Proceedings of Symposia in Applied Mathematics
Volume **77**, 2020
https://doi.org/10.1090/psapm/077/00681

The geometry of spectrahedra

Cynthia Vinzant

Abstract. Spectrahedra are convex, semialgebraic sets that are fundamental objects in the theory of semidefinite programming, and more generally matrix theory, convex optimization, and real algebraic geometry. Here we develop some of the relevant background from real algebraic and convex geometry, and explore its impacts on the geometry of spectrahedra and applications to cones of sums of squares and moments.

1. Introduction

A spectrahedron is the intersection of the cone of positive semidefinite matrices with an affine linear space. These appear in convex optimization as the feasible sets of semidefinite programs. In what follows, we investigate the geometry of spectrahedra using tools from real algebraic geometry and convexity and discuss implications for spectrahedra related to sums of squares. The geometry of spectrahedra depends heavily on the geometry of the cone of positive semidefinite matrices.

Definition 1.1. A real symmetric matrix A is *positive definite*, denoted $A \succ 0$, if all of its eigenvalues are positive. The matrix A is *positive semidefinite*, denoted $A \succeq 0$, if all of its eigenvalues are nonnegative.

There are several equivalent characterizations of positive semidefinite-ness. For an $N \times N$ real symmetric matrix A, the following conditions are equivalent:

(i) all eigenvalues of A are nonnegative,
(ii) all principal minors of A, $\det(A_{I,I})$ for $I \subseteq \{1, \ldots, N\}$, are nonnegative,
(iii) $\mathbf{v}^T A \mathbf{v} \geq 0$ for all $v \in \mathbb{R}^N$, and
(iv) there exists a matrix $B \in \mathbb{R}^{N \times k}$ where $k = \text{rank}(A)$ and

$$A \;=\; BB^T \;=\; (\langle \mathbf{r}_i, \mathbf{r}_j \rangle)_{1 \leq i,j \leq N} \;=\; \sum_{i=1}^{k} \mathbf{c}_i \mathbf{c}_i^T$$

where $\mathbf{r}_1, \ldots, \mathbf{r}_N$, $\mathbf{c}_1, \ldots, \mathbf{c}_k$ are the rows and columns of B.

Let $\mathcal{S}^N$ denote the vector space of real symmetric $N \times N$ matrices and $\mathcal{S}_+^N$ denote the set of $N \times N$ positive semidefinite matrices. The set $\mathcal{S}_+^N$ is a full-dimensional *convex semialgebraic cone* in the real vector space $\mathcal{S}^N \cong \mathbb{R}^{\binom{N+1}{2}}$. General convex

2010 *Mathematics Subject Classification.* Primary 90C22, 15B48, 14P10, 52A41.

Key words and phrases. Spectrahedra, positive semidefinite matrix, semidefinite program, sums of squares, moment problems.

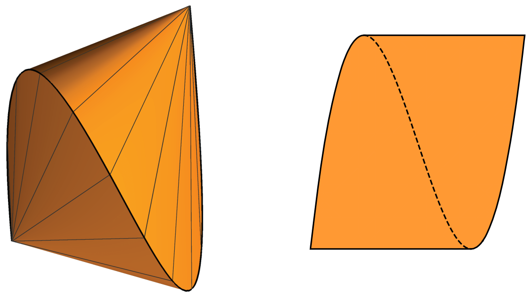

FIGURE 1. The spectrahedron in $\mathcal{S}^4$ defined by (1.1) and its image under the coordinate projection $(x, y, z) \mapsto (x, z)$.

and semialgebraic sets will be defined and discussed in Section 2. One example is the intersection of $\mathcal{S}_+^N$ with an affine linear subspace of $\mathcal{S}^N$.

DEFINITION 1.2. A *spectrahedron* is the intersection of the cone of positive semidefinite matrices with an affine linear space. The affine linear space can be parametrized as $\mathcal{L} = \{A(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^m\}$ where $A(\mathbf{x}) = A_0 + \sum_{i=1}^m x_i A_i$ is a linear matrix pencil given by $A_0, \ldots, A_m \in \mathcal{S}^N$. This linearly identifies the spectrahedron with

$$\mathcal{L} \cap \mathcal{S}_+^N \;\cong\; \left\{ \mathbf{x} \in \mathbb{R}^m \;:\; A(\mathbf{x}) = A_0 + \sum_{i=1}^m x_i A_i \succeq 0 \right\}.$$

We will refer to both sets (and any set linearly isomorphic to them) as spectrahedra.

EXAMPLE 1.3. Consider the three-dimensional affine linear space parametrized

$$A(x, y, z) \;=\; A_0 + xA_1 + yA_2 + zA_3 \;=\; \begin{pmatrix} 1 & x & y & z \\ x & 1 & x & y \\ y & x & 1 & x \\ z & y & x & 1 \end{pmatrix} \tag{1.1}$$

where (x, y, z) runs over $\mathbb{R}^3$ and the matrices $A_0, A_1, A_2, A_3 \in \mathcal{S}^4$ have the form $(A_k)_{ij} = 1$ if $|i - j| = k$ and 0 otherwise. The intersection of this affine space with the cone of 4×4 positive semidefinite matrices $\mathcal{S}_+^4$ gives the spectrahedron $S = \{(x, y, z) \in \mathbb{R}^3 : A(x, y, z) \succeq 0\}$ shown on the left in Figure 1. $\diamond$

EXAMPLE 1.4. Another important example of a spectrahedron is any polyhedron. A *polyhedron* is the set of points satisfying finitely many affine linear inequalities. That is, it is a set that can be written as

$$P = \left\{ \mathbf{x} \in \mathbb{R}^m : a_{0j} + \sum_{i=1}^m a_{ij} x_i \geq 0, \; j = 1, \ldots, N \right\},$$

where $a_{ij} \in \mathbb{R}$ for all $i = 0, \ldots, m$ and $j = 1, \ldots, N$. We can represent P as a spectrahedron using diagonal matrices. Indeed, for each $i = 0, \ldots, m$, let A_i be the $N \times N$ diagonal matrix with diagonal entries $(a_{i1}, \ldots, a_{iN})$. For any $\mathbf{x} \in \mathbb{R}^m$, the linear combination $A_0 + \sum_{i=1}^m A_i x_i$ is again a diagonal matrix, which is positive semidefinite if and only if its diagonal entries are nonnegative. This writes P as the spectrahedron,

$$P = \left\{ \mathbf{x} \in \mathbb{R}^m : A_0 + \sum_{i=1}^{m} A_i x_i \succeq 0 \right\}.$$

$\diamond$

Another important class of convex semialgebraic sets are *spectrahedral shadows*, which are the image of spectrahedra under linear projections. One can define these as sets that can be written in the form

$$\left\{ (x_1, \dots, x_m) \in \mathbb{R}^m : \ \exists (y_1, \dots, y_n) \in \mathbb{R}^n \text{ with } A_0 + \sum_{i=1}^{m} x_i A_i + \sum_{j=1}^{n} y_j B_j \succeq 0 \right\}$$

for some matrices $A_i, B_j \in \mathcal{S}^N$. For example, the projection of the spectrahedron S in Example 1.3 onto the coordinates (x, z) is shown on the right in Figure 1. As we will see, spectrahedral shadows are not always spectrahedra. This is a major distinction with polyhedra – the image of a polyhedron under linear projection is always a polyhedron.

Like the cone of positive semidefinite matrices, spectrahedra and spectrahedral shadows are convex, semialgebraic sets, and we can analyze both their algebraic and convex structure.

In what follows we build up some of the necessary background from real algebraic geometry and convexity (Section 2), introduce convex duality with a focus on spectrahedral cones (Section 3), examine the facial structure of the cone of positive semidefinite matrices and spectrahedra (Section 4), and end by discussing applications of this theory to sums of squares and nonnegative polynomials (Section 5).

2. Background

Throughout, we use the standard Euclidean inner product on $\mathbb{R}^n$, where for two vectors $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{y} = (y_1, \dots, y_n)$, $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i$. On $\mathcal{S}^N$, we use the trace inner product $\langle A, B \rangle = \mathrm{trace}(AB)$.

2.1. Real algebraic geometry and semialgebraic sets. The fundamental building blocks of real algebraic geometry are *semialgebraic sets*, which are sets defined by polynomial equations and inequalities. For more details on the subject, see [**3**, **10**]. Formally, we have the following:

Definition 2.1. A *basic closed semialgebraic set* in $\mathbb{R}^n$ is a set of the form

$$\{\mathbf{p} \in \mathbb{R}^n : g_1(\mathbf{p}) \geq 0, \dots, g_s(\mathbf{p}) \geq 0\}$$

where $g_1, \dots, g_s \in \mathbb{R}[x_1, \dots, x_n]$ are polynomials. A *semialgebraic set* is one that can be written as a finite boolean combination of basic closed semialgebraic sets, that is, one built from a finite collection of basic closed semialgebraic sets using finitely many complements, intersections, and unions.

Example 2.2. A disk and a square are both basic closed semialgebraic subsets of $\mathbb{R}^2$. Their union will be a closed semialgebraic set, but might not be a basic one. If D is the unit disk defined by $x^2+y^2 \leq 1$ and S is the square defined by $0 \leq x \leq 2$ and $-1 \leq y \leq 1$, then the union $D \cup S$ is semialgebraic, but not basic. These are shown is Figure 2.

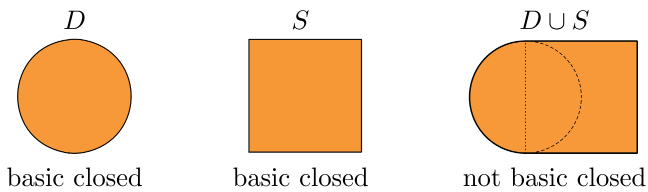

FIGURE 2. The semialgebraic sets of Example 2.2.

To see this, suppose that $D \cup S$ can be described by some set of polynomial inequalities, $g_1(x, y) \geq 0, \ldots, g_s(x, y) \geq 0$. Because this description is finite, there must be some polynomial g_i that changes sign and thus vanishes with odd multiplicity along the left half of the circle defined by $x^2 + y^2 = 1$. That is, for almost every choice of $a \in [-1, 1]$, the univariate polynomial $g_i(x, a) \in \mathbb{R}[x]$ has a root of odd multiplicity at $x = -\sqrt{1 - a^2}$. However, any polynomial in $\mathbb{R}[x, y]$ that vanishes on the left half of the circle $\{(-\sqrt{1 - a^2}, a) : a \in [-1, 1]\}$ must also vanish on the right half of the circle $\{(\sqrt{1 - a^2}, a) : a \in [-1, 1]\}$. In the language of algebraic geometry, the *Zariski closure* of the left half is the entire circle. In particular this is true for g_i and its derivatives $\left(\frac{\partial}{\partial x}\right)^k g_i$ with respect to x. Therefore g_i must also vanish with odd multiplicity on the right half of this circle, implying that g_i is negative at some point in $D \cup S$, and giving a contradiction. $\diamond$

The cone of positive semidefinite matrices is a basic closed semialgebraic set in $\mathcal{S}^N \cong \mathbb{R}^{\binom{N+1}{2}}$ defined by the non-negativity of principal minors. For example,

$$\mathcal{S}^2_+ = \left\{ \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \in \mathcal{S}^2 \ : \ a_{11}a_{22} - a_{12}^2 \geq 0, \ a_{11} \geq 0, \ a_{22} \geq 0 \right\}.$$

More generally, the principal minors $\det(A_{I,I})$ of a symmetric matrix A are polynomials in its entries and $\mathcal{S}^N_+$ is defined by the nonnegativity of all these polynomials. This gives an explicit description of $\mathcal{S}^N_+$ as a basic closed semialgebraic set.

The property of being a basic closed semialgebraic set is inherited by spectrahedra $\mathcal{S}^N_+ \cap \mathcal{L}$. (Note that we can impose an equality $\ell(\mathbf{x}) = 0$ by imposing the two inequalities $\ell(\mathbf{x}) \geq 0$ and $-\ell(\mathbf{x}) \geq 0$.) One can also see this in the coordinates $\mathbf{x} = (x_1, \ldots, x_m)$ of the parametrization $\mathcal{L} = \{A(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^m\}$. The spectrahedron $S = \{\mathbf{x} \in \mathbb{R}^m : A(\mathbf{x}) \succeq 0\}$ is defined by the nonnegativity of the principal minors of $A(\mathbf{x})$, which are polynomials in $\mathbf{x}$.

EXAMPLE 1.3 CONT'D. The determinant of the matrix $A(x, y, z)$ in (1.1) is a polynomial of degree four that factors into two quadratics. The semialgebraic description of S given by the principal minors of $A(x, y, z)$ can be simplified to

$$S = \{(x, y, z) \in \mathbb{R}^3 \ : \ (y + x)^2 \leq (z + 1)(x + 1), (y - x)^2 \leq (z - 1)(x - 1), x^2 \leq 1\}.$$

Its projection onto coordinates (x, z), seen in Figure 1, is the semialgebraic set

$$\pi(S) = \{(x, z) : -1 \leq z \leq 4x^3 - 3x, x \leq 1/2\} \cup \{(x, z) : 4x^3 - 3x \leq z \leq 1, -1/2 \leq x\}.$$

We can see that $\pi(S)$ is not basic closed using an argument similar to that of Example 2.2. In any polynomial description, $g_1 \geq 0, \ldots, g_s \geq 0$ of $\pi(S)$, some polynomial g_i would have to vanish with odd multiplicity along the curve given by $z = 4x^3 - 3x$, which forms part of the boundary of $\pi(S)$. However this curve

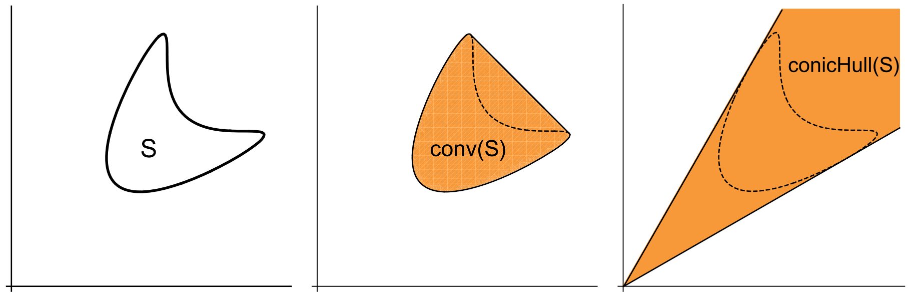

FIGURE 3. The convex and conic hull of a set S in $\mathbb{R}^2$.

also passes through the interior of $\pi(S)$, on which g_i can only vanish with even multiplicity. So no such representation can exist.

Since it is not basic closed, the convex semialgebraic set $\pi(S)$ cannot be written as a spectrahedron! $\diamond$

One fundamental theorem in real algebraic geometry is the Tarski-Seidenberg theorem, which states that the projection of a semialgebraic set is semialgebraic. As we see in Example 1.3, it is not always true that the projection of a basic closed semialgebraic set is basic.

2.2. Convex geometry and faces. Many important features of spectrahedra come from their structure as convex sets. We refer readers to [1] for more background on convexity and in particular [1, §II.12] for its relevance to the positive semidefinite cone and spectrahedra.

A set $S \subset \mathbb{R}^n$ is *convex* if for any points $\mathbf{x}, \mathbf{y} \in S$, the line segment joining them, $\{\lambda\mathbf{x} + (1-\lambda)\mathbf{y} : 0 \leq \lambda \leq 1\}$ is contained in S. We call S a *convex cone* if it is also invariant by nonnegative scaling, or equivalently if for any two points $\mathbf{x}, \mathbf{y} \in S$, the set of conic combinations $\{\lambda\mathbf{x} + \mu\mathbf{y} : \lambda, \mu \in \mathbb{R}_{\geq 0}\}$ is contained in S.

For example, $\mathcal{S}_+^N$ is a convex cone. To see this, suppose that $A, B \in \mathcal{S}_+^N$. Then for every nonzero vector $\mathbf{v} \in \mathbb{R}^N$, $\mathbf{v}^T A\mathbf{v}$ and $\mathbf{v}^T B\mathbf{v}$ are nonnegative. It follows that for any $\lambda, \mu \in \mathbb{R}_{\geq 0}$, $\mathbf{v}^T(\lambda A + \mu B)\mathbf{v} = \lambda\mathbf{v}^T A\mathbf{v} + \mu\mathbf{v}^T B\mathbf{v} \geq 0$, giving that the conic combination $\lambda A + \mu B$ is also positive semidefinite. One can check that the intersection of any convex sets is convex. Since $\lambda\mathbf{x} + (1-\lambda)\mathbf{y}$ is an affine linear combination of $\mathbf{x}$ and $\mathbf{y}$, any affine linear space is convex. Putting this together with the convexity of $\mathcal{S}_+^N$, we find that a spectrahedron $S = \mathcal{L} \cap \mathcal{S}_+^N$ is convex.

The *convex hull* of a set $S \subset \mathbb{R}^n$ is the smallest convex set containing S and its *conic hull* is the smallest convex cone containing S. See Figure 3. We can write these as

$$\mathrm{conv}(S) = \left\{ \sum_{i=1}^{k} \lambda_i \mathbf{p}_i : \; k \in \mathbb{N}, \; \mathbf{p}_i \in S, \; \lambda_i \geq 0, \; \sum_{i=1}^{k} \lambda_i = 1 \right\}, \text{ and}$$

$$\mathrm{conicHull}(S) = \left\{ \sum_{i=1}^{k} \lambda_i \mathbf{p}_i : \; k \in \mathbb{N}, \; \mathbf{p}_i \in S, \; \lambda_i \geq 0 \right\}.$$

In a convex set some points belong to the convex hull of others. This relationship gives rise to the facial structure of S. A *face* of S is a subset $F \subseteq S$ with the property that for any point in F and every way to write it as a (strict) convex combination of some elements of S, these elements must belong to F. That is, for every $\mathbf{p} \in F$, if $\mathbf{p} = \lambda\mathbf{x} + (1-\lambda)\mathbf{y}$ for some $\lambda \in (0,1)$ where $\mathbf{x}, \mathbf{y} \in S$, then $\mathbf{x}, \mathbf{y} \in F$. Note that the same condition with $\lambda = 0$ or 1 cannot impose any restrictions on the points $\mathbf{x}$ or $\mathbf{y}$, respectively, which is why we restrict to strict convex combinations with $\lambda \in (0,1)$. Both $\emptyset$ and S are always faces of S and we call these the *trivial* faces of S. A nonempty face F of S with $F \neq S$ is called *nontrivial*. An *extreme point* of S is a point $\mathbf{p} \in S$ that forms a singleton face of S. This is a point $\mathbf{p}$ with the property that every way of writing $\mathbf{p}$ as a strict convex combination $\lambda\mathbf{x} + (1-\lambda)\mathbf{y}$ of points $\mathbf{x}$, $\mathbf{y}$ in S with $\lambda \in (0,1)$ satisfies $\mathbf{x} = \mathbf{y} = \mathbf{p}$.

Krein-Milman Theorem. *A convex compact set in $\mathbb{R}^n$ is the convex hull of its extreme points.*

For any linear function ℓ on $\mathbb{R}^n$, the set of points F maximizing ℓ over S is a face of S. Indeed, if, for some $c \in \mathbb{R}$, the inequality $\ell(\mathbf{x}) \leq c$ holds for all $\mathbf{x} \in S$, then the set of points in S satisfying $\ell(\mathbf{x}) = c$ is a face of S. To see this, note that if $\ell(\mathbf{p}) = c$ and $\mathbf{p} = \lambda\mathbf{x} + (1-\lambda)\mathbf{y}$ for some $\lambda \in (0,1)$ and $\mathbf{x}, \mathbf{y} \in S$, then

$$c = \ell(\mathbf{p}) = \lambda\ell(\mathbf{x}) + (1-\lambda)\ell(\mathbf{y}) \geq \lambda c + (1-\lambda)c = c.$$

Equality of the left and right hand sides forces $\ell(\mathbf{x}) = \ell(\mathbf{y}) = c$.

We call faces of the form $\{\mathbf{x} \in S : \ell(\mathbf{x}) = c\}$ *exposed* and say that ℓ exposes the face. Note that the trivial faces of S are always exposed using the zero function $\ell(\mathbf{x}) = 0$. The set $\emptyset$ is the set of points satisfying $\ell(\mathbf{x}) = 1$ and S is the set of points satisfying $\ell(\mathbf{x}) = 1$. If ℓ is non-zero, then the set of points $\{\mathbf{x} \in \mathbb{R}^n : \ell(\mathbf{x}) = c\}$ forms an affine hyperplane H and $F = S \cap H$. Similarly, if $F = S \cap H$ is a face of S for some affine hyperplane, then it is the set of points maximizing ℓ where $\ell(\mathbf{x}) = c$ is the defining equation for H with signs chosen so that $\ell(\mathbf{x}) \leq c$ for all $\mathbf{x} \in S$. Therefore a nontrivial face of S is exposed if and only if it can be written as $S \cap H$ for some affine hyperplane.

Example 2.2 cont'd. All three sets from Example 2.2 are convex. The nontrivial faces of the disk D are the singleton sets consisting of a point on its boundary, each of which is exposed. For example, the point $(0,1)$ is exposed by the linear function y, as it equals the set of points maximizing y over $(x,y) \in D$. The square S has finitely many faces, namely $\emptyset$, its four vertices, its four edges, and S itself. These faces are all exposed, as highlighted in Figure 4.

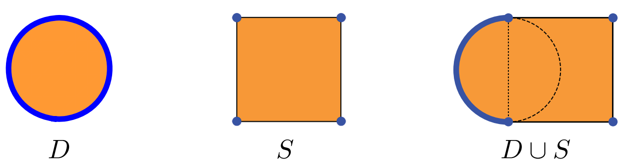

Figure 4. Extreme points of the convex sets in Example 2.2.

The facial structure of $D \cup S$ is more complicated. Its nontrivial faces consist of its extreme points, which are the points on the left half of the circle bounding D and two of the vertices of S as well as three edges of S. Note however that the points $(0,1)$ and $(0,-1)$ are not exposed on $D \cup S$. Up to scaling, the only linear function ℓ that maximizes $(0,1)$ over $D \cup S$ is $\ell(x,y) = y$, but every point on the top edge attains this maximum. As promised by the Krein-Milman Theorem, each set is the convex hull of its extreme points. ◇

In their seminal paper on spectrahedra, Ramana and Goldman show that all faces of a spectrahedron are exposed [13]. We discuss this theorem further and give a proof in Section 4.2. This gives another way of verifying that a convex semialgebraic set cannot be written as a spectrahedron.

Example 1.3 cont'd. The spectrahedron S in Figure 1 has a curve segment worth of extreme points, namely points of the form $(x, y, z) = (t, 2t^2 - 1, 4t^3 - 3t)$ for $t \in [-1, 1]$. As promised by the Krein-Milman Theorem, S is the convex hull of this curve segment. The other faces of S are the edges joining points on this curve to one of the end points $(\pm 1, 1, \pm 1)$, along with the trivial faces $\emptyset$ and S.

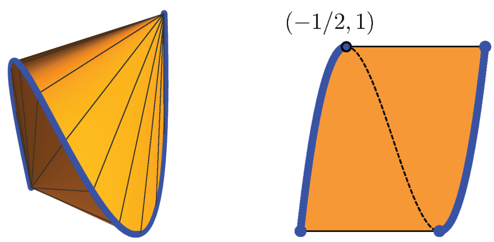

Figure 5. Extreme points of the convex sets in Figure 1.

The image $\pi(S)$ under projection $\pi(x, y, z) = (x, z)$ is again a compact, convex set and is the convex hull of its extreme points, shown in Figure 5. The projection $\pi(S)$ also has a non-exposed face. The point $(x, z) = (-1/2, 1)$ is a face of $\pi(S)$, as it cannot be written as a strict convex combination of other points in the set. Up to scaling, the only linear function maximized at this point is $\ell(x, z) = z$, which attains the maximum value $z = 1$. However the face of $\pi(S)$ exposed by this linear function is the line segment $\{(\lambda, 1) : \lambda \in [-1/2, 1]\}$. Therefore the singleton face $\{(-1/2, 1)\}$ of $\pi(S)$ is not exposed. This provides yet another reason that $\pi(S)$ cannot be written as a spectrahedron. ◇

We see here that one way to understand the structure of a convex set is through the affine linear functions that are nonnegative on it. The set of such functions form a convex set in their own right, known as the *dual*.

3. Convex duality

Given a set $S \subset \mathbb{R}^n$, we can define the *dual convex set* of S as

$$S^* \;=\; \{\mathbf{c} \in \mathbb{R}^n : 1 + \langle \mathbf{c}, \mathbf{x}\rangle \geq 0 \text{ for all } \mathbf{x} \in S\}.$$

One can check that S^* is a closed convex set that contains the origin. Moreover, if S is semialgebraic, then so is S^*. The definition gives a semialgebraic formula for S^* involving a quantifier. The theory of *quantifier elimination* over $\mathbb{R}$ (see e.g. [3]) then implies that S^* has a quantifier-free description as a semialgebraic set.

Some readers may instead be familiar with the *polar* of convex set, given by $S^\circ = \{\mathbf{c} \in \mathbb{R}^n : 1 \geq \langle \mathbf{c}, \mathbf{x}\rangle \text{ for all } \mathbf{x} \in S\}$. But one can check that these two sets are the negative of each other, $S^* = -S^\circ$.

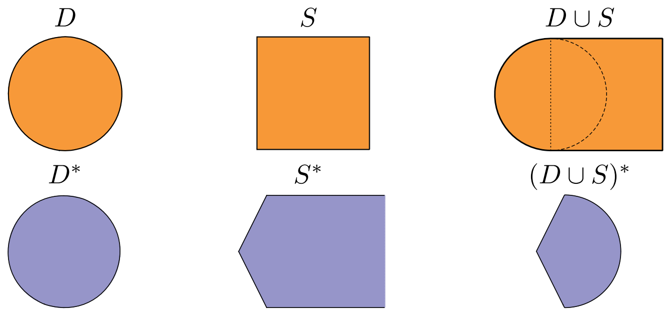

FIGURE 6. The convex sets of Example 2.2 and their duals.

From the definition, one can check that duality reverses inclusions and that the dual of a union of two convex sets is the intersection of their duals, i.e.

$$S \subseteq T \;\Rightarrow\; T^* \subseteq S^* \quad \text{and} \quad (S \cup T)^* \;=\; S^* \cap T^*$$

Moreover, if S is a closed convex set containing the origin, then $(S^*)^* = S$, see e.g. [1, Theorem IV.1.2].

EXAMPLE 2.2 CONT'D. Consider the duals of the convex sets in Example 2.2. The dual of each is semialgebraic. In this case, it so happens all three dual sets are basic closed semialgebraic sets, namely

$$\begin{aligned} D^* &= \{(a,b) \in \mathbb{R}^2 : a^2 + b^2 \leq 1\}, \\ S^* &= \{(a,b) \in \mathbb{R}^2 : -1 \leq b \leq 1, -2a - 1 \leq b \leq 2a + 1\}, \quad \text{and} \\ (D \cup S)^* = D^* \cap S^* &= \{(a,b) \in \mathbb{R}^2 : a^2 + b^2 \leq 1, -2a - 1 \leq b \leq 2a + 1\}. \end{aligned}$$

These sets are defined by eliminating (x,y) from the inequalities $ax + by + 1 \geq 0$ and $(x,y) \in D$, S, or $D \cup S$, respectively, and pictured in Figure 6.

It is instructive to imagine moving a point (a,b) in the dual set and the corresponding line $ax + by + 1 = 0$ in the plane of the original set, as is done for four points in Figure 7. For any (a,b) in $(D \cup S)^*$, the original, $D \cup S$, belongs to the halfspace $ax + by + 1 \geq 0$ bounded by this line, but when (a,b) belong to the *boundary* of the $(D \cup S)^*$, the line intersects $D \cup S$. By considering $((D \cup S)^*)^* = D \cup S$, we can also visualize a point (x,y) in $D \cup S$ as a valid inequality $ax + by + 1 \geq 0$ on points (a,b) in $D \cup S$, as seen in Figure 7. $\diamond$

One convenient interpretation of the dual convex set involves working with convex cones in one dimension higher. Formally, given a set $S \subseteq \mathbb{R}^n$, define the cone over S, denoted $\text{cone}(S)$, to be the convex cone in $\mathbb{R}^{n+1}$ defined by

$$\text{cone}(S) \;=\; \{(\lambda, \lambda \mathbf{x}) \;:\; \lambda \in \mathbb{R}_{\geq 0},\; \mathbf{x} \in S\}\,.$$

Note that we can recover S from its cone by restricting the points with first coordinate equal to one. Furthermore, the operation of coning over sets behaves nicely with duality.

PROPOSITION 3.1. *The operations of taking the cone over a set and taking its dual commute. That is, for a convex set $S \subseteq \mathbb{R}^n$, we have*

$$\text{cone}(S)^* \;=\; \overline{\text{cone}(S^*)}.$$

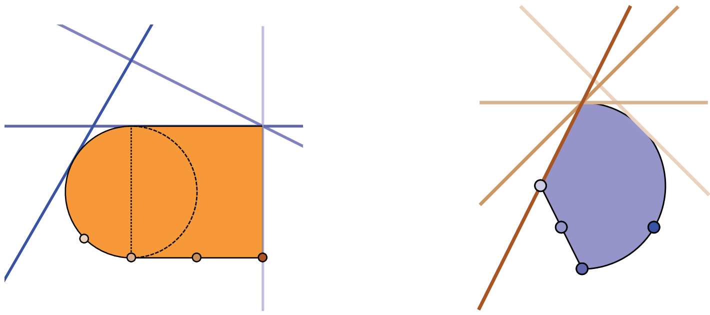

FIGURE 7. Point-line correspondences in a convex set and its dual.

PROOF. Consider a point $(c_0, \mathbf{c})$ in $\mathbb{R}^{n+1}$ with $c_0 > 0$.

$$\begin{aligned}
(c_0, \mathbf{c}) \in \text{cone}(S^*) \;&\Leftrightarrow\; \tfrac{1}{c_0}\mathbf{c} \in S^* \\
&\Leftrightarrow\; 1 + \langle \tfrac{1}{c_0}\mathbf{c}, \mathbf{y}\rangle \geq 0 \text{ for all } \mathbf{y} \in S \\
&\Leftrightarrow\; c_0 + \langle \mathbf{c}, \mathbf{y}\rangle \geq 0 \text{ for all } \mathbf{y} \in S \\
&\Leftrightarrow\; \lambda(c_0 + \langle \mathbf{c}, \mathbf{y}\rangle) = \langle (c_0, \mathbf{c}), (\lambda, \lambda\mathbf{y})\rangle \geq 0 \text{ for all } \mathbf{y} \in S \text{ and all } \lambda \geq 0 \\
&\Leftrightarrow\; \langle (c_0, \mathbf{c}), (x_0, \mathbf{x})\rangle \geq 0 \text{ for all } (x_0, \mathbf{x}) \in \text{cone}(S) \\
&\Leftrightarrow\; (c_0, \mathbf{c}) \in (\text{cone}(S))^*.
\end{aligned}$$

Finally we note that both $\text{cone}(S)^*$ and $\text{cone}(S^*)$ belong to the halfspace defined by $c_0 \geq 0$, so to finish, we need only worry about points with $c_0 = 0$.

Note that if $(0, \mathbf{c})$ belongs to $\text{cone}(S)^*$, then $1 + \langle \mathbf{c}, \lambda\mathbf{y}\rangle = \lambda(1/\lambda + \langle \mathbf{c}, \mathbf{y}\rangle) \geq 0$ for all $\mathbf{y} \in S$ and $\lambda \geq 0$. In particular, then $\epsilon + \langle \mathbf{c}, \mathbf{y}\rangle$ is also nonnegative for $\mathbf{y} \in S$ for $\epsilon > 0$. Then $(\epsilon, \mathbf{c})$ belongs to $\text{cone}(S)^*$ for every $\epsilon > 0$ and by the arguments above also $\text{cone}(S^*)$. Taking ϵ to zero then shows that $(0, \mathbf{c})$ belongs to the closure of this cone. Similarly, if $(0, \mathbf{c})$ belongs to the closure of $\text{cone}(S^*)$, then it is a limit of points in this cone with nonzero first coordinate, which must also belong to $\text{cone}(S)^*$. By definition, $\text{cone}(S)^*$ is closed, so it also contains these points. □

In particular, this shows that to understand duality of convex sets (up to issues of closure), it suffices to understand duality of convex cones, which is in many ways simpler. If K is a convex cone, then it is invariant under multiplication by nonnegative scalars. Then $\mathbf{c}$ belongs to K^* if and only if $1 + \langle \mathbf{c}, \lambda\mathbf{x}\rangle = 1 + \langle \lambda\mathbf{c}, \mathbf{x}\rangle$ is nonnegative for all $\mathbf{x} \in K$ and $\lambda \geq 0$. From this, we see that K^* is also a convex cone, known as the *dual cone* of K, and

$$K^* \;=\; \{\mathbf{c} \in \mathbb{R}^n : \langle \mathbf{c}, \mathbf{x}\rangle \geq 0 \text{ for all } \mathbf{x} \in K\}.$$

One can check that the following hold for any closed convex cones $C, K \subseteq \mathbb{R}^n$:

$$(C^*)^* = C \quad \text{and} \quad (C \cap K)^* = \overline{C^* + K^*}. \tag{3.1}$$

More formally, given a convex cone K in a real vector space V, its dual is the set of linear functions on V that are nonnegative on K,

$$K^* = \{\ell \in V^* : \ell(\mathbf{x}) \geq 0 \text{ for all } \mathbf{x} \in K\}.$$

Fixing an inner product identifies the vector space V with its dual V^*. For example, under the usual inner product $\langle \mathbf{x}, \mathbf{y}\rangle$ on $\mathbb{R}^n$, the nonnegative orthant is self-dual:

$$(\mathbb{R}^n_{\geq 0})^* \;=\; \{\mathbf{y} \in \mathbb{R}^n : \langle \mathbf{x}, \mathbf{y}\rangle \geq 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n_{\geq 0}\} \;=\; \mathbb{R}^n_{\geq 0}.$$

The self-duality of the nonnegative orthant plays an important role in many algorithms for linear programming. For semidefinite programming, the cone of positive semidefinite matrices plays a similar role.

Proposition 3.2. *Under the inner product $\langle X, Y\rangle = \text{trace}(XY)$, the cone of positive semidefinite matrices is self-dual:*

$$(\mathcal{S}^N_+)^* \;=\; \{Y \in \mathcal{S}^N : \langle X, Y\rangle \geq 0 \textit{ for all } X \in \mathcal{S}^N_+\} \;=\; \mathcal{S}^N_+.$$

Proof. If Y is positive semidefinite, then it can be written as BB^T for some real matrix B. If X is also positive semidefinite, then we can check that the inner product $\langle X, BB^T\rangle = \text{trace}(XBB^T) = \text{trace}(B^TXB)$ is nonnegative, since B^TXB is positive semidefinite. On the other hand, if Y is not positive semidefinite, then there is some vector $\mathbf{v} \in \mathbb{R}^N$ for which $\mathbf{v}^TY\mathbf{v} < 0$. The rank-one matrix $X = \mathbf{v}\mathbf{v}^T$ belong to $\mathcal{S}^N_+$, but its inner product with Y is negative: $\langle \mathbf{v}\mathbf{v}^T, Y\rangle = \mathbf{v}^TY\mathbf{v} < 0$. Therefore Y does not belong to the dual cone of $\mathcal{S}^N_+$. □

3.1. Dual operations: linear restriction and projection. Another important property of convex duality is how it interacts with linear projection and restriction. Let L be a m-dimensional linear subspace of $\mathbb{R}^n$. Given a basis $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ for L, consider the maps

$$\pi_L : \mathbb{R}^n \to \mathbb{R}^m \quad \text{given by} \quad \pi_L(\mathbf{v}) = (\langle \mathbf{a}_i, \mathbf{v}\rangle)_{i=1,\ldots,m}, \quad \text{and}$$

$$\pi_L^* : \mathbb{R}^m \to \mathbb{R}^n \quad \text{given by} \quad \pi_L^*(\mathbf{x}) = \sum_{i=1}^m x_i \mathbf{a}_i.$$

Explicitly, if A is an $m \times n$ matrix with rows $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$, then these maps are defined by $\pi_L(\mathbf{v}) = A\mathbf{v}$ and $\pi_L^*(\mathbf{x}) = A^T\mathbf{x}$. Note that π_L is surjective and the image is isomorphic to L. Its kernel consists of $L^\perp = \{\mathbf{v} \in \mathbb{R}^n : \langle \mathbf{u}, \mathbf{v}\rangle = 0 \text{ for all } \mathbf{u} \in L\}$. Similarly, the pullback π_L^* is injective and its image equals the linear space L.

Any linear function on $\mathbb{R}^m$ defines a map on $\mathbb{R}^n$ via the pullback π_L^*. Explicitly, given a linear function $\mathbf{y} \mapsto \langle \mathbf{x}, \mathbf{y}\rangle$ on $\mathbb{R}^m$, its pullback is given by $\mathbf{v} \mapsto \langle \mathbf{x}, \pi_L(\mathbf{v})\rangle$. Then, as the notation suggests, this linear function sends $\mathbf{v} \in \mathbb{R}^n$ to

$$\langle \mathbf{x}, \pi_L(\mathbf{v})\rangle = \sum_{i=1}^m x_i \langle \mathbf{a}_i, \mathbf{v}\rangle = \left\langle \sum_{i=1}^m x_i \mathbf{a}_i, \mathbf{v} \right\rangle = \langle \pi_L^*(\mathbf{x}), \mathbf{v}\rangle.$$

Under our identification of $(\mathbb{R}^n)^*$ with $\mathbb{R}^n$, this linear function is identified with the point $\pi_L^*(\mathbf{x})$.

Now let $C \subseteq \mathbb{R}^n$ be a convex cone and consider its image, $\pi_L(C)$ in $\mathbb{R}^m$. We would like to relate the dual cone of $\pi_L(C)$ to that of C. A linear function $\langle \mathbf{x}, \mathbf{y}\rangle$ is nonnegative for all $\mathbf{y} \in \pi_L(C)$ if and only if its pullback $\mathbf{v} \mapsto \langle \mathbf{x}, \pi_L(\mathbf{v})\rangle = \langle \pi_L^*(\mathbf{x}), \mathbf{v}\rangle$ is nonnegative for all $\mathbf{v} \in C$. This happens if and only if $\pi_L^*(\mathbf{x})$ belongs to the dual cone of C. See Figure 8. All together this gives that

$$(\pi_L(C))^* = \{\mathbf{x} \in \mathbb{R}^m : \pi_L^*(\mathbf{x}) \in C^*\} \quad \cong \quad C^* \cap L. \tag{3.2}$$

The last equality here uses that π_L^* is a linear isomorphism between $\mathbb{R}^m$ and L. This shows that intersection with a linear space is the dual operation to linear projection.

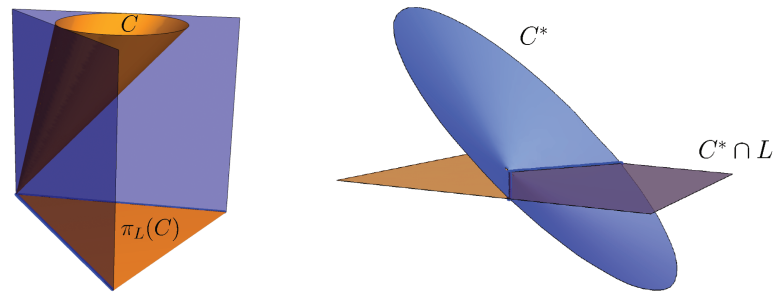

FIGURE 8. Nonnegative linear functions on a projection $\pi_L(C)$ come from nonnegative linear functions C that are constant along preimages. These constitute a linear slice of the dual cone C^*.

Recall that dualizing any convex cone K twice results in its closure, $(K^*)^* = \overline{K}$. Applying the dual to both sides of the equation above gives that

$$\overline{\pi_L(C)} = (\{\mathbf{x} \in \mathbb{R}^m : \pi_L^*(\mathbf{x}) \in C^*\})^* \quad \cong \quad (C^* \cap L)^*.$$

One way to interpret this statement is that the dual cone of a linear slice $C^* \cap L$ is isomorphic to the dual cone $C = (C^*)^*$ under projection. All together this gives the following:

PROPOSITION 3.3. *The dual cone of the image of C under linear map is linearly isomorphic to a linear section of C^*, namely*

$$(\pi_L(C))^* \quad \cong \quad C^* \cap L.$$

The dual cone of a linear section of C is the image of C^* under a linear map:

$$(C \cap L)^* \quad \cong \quad \overline{\pi_L(C^*)}.$$

Here $(C \cap L)^*$ is the convex dual of $(C \cap L)^*$ in the ambient space L. In short, projection and slicing are dual operations.

For the study of spectrahedra, an important special case is the cone of positive semidefinite matrices $C = C^* = \mathcal{S}_+^N$. If $\mathcal{L}$ is a linear space spanned by matrices $A_1, \ldots, A_m \in \mathcal{S}^N$, then we can consider the linear map

$$\pi_{\mathcal{L}} : \mathcal{S}^N \to \mathbb{R}^m \quad \text{given by } X \mapsto (\langle A_1, X\rangle, \ldots, \langle A_m, X\rangle).$$

A function $\ell(\mathbf{y}) = \langle \mathbf{x}, \mathbf{y}\rangle$ represented by $\mathbf{x} \in \mathbb{R}^m$ is nonnegative on $\pi_{\mathcal{L}}(\mathcal{S}_+^N)$ if and only if the function on $\mathcal{S}^N$ given by $X \mapsto \langle \pi_{\mathcal{L}}^*(\mathbf{x}), X\rangle = \langle \sum_{i=1}^m x_i A_i, X\rangle$ is nonnegative on $\mathcal{S}_+^N$. This happens exactly when the matrix $\sum_{i=1}^m x_i A_i$ is positive semidefinite. Therefore the dual cone of the projection is given by

$$\pi_{\mathcal{L}}(\mathcal{S}_+^N)^* \;=\; \left\{ \mathbf{x} \in \mathbb{R}^m \;:\; \sum_{i=1}^m x_i A_i \succeq 0 \right\} \;\cong\; \mathcal{S}_+^N \cap \mathcal{L}. \tag{3.3}$$

The image of $\mathcal{S}_+^N$ under the projection $\pi_{\mathcal{L}}$ may not be closed. For example, consider the image of $\mathcal{S}_+^2$ under the map $X \mapsto (X_{11}, X_{12})$. This contains points of the form $(X_{11}, X_{12}) = (\epsilon, 1)$ for every $\epsilon > 0$, coming from matrices with $X_{22} > 1/\epsilon$. However the point $(0, 1)$ does not belong to this image, as there is no positive semidefinite matrix X with $X_{11} = 0$ and $X_{12} > 0$.

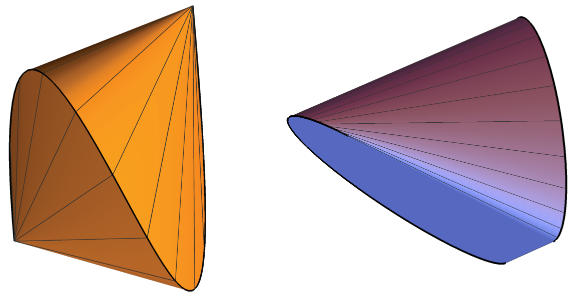

FIGURE 9. A spectrahedron and its dual convex body

However for any convex (not necessarily closed) cone $K \subset \mathbb{R}^n$, the dual of its dual equals its closure, i.e. $(K^*)^* = \overline{K}$. Taking the dual of the cones in (3.3), we find that the dual of a spectrahedral cone is the image of $\mathcal{S}_+^N$ under a linear map:

$$\overline{\pi_{\mathcal{L}}(\mathcal{S}_+^N)} \;=\; \left(\left\{\mathbf{x} \in \mathbb{R}^m \;:\; \sum_{i=1}^{m} x_i A_i \succeq 0\right\}\right)^* \;\cong\; (\mathcal{S}_+^N \cap \mathcal{L})^*. \tag{3.4}$$

It is also instructive to apply Proposition 3.3 to spectrahedral cones C and their images under linear maps. In particular, unlike spectrahedra, the class of spectrahedral shadows is closed under duality.

PROPOSITION 3.4. *The dual of a spectrahedral shadow is (the closure of) a spectrahedral shadow.*

PROOF. By Proposition 3.1, it suffices to work with a spectrahedral cone, i.e. the intersection of $\mathcal{S}_+^N$ with a linear space (rather than an affine linear space).

Consider the affine space $\mathcal{L}$ spanned by matrices $A_1, \ldots, A_n$ and the spectrahedral cone $C = \{\mathbf{x} \in \mathbb{R}^n \text{ s.t. } \sum_{i=1}^n A_i \succeq 0\} \cong \mathcal{S}_+^N \cap \mathcal{L}$. For $m < n$, we can examine the image of C under the projection $\pi_m : \mathbb{R}^n \to \mathbb{R}^m$ given by $\pi_m(\mathbf{x}) = (x_1, \ldots, x_m)$. The pullback $\pi_m^* : \mathbb{R}^m \to \mathbb{R}^n$ is given by $\pi_m^*(\mathbf{y}) = (\mathbf{y}, \mathbf{0})$. Then by (3.2),

$$(\pi_m(C))^* = \{\mathbf{y} \in \mathbb{R}^m : (\mathbf{y}, \mathbf{0}) \in C^*\}.$$

The dual cone C^* is the projection of the cone of positive semidefinite matrices: $C^* = \overline{\pi_{\mathcal{L}}(\mathcal{S}_+^N)} = \overline{\{(\langle X, A_i\rangle)_{i=1,\ldots,n} : X \in \mathcal{S}_+^N\}}$. Putting this together, we find that

$$(\pi_m(C))^* = \overline{\{(\langle X, A_i\rangle)_{i=1,\ldots,m} : X \in \mathcal{S}_+^N, \langle X, A_j\rangle = 0 \text{ for } j = m+1, \ldots, n\}}.$$

This is the image of the spectrahedron $\{X \in \mathcal{S}_+^N : \langle X, A_j\rangle = 0 \text{ for } j = m+1, \ldots, n\}$ under the linear map $X \mapsto (\langle X, A_i\rangle)_{i=1,\ldots m}$ up to closure. □

EXAMPLE 1.3 CONT'D. Let $A_0, A_1, A_2, A_3 \in \mathcal{S}^4$ be matrices with $(A_k)_{ij} = 1$ if $|i-j| = k$ and 0 otherwise for every $k = 0, 1, 2, 3$. The cone over the spectrahedron $C = \text{cone}(S)$ from Example 1.3 can be written as the intersection of the positive semidefinite cone with the linear space spanned by these matrices

$$C \;=\; \{(w, x, y, z) \in \mathbb{R}^4 : wA_0 + xA_1 + yA_2 + zA_3 \succeq 0\} \;\cong\; \mathcal{S}_+^4 \cap \mathcal{L}$$

where $\mathcal{L} = \text{span}_{\mathbb{R}}\{A_0, A_1, A_2, A_3\}$. Since the cone $\mathcal{S}_+^4$ is self-dual, the dual cone of C is given by the projection of $\mathcal{S}_+^4$ onto this linear space:

$$C^* \;=\; \{(\langle A_0, X\rangle, \langle A_1, X\rangle, \langle A_2, X\rangle, \langle A_3, X\rangle) \in \mathbb{R}^4 : X \in \mathcal{S}_+^4\} \;\cong\; \pi_{\mathcal{L}}(\mathcal{S}_+^4).$$

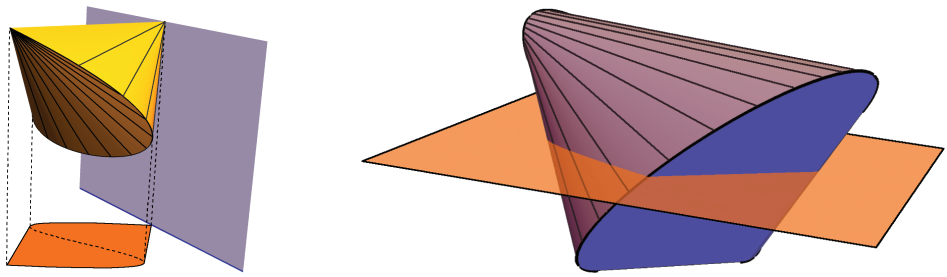

FIGURE 10. A projection of the spectrahedron from Example 1.3 and corresponding slice of its dual convex body.

To recover the original set S and its convex dual set S^*, we can restrict both sets to have first coordinate one, giving the convex sets

$$\begin{aligned} S &= \left\{(x,y,z) \in \mathbb{R}^3 : A_0 + xA_1 + yA_2 + zA_3 \succeq 0\right\} \text{ and} \\ S^* &= \left\{(\langle A_1, X\rangle, \langle A_2, X\rangle, \langle A_3, X\rangle) \in \mathbb{R}^3 : X \in \mathcal{S}^4_+, \langle A_0, X\rangle = 1\right\}, \end{aligned}$$

seen in Figure 9. Taking things one step further, we can consider the projection of S on to the coordinates (x, z),

$$\pi(S) = \left\{(x,z) \in \mathbb{R}^2 : \exists\, y \in \mathbb{R} \text{ with } A_0 + xA_1 + yA_2 + zA_3 \succeq 0\right\}$$

as shown in Figure 10.

To understand the dual convex body, we consider affine linear functions on $\mathbb{R}^3$ that come from affine linear functions on $\mathbb{R}^2$. These are exactly the functions $(x, y, z) \mapsto 1 + ax + by + cz$ where $(a, b, c) \in \mathbb{R}^3$ and $b = 0$. This lets us write the dual convex body as the intersection of S^* with the plane $b = 0$, giving

$$(\pi(S))^* = \left\{(\langle A_1, X\rangle, \langle A_3, X\rangle) \in \mathbb{R}^3 : X \in \mathcal{S}^4_+, \langle A_0, X\rangle = 1, \langle A_2, X\rangle = 0\right\}.$$

This is also the image of a spectrahedron under a linear map, namely the image of the spectrahedron $\{X \in \mathcal{S}^4_+ : \langle A_0, X\rangle = 1, \langle A_2, X\rangle = 0\}$ under the linear map $X \mapsto (\langle A_1, X\rangle, \langle A_3, X\rangle)$. Indeed, the self-duality of the cone of positive semidefinite matrices means that the class of projected spectrahedra is closed under duality. $\diamond$

One consequence of this section is that any linear function on the projection of a spectrahedron can be lifted to a linear function on the original spectrahedron. Therefore optimizing a linear function over a spectrahedral shadow can also be phrased as a semidefinite program, which we discuss in Section 3.2.

3.2. Convex optimization and semidefinite programming. A *conic optimization problem* has the form

$$\min_{\mathbf{x} \in K} \langle \mathbf{c}, \mathbf{x}\rangle \quad \text{such that} \quad \langle \mathbf{a}_i, \mathbf{x}\rangle = b_i \quad \text{for} \quad i = 1, \dots, m,$$

where K is a convex cone in $\mathbb{R}^n$, $\mathbf{c}, \mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$ and $b_1, \dots, b_m \in \mathbb{R}$. We call a point $\mathbf{x}$ *feasible* if it satisfies the constraints $\mathbf{x} \in K$ and $\langle \mathbf{a}_i, \mathbf{x}\rangle = b_i$ for all i. Many optimization problems can be reformulated as conic optimization problems. One benefit of this approach is that most of the complexity of the problem is packaged into the convex cone, rather than both the feasible region and objective function.

When the cone K is the nonnegative orthant, the feasible region $\{\mathbf{x} \in \mathbb{R}^n_{\geq 0} : \langle \mathbf{a}_i, \mathbf{x}\rangle = b_i \text{ for } i = 1, \dots, m\}$, is a *polyhedron* and corresponding conic optimization problem is called a *linear program*. When K is the cone of positive semidefinite matrices, the feasible region is a *spectrahedron* and the convex program is called a *semidefinite program*.

DEFINITION 3.5. A *semidefinite program* is the problem of minimizing a linear function over a spectrahedron, that is, a problem of the form

$$\min_{X\in\mathcal{S}^N_+}\langle C,X\rangle \quad \text{subject to } \langle A_i,X\rangle = b_i \text{ for } i=1,\ldots,m,$$

where $C, A_1,\ldots,A_m\in\mathcal{S}^N$ and $b_1,\ldots,b_m\in\mathbb{R}$.

The feasible region, $\{X\in\mathcal{S}^N_+ : \langle A_i,X\rangle = b_i \text{ for } i=1,\ldots,m\}$, of a semidefinite program is a spectrahedron, as it is the intersection of the positive semidefinite cone with the affine linear space of matrices satisfying the constraints.

Semidefinite programs are useful because of their ability to model a large range of problems and because they can be solved efficiently using interior point methods. See [**18**] for more details on algorithms and applications.

EXAMPLE 1.3 CONT'D. Let E_{ij} denote the 4×4 real symmetric matrix with (i,j)th and (j,i)th entry 1 and all other entries 0. The spectrahedron S given by the set of matrices $X\in\mathcal{S}^4_+$ satisfying

$$\langle E_{ii},X\rangle = 1 \text{ for all } i \text{ and } \langle E_{12}-E_{23},X\rangle = \langle E_{23}-E_{34},X\rangle = \langle E_{13}-E_{24},X\rangle = 0$$

is exactly the spectrahedron given in parametrized form in Example 1.3. Consider the matrix $C = 3E_{12}-E_{13}$. Minimizing the linear function $\langle C,X\rangle$ over S is equivalent to minimizing the linear function $6x-2z$ over

$$\left\{(x,y,z)\in\mathbb{R}^3 \;:\; \begin{pmatrix}1 & x & y & z\\ x & 1 & x & y\\ y & x & 1 & x\\ z & y & x & 1\end{pmatrix}\succeq 0\right\}.$$

The minimum equals $-4\sqrt{2}$, which is achieved by the point $(x,y,z) = (-\frac{1}{\sqrt{2}},0,\frac{1}{\sqrt{2}})$, at which point the matrix above is positive semidefinite and has rank two. $\diamond$

More generally, the optimal values of semidefinite programs are semialgebraic functions of the inputs, whose degrees are studied in [**11**, **14**].

Another fundamental benefit of conic programming is the ability to easily formulate a "dual problem" that provides lower bounds on the optimal value of the original problem (often called the *primal* problem):

$$\text{(Primal)} \qquad \min_{\mathbf{x}\in K}\langle \mathbf{c},\mathbf{x}\rangle \qquad \text{s.t.} \qquad \langle \mathbf{a}_i,\mathbf{x}\rangle = b_i \;\text{ for }\; i=1,\ldots,m$$

$$\text{(Dual)} \qquad \max_{\mathbf{y}\in\mathbb{R}^m}\sum_{i=1}^m b_iy_i \qquad \text{s.t.} \qquad \mathbf{c}-\sum_{i=1}^m \mathbf{a}_iy_i \;\in\; K^*$$

Here K^* denotes the dual cone of K.

Note that $\{\mathbf{c}-\sum_{i=1}^m \mathbf{a}_iy_i : \mathbf{y}\in\mathbb{R}^m\}$ parametrizes an affine linear space. So the dual problem still has form of optimizing a linear function over the intersection of a convex cone and an affine space.

THEOREM (Weak duality and complementary slackness). For any primal feasible $\mathbf{x}$ and dual feasible $\mathbf{y}$,

$$\langle \mathbf{c}, \mathbf{x} \rangle \;\geq\; \sum_{i=1}^{m} b_i y_i.$$

Moreover, if $\langle \mathbf{c}, \hat{\mathbf{x}} \rangle = \sum_{i=1}^{m} b_i \hat{y}_i$ for feasible points $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ then both are optimal.

PROOF. Suppose $\mathbf{x}, \mathbf{y}$ are feasible. Then

$$\begin{aligned} \langle \mathbf{c}, \mathbf{x} \rangle - \sum_{i=1}^{m} b_i y_i \;&=\; \langle \mathbf{c}, \mathbf{x} \rangle - \sum_{i=1}^{m} \langle \mathbf{a}_i, \mathbf{x} \rangle y_i \\ &=\; \langle \mathbf{c} - \sum_{i=1}^{m} y_i \mathbf{a}_i, \mathbf{x} \rangle \;\geq\; 0. \end{aligned}$$

Here the last inequality follows from the fact that $\mathbf{c} - \sum_{i=1}^{m} y_i \mathbf{a}_i \in K^*$ and $\mathbf{x} \in K$.

If $\langle \mathbf{c}, \hat{\mathbf{x}} \rangle = \sum_{i=1}^{m} b_i \hat{y}_i$, then $\langle \mathbf{c}, \hat{\mathbf{x}} \rangle$ is an upper bound for $\sum_i b_i y_i$ over all feasible $\mathbf{y}$. Since $\hat{\mathbf{y}}$ achieves this upper bound, it must be optimal. Similarly, $\sum_i b_i \hat{y}_i$ is a lower bound for $\langle \mathbf{c}, \mathbf{x} \rangle$, which is achieved by $\hat{\mathbf{x}}$. □

4. Facial structure, extreme points and ranks

In this section, we discuss in detail the facial structure and extreme points of spectrahedra. We also explore this theory in more detail for the Gram spectrahedron associated to a nonnegative univariate polynomial. To do this, we first need to understand the facial structure of $\mathcal{S}_+^N$.

4.1. Faces of the PSD cone. The facial structure of $\mathcal{S}_+^N$ is governed by linear subspaces parametrizing their kernels. This is discussed at length in [**1**, §II.12] and we go over some of the relevant proofs. In particular, we show the following bijection between linear subspaces of $\mathbb{R}^N$ and faces of the convex cone $\mathcal{S}_+^N$.

THEOREM 4.1. *For every linear subspace $L \subseteq \mathbb{R}^N$,*

$$\mathcal{F}_L \;=\; \left\{ A \in \mathcal{S}_+^N \;:\; L \subseteq \ker(A) \right\}$$

is a face of the cone $\mathcal{S}_+^N$. Moreover every face $\mathcal{S}_+^N$ has this form.

To prove this theorem, we need the following fundamental facts about positive semidefinite matrices.

LEMMA 4.2. *For $A, B \in \mathcal{S}_+^N$ and $\mathbf{v} \in \mathbb{R}^N$,*

(i) $\mathbf{v}^T A \mathbf{v} = 0$ *if and only if* $A\mathbf{v} = 0$,
(ii) $\langle A, B \rangle = 0$ *if and only if* $\operatorname{rowspan}(B) \subseteq \ker(A)$, *and*
(iii) $\ker(A + B) = \ker(A) \cap \ker(B)$.

PROOF. For (i), note that we can write the matrix A as a sum of rank-one matrices, $A = \sum_{i=1}^{d} \mathbf{u}_i \mathbf{u}_i^T$. Then

$$\mathbf{v}^T A \mathbf{v} \;=\; \sum_{j=1}^{d} \mathbf{v}^T \mathbf{u}_i \mathbf{u}_i^T \mathbf{v} \;=\; \sum_{i=1}^{d} (\langle \mathbf{u}_i, \mathbf{v} \rangle)^2.$$

As a sum of nonnegative terms, this equals zero if and only if each term $(\langle \mathbf{u}_i, \mathbf{v} \rangle)^2$ is zero. In particular, if $\mathbf{v}^T A \mathbf{v}$ is zero, then $\langle \mathbf{u}_i, \mathbf{v} \rangle = \mathbf{u}_i^T \mathbf{v}$ is zero for each i, implying that $A\mathbf{v} = \sum_i \mathbf{u}_i \mathbf{u}_i^T \mathbf{v}$ is zero. The other direction is clear.

(ii) Similarly if B is positive semidefinite, we can write $B = \sum_{j=1}^{e} \mathbf{v}_j \mathbf{v}_j^T$. Then

$$\langle A, B\rangle \;=\; \sum_{i=1}^{e} \langle A, \mathbf{v}_j \mathbf{v}_j^T\rangle \;=\; \sum_{j=1}^{e} \mathbf{v}_j^T A \mathbf{v}_j.$$

Since A is positive semidefinite, each term $\mathbf{v}_j^T A \mathbf{v}_j$ is nonnegative. The sum is zero if and only if each summand is. By part (i), this occurs if and only if $\mathbf{v}_j \in \ker(A)$ for all j. This shows that the rowspan of B is contained in the kernel of A.

Finally, for (iii), note that any vector in the kernel of both A and B is automatically contained in the kernel of their sum. For the reverse inclusion, suppose that $\mathbf{v}$ belongs to $\ker(A+B)$. In particular,

$$\mathbf{v}^T(A+B)\mathbf{v} \;=\; \mathbf{v}^T A \mathbf{v} + \mathbf{v}^T B \mathbf{v} \;=\; 0.$$

Since $A, B \in \mathcal{S}_+^N$, both summands are nonnegative, implying that they are both zero. Then by (i), $\mathbf{v}$ belongs to the kernel of both A and B. $\square$

Now we are ready to prove the theorem characterizing faces of $\mathcal{S}_+^N$.

Proof of Theorem 4.1. For the first statement, suppose that $A, B \in \mathcal{S}_+^N$ and $\lambda A + \mu B \in \mathcal{F}_L$ for some $\lambda, \mu > 0$. Then

$$L \;\subseteq\; \ker(\lambda A + \mu B) \;=\; \ker(\lambda A) \cap \ker(\mu B) \;=\; \ker(A) \cap \ker(B).$$

The subspace L belongs to the kernels of both A and B, meaning that $A, B \in \mathcal{F}_L$. This shows that $\mathcal{F}_L$ is a face of $\mathcal{S}_+^N$.

Now suppose that F is any face of $\mathcal{S}_+^N$. Let $X \in F$ be a matrix of maximal rank in F and let $L = \ker(X)$. First we claim that $F \subseteq \mathcal{F}_L$. To see this, let A be any element of F. Since F is a convex cone, $A + X$ also belongs to F and so, by assumption, the rank of $A + X$ cannot be larger than that of X. In particular, the dimension of the kernel of $A + X$ cannot be smaller than the dimension of $\ker(X)$. Lemma 4.2 then implies that $\ker(A) \cap \ker(X) = \ker(X)$, meaning that the kernel of A contains $L = \ker(X)$.

To show the reverse inclusion $\mathcal{F}_L \subseteq F$, let $B \in \mathcal{F}_L$. We claim that for sufficiently large $\lambda \in \mathbb{R}_+$, the matrix $A = \lambda X - B$ is positive semidefinite. To see this, consider the linear space $L^\perp$ consisting of vectors $\mathbf{v}$ with $\langle \mathbf{u}, \mathbf{v}\rangle = 0$ for all $\mathbf{u} \in L$ and the compact set

$$S = \left\{ \mathbf{v} \in L^\perp : ||\mathbf{v}||_2 = 1 \right\}.$$

Since X is positive semidefinite, the function $\mathbf{v} \mapsto \mathbf{v}^T X \mathbf{v}$ is nonnegative on S. Moreover, its minimum over X must be strictly positive, since no vector in S can belong to $L = \ker(X)$. Let m be this minimum. Similarly, the function $\mathbf{v} \mapsto \mathbf{v}^T B \mathbf{v}$ is nonnegative and attains some maximum M on S. For $\lambda > M/m$, the function $\mathbf{v} \mapsto \mathbf{v}^T(\lambda X - B)\mathbf{v}$ is positive on S. Indeed we check that

$$\mathbf{v}^T(\lambda X - B)\mathbf{v} = \lambda(\mathbf{v}^T X \mathbf{v}) - (\mathbf{v}^T B \mathbf{v}) \geq \lambda m - M > 0.$$

By scaling, we see that this function is nonnegative for all $\mathbf{v} \in L^\perp$. Note that any vector $\mathbf{x} \in \mathbb{R}^N$ can be written as $\mathbf{u} + \mathbf{v}$ where $\mathbf{u} \in L$ and $\mathbf{v} \in L^\perp$. If $A = \lambda X - B$, then, by assumption, $\mathbf{u}$ belongs to the kernel of A. This gives that

$$\mathbf{x}^T A \mathbf{x} \;=\; \mathbf{u}^T A \mathbf{u} + 2\mathbf{v}^T A \mathbf{u} + \mathbf{v}^T A \mathbf{v} \;=\; \mathbf{v}^T A \mathbf{v} \geq 0.$$

Now we have $A, B \in \mathcal{S}_+^N$ with $A + B = \lambda X \in F$. Since F is a face of $\mathcal{S}_+^N$, it follows that both A and B belong to F. This shows $\mathcal{F}_L \subseteq F$, as desired. $\square$

COROLLARY 4.3. Every face of $\mathcal{S}_+^N$ is exposed.

PROOF. By the above theorem, it suffices to show that for any subspace $L \subset \mathbb{R}^N$, the face $\mathcal{F}_L$ is exposed. Let B be a positive semidefinite matrix whose rowspan equals L. To construct such a matrix, we can, for example, take a set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_d$ spanning L and consider the positive semidefinite matrix $B = \sum_{i=1}^d \mathbf{v}_i \mathbf{v}_i^T$. By Lemma 4.2, a matrix $A \in \mathcal{S}_+^N$ satisfies $\langle A, B\rangle = 0$ if and only if $\operatorname{rowspan}(B) = L \subseteq \ker(A)$. This shows that the face $\mathcal{F}_L$ is exposed by the linear function $\langle \cdot, B\rangle$. $\square$

This theorem shows that every face of $\mathcal{S}_+^N$ is linearly isomorphic to $\mathcal{S}_+^k$ for some $k \leq N$. Given a linear space $L \subseteq \mathbb{R}^N$ of dimension d, consider an invertible $N \times N$ matrix U whose last d columns span L. That is, for any vector $\mathbf{x}$ of the form $(\mathbf{0}, \mathbf{v})$ with $\mathbf{v} \in \mathbb{R}^d$, $U\mathbf{v} \in L$. Then $X \mapsto U^T X U$ is an invertible linear transformation on $\mathcal{S}^N$ with the property that

$$A \in \mathcal{F}_L \quad \Leftrightarrow \quad U^T A U = \begin{pmatrix} \tilde{A} & 0 \\ 0 & 0 \end{pmatrix} \text{ for some } \tilde{A} \in \mathcal{S}_+^{N-d}.$$

This shows that $\mathcal{F}_L$ is linearly isomorphic to the cone $\mathcal{S}_+^{N-d}$.

4.2. Faces of spectrahedra. Many nice properties of $\mathcal{S}_+^N$ are inherited by spectrahedra. The first of which is that faces are characterized by kernels.

THEOREM 4.4. *Every face of a spectrahedron $S \subset \mathcal{S}_+^N$ has the form*

$$F \;=\; \{A \in S \;:\; L \subseteq \ker(A)\}$$

for some linear space $L \subseteq \mathbb{R}^N$. That is, $F = S \cap \mathcal{F}_L$, as in Theorem 4.1.

Here we need to refer to the *relative interior* of a convex set. Given a convex set $S \subset \mathbb{R}^n$, the *relative interior* of S, denoted $\operatorname{ri}(S)$, is the interior of S within its affine span. By [4, Theorem I.5.6], given a closed convex set S and a face F containing a point $\mathbf{x}$, F is the smallest face of S containing $\mathbf{x}$ if and only if $\mathbf{x}$ belongs to the relative interior of F.

LEMMA 4.5. *Given a closed convex set S and an affine linear space L in $\mathbb{R}^n$, every face F of $S \cap L$ has the form $\tilde{F} \cap L$ for some face $\tilde{F}$ of S. Moreover if $\tilde{F}$ is an exposed face of S then $\tilde{F} \cap L$ is an exposed face of $S \cap L$.*

PROOF. One can check that for every face $\tilde{F}$ of S, the intersection $\tilde{F} \cap L$ is a face of $S \cap L$ and we leave this to the reader.

Let $\mathbf{x}$ be a point in the relative interior of F and let $\tilde{F}$ be the smallest face of S containing $\mathbf{x}$. Then by [4, Theorem I.5.6], $\mathbf{x}$ belongs to the relative interior of F. We claim that $\mathbf{x}$ then belongs to the relative interior of $\tilde{F} \cap L$. Let $\mathbf{v}$ be a vector so that $\mathbf{x} + \lambda\mathbf{v}$ belongs to the affine span of $\tilde{F} \cap L$ for all $\lambda \in \mathbb{R}$. It suffices to show that for sufficiently small $\epsilon > 0$, $\mathbf{x} + \epsilon\mathbf{v}$ belongs to $\tilde{F} \cap L$. By assumption, it belongs to L. Moreover, the affine span of $\tilde{F} \cap L$ is contained in that of $\tilde{F}$, so $\mathbf{x} + \lambda\mathbf{v}$ belongs to the affine span of $\tilde{F}$ for all $\lambda \in \mathbb{R}$. Since $\mathbf{x}$ belongs to the relative interior of $\tilde{F}$, there is some $\epsilon > 0$ so that $\mathbf{x} + \epsilon\mathbf{v} \in \tilde{F}$, which proves the claim.

Since $\mathbf{x}$ belongs to both the relative interiors of the faces F and $\tilde{F} \cap L$ of $S \cap L$, they are both the smallest face of S containing $\mathbf{x}$ and therefore must be equal. Finally, if $\ell : \mathbb{R}^n \to \mathbb{R}$ is a linear function exposing the face $\tilde{F}$ of S, then the same linear function exposes the face F of $S \cap L$. $\square$

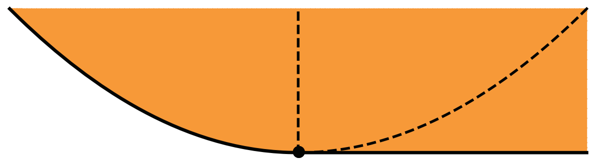

FIGURE 11. A non-exposed face of $\{(a,b) \in \mathbb{R}^2 : t^4 + at^2 + b \in P_{1,\leq 4}\}$.

PROOF OF THEOREM 4.4. A spectrahedron has the form $S = \mathcal{L} \cap \mathcal{S}_+^N$ for some affine linear space $\mathcal{L} \subset \mathcal{S}^N$. By Lemma 4.5, any face of S is the intersection of a face of $\mathcal{S}_+^N$ with $\mathcal{L}$. The characterization in Theorem 4.1 then completes the proof. $\square$

An immediate corollary of this is the theorem of Ramana and Goldman [13] that spectrahedra are facially exposed:

COROLLARY 4.6. Every face of a spectrahedron is exposed.

One interesting application is that the cone of nonnegative univariate polynomials of degree ≤ 4 is not a spectrahedron.

EXAMPLE 4.7. See also [2, Example 3.13]. Consider the convex cone $P_{1,\leq 4}$ of polynomials $f \in \mathbb{R}[t]$ of degree ≤ 4 that are globally nonnegative. We can show that $P_{1,\leq 4}$ is not a spectrahedron by exhibiting an affine linear section of it with a non-exposed face. Indeed, consider the affine linear space L of polynomials of $\mathbb{R}[t]_{\leq 4} \cong \mathbb{R}^5$ that are monic of degree 4, with no odd degree terms. That is $L = \{t^4 + at^2 + b : a, b \in \mathbb{R}\}$. Note that $f(t) = t^4 + at^2 + b$ is nonnegative for all $t \in \mathbb{R}$ if and only if the quadratic polynomial $p(t) = t^2 + at + b$ is nonnegative on $\mathbb{R}_{\geq 0}$. This happens if and only if $a, b \geq 0$ or $a^2 \leq 4b$. The point $(a,b) = (0,0)$ is a non-exposed face of the set $\{(a,b) : a,b \geq 0 \text{ or } a^2 \leq 4b\} \cong L \cap P_{1,\leq 4}$, seen in Figure 11. By Lemma 4.5, the polynomial $f = t^4$ corresponding to this point belongs to a non-exposed face of the convex cone $P_{1,\leq 4}$. Therefore $P_{1,\leq 4}$ cannot be written as a spectrahedron.

4.3. The Pataki range. The minimum of a linear function over a compact, convex set is always achieved at an extreme point. When this set is a spectrahedron, then the extreme point will be associated to a matrix of some rank. Such a point belongs to the boundary, and so this matrix necessarily drops rank. In fact it must satisfy much stronger inequalities. For spectrahedra obtained from generic affine linear spaces, the ranks must belong to what is often called the *Pataki range*, due to its introduction by Gabor Pataki in [12].

THEOREM 4.8 (Pataki range). *Let $A_0, \ldots, A_m \in \mathcal{S}^N$ be linearly independent over $\mathbb{R}$ and consider the spectrahedron $S = \{\mathbf{x} \in \mathbb{R}^m : A(\mathbf{x}) \succeq 0\}$ where $A(\mathbf{x}) = A_0 + \sum_{i=1}^m x_i A_i$. If $\hat{\mathbf{x}}$ is an extreme point of S and r is the rank of $A(\hat{\mathbf{x}})$, then*

$$\binom{r+1}{2} + m \leq \binom{N+1}{2}.$$

Furthermore, for generic $A_0, \ldots, A_m$, additionally $m \geq \binom{N-r+1}{2}$.

PROOF. Suppose that $\hat{\mathbf{x}}$ is an extreme point of S and r is the rank of $A(\hat{\mathbf{x}})$. Let L denote the kernel of $A(\hat{\mathbf{x}})$, which has dimension $N - r$. Then the matrix

$A(\hat{\mathbf{x}})$ belongs to the relative interior of the face $\mathcal{F}_L$ of $\mathcal{S}^N_+$. The intersection of $\mathcal{F}_L$ with the affine space parametrized by $A(\mathbf{x})$ is a face of S containing $A(\hat{\mathbf{x}})$, meaning that it must be just one point. Since $\mathcal{F}_L$ is linearly isomorphic to $\mathcal{S}^r_+$, a dimension count then shows that

$$\begin{aligned} m \;=\; \dim(\{A(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^m\}) \;&\leq\; \operatorname{codim}(\mathcal{F}_L) \\ &=\; \dim(\mathcal{S}^N) - \dim(\mathcal{F}_L) \;=\; \binom{N+1}{2} - \binom{r+1}{2}. \end{aligned}$$

For the second statement, note that the set $\mathcal{V}_r$ of matrices of rank $\leq r$ is a real variety of codimension $\binom{N-r+1}{2}$ in $\mathcal{S}^N$. In particular, a generic affine space of dimension $m < \binom{N-r+1}{2}$ does not intersect $\mathcal{V}_r$. Therefore for generic $A_0, \ldots, A_m$ where $m < \binom{N-r+1}{2}$, there is no $\mathbf{x} \in \mathbb{R}^m$ for which $A(\mathbf{x})$ has rank $\leq r$. □

For $N = 4$ and $m = 3$, this states that the rank r of any extreme point satisfies $\binom{r+1}{2} \leq \binom{4+1}{2} - 3 = 7$. This holds for $r = 1, 2, 3$, which tells us only that the 4×4 matrix is not full rank. For generic matrices, we find additionally that $3 \geq \binom{5-r}{2}$, which implies that $r \geq 2$. In this case the Pataki range is $\{2, 3\}$.

For higher m, the upper bound on r becomes more restrictive. If $N = 4$ and $m = 5$, then $\binom{r+1}{2} \leq \binom{4+1}{2} - 5 = 5$, giving $r = 1, 2$. In this case, there are *never* extreme points of corank one. Indeed, an asymptotic view shows that this is often the case. The dimension of $\mathcal{S}^N$ as a real vector space is $\binom{N+1}{2}$, which is the natural upper bound on m. If we consider $m \approx \frac{1}{2}\binom{N+1}{2}$, then the given upper bound on $\binom{r+1}{2}$ is $\approx \frac{1}{2}\binom{N+1}{2}$, which translates to an upper bound on r of $\approx N/\sqrt{2}$. Therefore the top rank of an extreme point can be much less than $N - 1$.

Example 1.3 cont'd. For our running example, we have $N = 4$ and $m = 3$. There is a curve segment worth of extreme points, $\mathbf{x} = (\cos(\theta), \cos(2\theta), \cos(3\theta))$ for $\theta \in [0, \pi]$. In the interior of this curve, for $\theta \in (0, \pi)$, the corresponding matrix has rank two and at the two endpoints, $\theta \in \{0, \pi\}$, the matrix has rank one.

These ranks satisfy Pataki's inequalities, but we see that their behavior is non-generic. In particular, there are extreme points of rank one and no extreme points of rank three. Taking a wider view though, this example can provide a good mental image of spectrahedra in higher dimensions – most of the points on the boundary are not extreme points and the spectrahedron is the convex hull of a low-dimensional part of its boundary. However, it still exhibits infinite families of faces. ◇

It is not difficult to show that each rank in the Pataki range appears as the rank of extreme points for some full-dimensional set of matrices $A_0, \ldots, A_m \in \mathcal{S}^N$. However examples achieving all of the ranks simultaneously were only discovered recently by Claus Scheiderer [**17**]. These examples come from *Gram spectrahedra* of sums of squares, which provide a fruitful class of examples on which to explore Pataki's inequalities.

4.4. Gram spectrahedra of univariate polynomials. Let $f \in \mathbb{R}[t]_{\leq 2d}$ be a positive polynomial of degree at most $2d$, and consider the spectrahedron

$$\operatorname{Gram}_+(f) \;=\; \left\{ G \in \mathcal{S}^{d+1}_+ \;:\; f = \begin{pmatrix} 1 & t & t^2 & \ldots & t^d \end{pmatrix} G \begin{pmatrix} 1 & t & t^2 & \ldots & t^d \end{pmatrix}^T \right\}.$$

As first noted by Choi, Lam, and Reznick [**5**], this spectrahedron provides a concise way of encapsulating the representations of the polynomial f as a sum of squares. This will be explored in greater detail in Section 5.

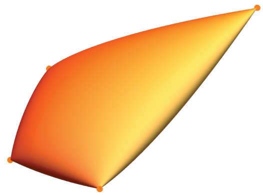

FIGURE 12. The Gram spectrahedron of $1 + t^6$ from Example 4.10.

As shown in [6], the ranks of the extreme points of $\mathrm{Gram}_+(f)$ are exactly the ranks belonging to the corresponding Pataki range. Each of the $2d+1$ coefficients of f impose linearly independent conditions on the matrix G. If the spectrahedron $\mathrm{Gram}_+(f)$ contains a positive definite matrix, then it has codimension $2d+1$ in $\mathcal{S}_+^{d+1}$ and dimension $m = \binom{d+2}{2} - (2d+1) = \binom{d}{2}$. For a generic positive polynomial $f \in \mathbb{R}[t]_{\leq 2d}$, the ranks r of the extreme points of $\mathrm{Gram}_+(f)$ satisfy

$$\binom{r+1}{2} \leq 2d+1 \quad \text{and} \quad \binom{d}{2} \geq \binom{d+2-r}{2}.$$

This gives that the Pataki interval is $2 \leq r \leq \lfloor(-1+\sqrt{16d+9})/2\rfloor$. The beautiful classical fact that every nonnegative polynomial in $\mathbb{R}[t]$ is a sum of *two* squares shows that the lowest rank two is always achieved. Indeed, if $f = g^2 + h^2$ where $g, h \in \mathbb{R}[t]_{\leq d}$, then we can write $g = \mathbf{m}_d^T \vec{\mathbf{g}}$ and $h = \mathbf{m}_d^T \vec{\mathbf{h}}$, where $\vec{\mathbf{g}}, \vec{\mathbf{h}} \in \mathbb{R}^{d+1}$ are the coefficient vectors of g and h, respectively, and $\mathbf{m}_d^T = \begin{pmatrix} 1 & t & t^2 & \dots & t^d \end{pmatrix}$. The positive semidefinite rank-two matrix $\vec{\mathbf{g}}\vec{\mathbf{g}}^T + \vec{\mathbf{h}}\vec{\mathbf{h}}^T$ belongs to $\mathrm{Gram}_+(f)$ and is an extreme point so long as f is not a single square.

In [17], Scheiderer shows that in fact *every* rank in the Pataki range is achieved by an extreme point of such a Gram spectrahedron.

THEOREM 4.9. [17] *Let $f \in \mathbb{R}[t]_{\leq 2d}$ be a generic, positive polynomial of degree $2d$. Then the ranks of the extreme points of $\mathrm{Gram}_+(f)$ are exactly the ranks belonging to the Pataki range.*

EXAMPLE 4.10 ($d = 3$). Consider the univariate polynomial $f = t^6 + 1$. The set of matrices $G \in \mathcal{S}^4$ for which $\mathbf{m}_3^T G \mathbf{m}_3 = f$ is an affine space of dimension $3 = \binom{5}{2} - 7$. Parametrizing this affine linear space by $(a, b, c) \in \mathbb{R}^3$ gives

$$\mathrm{Gram}_+(f) \;=\; \left\{ \begin{pmatrix} 1 & 0 & -a & -b \\ 0 & 2a & b & -c \\ -a & b & 2c & 0 \\ -b & -c & 0 & 1 \end{pmatrix} \in \mathcal{S}_+^4 : (a,b,c) \in \mathbb{R}^3 \right\}.$$

It contains four points of rank two given by $(a, b, c) = (0, 0, 0), \frac{1}{2}(1, \pm\sqrt{3}, 1), (2, 0, 2)$. The first corresponds to the expected decomposition as a sum of two squares $f = (1)^2 + (t^3)^2$. The last gives the representation $f = (1-2t^2)^2 + (2t-t^3)^2$.

The spectrahedron $\mathrm{Gram}_+(f)$ is shown in Figure 12. As this spectrahedron is not the convex hull of its four extreme points of rank-two, there must also be extreme points of higher rank. Indeed, the remainder of the boundary consists of extreme points of rank three. Therefore the spectrahedron $\mathrm{Gram}_+(f)$ achieves all ranks ($r = 2, 3$) of the Pataki range. ◇

5. Moment problems and sums of squares

One important application of the theory of semidefinite programming is to polynomial optimization, which is the problem of optimizing a polynomial function over a semialgebraic set. In many cases, this can be reformulated as a convex optimization over the convex cone of nonnegative polynomials. For example, to maximize the value of a polynomial $f(\mathbf{x})$ over $\mathbf{x} \in \mathbb{R}^n$, one could instead minimize the value of λ for which the polynomial $\lambda - f(\mathbf{x})$ belongs to the cone of nonnegative polynomials. This cone often proves intractable to work with, and so we consider a cone of "certifiably nonnegative" polynomials, namely sums of squares, in its place. The benefit (and a justification for a chapter on spectrahedra within these lecture notes) is that the cone of sums of squares is the image of the cone of positive semidefinite matrices under a linear map. Therefore one can approximate the solution to a polynomial optimization problem by solving a semidefinite program.

When these approximations are exact (i.e. when they give the correct solution for every objective function), the convex hull of the semialgebraic set being optimized over can be written as a spectrahedral shadow. This is true for many semialgebraic sets and was conjectured to hold for general semialgebraic sets in [**9**]. There are now several counterexamples to this conjecture, known as the Helton-Nie conjecture, discussed further below. These counterexamples come from various convex cones of nonnegative polynomials.

In the real vector space $\mathbb{R}[x_1, \ldots, x_n]_{\leq 2d}$ of real polynomials of degree $\leq 2d$, consider the two convex cones

$$\Sigma_{n,\leq 2d} = \left\{ \sum_{i=1}^{k} h_i^2 \ : \ k \in \mathbb{N}, \quad h_i \in \mathbb{R}[x_1, \ldots, x_n]_{\leq d} \right\}, \text{ and}$$
$$P_{n,\leq 2d} = \{ f \in \mathbb{R}[x_1, \ldots, x_n]_{\leq 2d} \ : \ f(\mathbf{p}) \geq 0 \text{ for all } \mathbf{p} \in \mathbb{R}^n \}.$$

The cone $\Sigma_{n,\leq 2d}$ of sums of squares is used as a tractable inner approximation of $P_{n,\leq 2d}$. Moreover, one can realize this as a projection of $\mathcal{S}^N_+$.

Proposition 5.1. *The convex cone $\Sigma_{n,\leq 2d}$ is the image of $\mathcal{S}^N_+$ under the map*

$$X \mapsto \mathbf{m}_d^T X \mathbf{m}_d,$$

where the entries of $\mathbf{m}_d$ form a basis for $\mathbb{R}[x_1, \ldots, x_n]_{\leq d}$ over $\mathbb{R}$ and $N = \binom{n+d}{d}$.

Proof. Suppose that $f \in \mathbb{R}[x_1, \ldots, x_n]_{\leq 2d}$ is a sum of squares $f = \sum_{i=1}^k g_i^2$ where $g_i \in \mathbb{R}[x_1, \ldots, x_n]_{\leq d}$. We can write each polynomial as $g_i = \vec{\mathbf{g}}_i^T \mathbf{m}_d$ for some real vector $\vec{\mathbf{g}}_i \in \mathbb{R}^N$. Consider the positive semidefinite matrix $G = \sum_{i=1}^k \vec{\mathbf{g}}_i \vec{\mathbf{g}}_i^T$ in $\mathcal{S}^N$. Then

$$\mathbf{m}_d^T G \mathbf{m}_d \ = \ \sum_{i=1}^{k} \mathbf{m}_d^T \vec{\mathbf{g}}_i \vec{\mathbf{g}}_i^T \mathbf{m}_d \ = \ \sum_{i=1}^{k} (\vec{\mathbf{g}}_i^T \mathbf{m}_d)^2 \ = \ \sum_{i=1}^{k} g_i^2,$$

showing that $f = \sum_{i=1}^k g_i^2$ belongs to the image of $\mathcal{S}^N_+$.

Conversely, suppose $f = \mathbf{m}_d^T G \mathbf{m}_d$ for some positive semidefinite matrix G in $\mathcal{S}^N_+$. We can write $G = \sum_{i=1}^k \vec{\mathbf{g}}_i \vec{\mathbf{g}}_i^T$ for some vectors $\vec{\mathbf{g}}_i \in \mathbb{R}^N$. Then the above equation shows that f is a sum of squares $\sum_i g_i^2$ where $g_i = \vec{\mathbf{g}}_i^T \mathbf{m}_d$. □

By definition $P_{n,\leq 2d}$ is the cone of polynomials on which all linear functions $\{\mathrm{ev}_{\mathbf{p}} : \mathbf{p} \in \mathbb{R}^n\}$ are nonnegative, where for $f \in \mathbb{R}[x_1, \ldots, x_n]_{\leq 2d}$, $\mathrm{ev}_{\mathbf{p}}(f) = f(\mathbf{p})$.

Therefore the cone $P_{n,\le 2d}$ is dual to the conical hull of the set of these point evaluations. Taking duals again, we find that

$$P^*_{n,\le 2d} = \overline{\text{conicHull}\left\{\text{ev}_{\mathbf{p}} \in \mathbb{R}[x_1,\dots,x_n]^*_{\le 2d} : \mathbf{p} \in \mathbb{R}^n\right\}}.$$

By a classical theorem of Haviland [8], this can also be seen as the set of linear functions $f \mapsto \int f\, d\mu$ obtained by integration with respect to a nonnegative Borel measure on $\mathbb{R}^n$. In order to specify this linear function, it suffices to give its values on the monomial basis $\{x_1^{\alpha_1}\cdots x_n^{\alpha_n} : \alpha \in \mathbb{N}^n, |\alpha| \le 2d\}$ of $\mathbb{R}[x_1,\dots,x_n]_{\le 2d}$. Here $|\alpha|$ denotes $\alpha_1 + \ldots + \alpha_n$. Therefore we can identify $P^*_{n,\le 2d}$ with the cone of *moments*

$$\begin{aligned} P^*_{n,\le 2d} &\cong \overline{\left\{\left(\int x_1^{\alpha_1}\cdots x_n^{\alpha_n}\; d\mu\right)_{|\alpha|\le 2d} \;:\; \mu \text{ a nonnegative Borel measure on } \mathbb{R}^n\right\}} \\ &= \overline{\text{conicHull}\left\{(p_1^{\alpha_1}\cdots p_n^{\alpha_n})_{|\alpha|\le 2d} \;:\; \mathbf{p} \in \mathbb{R}^n\right\}}. \end{aligned}$$

Since duality reverses inclusion, the dual cone to the sums of squares cone is an *outer* approximation of this cone

$$P^*_{n,\le 2d} \;\subseteq\; \Sigma^*_{n,\le 2d}.$$

Since the sums of squares cone is the image of the cone of positive semidefinite matrices under a linear map, it follows from Proposition 3.3 that $\Sigma^*_{n,\le 2d}$ is a spectrahedron. If the nonnegative measure μ is a probability measure, that is $\int 1\; d\mu = 1$, then the linear function $\ell \in P^*_{n,\le 2d}$ given by $\ell(f) = \int f d\mu$ is the *expectation* of f. Elements of the cone $\Sigma^*_{n,\le 2d}$ that take value one on the constant polynomial $1 \in \mathbb{R}[x_1,\dots,x_n]_{\le 2d}$ are sometimes called *pseudo-expectations*.

Every nonnegative polynomial in one variable is a sum of squares, which means that the dual cones $\Sigma^*_{1,\le 2d}$ and $P^*_{1,\le 2d}$ are equal.

Example 5.2. ($n = 1$, $d = 2$) Choosing the basis $\{1, t, t^2, t^3, t^4\}$ for $\mathbb{R}[t]_{\le 4}$ identifies the point evaluation ev_p for $p \in \mathbb{R}$ with the point $(1, p, p^2, p^3, p^4) \in \mathbb{R}^5$ and identifies the cone $P^*_{1,\le 2d}$ with the conic hull of these points. On the other hand, the cone $\Sigma_{1,\le 4}$ is the image of the cone $\mathcal{S}^3_+$ under the linear map

$$X \;\mapsto\; \begin{pmatrix}1 & t & t^2\end{pmatrix} X \begin{pmatrix}1 & t & t^2\end{pmatrix}^T \;=\; \sum_{k=0}^{4} \langle A_k, X\rangle t^k$$

where A_k is the real symmetric matrix with entries $(A_k)_{ij} = 1$ for $i + j - 2 = k$ and $(A_k)_{ij} = 0$ otherwise. The dual cone $\Sigma^*_{1,\le 4}$ is then linearly isomorphic to the intersection of the linear space $\text{span}_{\mathbb{R}}\{A_0,\dots,A_4\}$ and the positive semidefinite cone $\mathcal{S}^3_+$. All together this represents the conic hull of point evaluations as a spectrahedron:

$$\overline{\text{conicHull}\{(1,p,p^2,p^3,p^4) : p \in \mathbb{R}\}} \;=\; \left\{(y_0,\dots,y_4) \in \mathbb{R}^5 : \begin{pmatrix} y_0 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{pmatrix} \succeq 0\right\}.$$

$\diamond$

From the equality of $\Sigma^*_{1,\le 2d}$ and $P^*_{1,\le 2d}$, we immediately get that the convex hull of any curve parametrized by univariate polynomials is a spectrahedral shadow.

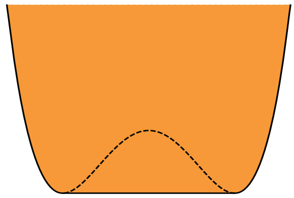

Figure 13. The convex hull of the curve in Example 5.3.

Example 5.3. Consider the curve in $\mathbb{R}^2$ parametrized by $\gamma(t) = (t, t^4 - 2t^2 + 1)$. The convex hull of this curve is shown in Figure 13. This set is not a spectrahedron, as it has a non-exposed face, but it is a spectrahedral shadow:

$$\operatorname{conv}\{\gamma(t) : t \in \mathbb{R}\} = \left\{ (a,b) \in \mathbb{R}^2 : \exists y_2, y_3 \in \mathbb{R} \text{ s.t. } \begin{pmatrix} 1 & a & y_2 \\ a & y_2 & y_3 \\ y_2 & y_3 & b + 2y_2 - 1 \end{pmatrix} \succeq 0 \right\}.$$

To see this, note that the convex hull $\operatorname{conv}\{\gamma(t) : t \in \mathbb{R}\}$ equals is image of the convex hull $\operatorname{conv}\{(1, t, t^2, t^3, t^4) : t \in \mathbb{R}\} = \Sigma^*_{1,\leq 4} \cap \{y_0 = 1\}$ under the affine linear map $(1, y_1, y_2, y_3, y_4) \mapsto (y_1, y_4 - 2y_2 + 1)$. Solving $y_1 = a$ and $b = y_4 - 2y_2 + 1$ for y_1 and y_4 gives the representation above. $\diamond$

In fact, sums of squares give a good approximation for nonnegative polynomials on any curve. Considering linear polynomials that are nonnegative on a variety as sums of squares gives the following:

Theorem 5.4 ([**15**]). *The convex hull of any real algebraic curve is a spectrahedral shadow.*

This was originally evidence for the Helton-Nie conjecture that every convex semialgebraic set could be written as the projection of a spectrahedron. This conjecture was disproved by Scheiderer in 2016.

Theorem 5.5 ([**16**]). *Not every convex, semialgebraic set is a spectrahedral shadow. In particular, the set of nonnegative polynomials of degree ≤ 6 in three variables, $P_{3,\leq 6}$ is not a spectrahedral shadow.*

Building off of these techniques, Hamza Fawzi showed that $P_{2,\leq 6}$ and $P_{3,\leq 4}$ are also not spectrahedral shadows [**7**]. From this, one can show that the cone $P_{n,\leq 2d}$ is not a spectrahedral shadow for all $(n, 2d)$ with $n \geq 2$ and $2d \geq 6$ or with $n \geq 3$ and $2d \geq 4$. Moreover, these are the minimal cases of $(n, 2d)$ in which this is possible.

For $d = 2$, $n = 1$, and $(n, 2d) = (2, 4)$, a classical theorem of Hilbert states that every nonnegative polynomial in $\mathbb{R}[x_1, \dots, x_n]_{\leq 2d}$ can be written as a sum of squares, i.e. $\Sigma_{n,\leq 2d} = P_{n,\leq 2d}$. By Proposition 5.1, $\Sigma_{n,\leq 2d}$ is the image of $\mathcal{S}^N_+$ under a linear map. Therefore in the Hilbert cases, it follows that the cone $\Sigma_{n,\leq 2d} = P_{n,\leq 2d}$ is the image of $\mathcal{S}^N_+$ under a linear map. Using Proposition 3.3, it follows that $\Sigma^*_{n,\leq 2d} = P^*_{n,\leq 2d}$ is a spectrahedron. All together, this provides a complete characterization of the pairs $(n, 2d)$ for which the cone of nonnegative polynomials $P_{n,\leq 2d}$ is a spectrahedron or spectrahedral shadow.

Proposition 5.6. *For $n, d \in \mathbb{Z}_+$, consider the cones $P_{n,\leq 2d}$ of nonnegative polynomials in $\mathbb{R}[x_1, \dots, x_n]_{\leq 2d}$ and its dual cone $P^*_{n,\leq 2d}$. Then*

- *both $P_{n,\leq 2}$ and $P^*_{n,\leq 2}$ are spectrahedra,*
- *for $n = 1$ and $(n, 2d) = (2, 4)$, $P^*_{n,\leq 2d}$ is a spectrahedron and $P_{n,\leq 2d}$ is a spectrahedral shadow but not a spectrahedron, and*
- *in all other cases, neither $P_{n,\leq 2d}$ nor $P^*_{n,\leq 2d}$ is a spectrahedral shadow.*

PROOF. For the first, notice that any quadratic polynomials $q \in \mathbb{R}[x_1, \ldots, x_n]_{\leq 2}$ can be uniquely represented as $(1, \mathbf{x})Q(1, \mathbf{x})^T$ for a real symmetric matrix $Q \in \mathcal{S}^{n+1}$. Moreover q is nonnegative if and only if Q is positive semidefinite. Therefore $P_{n,\leq 2}$ is linearly isomorphic to the cone $\mathcal{S}^{n+1}_+$, and by the self-duality of $\mathcal{S}^{n+1}_+$, so is $P^*_{n,\leq 2}$.

For $n = 1$ and $(n, 2d) = (2, 4)$, the cones $\Sigma_{n,\leq 2d}$ and $P_{n,\leq 2d}$ are equal. As explained in detail above for $(n, d) = (1, 2)$, the cone $\Sigma_{n,\leq 2d}$ is the image of the positive semidefinite cone under a linear map, meaning that the dual cone $\Sigma^*_{n,\leq 2d} = P^*_{n,\leq 2d}$ is spectrahedral. For $n \geq 1$ and $d \geq 2$, the cone $P_{1,\leq 4}$ is linearly isomorphic to a face of $P_{n,\leq 2d}$, so Example 4.7 shows that $P_{n,\leq 2d}$ has a non-exposed face and therefore cannot be written as a spectrahedron.

Finally the results of Scheiderer [**16**] and Fawzi [**7**] show that in all other cases, $P_{n,\leq 2d}$ is not a spectrahedral shadow. Because the class of spectrahedral shadows is closed under duality, as seen in Proposition 3.4, neither is the dual cone $P^*_{n,\leq 2d}$. □

References

[1] Alexander Barvinok, *A course in convexity*, Graduate Studies in Mathematics, vol. 54, American Mathematical Society, Providence, RI, 2002. MR1940576

[2] Grigoriy Blekherman, Pablo A. Parrilo, and Rekha R. Thomas (eds.), *Semidefinite optimization and convex algebraic geometry*, MOS-SIAM Series on Optimization, vol. 13, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, 2013. MR3075433

[3] Jacek Bochnak, Michel Coste, and Marie-Françoise Roy, *Real algebraic geometry*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 36, Springer-Verlag, Berlin, 1998. Translated from the 1987 French original; Revised by the authors. MR1659509

[4] Arne Brøndsted, *An introduction to convex polytopes*, Graduate Texts in Mathematics, vol. 90, Springer-Verlag, New York-Berlin, 1983. MR683612

[5] M. D. Choi, T. Y. Lam, and B. Reznick, *Sums of squares of real polynomials*, K-theory and algebraic geometry: connections with quadratic forms and division algebras (Santa Barbara, CA, 1992), Proc. Sympos. Pure Math., vol. 58, Amer. Math. Soc., Providence, RI, 1995, pp. 103–126. MR1327293

[6] Lynn Chua, Daniel Plaumann, Rainer Sinn, and Cynthia Vinzant, *Gram spectrahedra*, Ordered algebraic structures and related topics, Contemp. Math., vol. 697, Amer. Math. Soc., Providence, RI, 2017, pp. 81–105, DOI 10.1090/conm/697/14047. MR3716067

[7] H. Fawzi, *The set of separable states has no finite semidefinite representation except in dimension* 3×2, preprint available at `arXiv:1905.02575`, 2019.

[8] E. K. Haviland, *On the Momentum Problem for Distribution Functions in More Than One Dimension. II*, Amer. J. Math. **58** (1936), no. 1, 164–168, DOI 10.2307/2371063. MR1507139

[9] J. William Helton and Jiawang Nie, *Sufficient and necessary conditions for semidefinite representability of convex hulls and sets*, SIAM J. Optim. **20** (2009), no. 2, 759–791, DOI 10.1137/07070526X. MR2515796

[10] Murray Marshall, *Positive polynomials and sums of squares*, Mathematical Surveys and Monographs, vol. 146, American Mathematical Society, Providence, RI, 2008. MR2383959

[11] Jiawang Nie, Kristian Ranestad, and Bernd Sturmfels, *The algebraic degree of semidefinite programming*, Math. Program. **122** (2010), no. 2, Ser. A, 379–405, DOI 10.1007/s10107-008-0253-6. MR2546336

[12] Gábor Pataki, *On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues*, Math. Oper. Res. **23** (1998), no. 2, 339–358, DOI 10.1287/moor.23.2.339. MR1626662

[13] Motakuri Ramana and A. J. Goldman, *Some geometric results in semidefinite programming*, J. Global Optim. **7** (1995), no. 1, 33–50, DOI 10.1007/BF01100204. MR1342934

[14] Kristian Ranestad, *Algebraic degree in semidefinite and polynomial optimization*, Handbook on semidefinite, conic and polynomial optimization, Internat. Ser. Oper. Res. Management Sci., vol. 166, Springer, New York, 2012, pp. 61–75, DOI 10.1007/978-1-4614-0769-0_3. MR2894691

[15] Claus Scheiderer, *Semidefinite representation for convex hulls of real algebraic curves*, SIAM J. Appl. Algebra Geom. **2** (2018), no. 1, 1–25, DOI 10.1137/17M1115113. MR3755651

[16] Claus Scheiderer, *Spectrahedral shadows*, SIAM J. Appl. Algebra Geom. **2** (2018), no. 1, 26–44, DOI 10.1137/17M1118981. MR3755652

[17] C. Scheiderer, *Extreme points of gram spectrahedra of binary forms*, preprint available at arXiv:1802.05513, 2018.

[18] Lieven Vandenberghe and Stephen Boyd, *Semidefinite programming*, SIAM Rev. **38** (1996), no. 1, 49–95, DOI 10.1137/1038003. MR1379041

Department of Mathematics, North Carolina State University, Raleigh, North Carolina