A^* and B^* and using the relation $G^{**} = G$ and the inclusion-reversing property, we see that $\operatorname{conv}(A \cup B)^* \supseteq A^* \cap B^*$.

Now let $x \in \operatorname{conv}(A \cup B)^*$. Then $x^T a \leq 1$ for all $a \in A$ and $x^T b \leq 1$ for all $b \in B$. Thus $\operatorname{conv}(A \cup B)^* \subseteq A^* \cap B^*$. Applying this result to A^* and B^*, and using $G^{**} = G$ and the inclusion-reversing property again, we see that $\operatorname{conv}(A^* \cup B^*) \supseteq (A \cap B)^*$. Thus the theorem is proved. $\square$

To each exposed face of a convex body, one can associate a corresponding exposed *conjugate face* of the polar body. This notion of conjugacy is well-behaved — the conjugate of the conjugate of an exposed face is the face itself, and conjugacy is an inclusion-reversing bijection from the lattice of exposed faces of a convex body to the lattice of exposed faces of its polar body.

Definition 2.8.5. Fix a convex body G. Let $F \subseteq G$. The *conjugate face* of F is the face $\hat{F}$ of G^* given by

$$\hat{F} = \{x \in G^* : \langle x, y \rangle = 1 \text{ for all } y \in F\}.$$

Proposition 2.8.6. *Let F be a face of G, and let $x \in \operatorname{relint} F$. Then $\hat{F} = \{y \in G^* : \langle x, y \rangle = 1\}$.*

Proof : One direction is clear, namely that $\hat{F} \subseteq \{y \in G^* : \langle x, y \rangle = 1\}$. Now suppose that y is such that $\langle x, y \rangle = 1$ but $y \notin \hat{F}$. Let $z \in G$ be such that $\langle z, y \rangle \neq 1$. since $y \in G^*$ and $z \in G$, we see that $\langle z, y \rangle < 1$. Let x' be some point of F on the ray $\{x + \lambda(x - z) : \lambda > 0\}$. (Such a point exists since x lies in the relative interior of F.) We observe that, for some positive λ,

$$\langle x', y \rangle = \langle x, y \rangle + \lambda \langle x - z, y \rangle > \langle x, y \rangle.$$

Thus $y \notin G^*$, a contradiction from which the containment $\hat{F} \supseteq \{y \in G^* : \langle x, y \rangle = 1\}$ follows. $\square$

Proposition 2.8.7. *Let F be a proper face of the convex body G. Then $\hat{F}$ is a nonempty exposed face of G^* and $\hat{\hat{F}}$ is the smallest exposed face of G containing F.*

Proof : Let $x \in \operatorname{relint} F$. Then $x \in \operatorname{bd} G$. Since $G^{**} = G$, there is a point y of G^* for which $\langle x, y \rangle = 1$. Thus $\hat{F}$ is nonempty. Exposure of $\hat{F}$ then follows from the preceding proposition.

Let $x \in F$ and let $y \in \hat{F}$. Then $\langle x, y \rangle = 1$, so $x \in \hat{\hat{F}}$. Thus $\hat{\hat{F}}$ is an exposed face of G containing F. Let $E = \{x \in G : \langle a, x \rangle = 1\}$ be an exposed face of G that contains F. Then $a \in \hat{F}$. Consequently, $\hat{\hat{F}} \subseteq \{x \in G : \langle a, x \rangle = 1\} = E$. This proves that $\hat{\hat{F}}$ is the smallest exposed face of G containing F, as desired. □

Definition 2.8.8. Let cone^d be the set of pointed closed convex cones with interior in $\mathbb{R}^d$. Let $K \in \mathrm{cone}^d$. Define the *dual cone* of K to be

$$K^* = \{x \in \mathbb{R}^d : \langle x, y \rangle \geq 0 \text{ for all } y \in K\}.$$

Proposition 2.8.9. *If $K \in \mathrm{cone}^d$, then $K^* \in \mathrm{cone}^d$. Furthermore, $K^{**} = K$.*

PROOF : This proof is so similar to the analogous proof for convex bodies that we omit it. □

Proposition 2.8.10. *Let A and B be in cone^d. Then $(A \cap B)^* = \mathrm{cl}(A^* + B^*)$ and $(\mathrm{cl}(A + B))^* = A \cap B$.*

PROOF : Omitted. □

2.9 Automorphisms of convex sets

Definition 2.9.1. Let G be a convex subset of $\mathbb{R}^d$. The *automorphism group* of G, denoted $\mathrm{Aut}\, G$, is the set of invertible affine maps T such that $TG = G$.

Theorem 2.9.2. *The automorphism group of a closed, convex set is closed in the group of invertible affine maps.*

PROOF : Let G be a closed, convex set. Let $\{T_n\}_{n=1}^{\infty}$ be a sequence of linear maps in $\mathrm{Aut}\, G$ that tends to the invertible linear map T. For each $x \in G$, we see that $Tx = \lim_{n \to \infty} T_n x$; since G is closed, this limit lies in G. Thus T maps points of G to points of G — that is, $TG \subseteq G$.

By continuity of inverses, we see that $T^{-1} = \lim_{n\to\infty} T_n^{-1}$. One can then apply the argument above to show that $T^{-1}G \subseteq G$. It follows that $G \subseteq TG$; equality is therefore proven. Thus $T \in \operatorname{Aut} G$ as well. $\square$

In particular, the automorphism group of a closed, convex set is a matrix Lie group.

It is clear that any automorphism of G must fix the centroid of G. We can prove more, however: every automorphism of G must fix the centre of the *Löwner-John ellipsoid* of G, which is the smallest ellipsoid containing G.

Theorem 2.9.3. $\det^{1/d}$ *is concave on* Σ^d_{++}.

Proof : We note that det is hyperbolic in direction I. Thus, from Gårding's inequality 4.1.11, we immediately see that $\det^{1/d}$ is concave. $\square$

Theorem 2.9.4. *Let A and B be in Σ^d_{++}. If A and B are not scalar multiples of one another, then* $\det^{1/d}$ *is strictly concave on the line segment between A and B.*

Proof : If $\det^{1/d}$ is not strictly concave on $[A, B]$, then there is a nonempty subinterval of $[A, B]$ on which $\det^{1/d}$ is linear. Since $\det^{1/d}$ is the dth root of a polynomial, this implies that det has a root of order d somewhere on the line through A and B. The only root of order d that det has is at zero; therefore A and B are scalar multiples of one another. This is a contradiction; $\det^{1/d}$ must be strictly concave. $\square$

The following theorem is classical and due to Fritz John [23].

Theorem 2.9.5 (Fritz John's theorem [23])**.** *Let G be a compact subset of $\mathbb{R}^d$ with interior. There exists a unique minimum-volume ellipsoid containing G.*

Proof : Let

$$X = \{(A, c) : ||Ax - c|| \leq 1 \text{ for all } x \in G;\ A \in \Sigma^d_+;\ c \in \mathbb{R}^d\}.$$

We prove that X is convex. Let (A_1, c_1) and (A_2, c_2) be in X, and let $\lambda \in [0, 1]$. If $x \in G$, then

$$||(\lambda A_1 + (1-\lambda)A_2)x - \lambda c_1 - (1-\lambda)c_2|| = ||\lambda(A_1x - c_1) + (1-\lambda)(A_2x - c_2)||$$
$$\leq ||\lambda(A_1x - c_1)|| + ||(1-\lambda)(A_2x - c_2)|| \leq \lambda||A_1x - c_1|| + (1-\lambda)||A_2x - c_2|| \leq 1.$$

Thus X is convex. By Gårding's inequality, $\det^{1/d}$ is concave on Σ_+^d and hence on X.

One can check that the level sets of det in X are bounded; it follows that det has a maximiser in X. Suppose that det has two distinct maximisers, say (A_1, c_1) and (A_2, c_2), in X. We note that

$$\det \frac{A_1 + A_2}{2} = \frac{1}{2^d} \det(I + A_1^{-1/2} A_2 A_1^{-1/2}).$$

Let $\lambda_1, \ldots, \lambda_d$ be the eigenvalues of $A_1^{-1/2} A_2 A_1^{-1/2}$; not all of these are one since $\{A_1, A_2\}$ is a linearly independent set. By the Cauchy-Schwarz inequality,

$$\frac{1}{2^d} \det(I + A_1^{-1/2} A_2 A_1^{-1/2}). = \frac{1}{2^d} \prod_{i=1}^{d} (1 + \lambda_i) > \frac{1}{2^d} \prod_{i=1}^{d} 2\lambda_i = \prod_{i=1}^{d} \lambda_i = 1,$$

a contradiction to the maximality of $\det A_1$ and $\det A_2$.

Suppose that $E = \{x \in \mathbb{R}^d : ||Ax - c|| \le 1\}$ is an ellipsoid containing G. Note that $Ax - c = A(x - A^{-1}c)$; it quickly follows that the volume of E is $1/\det A$. Note also that $x \in E$ if and only if

$$(x - A^{-1}c)^T A^T A (x - A^{-1}c) \le 1.$$

Let B be the unique positive definite square root of $A^T A$. Then $x \in E$ if and only if

$$||Bx - BA^{-1}c|| \le 1.$$

We may therefore assume that the A in the definition of E is positive definite, and hence that (A, c) lies in X. Thus the unique maximiser of det corresponds to the unique minimum- volume ellipsoid containing G. Fritz John's theorem is therefore proved. $\square$

We note that any automorphism for a convex body G must be an automorphism of its Löwner-John ellipsoid. We note that the automorphism group of an ellipsoid centred at the origin is simultaneously similar to the orthogonal group, with the similarity being given by any linear map sending the unit sphere to the ellipsoid. Thus we are justified in considering only convex bodies whose Löwner- John ellipsoid is the unit ball centred at the origin.

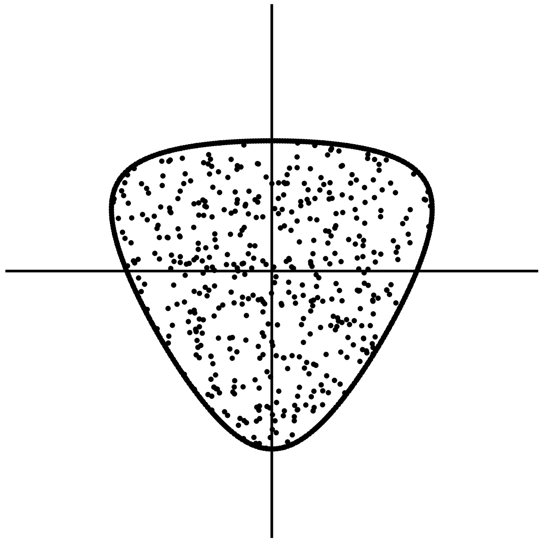

Figure 2.3: The set G from Example 2.9.7.

The automorphism group of any such body will be a (compact) subgroup of the orthogonal group $O(d)$.

We now compute the automorphism groups of a few convex bodies.

Example 2.9.6 (The simplex)**.** Let G be a regular d-simplex centred at the origin such that each vertex is at distance one from the origin. Then the Löwner-John ellipsoid of G is the unit ball centred at the origin. G has $d+1$ points at distance one from the origin, so every automorphism must permute these $d+1$ points. Furthermore, a permutation of the $d+1$ vertices of G defines a unique affine map. It can be checked that every such map is an automorphism of G.

Example 2.9.7. We study the convex body in Figure 2.3. Let

$$G = \left\{ (x,y) \in \mathbb{R}^2 : (y-1)\left(\frac{\sqrt{3}}{2}x - \frac{1}{2}y - 1\right)\left(\frac{-\sqrt{3}}{2}x - \frac{1}{2}y - 1\right) \leq \frac{-1}{2} \right\}$$

The maps R of rotation by 120 degrees about the origin and S of reflection about the y-axis defined by

$$R(x,y) = \left(\frac{-x}{2} + \frac{\sqrt{3}y}{2}, \frac{-\sqrt{3}x}{2} - \frac{y}{2}\right)$$

and $S(x, y) = (-x, y)$ are linear. It can be checked using the definition of G that $R \in \operatorname{Aut} G$ and $S \in \operatorname{Aut} G$. It follows that R must be an automorphism of the minimum-volume ellipsoid containing G and hence that the minimum-volume ellipsoid containing G is a disc centred at the origin. G has three farthest points from the origin; they lie along the rays at 30, 150, and 270 degrees. Every automorphism of G must permute these three farthest points. Since these three farthest points form an affinely independent set, an automorphism of G is defined by its action on the three farthest points. It follows that $\operatorname{Aut} G$ is simply the group generated by R and S — it is isomorphic to the dihedral group with six elements.

We now consider the automorphism group of the cone

$$K := \left\{ \begin{pmatrix} \lambda x \\ \lambda \end{pmatrix} : x \in G, \lambda \geq 0 \right\}$$

for some fixed convex body G whose Löwner-John ellipsoid is the unit ball centred at the origin. We can naturally identify $\operatorname{Aut} G$ with a subgroup of $\operatorname{Aut} K$ by the map

$$T \mapsto \begin{pmatrix} T & 0 \\ 0 & 1 \end{pmatrix}$$

Proposition 2.9.8. $\operatorname{Aut} G$ *is the subgroup of* $\operatorname{Aut} K$ *that maps the hyperplane* $\{(x, \lambda) : \lambda = 1\}$ *to itself.*

PROOF : Trivial. □

One might conjecture that $\operatorname{Aut} G$ is a normal subgroup, or even a direct summand, of $\operatorname{Aut} K$. This is not the case, as the following example shows:

Example 2.9.9. Let $G = [-1, 1] \subseteq \mathbb{R}$. Then

$$K = \{(x, y)^T : x + y \geq 0 \text{ and } y - x \geq 0\}.$$

The map $T(x, y) = (-x, y)$ is in $\operatorname{Aut} G$. The map $T'(x, y) = (3x + y, x + 3y)$ is in $\operatorname{Aut} K$. Unfortunately, $T'^{-1}(T(T'(1, 1))) = T'^{-1}(T(4, 4)) = T'^{-1}(-4, 4) = (-2, 2)$. Thus $T'^{-1}TT'$ is clearly not a member of $\operatorname{Aut} G$.

I believe, but I cannot prove, the following result:

Conjecture 2.9.10. $\operatorname{Aut} K$ is the semidirect product of some subgroup of $\operatorname{Aut} K$ with $\operatorname{Aut} G$. That is, there is a normal subgroup X of $\operatorname{Aut} K$ such that every element of $\operatorname{Aut} K$ may be written uniquely as xg for some $x \in X$ and $g \in \operatorname{Aut} G$.

I further suspect that this subgroup X is isomorphic to $\mathbb{R}^n$ under addition for some n.

One can impose conditions on $\operatorname{Aut} K$ and $\operatorname{Aut} G$ so that this result becomes almost trivial. Namely, if in each connected component of $\operatorname{Aut} K$ there is a unique element of $\operatorname{Aut} G$, then $\operatorname{Aut} K$ is the product of the connected component of the identity in $\operatorname{Aut} K$ with $\operatorname{Aut} G$.

2.10 Homogeneous and symmetric cones

Definition 2.10.1. Let K be a convex cone. We say that K is *homogeneous* if $\operatorname{Aut} K$ acts transitively on $\operatorname{int} K$. That is, if, for all x and y in $\operatorname{int} K$ there exists a $T \in \operatorname{Aut} K$ for which $Tx = y$.

We say that K is *symmetric* if K is homogeneous and $K = K^*$.

Symmetric cones have a very deep connection with the theory of Jordan algebras. The cone of squares in a Euclidean Jordan algebra is a symmetric cone. Moreover, every symmetric cone arises as the cone of squares in a Euclidean Jordan algebra. Using the classification theorem for Euclidean Jordan algebras established by P. Jordan, J. von Neumann, and E. Wigner [24], one can prove that every symmetric cone arises as a direct sum of members of five simple classes of cone — the second-order cones, the symmetric positive-semidefinite real matrices of a particular size, the Hermitian positive-semidefinite complex matrices of a particular size, the Hermitian positive-semidefinite quaternionic matrices of a particular size, and the Hermitian 3×3 matrices over the octonions. It quickly follows from this result that each symmetric cone arises as a slice of a semidefinite cone. J. Faraut and A. Korányi have written a book [13] on the theory of symmetric cones which contains proofs of the above results and many further details about symmetric cones.

It is a recent theorem, discovered independently by C. B. Chua and L. Faybusovich ([11]; [14]), that every homogeneous cone arises as the intersection of a semidefinite cone with an affine space. Their proofs rely upon the machinery of T-algebras, developed by E. B. Vinberg ([40]; [39]), which give an algebraic characterisation of the structure of a homogeneous cone.

We now attempt to give a taste of the theory of homogeneous cones. We closely follow the presentation of Truong and Tunçel [37].

Definition 2.10.2. Let $K \subseteq \mathbb{R}^d$ be a closed convex cone. A *K-symmetric bilinear form* on $\mathbb{R}^n$ is a bilinear map $T : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^d$ such that:

- $T(u, v) = T(v, u)$ for all u and v in $\mathbb{R}^n$.
- $T(u, u) \in K$ for all $u \in \mathbb{R}^n$.
- If $T \neq 0$, then $T(u, u) = 0$ implies $u = 0$.

We note that if $d = 1$ and K is the nonnegative real axis, then the K-symmetric bilinear forms on $\mathbb{R}^n$ are precisely the inner products on $\mathbb{R}^n$ together with the zero map. We shall continue our development, however, bearing in mind a less trivial example:

Example 2.10.3. Let K be the cone Σ^n_+ (in the ambient space Σ^n, so $d = \binom{n}{2}$). Let T be the bilinear map on $\mathbb{R}^{k\times n}$ given by $T(U, V) = \frac{1}{2}(UV^T + VU^T)$. It is immediate that T is symmetric and bilinear. If $U \neq 0$, then $T(U, U) = UU^T$, which is certainly positive semidefinite. Further, if $U \neq 0$, then there is a $u \notin \ker U$; we compute $u^T T(U, U)u = u^T UU^T u = ||Uu||^2 \neq 0$. Thus T is a K-bilinear form.

Definition 2.10.4. Let $K \subseteq \mathbb{R}^d$ be a closed convex cone and let T be a K-symmetric bilinear form on $\mathbb{R}^n$. The *Siegel domain* of K and T is

$$SD(K, T) := \{(x, v) \in \mathbb{R}^d \times \mathbb{R}^n : x - T(v, v) \in K\}.$$

The *Siegel cone* of K and T is

$$SC(K, T) := \mathrm{cl}\{(t, x, v) \in \mathbb{R}_{++} \times \mathbb{R}^d \times \mathbb{R}^n : tx - T(v, v) \in K\}.$$

If T is an inner product on $\mathbb{R}^n$ and $K = \mathbb{R}_+$ (i.e. $d = 1$), then $SD(K, T)$ is the paraboloid

$$\{(x, y) \in \mathbb{R} \times \mathbb{R}^n : x \geq ||y||^2\}$$

and $SC(K,T)$ is the set

$$\mathrm{cl}\{(t,x,y) \in \mathbb{R}_{++} \times \mathbb{R}_{+} \times \mathbb{R}^n : tx \geq ||y||^2\}.$$

Example 2.10.5. We continue with Example 2.10.3. We compute

$$SD(K,T) = \{(M,U) \in \Sigma^n \times \mathbb{R}^{k \times n} : M - UU^T \succeq 0\}$$

and

$$SC(K,T) = \mathrm{cl}\{(t,M,U) \in \mathbb{R}_{++} \times \Sigma^n \times \mathbb{R}^{k \times n} : tM - UU^T \succeq 0\}.$$

A striking theorem of E. Vinberg [40] states that if K is homogeneous and T is a K-symmetric bilinear form satisfying some mild conditions, then the Siegel cone $SC(K,T)$ is also a homogeneous cone. An even more striking theorem of S. Gindikin [16] states that *every* homogeneous cone in at least two dimensions arises as the Siegel cone of a homogeneous cone K and a K-symmetric bilinear form satisfying the same mild conditions. Truong and Tunçel [37] use this recursive construction to much effect, proving numerous geometric results about the boundary structure of homogeneous cones without recourse to the semidefinite representation theorem of Chua and Faybusovich. Among them are a characterisation of the extreme rays of a Siegel cone and its dual and a proof that that homogeneous cones are facially exposed.

Chapter 3

Polynomials in one variable

In Section 3.1, we will present a certain useful symmetric function of the roots of a polynomial called the discriminant, and discuss some of the information it uncovers. In Section 3.2, we will show that the discriminant can be computed knowing only the coefficients of a polynomial, and then use the same technique to characterise completely when the roots of a polynomial with real coefficients are all real or all nonnegative.

3.1 The discriminant

Let $p \in \mathbb{R}[x]$, and let n be the degree of p. Let $p_0, \ldots, p_n$ be the coefficients of p, so that

$$p(x) = \sum_{i=0}^{n} p_i x^i.$$

Definition 3.1.1. Let the complex roots of p be $r_1, \ldots, r_n$. Define the *discriminant* of p by

$$\operatorname{disc} p = p_n^{2n-2} \prod_{i=1}^{n} \prod_{j=1}^{i-1} (r_i - r_j)^2.$$

Example 3.1.2. Let $p(x) = ax^2 + bx + c$. Then the roots of p are famously

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

The discriminant of p is then

$$a^2 \left(\frac{-b + \sqrt{b^2 - 4ac}}{2a} - \frac{-b - \sqrt{b^2 - 4ac}}{2a} \right)^2 = a^2 \frac{b^2 - 4ac}{a^2} = b^2 - 4ac.$$

The discriminant tells many things about the roots of a polynomial. In particular,

Proposition 3.1.3. *Let $p \in \mathbb{R}[x]$. Then*

(a). $\operatorname{disc} p = 0$ *if and only if p has a multiple root.*

(b). If p only has real roots, then $\operatorname{disc} p \geq 0$.

(c). $\operatorname{disc} p < 0$ *if and only if p has an odd number of pairs of nonreal roots and no multiple roots.*

Proof :

(a). Trivial.

(b). $\operatorname{disc} p$ is the square of a nonzero real number.

(c). Assume, without loss of generality, that p is monic. Say p has real roots $r_1, r_2, \ldots, r_k$ and nonreal roots $c_1, \bar{c}_1, c_2, \bar{c}_2, \ldots, c_m, \bar{c}_m$. We note that

$$\prod_{i=1}^{k} \prod_{j=1}^{i-1} (r_i - r_j)^2 > 0$$

since $(x - r_1)(x - r_2) \ldots (x - r_n)$ is a polynomial with only real roots and no repeated roots and this product is its discriminant. We also note that

$$\prod_{i=1}^{m} \prod_{j=1}^{k} (c_i - r_j)^2 (\bar{c}_i - r_j)^2 = \prod_{i=1}^{m} \prod_{j=1}^{k} ((c_i - r_j)\overline{(c_i - r_j)})^2$$

$$= \prod_{i=1}^{m} \prod_{j=1}^{k} |c_i - r_j|^4 > 0.$$

Similarly,

$$\prod_{i=1}^{m}\prod_{j=1}^{i-1}(c_i - c_j)^2(c_i - \bar{c}_j)^2(\bar{c}_i - c_j)^2(\bar{c}_i - \bar{c}_j)^2$$
$$= \prod_{i=1}^{m}\prod_{j=1}^{i-1}\Big((c_i - c_j)\overline{(c_i - c_j)}(c_i - \bar{c}_j)\overline{(c_i - \bar{c}_j)}\Big)^2$$
$$= \prod_{i=1}^{m}\prod_{j=1}^{i-1}|c_i - c_j|^4|c_i - \bar{c}_j|^4 > 0.$$

The remaining factors in $\operatorname{disc} p$ are

$$\prod_{i=1}^{m}(c_i - \bar{c}_i)^2$$

Since $c_i - \bar{c}_i$ is purely imaginary for each i, its square is always negative. We take the product of m negative things; the product is negative if and only if m is odd. This establishes the desired result.

□

3.2 Vandermonde and Hankel matrices; Newton sums

The discriminant, in the form given in the previous section, does not appear particularly useful — it appears that one must compute all of the roots of a polynomial before one gets the small amount of summary information the discriminant provides. Vandermonde and Hankel matrices, the theory of which we develop below, allow one to compute the discriminant in terms of the *coefficients* of the polynomial. The minors of the Hankel matrix will also provide some information about the roots of a polynomial to which the discriminant alone is blind.

The following famous result computes the "Vandermonde determinant." It finds particular use in polynomial interpolation; since the Vandermonde matrix is horribly ill-conditioned, one cannot solve a system involving a Vandermonde matrix directly. However, Cramer's rule and this theorem together result in a more numerically stable formula due to Lagrange.

Theorem 3.2.1. *Let* $\mathrm{Vand}(r_1, \ldots, r_n)$ *be the Vandermonde matrix*

$$\begin{pmatrix} 1 & r_1 & \ldots & r_1^{n-1} \\ 1 & r_2 & \ldots & r_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & r_n & \ldots & r_n^{n-1} \end{pmatrix}.$$

Then

$$\det \mathrm{Vand}(r_1, \ldots, r_n) = \prod_{i=1}^{n} \prod_{j=1}^{i-1} (r_i - r_j).$$

PROOF : We note that

$$\begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ -1 & 1 & 0 & \ldots & 0 \\ -1 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & \ldots & 1 \end{pmatrix} \begin{pmatrix} 1 & r_1 & \ldots & r_1^{n-1} \\ 1 & r_2 & \ldots & r_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & r_n & \ldots & r_n^{n-1} \end{pmatrix} = \begin{pmatrix} 1 & r_1 & \ldots & r_1^{n-1} \\ 0 & r_2 - r_1 & \ldots & r_2^{n-1} - r_1^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & r_n - r_1 & \ldots & r_n^{n-1} - r_1^{n-1} \end{pmatrix}.$$

and

$$\begin{pmatrix} r_2 - r_1 & \ldots & r_2^{n-1} - r_1^{n-1} \\ \vdots & \ddots & \vdots \\ r_n - r_1 & \ldots & r_n^{n-1} - r_1^{n-1} \end{pmatrix} \begin{pmatrix} 1 & -2r_1 & 3r_1^2 & -4r_1^3 & \ldots & (-1)^{n-2}\binom{n-2}{n-3} r_1^{n-2} \\ 0 & 1 & -3r_1 & 6r_1^2 & \ldots & (-1)^{n-3}\binom{n-2}{n-4} r_1^{n-3} \\ 0 & 0 & 1 & -4r_1 & \ldots & (-1)^{n-4}\binom{n-2}{n-5} r_1^{n-4} \\ 0 & 0 & 0 & 1 & \ldots & (-1)^{n-5}\binom{n-2}{n-6} r_1^{n-5} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 1 \end{pmatrix}$$

$$= \begin{pmatrix} r_2 - r_1 & (r_2 - r_1)^2 & \ldots & (r_2 - r_1)^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ r_n - r_1 & (r_n - r_1)^2 & \ldots & (r_n - r_1)^{n-1} \end{pmatrix}.$$

The determinant of this last matrix is simply

$$(r_2 - r_1) \dots (r_n - r_1) \det \mathrm{Vand}(r_2 - r_1, \dots, r_n - r_1)$$
$$= (r_2 - r_1) \dots (r_n - r_1) \det \mathrm{Vand}(r_2, \dots, r_n)$$

Induction on n gives the desired result. □

Remark 3.2.2. An easier proof, suggested to me by D. Wagner, proceeds as follows: Note that $\mathbb{Z}[r_1, \dots, r_n]$ is a unique factorisation domain in which $\det \mathrm{Vand}(r_1, \dots, r_n)$ lies. We note that $r_i - r_j$, for $i \neq j$, divides det Vand$(r_1, \dots, r_n)$. Since $\{r_i - r_j : 1 \leq j < i \leq n\}$ is a set of pairwise coprime elements of $\mathbb{Z}[r_1, \dots, r_n]$, it follows that

$$\prod_{i=1}^{n} \prod_{j=1}^{i} (r_i - r_j) \mid \det \mathrm{Vand}(r_1, \dots, r_n).$$

The degrees of the two sides match, so

$$\det \mathrm{Vand}(r_1, \dots, r_n). = \alpha \prod_{i=1}^{n} \prod_{j=1}^{i} (r_i - r_j)$$

for some $\alpha \in \mathbb{Z}$. The coefficient of $r_1^0 r_2^1 \dots r_n^{n-1}$ on the left-hand side is one and on the right-hand side is α, so $\alpha = 1$ and the theorem is proved.

Corollary 3.2.3. *Let $r_1, \dots, r_n$ be the roots of a monic polynomial p. Then*

$$\operatorname{disc} p = (\det \mathrm{Vand}(r_1, \dots, r_n))^2 = \det(\mathrm{Vand}^T(r_1, \dots, r_n)\,\mathrm{Vand}(r_1, \dots, r_n)).$$

We now develop some tools to make this formula more philosophically satisfying. Namely, we shall show that the right-hand matrix contains entries that can be computed from the coefficients of p.

Definition 3.2.4. The kth $n \times n$ *elementary Hankel matrix* is the matrix Hank_k with entries

$$(\mathrm{Hank}_k)_{ij} = \begin{cases} 1 & \text{if } i + j = k \\ 0 & \text{otherwise.} \end{cases}$$

(There is a slight awkwardness hidden in our choice of notation — the first nonzero $n \times n$ elementary Hankel matrix is Hank_2 and the last is Hank_{2n}.)

The kth *Newton sum* of a polynomial with roots at $r_1, \ldots, r_n$ is

$$s_k = \sum_{i=1}^{n} r_i^k.$$

We overload notation slightly and let $\mathrm{Hank}(p)$ be the $\deg p \times \deg p$ matrix

$$\mathrm{Hank}(p) = \sum_{i=2}^{2 \deg p} s_{i-2} \, \mathrm{Hank}_i \, .$$

Let $r_1, \ldots, r_n$ be the roots of some polynomial p. We note that the (i, j) entry of $\mathrm{Vand}^T(r_1, \ldots, r_n) \, \mathrm{Vand}(r_1, \ldots, r_n)$ is precisely the $(i + j - 2)$th Newton sum. That is,

$$\left(\mathrm{Vand}^T(r_1, \ldots, r_n) \, \mathrm{Vand}(r_1, \ldots, r_n)\right) = \sum_{i=2}^{2n} s_{i-2} \, \mathrm{Hank}_i = \mathrm{Hank}(p).$$

The Newton sums of a polynomial can be computed in terms of the coefficients of the polynomial, as we show next:

Proposition 3.2.5. *Let p be a monic polynomial of degree d. Let the roots of p be $r_1, \ldots, r_d$, and let p_i be the coefficient of x^i in p. Then the p_i and s_i are related by the following system of equations, called the Newton identities:*

$$\begin{aligned}
p_d &= 1 \\
p_{d-1} &= -s_1 p_d \\
p_{d-2} &= -\frac{1}{2}\left(s_1 p_{d-1} + s_2 p_d\right) \\
p_{d-3} &= -\frac{1}{3}\left(s_1 p_{d-2} + s_2 p_{d-1} + s_3 p_d\right) \\
&\vdots \\
p_{d-k} &= -\frac{1}{k}\left(\sum_{i=1}^{k} s_i p_{d-k+i}\right).
\end{aligned}$$

Proof : The first equation is clear since p is monic; we shall ignore it. A direct combinatorial proof of the remaining identities exists, but we prefer to present instead an elegant argument given by S. Basu, R. Pollack, and M.-F. Roy [3].

Let R be a real number larger than the magnitude of any of the roots of p. We shall perform calculations in the region $\{x \in \mathbb{C} : |x| > R\}$.

Note that $p(x) = \prod_{i=1}^{d}(x - r_i)$. Doing logarithmic differentiation, we see that $p'(x)/p(x) = \sum_{i=1}^{d}(x - r_i)^{-1}$. We observe $(x - r_i)^{-1} = x^{-1}\sum_{j=0}^{\infty}(\frac{r_i}{x})^j$. Thus,

$$\frac{p'(x)}{p(x)} = \sum_{i=1}^{d}\sum_{j=0}^{\infty}\frac{r_i^j}{x^{j+1}} = \sum_{j=0}^{\infty}\sum_{i=1}^{d}\frac{r_i^j}{x^{j+1}} = \sum_{j=0}^{\infty}\frac{\sum_{i=1}^{d} r_i^j}{x^{j+1}} = \sum_{j=0}^{\infty}\frac{s_j}{x^{j+1}}.$$

Multiplying both sides by $p(x)$, we see that

$$\begin{aligned} p'(x) &= \left(\sum_{j=0}^{\infty}\frac{s_j}{x^{j+1}}\right)\left(\sum_{i=0}^{d} p_i x^i\right) \\ &= \sum_{j=0}^{\infty}\sum_{i=0}^{d}\frac{p_i s_j}{x^{j-i+1}} = \sum_{k=-\infty}^{d}\left(\sum_{i=\max(0,k)}^{d} p_i s_{i-k}\right)x^{k-1}. \end{aligned}$$

We shall next prove that the negative-degree terms sum to zero. The promised Newton identities will then follow by extracting coefficients in the formula above.

We take the dth derivative of each side; since $p'(x)$ is a polynomial of degree $d-1$, the left-hand side is zero. All of the positive-degree terms on the right-hand side also vanish, since each one is of degree at most $d-1$, leaving only the negative-degree terms. Thus, we have an equation of the form

$$0 = \sum_{i=1}^{\infty} a_i x^{-i}$$

that holds for any x outside a sufficiently large disc. Letting $f(x) = \sum_{i=1}^{\infty} a_i x^{-i}$, we see that the function $f(1/x) = \sum_{i=1}^{\infty} a_i x^i$ is analytic on some disc containing the origin and is equal to zero on that disc. It follows that f is zero; the Newton identities therefore follow. $\square$

Corollary 3.2.6. *Given the coefficients of a polynomial, one can exactly and efficiently compute its Newton sums. Likewise, given the Newton sums of a monic polynomial, one can exactly and efficiently compute its coefficients.*

Proposition 3.2.7. $\operatorname{Hank}(p)$ *is a real symmetric matrix.*

Theorem 3.2.8. *Let $p \in \mathbb{R}[x]$. Then p only has real roots if and only if* $\operatorname{Hank}(p)$ *is positive semidefinite.*

PROOF : Let $r_1, \ldots, r_k$ be the roots of p. If p only has real roots, then

$$\operatorname{Vand}^T(r_1, \ldots, r_k) = \operatorname{Vand}^*(r_1, \ldots, r_k).$$

It quickly follows that

$$\operatorname{Hank}(p) = \operatorname{Vand}^T(r_1, \ldots, r_k) \operatorname{Vand}(r_1, \ldots, r_k)$$

is positive semidefinite.

Now suppose that p has some conjugate pair of complex roots; without loss of generality, say r_1 and r_2. Also assume that all of the roots of p are distinct. Let A and B be real matrices such that

$$\operatorname{Vand}(r_1, \ldots, r_k) = A + Bi.$$

We observe that

$$\begin{aligned}\operatorname{Hank}(p) &= \operatorname{Vand}^T(r_1, \ldots, r_k) \operatorname{Vand}(r_1, \ldots, r_k) \\ &= (A^T + B^T i)(A + Bi) = A^T A + i(A^T B + B^T A) - B^T B.\end{aligned}$$

$\operatorname{Hank}(p)$ is a real symmetric matrix, so its imaginary part is zero; therefore, $\operatorname{Hank}(p) = A^T A - B^T B$. We note that A has kernel since its first two rows are equal; let x be a real vector in $\ker A$. Since $\operatorname{Vand}(r_1, \ldots, r_k)$ is nonsingular, we see that $Bx \neq 0$. Thus, $x^T \operatorname{Hank}(p) x = -x^T B^T B x < 0$. Hence $\operatorname{Hank}(p)$ is not positive semidefinite if p has distinct roots and at least one pair of nonreal roots.

Now suppose that p has some nonreal roots and some repeated roots. Let $q = p/\gcd(p, p')$. (Then q has roots exactly where p has roots and q only has simple roots.) By the above argument, $\operatorname{Hank}(q)$ is not positive semidefinite. Let $s_1, \ldots, s_l$ be the roots of q, and suppose the roots of p are ordered so that $r_i = s_i$ for all i with $1 \le i \le l$. Let E be the $l \times k$ matrix whose (i, j) entry is one if $i = j$ and zero otherwise. Then

$$\mathrm{Vand}(s_1, \ldots, s_l) = E\,\mathrm{Vand}(r_1, \ldots, r_k)E^T.$$

Thus

$$\begin{aligned}\mathrm{Hank}(p) &= E\,\mathrm{Vand}(s_1, \ldots, s_l)E^T E\,\mathrm{Vand}^T(s_1, \ldots, s_l)E^T \\ &= E\,\mathrm{Vand}(s_1, \ldots, s_l)\,\mathrm{Vand}^T(s_1, \ldots, s_l)E^T = E\,\mathrm{Hank}(q)E^T.\end{aligned}$$

Since E^T is a surjection and $\mathrm{Hank}(q)$ is not positive semidefinite, it follows that $\mathrm{Hank}(p)$ is not positive semidefinite. This proves the desired result. □

A related result is

Theorem 3.2.9. *Let p be a real polynomial with only real roots. Then all of p's roots are nonpositive if and only if all of p's coefficients are nonnegative.*

Proof : The coefficients of $\prod_{i=1}^n (x + r_i)$ are simply the elementary symmetric functions of the r_i. The elementary symmetric functions map nonnegative numbers to nonnegative numbers; it follows that if all of p's roots are nonpositive then all of p's coefficients are nonnegative.

Now suppose that p is a real polynomial that has nonnegative coefficients. Then, if $x > 0$, $\sum_{i=0}^n p_i x^i > p_n x^n > 0$. □

Corollary 3.2.10. *If $p \in \mathbb{R}[x]$, then p has only nonnegative real roots if and only if $\mathrm{Hank}(p(x^2))$ is positive semidefinite.*

3.3 Interlacing and the Gauss-Lucas theorem

Theorem 3.3.1 (Interlacing)**.** *If $p(x)$ is a polynomial with roots $r_1, \ldots, r_n$, all real, then $p'(x)$ only has real roots. Further, if $q_1, \ldots, q_{n-1}$ are the roots of $p'(x)$, these roots may be ordered so that*

$$r_1 \le q_1 \le r_2 \le \ldots \le r_{n-1} \le q_{n-1} \le r_n.$$

PROOF : We observe that $p(x)$ is a scalar multiple of $(x-r_1)(x-r_2)\dots(x-r_n)$. We assume, without loss of generality, that p in fact equals $(x-r_1)\dots(x-r_n)$. We compute $p'(x)=\sum_{i=1}^n(x-r_1)(x-r_2)\dots\widehat{(x-r_i)}\dots(x-r_n)$. Suppose first that all of the roots of p are distinct. We observe that

$$p'(r_i)=(x-r_1)\dots\widehat{(x-r_i)}\dots(x-r_n).$$

This is positive if $n-i$ is even and negative if $n-i$ is odd. It follows from the intermediate value theorem that p' has a root in (r_i,r_{i+1}) for each i with $1\le i\le n-1$. These are all of the roots of p' and the theorem is proved when the roots of p are all distinct.

Now suppose that p has repeated roots. p is then the limit a sequence p_n of polynomials with only simple roots. Since differentiation is continuous and the roots of p_n' interlace the roots of p_n for each n, it follows from continuity of roots that the roots of p' interlace the roots of p. $\square$

Theorem 3.3.2 (Gauss-Lucas). *If $p(x)$ is a complex polynomial whose roots are $r_1,\dots,r_n$, then the roots of $p'(x)$ are contained in* $\mathrm{conv}(r_1,\dots,r_n)$.

PROOF : We observe that, for x not a root of p,

$$p'(x)=\left(\sum_{i=1}^n\frac{1}{x-r_i}\right)p(x).$$

Thus, if $p'(x)=0$, then either x is a root of p or

$$0=\sum_{i=1}^n\frac{1}{x-r_i}=\sum_{i=1}^n\frac{\bar{x}-\bar{r}_i}{|x-r_i|^2}=\sum_{i=1}^n\frac{1}{|x-r_i|^2}(\bar{x}-\bar{r}_i).$$

Rearranging the last sum, we find that

$$\left(\sum_{i=1}^n\frac{1}{|x-r_i|^2}\right)\bar{x}=\sum_{i=1}^n\frac{1}{|x-r_i|^2}\bar{r}_i.$$

For $i\in\{1,\dots,n\}$, let

$$\lambda_i=\frac{1/|x-r_i|^2}{\sum_{i=1}^n 1/|x-r_i|^2}.$$

Then $\lambda_i \geq 0$ for each i and $\sum_{i=1}^n \lambda_i = 1$. Further, we have

$$\bar{x} = \sum_{i=1}^{n} \lambda_i \bar{r}_i.$$

Thus $\bar{x}$ is a convex combination of $\bar{r}_1, \ldots, \bar{r}_n$; the Gauss-Lucas theorem follows after complex conjugation. $\square$

Definition 3.3.3. We overload some terminology and call univariate real polynomials that only have real roots *univariate hyperbolic polynomials*, (Note that we have dropped the condition that the polynomial be homogeneous.)

Let p and q be univariate hyperbolic polynomials. We say that p and q are *in proper position* and write $p \ll q$ if $(p'q - q'p)(x) \leq 0$ for all real x.

Two polynomials are in proper position iff their zeroes interlace:

Proposition 3.3.4. *Let p and q be in $\mathbb{R}[x]$. Let the roots of p be $r_1, \ldots, r_a$, and let the roots of q be $s_1, \ldots, s_b$. If $p \ll q$, then either:*

(a). $a = b$ and $r_1 \leq s_1 \leq r_2 \leq s_2 \leq \ldots \leq r_a \leq s_b$, or

(b). $a = b + 1$ and $r_1 \leq s_1 \leq r_2 \leq s_2 \leq \ldots \leq s_b \leq r_a$, or

(c). $a = b - 1$ and $s_1 \leq r_1 \leq s_2 \leq r_2 \leq \ldots \leq r_a \leq s_b$, or

(d). $a = b$ and $s_1 \leq r_1 \leq s_2 \leq r_2 \leq \ldots \leq s_b \leq r_a$.

Proof : $p'q - q'p = q^2(p/q)'$. If p/q has two consecutive poles that are not separated by a root, its derivative clearly must change sign between the two poles. Likewise, if p/q has two consecutive roots that are not separated by a pole, its derivative must change sign between the two roots. Thus the number of roots of p and the number of roots of q must differ by at most one and satisfy one of the four strings of inequalities listed. $\square$

The following fundamental result is attributed to Obreschkoff [30]. Though our argument is long, this theorem is not especially hard to prove.

Theorem 3.3.5 (Obreschkoff's theorem)**.** *Letp and q are univariate hyperbolic polynomials. The polynomial $\alpha p + \beta q$ is univariate hyperbolic for all real α and β if and only if $p \ll q$ or $q \ll p$.*

Proof : The following argument is overcomplicated. There must be a better proof, but I do not know it.

We first prove that if $p \ll q$ or $q \ll p$ then $\alpha p + \beta q$ is univariate hyperbolic for all real α and β. Fix α and β.

We examine only the case where $\gcd(p, q) = 1$. This, together with interlacing, excludes p and q from having repeated roots. If $\gcd(p, q) \neq 1$ and $p \ll q$, then $p/\gcd(p, q) \ll q/\gcd(p, q)$; further, $\alpha p + \beta q = \gcd(p, q)(\alpha p/\gcd(p, q) + \beta q/\gcd(p, q))$ is univariate hyperbolic if and only if $\alpha p/\gcd(p, q) + \beta q/\gcd(p, q)$ is univariate. Thus there is no loss in generality caused by assuming $\gcd(p, q) = 1$.

Let the roots of p be $r_1 < r_2 < \ldots < r_n$ and the roots of q be $s_1 < s_2 < \ldots < s_m$. By the preceding theorem, we may order the roots of p and q so that (without loss of generality) $r_1 \leq s_i \leq r_{i+1}$ for each i. We also assume without loss of generality that α and β are positive and that p is negative to the left of r_1. There are still eight cases, however: q can be either positive or negative left of s_1, $\deg p$ could be either even or odd, and $\deg q$ could either equal $\deg p$ or be one smaller. We shall tackle exactly one of these cases and refer to our excellent Figure 3.1 for the remaining cases.

Assume that $\deg p = \deg q = n$ is even and that q is positive left of s_1. Then p is negative on $(-\infty, r_1)$, (r_2, r_3), …, (r_n, ∞), positive on (r_1, r_2), …, (r_{n-1}, r_n). q is negative on (s_1, s_2), (s_3, s_4), …, (s_{n-1}, s_n) and positive on $(-\infty, s_1)$, (s_2, s_3), …, (s_n, ∞). It follows that $\alpha p + \beta q$ is positive on (r_1, s_1), (r_3, s_3), …, (r_{n-1}, s_{n-1}) and negative on (r_2, s_2), (r_4, s_4), …, (r_n, s_n). Thus, by the intermediate value theorem, $\alpha p + \beta q$ has roots in (s_1, r_2), (s_3, r_4), …, (s_{n-1}, r_n). As such, $\alpha p + \beta q$ has at least $n - 1$ real roots. But nonreal roots must come in pairs; $\alpha p + \beta q$ has all n roots real. The theorem is therefore proven in this case, and we choose to omit the proofs for the seven other cases.

Suppose $\gcd(p, q) \neq 1$. Let $\hat{p} := p/\gcd(p, q)$ and $\hat{q} := q/\gcd(p, q)$. The above argument shows that $\alpha\hat{p} + \beta\hat{q}$ is univariate hyperbolic for any real α and β. Since $\gcd(p, q)|p$, $\gcd(p, q)$ is also univariate hyperbolic, it follows

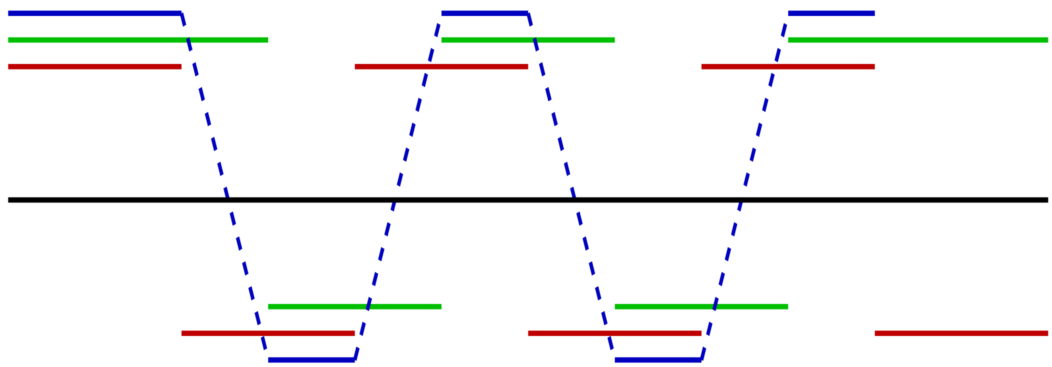

Figure 3.1: A proof that $p+q$ is univariate hyperbolic in the case where p and q both have positive leading coefficient, $\deg p = 6$, $\deg q = 5$, and the zeroes of p and q interlace. The red line shows the sign of p, while the green line shows the sign of q. The solid blue line indicates the sign of $p+q$ where it can be deduced from the signs of p and q. Continuity of $p+q$ implies that a root of $p+q$ must lie at an abscissa in each dashed blue line segment; there are at least $\deg(p+q)-1$ such segments.

that $(\gcd(p,q))(\alpha\hat{p}+\beta\hat{q}) = \alpha p+\beta q$ is univariate hyperbolic. This proves the backward direction of Obreschkoff's theorem for arbitrary choices of p and q.

We now prove that if $\alpha p + \beta q$ is univariate hyperbolic for all choices of real α and β then either $p \ll q$ or $q \ll p$. We shall assume that $\gcd(p,q) = 1$; no generality is lost. Let p's roots be $r_1, \ldots, r_n$. Suppose, for the sake of contradiction, that the roots of p and q do not interlace. Then there is a pair of roots of one polynomial, say r_k and r_{k+1} of p, such that q has the same sign on $[r_k, r_{k+1}]$. Assume, for now, that $r_k < r_{k+1}$. Without loss of generality, suppose q is positive on $[r_k, r_{k+1}]$ and p is positive on (r_k, r_{k+1}).

Let $a_i(t)$ be the ith smallest root of $P_t := (1-t)p - tq$. We recall that a_i is a continuous function of t. We note that $a_k(0) = r_k$ and $a_{k+1}(0) = r_{k+1}$. We also note that, if $0 < t \le 1$, then $P_t(r_k) < 0$ and $P_t(r_{k+1}) < 0$. Fix $c \in (r_k, r_{k+1})$. Then there is an $0 < \epsilon < 1$ for which $P_\epsilon(c) > 0$. But $P_\epsilon(r_k) < 0$ and $P_\epsilon(r_{k+1}) < 0$. Thus, P_ϵ has at least two roots in (r_k, r_{k+1}). Since neither r_k nor r_{k+1} is a root of P_t when $0 < t \le 1$, P_t must have at least two roots in (r_k, r_{k+1}) for all $t \in (0,1]$. Sadly, $P_1 = q$ has no roots in (r_k, r_{k+1}), a contradiction from which we conclude that either p or q has a multiple root or the roots of p and q interlace. $\square$

3.4 Convexity results

The following result on univariate polynomials will quickly imply Gårding's inequality for hyperbolic polynomials.

Theorem 3.4.1 ([15]). *Let p be a polynomial of degree n with only real roots. If p is positive on the interval (a, b), then $p^{1/n}$ is concave on (a, b).*

Proof : Let t be a point at which p is positive.

Let c and $r_1, \ldots, r_n$ be such that

$$p(x) = c \prod_{i=1}^{n} (x - r_i).$$

Then, if x is not a root of p,

$$p'(x) = c \left(\sum_{i=1}^{n} \frac{1}{x - r_i} \right) p(x)$$

and

$$p''(x) = c \left(- \left(\sum_{i=1}^{n} \frac{1}{(x - r_i)^2} \right) p(x) + \left(\sum_{i=1}^{n} \frac{1}{x - r_i} \right)^2 p(x) \right).$$

We note that

$$\frac{d}{dt} p(t)^{1/n} = \frac{1}{n} p'(t) p(t)^{1/n - 1}$$

and

$$\begin{aligned} \frac{d^2}{dt^2} p(t)^{1/n} = \frac{1}{n} \left(p''(t) p(t)^{1/n-1} + (1/n - 1) p(t)^{1/n-2} (p'(t))^2 \right) \\ = \frac{p(t)^{1/n-2}}{n^2} \left(n p''(t) p(t) + (1 - n)(p'(t))^2 \right) \end{aligned}$$

We see that $\frac{p(t)^{1/n}}{n^2}$ is positive whenever $t \neq 0$. We now apply the above calculations of p' and p'' and find

$$\frac{1}{c}n^2 p(t)^{-1/n}\frac{d^2}{dt^2}p(t)^{1/n} = n\left(\sum_{i=1}^{n}\frac{1}{t-r_i}\right)^2 - n\left(\sum_{i=1}^{n}\frac{1}{(t-r_i)^2}\right)$$
$$+ (1-n)\left(\sum_{i=1}^{n}\frac{1}{t-r_i}\right)^2 = \left(\sum_{i=1}^{n}\frac{1}{t-r_i}\right)^2 - n\left(\sum_{i=1}^{n}\frac{1}{(t-r_i)^2}\right)$$

Now let $v \in \mathbb{R}^n$ be the vector whose ith entry is $1/(t-r_i)$. We note that, by the Cauchy-Schwarz inequality,

$$\left(\sum_{i=1}^{n}\frac{1}{t-r_i}\right)^2 = \langle \bar{e}, v\rangle^2 \leq ||\bar{e}||^2||v||^2 = n\sum_{i=1}^{n}\frac{1}{(t-r_i)^2}.$$

This establishes that

$$\frac{d^2}{dt^2}p(t)^{1/n} \leq 0.$$

A theorem of elementary calculus reminds us that a function whose second derivative is nonpositive everywhere on an interval must be concave, proving the desired result. □

Chapter 4

Boundary structure of hyperbolic cones

Let $p \in \mathbb{R}[x_1, \ldots, x_d]$. We begin in Section 1 by recalling the meaning of the phrase "p is hyperbolic in direction e" and proving some fundamental facts about the cones associated with hyperbolic polynomials. In Section 2, we study a selection of techniques for constructing hyperbolic polynomials from other hyperbolic polynomials and prove the main theorem of this thesis, namely that the intersection of all derivative cones of a nontrivial hyperbolic cone yield the cone itself. We also discuss some recent work by J. Borcea, P. Brändén, and B. Shapiro toward classifying all linear maps on the space of real polynomials in d variables that preserve a related notion called *real stability*. In Section 3, we prove another fundamental result of J. Renegar, namely that all hyperbolic cones are facially exposed. In Section 4, we collect a few basic properties of *strictly hyperbolic polynomials*, and we conclude with a brief discussion of the set of extreme rays of hyperbolic cones in Section 5.

4.1 Definitions and background

We follow quite closely Section 2 of Renegar's paper [34].

Definition 4.1.1. p is said to be *hyperbolic* in direction $e \in \mathbb{R}^d$ if $p(e) > 0$

and, for all $x \in \mathbb{R}^d$, the univariate polynomial $\lambda \mapsto p(x + \lambda e)$ only has real roots. e is called a *direction of hyperbolicity* of p.

Another definition, more popular when nonhomogeneous polynomials are to be considered, is:

Proposition 4.1.2. *Let p be a homogeneous polynomial in d variables. p is hyperbolic in direction $e \in \mathbb{R}^d$ if and only if there exists a $\tau_0 \in \mathbb{R}$ such that, whenever $\tau < \tau_0$, $p(x + \tau ie)$ is never zero.*

PROOF : Suppose p is hyperbolic in direction e. Take $\tau_0 = 0$. Fix $x \in \mathbb{R}^d$. The polynomial $\lambda \mapsto p(x + \lambda e)$ only has real roots — in particular, it has no roots with negative imaginary part. This proves the " $\Longrightarrow$ " direction.

Now let τ_0 be such that, whenever $\tau < \tau_0$ and $x \in \mathbb{R}^d$, $p(x + \tau ie)$ is never zero. Let $q_x(\lambda) := p(x + \lambda e)$. We observe that if λ is a root of q_x, then, by homogeneity, $\mu\lambda$ is a root of $q_{\mu x}$. It follows that $p(x + \tau ie)$ is never zero whenever $\tau < 0$. Since p is a polynomial with real coefficients, $p(x + \tau ie)$ is also never zero whenever $\tau > 0$. It follows that p only has real roots. $\square$

Hereafter, assume that $p \in \mathbb{R}[x_1, \ldots, x_d]$ is hyperbolic in direction $e \in \mathbb{R}^d$.

Definition 4.1.3. Let $\Lambda_{++}(p, e)$ be the connected component of e in the set $\{x \in \mathbb{R}^d : p(x) \neq 0\}$. The *hyperbolicity cone* of p in direction e is $\Lambda_+(p, e) = \operatorname{cl} \Lambda_{++}(p, e)$.

Let K be a convex cone. Then K is said to be *hyperbolic* if K is the hyperbolicity cone of some hyperbolic polynomial in a direction of hyperbolicity.

We motivate our study of hyperbolic polynomials largely by considering the following three examples.

Example 4.1.4.

- *The nonnegative orthant*
  Let $p(x_1, \ldots, x_n) = x_1 x_2 \ldots x_n$. Then p is hyperbolic in direction $\bar{e}$ and its hyperbolicity cone is the nonnegative orthant.
- *The second-order cone*
  Let $p(x_1, \ldots, x_n) = x_2^2 + \ldots + x_n^2 - x_1^2$. Then p is hyperbolic in the direction e_1, and the hyperbolicity cone of p in direction e_1 is the cone
  $$\{(x_1, \ldots, x_n)^T : x_2^2 + \ldots + x_n^2 \geq x_1^2,\ x_1 \geq 0\}.$$

- *The positive semidefinite cone*

  Identify the $n \times n$ symmetric real matrices with $\mathbb{R}^{\binom{n}{2}}$ in the usual way. The polynomial $p = \det$ is hyperbolic in the direction $e = I$, and its hyperbolicity cone is the cone of $n \times n$ real symmetric positive semidefinite matrices.

Definition 4.1.5. Motivated by the example of the semidefinite cone, we define the *eigenvalues* of a point $x \in \mathbb{R}^d$ to be $\{\lambda \in \mathbb{R} : p(x + \lambda e) = 0\}$.

Theorem 4.1.6 ([15])**.** *The cone Λ_{++} is precisely the set of points whose eigenvalues are all positive, and Λ_+ is the set of points whose eigenvalues are all nonnegative.*

Proof : Let $x \in \Lambda_{++}$. Since the roots of a polynomial vary continuously with its coefficients, we see that λ_{min} is continuous on Λ_{++}. Since λ_{min} is never zero on Λ_{++} and Λ_{++} is connected, we conclude that $\lambda_{min}(x) > 0$. Thus, every point of Λ_{++} has only positive eigenvalues.

Now let x be a point with only positive eigenvalues. Then $x + \lambda e$ is also a point with only positive eigenvalues for any $\lambda \geq 0$. By homogeneity, then, so is $x/\lambda + e$. The arc from x to $x + e$ given by $\gamma_1(t) = x + 2te$ for $t \in [0, \frac{1}{2}]$ and the arc from $x + e$ to e given by $\gamma_2(t) = 2(1-t)x + e$ for $t \in [\frac{1}{2}, 1]$ can be concatenated to form an arc from x to e on which all points have only positive eigenvalues. Thus, x is in the same connected component of $\{y : p(y) \neq 0\}$ as e.

Let x have nonnegative eigenvalues. Then, for any $\epsilon > 0$, $x + \epsilon e$ has positive eigenvalues and therefore belongs to Λ_{++}. This proves that all points with only nonnegative eigenvalues are contained in Λ_+. Now suppose $x \in \Lambda_+$ has a negative eigenvalue. By continuity of roots, there is an $\epsilon > 0$ such that every point within ϵ of x has a negative eigenvalue. Then no point within ϵ of x is in Λ_{++}; it follows that $x \notin \operatorname{cl} \Lambda_{++} = \Lambda_+$. $\square$

Corollary 4.1.7.

$$\Lambda_{++}(p, e) = \{x \in \mathbb{R}^d : \operatorname{Hank} p(x - \lambda^2 e) \succ 0\}$$

and

$$\Lambda_+(p, e) = \{x \in \mathbb{R}^d : \operatorname{Hank} p(x - \lambda^2 e) \succeq 0\}.$$

Corollary 4.1.8. *$\Lambda_{++}(p, e)$ is the set of points $x \in \mathbb{R}^d$ such that $p(x + \lambda e)$ has nonnegative coefficients.*

Theorem 4.1.9 ([15])**.** *If $x \in \Lambda_{++}$, then p is hyperbolic in direction x.*

Proof [34]: We need only show that, for $y \in \mathbb{R}^d$, the polynomial $\lambda \mapsto p(y + \lambda x)$ only has real roots.

Fix $\alpha > 0$. We shall show that all roots of $\lambda \mapsto p(\alpha i e + \lambda x + s y)$ have negative imaginary part for every $s \in [0, \infty)$. From this and continuity of roots, it quickly follows that $\lambda \mapsto p(\lambda x + s y)$ only has real roots.

When $s = 0$, this follows by hyperbolicity of p; all roots of $\lambda \mapsto p(\alpha i e + \lambda x)$ are merely αi times a root of $\lambda \mapsto p(e + \lambda x)$ — and all roots of this polynomial are negative.

Now suppose that there is some positive s for which $\lambda \mapsto p(\alpha i e + \lambda x + s y)$ has a root of nonnegative imaginary part. By continuity of roots, there is a least such s; at this s, $\lambda \mapsto p(\alpha i e + \lambda x + s y)$ has a real root; let μ be this root. We observe that αi is a root of $\beta \mapsto p(\beta e + \mu x + s y)$. Sadly, $\mu x + s y$ is a real vector; hyperbolicity in direction e permits this polynomial no nonreal roots.

It follows that there is no positive s for which $\lambda \mapsto p(\alpha i e + \lambda x + s y)$ has a root of nonnegative imaginary part. This is true for any $\alpha > 0$; we let α tend to zero and apply continuity of roots again to find that $\lambda \mapsto p(\lambda x + s y)$ has no roots of positive imaginary part. But p has real coefficients, so p also has no roots of negative imaginary part. Since y was arbitrary, it follows that p is hyperbolic. $\square$

Corollary 4.1.10. *If $f \in \Lambda_{++}(p, e)$, then, for any $x \in \mathbb{R}^d$, the polynomial $\lambda \mapsto p(f + \lambda x)$ only has real roots.*

We first prove an inequality due to L. Gårding. It will follow from this inequality that all hyperbolic cones are convex.

Theorem 4.1.11 (Gårding's inequality [15])**.** *If p is hyperbolic in direction e, then $p^{1/n}$ is concave on $\Lambda_{++}(p, e)$.*

Proof : Let $x_0 \in \Lambda_{++}(p,e)$. Let ϵ be such that $B(x_0,\epsilon) \subseteq \Lambda_{++}(p,e)$. Let $v \in \mathbb{R}^d$ have norm one. Let $g(t) = p(x_0 + tv)$. Then, by 3.4.1, g is concave on $(-\epsilon, \epsilon)$. Since this holds for any $v \in \mathbb{R}^d$ of norm one, we see that $p^{1/n}$ is concave on $B(x_0, \epsilon)$. For any $x_0 \in \Lambda_{++}(p,e)$, one can find an $\epsilon > 0$ for which $p^{1/n}$ is concave on $B(x_0,\epsilon)$. Thus $p^{1/n}$ is concave on $\Lambda_{++}(p,e)$. $\square$

Theorem 4.1.12 ([15])**.** *Hyperbolic cones are convex.*

Proof : Omitted; this follows easily from 4.1.11. $\square$

Remark 4.1.13. The dual cone of a hyperbolic cone is not, in general, a hyperbolic cone. Let

$$K = \{(x,y,z)^T \in \mathbb{R}^3 : x^2 + y^2 \leq z^2, x^2 - 2yz + z^2 \geq 0\}.$$

This is a hyperbolic cone — it is the hyperbolicity cone of $p(x,y,z) = (z^2 - x^2 - y^2)(z^2 - 2yz + x^2)$ in the direction $(0,0,1)^T$. Its dual, unfortunately, is the cone

$$K^* = \{(x,y,z)^T \in \mathbb{R}^3 : (x/z, y/z) \in B(0,1) + [(0,0)^T, (1,0)^T], z \geq 0\}.$$

We shall prove later that every hyperbolic cone is facially exposed. Since K^* is not facially exposed, K^* cannot be a hyperbolic cone.

Pictures of the $z = 1$ slices of these cones are given in Figures 4.1 and 4.2.

A major open problem is

Problem 4.1.14 (The generalised Lax conjecture)**.** If p is hyperbolic in direction e, does there exist a dimension D and a subspace S of Σ^D so that $\Lambda_+(p,e)$ is linearly isomorphic to $S \cap \Sigma_+^D$?

Remark 4.1.15. When $d = 2$, one can factorise the zero set of a hyperbolic polynomial as a product of linear functions and quickly find a semidefinite representation; the generalised Lax conjecture trivially holds in two dimensions. The case $d = 3$ was recently proven by Lewis, Parrilo, and Ramana [26] using results of Helton and Vinnikov [21] and Vinnikov [41]. In higher dimensions, however, the conjecture is still open.

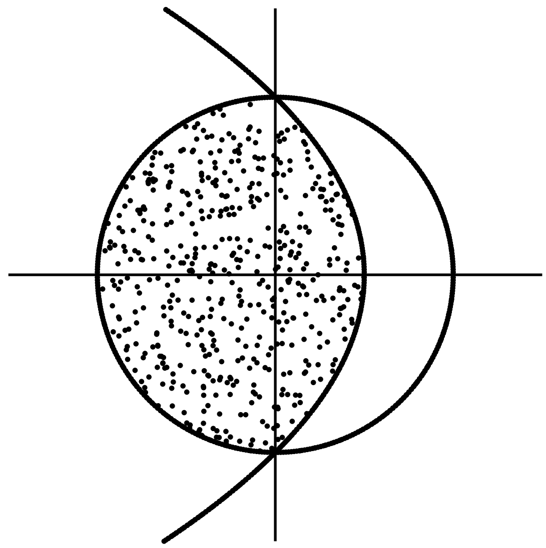

Figure 4.1: The set $\{(x, y)^T \in \mathbb{R}^2 : x^2 + y^2 \leq 1, x^2 - 2y + 1 \geq 0\}$.

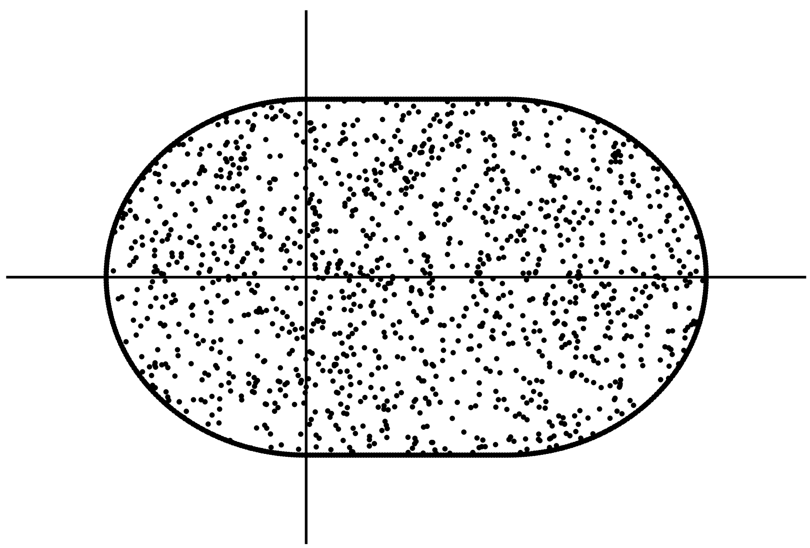

Figure 4.2: The set $B(0, 1) + [(0, 0), (1, 0)]$.

The preceding two corollaries appear to suggest a purely algebraic proof strategy for the generalised Lax conjecture — given the coefficients of a hyperbolic polynomial, find a linear map from $\mathbb{R}^d$ to Σ^d mapping exactly those points in Λ_+ to the positive semidefinite cone. Sadly, difficulties quickly arise.

If one wishes to apply 4.1.7 directly, one needs to find a linear map from $\mathbb{R}^d$ to Σ^d mapping a vector x to a matrix similar to $\operatorname{Hank} p(x - \lambda^2 e)$. Since $\operatorname{Hank} p(x - \lambda^2 e)$ has exactly one entry of maximum total degree and hence any similar matrix will have at least one entry of that degree, this approach is doomed.

A different attack, aimed instead at proving that every hyperbolic cone is a linear image of a slice of a semidefinite cone, is as follows: Consider the polynomial p as a univariate polynomial in the direction e whose coefficients are polynomials on $\mathbb{R}^d$. First, find a way to represent the coefficients of p using linear matrix inequalities that will move no roots across the imaginary axis. Then, relate the coefficients of p, as a function of x, to the Newton sums, again as a function of x, in a semidefinite-representable way. Then form a Hankel matrix from the Newton sums and constrain it to be positive semidefinite.

Two difficult questions arise. First, how does one represent the coefficients of p using linear matrix inequalities without scattering p's roots all over the complex plane? Recent work of J. Borcea, P. Brändén, and B. Shapiro [5] has characterised all differential operators with polynomial coefficients that leave the roots of all real-rooted polynomials on the real line.

Second, the equations relating the Newton sums to the coefficients of a polynomial are unpleasantly nonlinear. How can one write linear matrix inequalities in the coefficients that produce values for Newton sums that give rise to a positive semidefinite Hankel matrix exactly when the true Newton sums do?

The first question seems more fundamental than the second. Using 3.2.9, we see that a real-rooted real polynomial only has nonpositive real roots iff its coefficients are all positive. We shall therefore assume that $e = e_1$ and extract coefficients:

$$[\lambda^k]p(x+\lambda e) = [\lambda^k]\sum_{i=0}^{n} p_i(x_2,\ldots,x_n)(x_1+\lambda)^i = \sum_{i=0}^{n} p_i(x_2,\ldots,x_n)\binom{i}{k}x_1^{i-k}.$$

Thus, it is sufficient to find that some matrix similar or conjugate to

$$\mathrm{Diag}(\sum_{i=0}^{n} p_i(x_2,\ldots,x_n)\binom{i}{k}x_1^{i-k} : 0 \le k \le n)$$

is a linear image of $x_1,\ldots,x_n$. This, too, is hopeless for reasons involving the degrees of polynomials, and so we must use positive semidefiniteness constraints to enforce what is necessary. I do not see how to proceed.

The central failing of the above proof techniques is that they use only results that relate the roots of a polynomial to its coefficients — they leave relatively undisturbed the question of their semidefinite representability.

4.2 New hyperbolic polynomials from old

The first three theorems are folklore. The first states that, if two hyperbolic cones share an interior point, their intersection is hyperbolic.

Theorem 4.2.1. *If p and q are polynomials hyperbolic in direction e, then pq is also a polynomial hyperbolic in direction e. Furthermore, $\Lambda_+(pq,e) = \Lambda_+(p,e) \cap \Lambda_+(q,e)$.*

Proof : Since $p(e) > 0$ and $q(e) > 0$, it follows that $(pq)(e) = p(e)q(e) > 0$. Fix $x \in \mathbb{R}^d$. Then $\lambda \mapsto p(x+\lambda e)$ and $\lambda \mapsto q(x+\lambda e)$ only have real roots; it follows that $\lambda \mapsto (pq)(x+\lambda e) = p(x+\lambda e)q(x+\lambda e)$ only has real roots. Thus, pq is hyperbolic.

Further, $\lambda \mapsto p(x+\lambda e)$ and $\lambda \mapsto q(x+\lambda e)$ only have nonpositive real roots if and only if $\lambda \mapsto p(x+\lambda e)q(x+\lambda e) = (pq)(x+\lambda e)$ only has nonpositive real roots. From this it follows that $\Lambda_+(pq,e) = \Lambda_+(p,e) \cap \Lambda_+(q,e)$. □

The next theorem states that direct sums of hyperbolic polynomials are hyperbolic.

Theorem 4.2.2. *If $p \in \mathbb{R}[x_1, \ldots, x_p]$ and $q \in \mathbb{R}[y_1, \ldots, y_q]$ are hyperbolic in directions e_p and e_q respectively, then the polynomial $p \oplus q \in \mathbb{R}[x_1, \ldots, x_p, y_1, \ldots, y_q]$ defined by $(p \oplus q)(x, y) = p(x)q(y)$ is hyperbolic in direction $e := e_p \oplus e_q$. Furthermore, $\Lambda_+(p \oplus q, e) = \Lambda_+(p, e_p) \oplus \Lambda_+(q, e_q)$.*

Proof : We note that $(p \oplus q)(x, y) = 0$ if and only if either $p(x) = 0$ or $q(y) = 0$. Thus, $(p \oplus q)(x + \lambda e_p, y + \lambda e_q) = 0$ if and only if $p(x + \lambda e_p) = 0$ or $q(y + \lambda e_q) = 0$. The theorem quickly follows. $\square$

The directional derivative of a hyperbolic polynomial in a direction of hyperbolicity is also hyperbolic; further, the cone generated by the directional derivative contains the original cone.

Theorem 4.2.3 ([34])**.** *If p is hyperbolic in direction e, then $p'_e = (\nabla p)(\cdot)[e]$ is also hyperbolic in direction e. Furthermore, $\Lambda_+(p'_e, e) \supseteq \Lambda_+(p, e)$.*

Proof : Apply the interlacing theorem for roots of a polynomial. $\square$

In a certain sense (of limited utility), the derivative cones completely describe $\Lambda_+(p, e)$. We make this precise:

Theorem 4.2.4. *Suppose that $\deg p \geq 2$. $\Lambda_+(p, e)$ is the intersection of all of its derivative cones — that is,*

$$\Lambda_+(p, e) = \cap_{e' \in \Lambda_{++}} \Lambda_+(p'_{e'}, e).$$

Proof : Fix $z \in \operatorname{bd} \Lambda_+(p, e)$.

We first dispatch two degenerate cases:

- If $(\nabla p)(z) = 0$, then z is also in the boundary of any derivative cone; it therefore must be in the boundary of the intersection of all derivative cones. Thus assume $(\nabla p)(z) \neq 0$.
- If $\Lambda_+(p, e)$ has nontrivial lineality space L — that is, the biggest linear space L contained in $\Lambda_+(p, e)$ has positive dimension — work instead with the restriction of p to $L^\perp$ and the projection of e onto $L^\perp$.

It suffices to show that there are points arbitrarily close to z in the two-dimensional linear subspace spanned by z and $(\nabla p)^T(z)$ that are excluded from some derivative cone. In particular, it is sufficient to establish this result in the case where the dimension of the ambient space, d, is two. We therefore focus only on this case, and we assume without loss of generality that $z = (1,0)^T$. and that $(\nabla p)(z) = (0, c)$ for some positive c.

Since $\Lambda_+(p, e)$ is linefree, $\deg p \geq 2$. Thus, p splits as

$$p(x, y) = y \prod_{i=1}^{n-1} (x + a_i y).$$

The gradient of p is then

$$(\nabla p)(x, y) = \begin{pmatrix} y \sum_{i=1}^{n-1} \prod_{j \neq i} (x + a_j y) \\ \prod_{i=1}^{n-1} (x + a_i y) + \sum_{i=1}^{n-1} a_i \prod_{j \neq i} (x + a_j y) \end{pmatrix}^T.$$

Evaluated at $(1, -\epsilon)$, this is

$$(\nabla p)(1, -\epsilon) = \begin{pmatrix} -\epsilon \sum_{i=1}^{n-1} \prod_{j \neq i} (1 - \epsilon a_j) \\ \prod_{i=1}^{n-1} (1 - \epsilon a_i) + \sum_{i=1}^{n-1} a_i \prod_{j \neq i} (1 - \epsilon a_j) \end{pmatrix}^T.$$

The component of this gradient in the first coordinate is negative; it follows that $p'_f(1, -\epsilon) < 0$ for some f in $\Lambda_{++}(p, e)$ near $(1,0)^T$, and hence that $(1, -\epsilon)^T$ is excluded from some derivative cone. The result is therefore proven. □

We now sketch some of the work of Borcea, Brändén, and Shapiro characterising the differential operators with polynomial coefficients that preserve hyperbolicity.

Recall that if p and q are in $\mathbb{R}[x]$ and only have real roots then $p \ll q$ if $p'(x)q(x) - q'(x)p(x) \geq 0$ for all x — that is, if the zeroes of p and q interlace. We present an analogous notion for multivariate polynomials.

Definition 4.2.5. Now let p and q be multivariate polynomials hyperbolic in the same direction e. We say that p and q are *in proper position* and write $p \ll q$ if $\lambda \mapsto p(x + \lambda e) \ll \lambda \mapsto q(x + \lambda e)$ for all $x \in \mathbb{R}^d$.

We now study a generalisation of Obreschkoff's theorem to hyperbolic polynomials in many variables. This generalisation is due to Borcea, Brändén, and Shapiro [5], but we shall refer to it as "Obreschkoff's theorem" regardless.

Theorem 4.2.6 (Obreschkoff's theorem)**.** *Let p and q be hyperbolic in direction e. The polynomial $\alpha p + \beta q$ is hyperbolic for all real α and β if and only if $p \ll q$ or $q \ll p$.*

Proof [5]: The backward direction quickly follows from Obreschkoff's theorem in a single variable. We therefore consider only the forward direction. By Obreschkoff's theorem in a single variable, for each $x \in \mathbb{R}^d$ we have either $p(x + \lambda e) \ll q(x + \lambda e)$ or $q(x + \lambda e) \ll p(x + \lambda e)$. If at least one of these is true for all x, then it is proven that $p \ll q$ or $q \ll p$. Thus suppose that there are x_1 and x_2 such that $p(x_1 + \lambda e) \ll q(x_1 + \lambda e)$, $q(x_2 + \lambda e) \ll p(x_2 + \lambda e)$, but $q(x_1 + \lambda e) \not\ll p(x_1 + \lambda e)$ and $p(x_2 + \lambda e) \not\ll q(x_2 + \lambda e)$.

Pick an arc from x_1 to x_2 that evades the line through origin in direction e; by continuity, there is an x_3 on this arc for which $p(x_3 + \lambda e) \ll q(x_3 + \lambda e) \ll p(x_3 + \lambda e)$. Then $p(x_3 + \lambda e)$ is a scalar multiple of $q(x_3 + \lambda e)$; say $p(x_3 + \lambda e) = \mu q(x_3 + \lambda e)$. The polynomial $p - \mu q$ is, by hypothesis, hyperbolic in direction e. Unfortunately, it is zero on the line through x_3 in direction e; we calculate

$$(p - \mu q)(e) = \lim_{t\to 0}(p - \mu q)(tx_3 + e) = \lim_{t\to 0} t^n (p - \mu q)(x_3 + e/t) = \lim_{t\to 0} t^n \cdot 0 = 0.$$

Thus $p - \mu q$ is not in fact hyperbolic in direction e. Since $q(x_1 + \lambda e) \not\ll q(x_1 + \lambda e)$, however, we see that p is not a scalar multiple of q — that is, $p - \mu q$ is not identically zero. This is a contradiction from which we conclude that $p \ll q$ or $q \ll p$. $\square$

Corollary 4.2.7 ([5])**.** *Let p be hyperbolic in direction e. The set*

$$\{q \text{ hyperbolic in direction } e : q \ll p\}$$

is a convex cone.

G. Pólya and I. Schur classified, in 1917, the set of univariate hyperbolicity preservers λ that act on polynomials by $\lambda p = \sum_{i=0}^{\infty} \lambda_i p_i x^i$. They proved the following theorem:

Theorem 4.2.8 (Pólya-Schur multiplier theorem [31])**.** *Let $\lambda : \mathbb{N} \to \mathbb{R}$ be a sequence of real numbers. Let λ act on $\mathbb{R}[x]$ as above. Define a function on $\mathbb{C}$ by*

$$\Phi(z) = \sum_{k=0}^{\infty} \frac{\lambda_k z^k}{k!}.$$

Then the following are equivalent:

(a). *λ is such that, if p is univariate hyperbolic, then $\lambda(p)$ is either zero or univariate hyperbolic.*

(b). *Φ is entire and is the limit, in the topology of uniform convergence on compacts, of a sequence of polynomials with only real roots of the same sign.*

(c). *One can write either $\Phi(z)$ or $\Phi(-z)$ in the form*

$$Ce^{\alpha z}z^n\prod_{i=0}^{\infty}(1+a_iz)$$

where $C \in \mathbb{R}$, $\alpha \geq 0$, $n \in Z$, and $a_i \geq 0$ for all i.

(d). *For every natural n, the polynomial $\lambda((1+z)^n)$ is univariate hyperbolic with all zeroes of the same sign.*

Definition 4.2.9 ([5])**.** Let $p \in \mathbb{R}[x_1, \ldots, x_d]$. p is said to be *real stable* if, for all $x \in \mathbb{R}^d$ and $e \in \mathbb{R}^d_{++}$, the univariate polynomial $p(x + \lambda e)$ only has real roots.

The main theorem in Borcea, Brändén, and Shapiro's paper is the following, which characterises the *differential* operators with *polynomial* coefficients that preserve a certain modified notion of hyperbolicity.

Theorem 4.2.10 ([5])**.** *Let $T : \mathbb{R}[x_1, \ldots, x_d] \to \mathbb{R}[x_1, \ldots, x_d]$ be defined by*

$$T(p) = \sum_{\alpha\in\mathbb{N}^d} t_\alpha \partial^\alpha(p)$$

where $t_\alpha \in \mathbb{R}[x_1, \ldots, x_d]$ for all α and is zero for all but finitely many α. (That is, T is a general finite-order linear partial differential operator with polynomial coefficients.)

Let $q \in \mathbb{R}[x_1, \ldots, x_d][w_1, \ldots, w_d]$ be defined by $q(w) = \sum_{\alpha\in\mathbb{N}^d} t_\alpha(-w)^\alpha$. Then q is real stable if and only if T maps all real stable polynomials to real stable polynomials.

4.3 Hyperbolic cones are facially exposed

We prove a result of Renegar, namely that all hyperbolic cones are facially exposed.

Proposition 4.3.1 ([34]). *Let $x \in \operatorname{bd} \Lambda_+$ and let y be a tangent direction to* $\operatorname{bd} \Lambda_+$ *at x. Then $(Hp)(x)[y, y] \leq 0$.*

PROOF : Λ_+ is convex. □

Theorem 4.3.2 ([34]). *Let $x \in \operatorname{bd} \Lambda_+$ be a point where $(\nabla p)(x) \neq 0$. Let y be a tangent direction to* $\operatorname{bd} \Lambda_+$ *at x. Either $p(x + \lambda t) = 0$ for all λ or $(Hp)(x)[y, y] < 0$.*

PROOF [34]: Suppose that $(Hp)(x)[y, y] = 0$ yet $p(x + \lambda t)$ is not identically zero. We shall restrict our attention to the one-dimensional slices $S_\alpha = \{x + \alpha e + \beta y : \beta \in \mathbb{R}\}$ — let p_α be the restriction of p to S_α.

We note that zero is a root of multiplicity at least three of p_0, since p and its second derivatives are zero at x and y is a tangent direction at x. Choose $\delta > 0$ so that 0 is the only root of p_0 within δ of 0. Then choose $\epsilon > 0$ so that p_λ has at least three roots within δ of zero whenever $0 \leq \lambda \leq \epsilon$ but no root at distance δ from 0. (This is possible by continuity of roots.)

Since $x + \lambda e \in \Lambda_{++}$ whenever $\lambda > 0$, it follows by hyperbolicity of p that p_λ only has real roots whenever $\lambda > 0$. (Namely, because $p(x + \lambda e + \beta y) = \frac{1}{\beta^n} p(\frac{x+\lambda e}{\beta} + y)$ and the right-hand side is a slice of p in direction $x + \lambda e$.)

Since 0 is not a root of p_ϵ, it follows that p_ϵ has at least two roots either in $(-\delta, 0)$ or in $(0, \delta)$; without loss of generality assume $(0, \delta)$.

The only point along the path $P = [x, x + \delta y][x + \delta y, x + \delta y + \epsilon e]$ at which p is zero is x. For each point z on P, define $q_z(t) = p(z + t(x + \epsilon e - z))$. It follows from continuity of roots that, whenever $z \in P \setminus \{x\}$, q_z has two roots in $(0, 1)$. It follows, again by continuity of roots, that q_x has two roots in $[0, 1]$. Sadly, $q_x(t) = p(x + t\epsilon e)$; for $t \in (0, 1]$, the point $x + t\epsilon e$ is in Λ_{++}. It follows that q_x has two roots at zero. This quickly implies that the gradient of p at x is zero, a contradiction. □

Theorem 4.3.3 ([34]). *Let $\Lambda_+(p,e)$ be pointed and have interior. Let $x \in \operatorname{bd} \Lambda_+(p'_e,e) \setminus \Lambda_+(p,e)$ be such that $(\nabla p'_e)(x) \neq 0$ and let $h \perp (\nabla p'_e)(x)$. Then either h is a scalar multiple of x or $(Hp'_e)(x)[h,h] < 0$.*

Proof [34]: Suppose otherwise. Then there is an $x \in \operatorname{bd} \Lambda_+(p'_e,e) \setminus \Lambda_+(p,e)$ for which $(\nabla p'_e)(x) \neq 0$ and an $h \perp (\nabla p'_e)(x)$ such that h is not a scalar multiple of x and $(Hp'_e)(x)[h,h] = 0$. By Theorem 4.3.2, then, $p'_e(x+\lambda h) = 0$ for all $\lambda \in \mathbb{R}$. By homogeneity, then, p'_e is zero on the linear space spanned by x and h; we may select h so that the line through x and $x+h$ intersects $\Lambda_+(p'_e,e)$ in a line segment. Let x_1 and x_2 be the ends of this line segment. Since $\Lambda_+(p,e)$ is convex, at least one of x_1 and x_2 must lie outside $\Lambda_+(p,e)$; without loss of generality, say x_1. We note that the line segment $[x_1,x_2]$ is contained within $\Lambda_+(p'_e,e)$.

It is clear (from continuity of roots applied to $p'_e(\lambda x_1 + (1-\lambda)x_2)$ when $\lambda < 0$) that $p'_e(x_1+\lambda e)$ has a root of multiplicity $k \geq 2$ at zero and no positive roots. By interlacing, $p(x_1+\lambda e)$ must also have a root at zero; since differentiation reduces the order of roots, this root must be of order $k+1$. Sadly, then, $p(x_1+\lambda e)$ has no positive roots and hence $x_1 \in \Lambda_+(p,e)$. This is a contradiction from which the theorem follows. $\square$

Theorem 4.3.4 ([34]). *Hyperbolic cones are facially exposed.*

Proof : We prove this first for cones that contain no line.

We induct on $\deg p$. If $\deg p = 1$, it is clear that $\Lambda_+(p,e)$ is facially exposed. Thus assume $\deg p > 1$. Let F be a face of $\Lambda_+(p,e)$ and let x be in its relative interior. If $(\nabla p)(x) = 0$, then F is also a face of $\Lambda_+(p'_e,e)$, hence an exposed face of $\Lambda_+(p'_e,e)$, hence an exposed face of $\Lambda_+(p,e)$.

Thus assume that $(\nabla p)(x) \neq 0$. Let F' be the exposed face of $\Lambda_+(p,e)$ defined by

$$F' = \{y \in \Lambda_+(p,e) : (\nabla p)(x)y = 0\}.$$

F' certainly contains F. If $F' \neq F$, then there is a direction h and an $\epsilon > 0$ for which $x + \epsilon h \in F' \setminus F$. But then $p = 0$ on the line $\{x + \lambda h : \lambda \in \mathbb{R}\}$; since p only has a simple root at x, it follows from continuity of roots that $x - \delta h \in \Lambda_+(p,e)$ for some $\delta > 0$, Thus F was not a face, a contradiction. Thus, all linefree hyperbolic cones are facially exposed.

Suppose $\Lambda_+(p,e)$ contains a line. Let L be the largest linear space contained in $\Lambda_+(p,e)$. $\Lambda_+(p,e) \cap L^\perp$ contains no line, and the faces of $\Lambda_+(p,e)$ are exactly the orthogonal direct sums of the faces of $\Lambda_+(p,e) \cap L^\perp$ with L; their exposure quickly follows. □

4.4 Strictly hyperbolic polynomials

Definition 4.4.1 ([22])**.** Let $p \in \mathbb{R}[x_1, \ldots, x_d]$. p is said to be *strictly hyperbolic* in direction e if p is hyperbolic in direction e and $(\nabla p)(x) \neq 0$ iff $x \neq 0$.

Proposition 4.4.2. *p is strictly hyperbolic in direction e if* $\operatorname{Hank} p(x + \lambda e)$ *is positive definite whenever x is not a scalar multiple of e.*

Proof : If $\operatorname{Hank} p(x + \lambda e)$ is positive semidefinite whenever x is not a scalar multiple of e, continuity implies that $\operatorname{Hank} p(x+\lambda e)$ is always positive semidefinite and hence that p is hyperbolic in direction e. If $\operatorname{Hank} p(x + \lambda e)$ is positive definite whenever x is not a scalar multiple of e, then $p(x + \lambda e)$ has no repeated roots unless x is a scalar multiple of e. Since $(\nabla p)(x) = 0$ iff x is a repeated root of p, the theorem follows. □

W. Nuij proved a number of fundamental results about strictly hyperbolic polynomials through surprisingly elegant means. We reproduce three of them below.

Theorem 4.4.3 ([29])**.** *Let $\mathcal{P}_n$ be the space of polynomials of degree n.*

(a). The set of strictly hyperbolic polynomials of degree n is open in the set of all polynomials of degree n.

(b). Every polynomial hyperbolic in direction e is a limit of polynomials strictly hyperbolic in direction e.

(c). The set of polynomials $p \in \mathcal{P}_n$ strictly hyperbolic in direction e such that $p(e) = 1$ is simply connected.

Proof [29]: Assume, without loss of generality, that e is the direction (1, 0, 0, …, 0)^T. Name the variables, in order, $x_1, \ldots, x_n$. Define a function $T_{k,s} : \mathcal{P}_n \to \mathcal{P}_n$ by

$$T_{k,s}(p) = p + sx_k \frac{\partial p}{\partial x_1}.$$

We note that $T_{k,s}$ is a linear operator. It follows from interlacing and Obreschkoff's theorem that, if p is hyperbolic and $2 \leq k \leq d$, then $T_{k,s}(p)$ is hyperbolic.

We prove (a) by purely topological means. Let S be the set of all points x such that $x_1 = 0$ and $x_2 + \ldots + x_n = 1$. Then p is strictly hyperbolic if $M(x) = \operatorname{Hank} p(x + \lambda e)$ is positive definite whenever $x \in S$. The entries of $\operatorname{Hank} p(x + \lambda e)$ are continuous functions of x, so M is continuous. The set of strictly hyperbolic polynomials in direction e, then, is the preimage under M of the set of positive definite matrices; it is therefore open. This proves (a).

Suppose that x is such that $\lambda \mapsto p(x + \lambda e)$ has a root of order ρ at zero. Then, if $sx_k \neq 0$, $\lambda \mapsto T_{k,s}(p)(x + \lambda e)$ as a root of of order at most $\rho - 1$ at zero. Thus,

$$T^n_{2,s} T^n_{3,s} \ldots T^n_{d,s} p$$

only has simple roots other than zero. One recovers p by letting s tend to zero, proving (b).

We omit the proof of (c). $\square$

Part (b) is the most interesting for our purposes; it allows us to prove things about hyperbolic cones that do not deeply rely upon the boundary structure by proving such results about strictly hyperbolic polynomials and then appealing to continuity.

4.5 Extreme rays of hyperbolic cones

A complete characterisation of the extreme rays of a hyperbolic cone that appeals only to properties of the defining polynomial and direction currently seems slightly out-of-reach, but we can make some useful progress in this direction. We begin by recording a theorem of J. Renegar that appears to be the only previously-known result toward characterising the extreme rays of a hyperbolic cone.

Theorem 4.5.1 ([34])**.** *Let $\Lambda_+(p, e)$ be linefree and suppose that $d \geq 3$. If $x \in \operatorname{bd} \Lambda_+(p'_e, e) \setminus \Lambda_+(p, e)$, then $(\nabla p'_e)(x) \neq 0$ and x defines an extreme ray of $\Lambda_+(p'_e, e)$.*

PROOF : By 4.3.3, the boundary of $\Lambda_+(p'_e, e)$ has positive curvature in all but one direction at x; it follows that x defines an extreme ray. $\square$

One wonders how useful the above theorem is for characterising the extreme rays of a hyperbolic cone. We observe that it is very useful for derivatives of strictly hyperbolic polynomials:

Theorem 4.5.2. *Suppose p is strictly hyperbolic in direction e and is such that $\Lambda_+(p, e)$ contains no line. If $x \in \operatorname{bd} \Lambda_+(p'_e, e)$, then x is an extreme ray of $\Lambda_+(p'_e, e)$.*

PROOF : We note that, if $x \in \operatorname{bd} \Lambda_+(p'_e, e)$, then $x \notin \operatorname{bd} \Lambda_+(p, e)$ since p has no multiple roots. Theorem 4.5.1 applies and yields the desired result. $\square$

References

[1] Michael Atiyah, Raoul Bott, and Lars Gårding. Lacunas for hyperbolic differential operators I. *Acta Mathematica*, 128:109–189, 1972.

[2] Michael Atiyah, Raoul Bott, and Lars Gårding. Lacunas for hyperbolic differential operators with constant coefficients II. *Acta Mathematica*, 131:145–206, 1973.

[3] Saugata Basu, Richard Pollack, and Marie-Françoise Roy. *Algorithms in real algebraic geometry*, volume 10 of *Algorithms and Computation in Mathematics*. Springer-Verlag, 2005.

[4] Heinz H. Bauschke, Osman Güler, Adrian S. Lewis, and Hristo S. Sendov. Hyperbolic polynomials and convex analysis. *Canadian Journal of Mathematics*, 53:470–488, 2001.

[5] Julius Borcea, Petter Brändén, and Boris Shapiro. Classification of hyperbolicity and stability preservers: the multivariate Weyl algebra case. *arXiv:math/0606360*.

[6] Julius Borcea, Petter Brändén, and Boris Shapiro. Applications of stable polynomials to mixed determinants: Johnson's conjectures, unimodality and symmetrized Fischer products. *Duke Journal of Mathematics*, to appear, 2008.

[7] Petter Brändén. Polynomials with the half-plane property and matroid theory. *arXiv:0605678*.

[8] Petter Brändén, Julius Borcea, and Thomas Liggett. Negative dependence and the geometry of polynomials. *arXiv:0707.2840*.

[9] Constantine Carathéodory. Über den Variabilitätsbereich der Fourierschen Konstanten von positiven harmonischen Funktionen. *Rend. Circ. Mat. Palermo*, 32:193–217, 1911.

[10] Young-Bin Choe, James G. Oxley, Alan D. Sokal, and David G. Wagner. Homogeneous multivariate polynomials with the half-plane property. *Advances in Applied Mathematics*, 32:88–187, 2004.

[11] Chek Beng Chua. Relating homogeneous cones and positive definite cones via T-algebras. *SIAM Journal on Optimization*, 14:500–506, 2003.

[12] Chek Beng Chua and Levent Tunçel. Invariance and efficiency of convex representations. *Math. Programming*, 111:113–140, 2008.

[13] Jacques Faraut and Adam Korányi. *Analysis on symmetric cones*. Oxford University Press, 1994.

[14] Leonid Faybusovich. On Nesterov's approach to semi-infinite programming. *Acta Applicandae Mathematicae*, 74:195–215, 2002.

[15] Lars Gårding. An inequality for hyperbolic polynomials. *Journal of Mathematics and Mechanics*, 8:957–965, 1959.

[16] Semën G. Gindikin. *Tube domains and the Cauchy problem*. Number 111 in Translation of Mathematical Monographs. American Mathematical Society, 1992.

[17] Osman Güler. Hyperbolic polynomials and interior point methods for convex programming. *Mathematics of Operations Research*, 22:350–377, 1997.

[18] Leonid Gurvits. Hyperbolic polynomials approach to Van der Waerden/Schrijver-Valiant like conjectures: sharper bounds, simpler proofs and algorithmic applications. In *Symposium on the Theory of Computation 2006*, pages 417–426, 2006.

[19] Eduard Helly. Über Mengen konvexer Körper mit gemeinschaftlichen Punkten. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 22:175–176, 1923.

[20] J. William Helton and Jiawang Nie. Sufficient and necessary conditions for semidefinite representability of convex hulls and sets. *arXiv:0705.4068v5.*

[21] J. William Helton and Victor Vinnikov. Linear matrix inequality representation of sets. *Communications on Pure and Applied Mathematics*, pages 654–674, 2002.

[22] Lars Hörmander. *Analysis of linear partial differential operators II*, volume 256 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, 1983.

[23] Fritz John. Extremum problems with inequalities as subsidiary conditions. In *Studies and essays presented to R. Courant on his 60th birthday*. Interscience, 1948.

[24] Pascual Jordan, John von Neumann, and Eugene Wigner. On an algebraic generalization of the quantum mechanical formalism. *Annals of Mathematics*, 35:29–64, 1934.

[25] Nikolai V. Krylov. On the general notion of fully nonlinear second-order elliptic equations. *Transactions of the American Mathematical Society*, 347:857–895, 1995.

[26] Adrian S. Lewis, Pablo A. Parrilo, and Motakuri V. Ramana. The Lax conjecture is true. *Proceedings of the American Mathematical Society*, 133:2495–2499, 2005.

[27] Yurii Nesterov. Squared functional systems and optimization problems. In Hans Frenk, Kees Roos, Tamás Terlaky, and Shuzhong Zheng, editors, *High performance optimization*, number 33 in Applied Optimization, pages 405–440. Kluwer Academic Publishers, 2000.

[28] Yurii Nesterov and Arkadii Nemirovskii. *Interior-point polynomial algorithms in convex programming.* Number 13 in Studies in Applied Mathematics. Society for Industrial and Applied Mathematics, 1994.

[29] Wim Nuij. A note on hyperbolic polynomials. *Mathematica Scandinavica*, 23:69–72, 1968.

[30] Nikola Obreschkoff. *Verteilung und Berechnung der Nullstellen reeller Polynome*. VEB Deutscher Verlag der Wissenschaften, 1963.

[31] Gergely Pólya and Issai Schur. Über zwei Arten vin Faktorenfolgen in der Theorie der algebraischen Gleischungen. *Journal fur die Reine und angewandte Mathematik*, 114:89–113, 1913.

[32] Johann Radon. Mengen konvexer Körper, die einen gemeinsamen Punkt enthalten. *Mathematische Annalen*, 83:113–115, 1921.

[33] Qazi Ibadur Rahman and Gerhard Schmeisser. *Analytic theory of polynomials*. London Mathematics Society, 2002.

[34] James Renegar. Hyperbolic programs, and their derivative relaxations. *Foundations of Computational Mathematics*, 6(1):59–79, 2006.

[35] Rolf Schneider. *Convex bodies: the Brunn-Minkowski theory*. Cambridge University Press, 1992.

[36] N. Z. Shor. *Metody minimizatsii nedifferentstruemykh funktsii i ikh prilozheniya (*Minimization methods for nondifferentiable functions and their applications*)*. Naukova Dumka, 1979.

[37] Van Anh Truong and Levent Tunçel. Geometry of homogeneous convex cones, duality mapping, and optimal self-concordant barriers. *Mathematical Programming*, 100:295–316, 2005.

[38] Levent Tunçel. *Polyhedral and semidefinite programming methods in combinatorial optimization*. Fields Institute monograph series. American Mathematical Society, to appear.

[39] Ernest B. Vinberg. The theory of homogeneous cones. *Transactions of the Moscow Mathematical Society*, 12:340–403, 1963.

[40] Ernest B. Vinberg. The structure of the group of automorphisms of a homogeneous convex cone. *Transactions of the Moscow Mathematical Society*, 13:63–93, 1965.

[41] Victor Vinnikov. Self-adjoint determinantal representations of real plane curves. *Mathematische Annalen*, pages 453–479, 1993.

[42] David G. Wagner and Yehua Wei. A criterion for the half-plane property. *arXiv:0709.1269*.

[43] G. B. Yudin and Arkadii S. Nemirovskii. Informational complexity and effective methods for the solution of convex extremal problems (russian). *Ekonomika i Matematicheskie Metody*, 12:357–369, 1978.