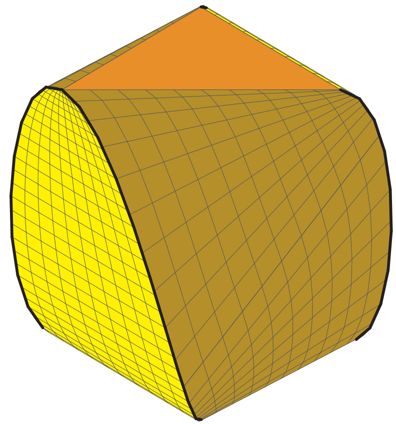
Orbitopes

28th Friends of Mathematics Lecture
Kansas State University
27 April 2010

Frank Sottile

sottile@math.tamu.edu

Work with
Raman Sanyal and Bernd Sturmfels
University of California, Berkeley



Orbitopes

or·bi·tope |ˈɔr bi toʊp|

noun

1 the convex hull of an orbit of a compact group acting linearly on
a vector space.

2 highly symmetric convex bodies that have appeared in
many areas of mathematics and its applications,
but have just begun to attract systematic attention.

3 subject of this talk.

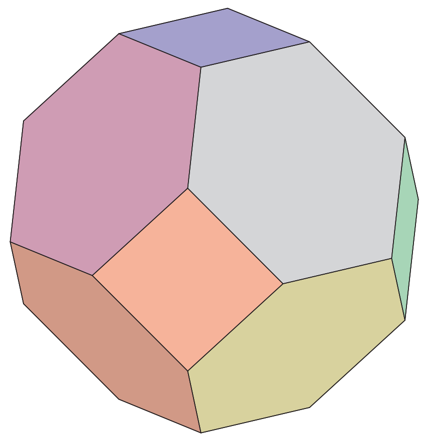
Orbitopes

Throughout, G will be a real compact linear algebraic group.

E.g. $G = SO(d)$, the special orthogonal group, or
$G = U(d)$, the unitary group.

Let V be a finite-dimensional real vector space on which G acts.
The *orbit* of G through $v \in V$ is $G \cdot v = \{g \cdot v \mid g \in G\}$, an algebraic manifold. The *orbitope* $\mathcal{O}_v$ of G through $v \in V$ is the convex hull of $G \cdot v$, a convex semi-algebraic set.

Orbitopes of finite groups G include the beautifully symmetric Platonic and Archimedean solids, such as the permutahedron for S_4, at right.
We will be interested in orbitopes for continuous groups.



Low-dimensional connected Orbitopes

$d = 1$: There are no one-dimensional orbitopes because $SO(1)$ is a point.

$d = 2$: The only orbitopes in $\mathbb{R}^2$ are discs.

$d = 3$: The only orbitopes in $\mathbb{R}^3$ for connected groups are balls.

$d = 4$: $SO(4)$ has connected subgroups G of dimension 1, 2, 3 and 6.

$\dim G = 6$, then $G = SO(4)$ whose orbitopes are balls in $\mathbb{R}^4$.

$\dim G = 3$, then $G = SU(2)$ acting as the unit quaternions on $\mathbb{H} = \mathbb{R}^4$, and the orbitopes are again balls.

$\dim G = 2$, then $G \simeq SO(2) \times SO(2)$ and orbitopes are products of two discs.

$\dim G = 1$, then $G \simeq SO(2)$ with orbitopes

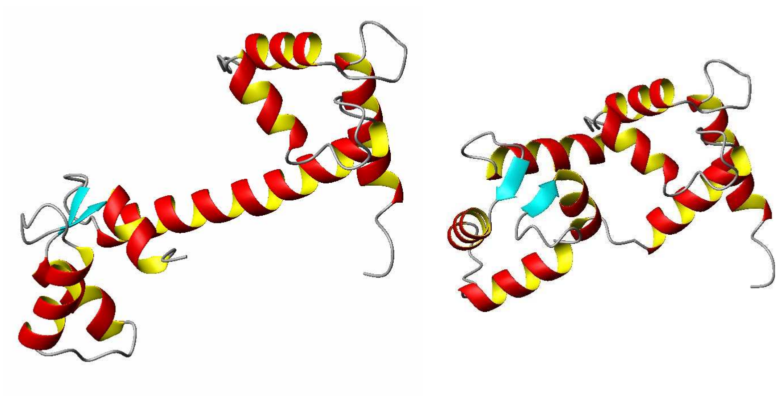
$$\text{conv}\{(\cos(p\theta), \sin(p\theta), \cos(q\theta), \sin(q\theta)) \in \mathbb{R}^4 \mid \theta \in [0, 2\pi]\}\ ,$$

which were introduced one century ago by Carathéodory.

Two examples of Orbitopes

⇒ The Harvey-Lawson (& Morgan, Bryant, Mackenzie...) theory of calibrated geometry for minimal submanifolds amounts to identifying faces of the *Grassmann orbitope*, which is the convex hull the $SO(n)$ orbit of a decomposable tensor in $\wedge_k \mathbb{R}^d$.

⇒ Motivated by protein structure reconstruction, Longinetti, Sgheri and I studied $SO(3)$ orbitopes in the space of symmetric 3×3 tensors.



Three perspectives on Orbitopes

While orbitopes have been studied previously in different areas of mathematics from different perspectives, their systematic study is now warranted from the new perspective of convex algebraic geometry, which combines three areas of mathematics, leading to several motivating questions.

Classical convexity : Determine the faces, face lattice, dual bodies, and Carathéodory numbers of orbitopes.

Algebraic geometry : Describe the Zariski boundary of an orbitope, its equation, and Whitney stratification.

Optimization : Can the orbitope be represented as a spectrahedron? How can one efficiently optimize over an orbitope?

Spectrahedra as noncommutative polytopes

A polyhedron P has a facet description

$$P = \{x \in \mathbb{R}^d \mid x_1 a_1 + x_2 a_2 + \cdots + x_d a_d + b \geq 0\} ,$$

where $a_i, b \in \mathbb{R}^n$, and $\geq$ is componentwise comparison.

A symmetric matrix $X \in \mathbb{R}^{n\times n}$ (or hermitian $X \in \mathbb{C}^{n\times n}$) is *positive semi-definite (PSD), $X \succeq 0$*, if all its eigenvalues are nonnegative, equivalently if all symmetric minors are nonnegative. PSD matrices form a convex cone.

A *spectrahedron* is set consisting of those $x \in \mathbb{R}^d$ such that

$$x_1 A_1 + x_2 A_2 + \cdots + x_d A_d + B \succeq 0 ,$$

where A_i, B are symmetric (or hermitian) matrices. A polyhedron is a spectrahedron when A_i, B mutually commute (are diagonal).

Optimization

Polytopes are the natural domains of linear programs which are a critically important class of optimization problems.

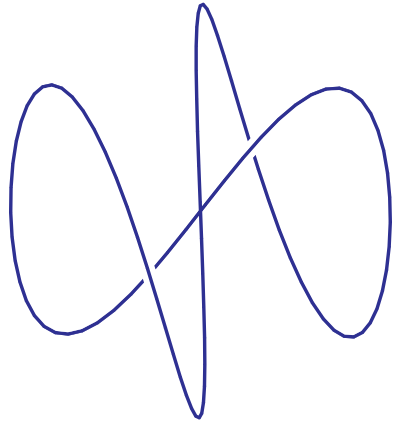
Semi-definite programming is a recent extension of linear programming. It gives efficient algorithms for optimizing linear functions over spectrahedra.

Since linear functions pull back along (linear) projections, semi-definite programming gives efficient algorithms for optimizing linear functions over projections of spectrahedra.

Because of this, it is important to understand which convex semi-algebraic sets are spectrahedra or projections of spectrahedra, together with spectrahedral representations. This question about the structure of convex semi-algebraic sets drives the field of convex algebraic geometry.

Convexity

The convex hull of this trigonometric curve

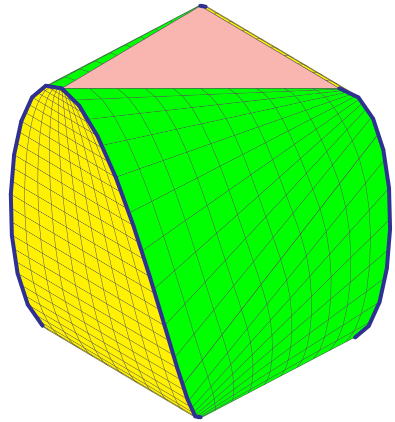


$$(\cos(\theta), \sin(2\theta), \cos(3\theta))$$

Convexity

The convex hull of this trigonometric curve has boundary consisting of two families of line segments (yellow and green) and two triangles. The extreme points are that part of the curve lying in the boundary.

The triangle edges are special—they are not exposed by any linear functional.

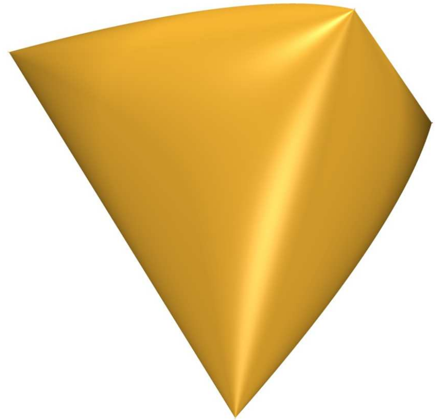


$$(\cos(\theta), \sin(2\theta), \cos(3\theta))$$

Each point lies in a convex hull of at most three extreme points (it is fibered by triangular and rectangular slices), so it has *Carathéodory number* 3.

It is not a spectrahedron, as it has non-exposed faces, but it is a projection of a spectrahedron (the Carathéodory orbitope C_3), as we'll see.

Algebraic Geometry



This convex body is the bounded component of the complement of a singular cubic surface.
Its faces are the four singular points, the six edges between them, and every other point in its boundary is extreme. All are exposed.

This is a spectrahedron (hyperplane section of Carathéodory orbitope C_2).

Its Zariski boundary is the cubic surface, while the boundary of the previous body is a reducible hypersurface of degree 21, with the green ruled surface having degree 16.

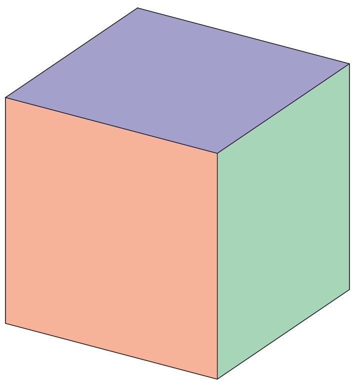
Polarity

The optima of all linear functions on a convex body $\mathcal{O} \subset V$ are encoded in its *polar body*

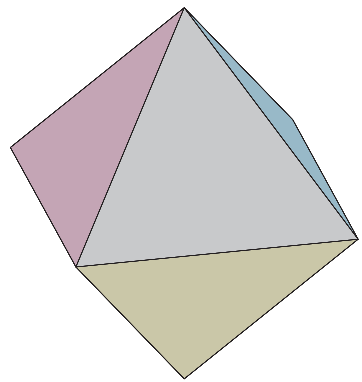
$$\mathcal{O}^\circ = \{\ell \in V^* \;:\; \ell|_{\mathcal{O}} \leq 1\}.$$

When $\mathcal{O}$ is centrally symmetric, it is the ball of a norm on V and its polar $\mathcal{O}^\circ$ is the ball of the dual norm.

The Whitney stratification of the boundary of $\mathcal{O}$ is expected to correspond to the stratification of the boundary of $\mathcal{O}^\circ$. This correspondence is exact and inclusion-reversing for polyhedra.



$\Longleftrightarrow$



Carathéodory orbitopes

The convex hull of the trigonometric moment curve,

$$\{(\cos(\theta), \sin(\theta), \cos(2\theta), \ldots, \sin(d\theta)) \mid \theta \in [0, 2\pi)\}\,,$$

in $\mathbb{R}^{2d}$ is the *Carathéodory orbitope C_d*, studied by Carathéodory in 1907.

It is an orbitope, as $\mathbb{R}^{2d} = (\mathbb{R}^2)^d$ is a representation of the circle group $SO(2)$ where a rotation matrix acts on the nth factor via its nth power,

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}^n = \begin{pmatrix} \cos(n\theta) & -\sin(n\theta) \\ \sin(n\theta) & \cos(n\theta) \end{pmatrix}$$

Every orbitope of $SO(2)$ is a coordinate projection of some Carathéodory orbitope, and convex hulls of trigonometric curves are projections of Carathéodory orbitopes.

Spectrahedral representation

By classical results on positive trigonometric polynomials, C_d is those $(x_1, \ldots, x_d) \in \mathbb{C}^d = (\mathbb{R}^2)^d$ such that the Toeplitz matrix is PSD,

$$\begin{pmatrix} 1 & x_1 & x_2 & \cdots & x_d \\ \overline{x_1} & 1 & x_1 & \cdots & x_{d-1} \\ \overline{x_2} & \overline{x_1} & 1 & & \vdots \\ \vdots & \vdots & & \ddots & x_1 \\ \overline{x_d} & \overline{x_{d-1}} & \cdots & \overline{x_1} & 1 \end{pmatrix} \succeq 0\,.$$

Theorem. *C_d is a neighborly, simplicial convex body whose faces are in inclusion preserving correspondence with sets of at most d points on $SO(2)$. It has Carathéodory number $d+1$.*

Thus $SO(2)$-Orbitopes are projections of spectrahedra, but not in general spectrahedra—they have non-exposed faces, by work of Smilansky and Barvinok-Novik.

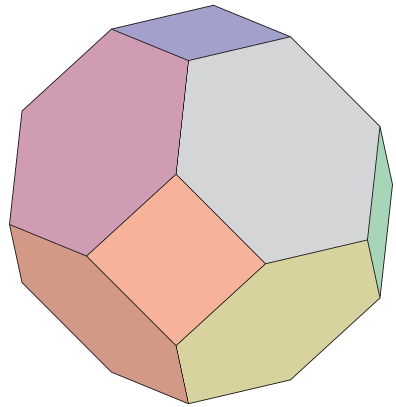
Symmetric Schur-Horn Orbitopes

Permutahedra are orbitopes for the symmetric group S_d. Specifically, let D be the diagonal $d \times d$ matrices of trace zero. Given $p, q \in D$, let $\lambda(p)$ be the components of p in nonincreasing order, and write $q \trianglelefteq p$ if

$$\lambda(q)_1 + \cdots + \lambda(q)_k \leq \lambda(p)_1 + \cdots + \lambda(p)_k, \quad k = 1, \ldots, d-1.$$

Then the permutahedron through $p \in D$ is

$$\Pi_p \;=\; \{q \in D \mid q \trianglelefteq p\}.$$



Let $S_2\mathbb{R}^d$ be the space of symmetric $d \times d$ matrices with trace zero, an irreducible representation of $SO(d)$ acting by conjugation.

The *symmetric Schur-Horn orbitope* $\mathcal{O}_M$ through $M \in S_2\mathbb{R}^d$ is the convex hull of the orbit $SO(d) \cdot M$.

Non-commutative permutahedra

Schur-Horn Theorem. *Given $M \in S_2\mathbb{R}^d$ with diagonal D$(M) \in D$ and eigenvalues $\lambda(M)$, we have D$(M) \trianglelefteq \lambda(M)$. In fact, D$(\mathcal{O}_M) = \Pi_{\lambda(M)}$ and $\Pi_{\lambda(M)}$ is the intersection of $\mathcal{O}_M$ with the diagonal.*

Corollary. $\mathcal{O}_M = \{A \in S_2\mathbb{R}^d \mid \lambda(A) \trianglelefteq \lambda(M)\}$.

This implies a complete facial description of $\mathcal{O}_M$ (similar to that of $\Pi_{\lambda(M)}$), as well as a spectrahedral representation using Lie algebra Schur functors (a.k.a *additive compound matrices*).

Let $\mathcal{L}_k \colon \mathfrak{gl}(\mathbb{R}^d) \longrightarrow \mathfrak{gl}(\wedge_k\mathbb{R}^d)$ be the induced map on Lie algebras. The eigenvalues of $\mathcal{L}_k(M)$ are sums of k distinct eigenvalues of M.

If $l_k(M)$ is the sum of the k largest eigenvalues of M,

$$\mathcal{O}_M = \{A \in S_2\mathbb{R}^d \mid l_k(M)I_{\binom{d}{k}} - \mathcal{L}_k(A) \succeq 0,\ k = 1, \ldots, d-1\}.$$

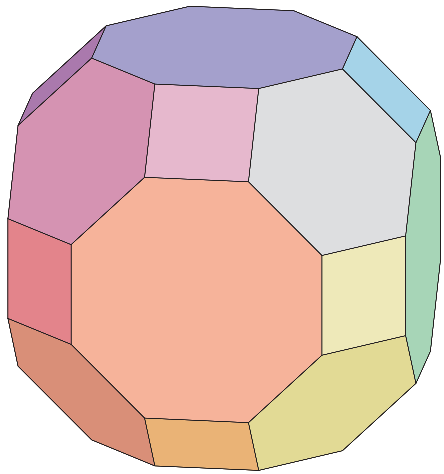
Skew-symmetric Schur-Horn Orbitopes

$SO(d)$ also acts on skew-symmetric $d \times d$ matrices, $\wedge_2\mathbb{R}^d$. A matrix $N \in \wedge_2\mathbb{R}^d$ may be conjugated to one of the form

$$\begin{pmatrix} & \Lambda \\ -\Lambda & \end{pmatrix} \text{ or } \begin{pmatrix} & & \Lambda \\ & 0 & \\ -\Lambda & & \end{pmatrix}$$

depending on the parity of d. The diagonal matrix $\Lambda = \Lambda(N)$ is the *skew diagonal* of N—it plays the role here of diagonal matrices.

The role of the permutahedron is played by the $B_{\lfloor \frac{d}{2} \rfloor}$-permutahedron. The orbitope $\mathcal{O}_N$ for $N \in \wedge_2\mathbb{R}^d$ has inequalities and faces corresponding to those from the $B_{\lfloor \frac{d}{2} \rfloor}$-permutahedron.



Tautological Orbitopes

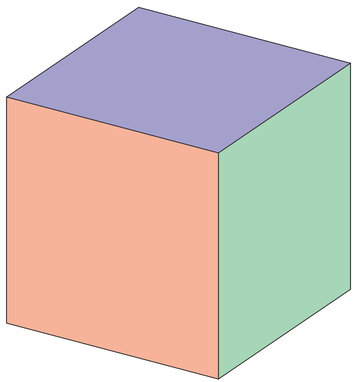
A compact group G acting on V acts by left translation on $\mathfrak{gl}V$. Orbitopes for this action are called *tautological orbitopes*.

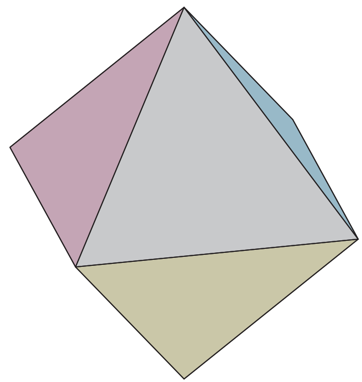
For $O(d) \subset \mathfrak{gl}_d\mathbb{R}$ with diagonal projection D,

$$\mathrm{D}(\mathrm{conv}(O(d))) = [-1, +1]^d, \text{ the } d\text{-cube.}$$

The *nuclear norm* of $A \in \mathfrak{gl}_d\mathbb{R}$ is the sum of its singular values, and the operator norm is its maximal singular value.

Theorem. $\mathrm{conv}(O(d))$ is the operator norm ball. Its polar $\mathrm{conv}(O(d))^\circ$ is the nuclear norm ball. Both objects are spectrahedra with diagonal projection the d-cube and d-dimensional octahedron (crosspolytope). The polytopes control the facial structure of the orbitopes as before.





Veronese Orbitopes

The Veronese map

$$\nu_m \ : \ \mathbb{R}^d \ \longrightarrow \ \mathsf{Sym}_m\mathbb{R}^d \simeq \mathbb{R}^{\binom{d+m-1}{m-1}}$$

has image the set of decomposable tensors.

The *Veronese orbitope* $\mathcal{V}_{d,m}$ is the convex hull of an orbit of decomposable tensors, which we may take to be $\nu_m(\mathbb{S}^{d-1})$.

When $m = 2n$ is even, the cone over the coorbitope $\mathcal{V}^\circ_{d,2n}$ is the cone of non-negative d-ary forms of degree $2n$. These cones are nearly unknowable, except when $d = 3$ and $2n = 4$.

$\Rightarrow$ This suggests that understanding orbitopes $\mathcal{O}$ and their polars $\mathcal{O}^\circ$ will be at least as hard as understanding positive polynomials.

Ternary Quartics

The Veronese orbitope $\mathcal{V}_{3,4}$ is a 14-dimensional convex body. Since non-negative ternary quartics are sums of squares, $\mathcal{V}^{\circ}_{3,4}$ is a projection of a spectrahedron—but not a spectrahedron, as Blekherman showed it has non-exposed faces.

The boundary of $\mathcal{V}^{\circ}_{3,4}$ is the irreducible hypersurface of degree 27 cut out by the discriminant of the ternary quartic.

Theorem. (Reznick) $\mathcal{V}_{3,4}$ is a spectrahedron. It equals those λ_{abc} such that

$$\begin{pmatrix} \lambda_{400} & \lambda_{220} & \lambda_{202} & \lambda_{310} & \lambda_{301} & \lambda_{211} \\ \lambda_{220} & \lambda_{040} & \lambda_{022} & \lambda_{130} & \lambda_{121} & \lambda_{031} \\ \lambda_{202} & \lambda_{022} & \lambda_{004} & \lambda_{112} & \lambda_{103} & \lambda_{013} \\ \lambda_{310} & \lambda_{130} & \lambda_{112} & \lambda_{220} & \lambda_{211} & \lambda_{121} \\ \lambda_{301} & \lambda_{121} & \lambda_{103} & \lambda_{211} & \lambda_{202} & \lambda_{112} \\ \lambda_{211} & \lambda_{031} & \lambda_{013} & \lambda_{121} & \lambda_{112} & \lambda_{022} \end{pmatrix} \succeq 0\,.$$

and $\lambda_{400} + \lambda_{040} + \lambda_{004} + 2\lambda_{220} + 2\lambda_{202} + 2\lambda_{022} = 1$.

Thank You!