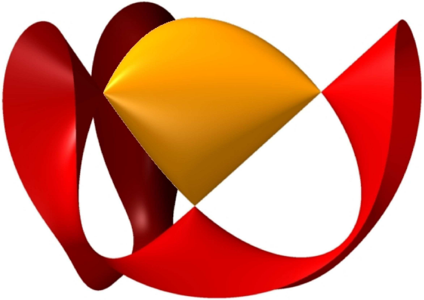
The Geometry of Semidefinite Programming

Bernd Sturmfels
UC Berkeley



Positive Semidefinite Matrices

For a real symmetric $n \times n$-matrix A the following are equivalent:

- All n eigenvalues of A are positive real numbers.
- All 2^n principal minors of A are positive real numbers.
- Every non-zero vector $x \in \mathbb{R}^n$ satisfies $x^T A \cdot x > 0$.

A matrix A is *positive definite* if it satisfies these properties, and it is *positive semidefinite* if the following equivalent properties hold:

- All n eigenvalues of A are non-negative real numbers.
- All 2^n principal minors of A are non-negative real numbers.
- Every vector $x \in \mathbb{R}^n$ satisfies $x^T A \cdot x \geq 0$.

The set of all positive semidefinite $n \times n$-matrices is a convex cone of full dimension $\binom{n+1}{2}$. It is closed and semialgebraic.

The interior of this cone consists of all positive definite matrices.

Semidefinite Programming

A *spectrahedron* is the intersection of the cone of positive semidefinite matrices with an affine-linear space. Its algebraic representation is a linear combination of symmetric matrices

$$A_0 + x_1 A_1 + x_2 A_2 + \cdots + x_m A_m \ \succeq \ 0 \qquad (*)$$

Engineers call this is a *linear matrix inequality*.

Semidefinite Programming

A *spectrahedron* is the intersection of the cone of positive semidefinite matrices with an affine-linear space. Its algebraic representation is a linear combination of symmetric matrices

$$A_0 + x_1 A_1 + x_2 A_2 + \cdots + x_m A_m \ \succeq \ 0 \qquad (*)$$

Engineers call this is a *linear matrix inequality*.

Semidefinite programming is the computational problem of maximizing a linear function over a spectrahedron:

$$\mathrm{Maximize} \ \ c_1 x_1 + c_2 x_2 + \cdots + c_m x_m \ \text{ subject to } \ (*)$$

Example: *The smallest eigenvalue of a symmetric matrix* A *is the solution of the SDP* Maximize x subject to $A - x \cdot \mathrm{Id} \succeq 0$.

Convex Polyhedra

Linear programming is semidefinite programming for diagonal matrices. If $A_0, A_1, \ldots, A_m$ are diagonal $n \times$n-matrices then

$$A_0 + x_1 A_1 + x_2 A_2 + \cdots + x_m A_m \succeq 0$$

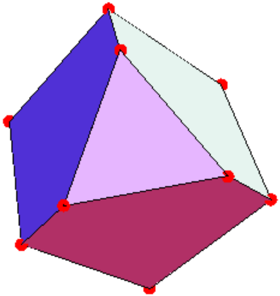
translates into a system of n linear inequalities in the m unknowns.

Convex Polyhedra

Linear programming is semidefinite programming for diagonal matrices. If $A_0, A_1, \ldots, A_m$ are diagonal $n\times$n-matrices then

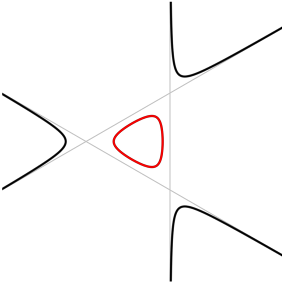
$$A_0 + x_1 A_1 + x_2 A_2 + \cdots + x_m A_m \ \succeq \ 0$$

translates into a system of n linear inequalities in the m unknowns. A spectrahedron defined in this manner is a convex polyhedron:



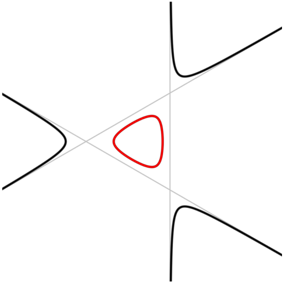
Pictures in Dimension Two

Here is a picture of a spectrahedron for $m = 2$ and $n = 3$:

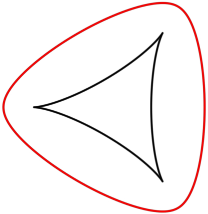


Pictures in Dimension Two

Here is a picture of a spectrahedron for $m = 2$ and $n = 3$:



Duality is important in convex optimization:



Example: Multifocal Ellipses

Given m points $(u_1, v_1), \ldots, (u_m, v_m)$ in the plane $\mathbb{R}^2$, and a radius $d > 0$, their *m-ellipse* is the convex algebraic curve

$$\left\{ (x, y) \in \mathbb{R}^2 \,:\, \sum_{k=1}^{m} \sqrt{(x-u_k)^2 + (y-v_k)^2} \,=\, d \right\}.$$

The 1-ellipse and the 2-ellipse are algebraic curves of degree 2.

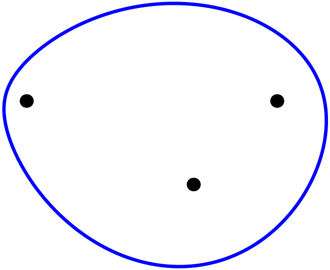
Example: Multifocal Ellipses

Given m points $(u_1, v_1), \ldots, (u_m, v_m)$ in the plane $\mathbb{R}^2$, and a radius $d > 0$, their *m-ellipse* is the convex algebraic curve

$$\left\{ (x, y) \in \mathbb{R}^2 \, : \, \sum_{k=1}^{m} \sqrt{(x-u_k)^2 + (y-v_k)^2} \, = \, d \right\}.$$

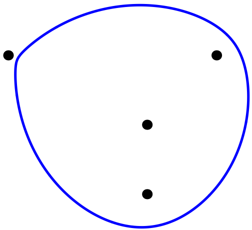
The 1-ellipse and the 2-ellipse are algebraic curves of degree 2.

The 3-ellipse is an algebraic curve of degree 8:

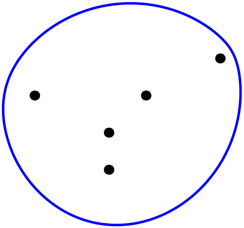


2, 2, 8, 10, 32, ...

The 4-ellipse is an algebraic curve of degree 10:

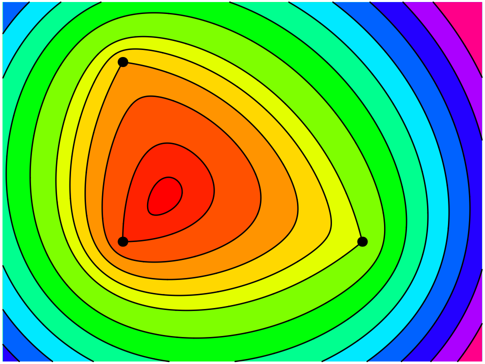


The 5-ellipse is an algebraic curve of degree 32:



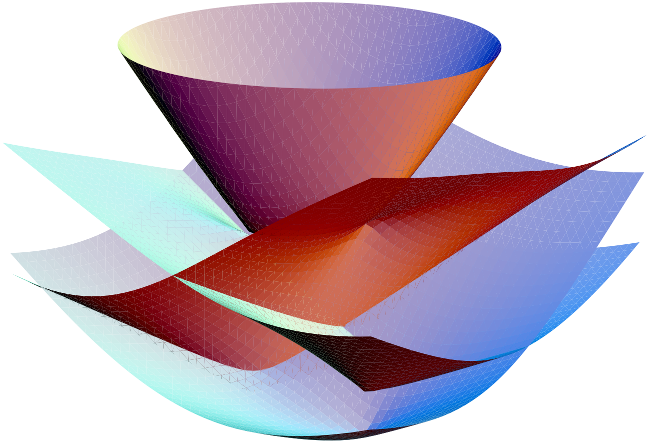
Concentric Ellipses

What is the algebraic degree of the m-ellipse?
How to write its equation?



What is the smallest radius d for which the m-ellipse is non-empty? How to compute the Fermat-Weber point?

3D View



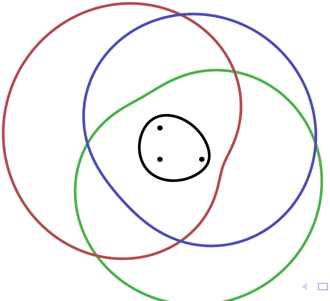
$$\mathcal{C} \;=\; \left\{(x, y, d) \in \mathbb{R}^3 \,:\, \sum_{k=1}^{m} \sqrt{(x-u_k)^2 + (y-v_k)^2} \;\leq\; d \,\right\}.$$

Ellipses are Spectrahedra

The 3-ellipse with foci $(0,0),(1,0),(0,1)$ has the representation

$$\begin{bmatrix} d+3x-1 & y-1 & y & 0 & y & 0 & 0 & 0 \\ y-1 & d+x-1 & 0 & y & 0 & y & 0 & 0 \\ y & 0 & d+x+1 & y-1 & 0 & 0 & y & 0 \\ 0 & y & y-1 & d-x+1 & 0 & 0 & 0 & y \\ y & 0 & 0 & 0 & d+x-1 & y-1 & y & 0 \\ 0 & y & 0 & 0 & y-1 & d-x-1 & 0 & y \\ 0 & 0 & y & 0 & y & 0 & d-x+1 & y-1 \\ 0 & 0 & 0 & y & 0 & y & y-1 & d-3x+1 \end{bmatrix}$$

The ellipse consists of all points (x,y) where this symmetric 8×8-matrix is positive semidefinite. Its boundary is a curve of degree eight:



2, 2, 8, 10, 32, 44, 128, ...

Theorem: *The polynomial equation defining the m-ellipse has degree* 2^m *if m is odd* *and degree* $2^m - \binom{m}{m/2}$ *if m is even. We express this polynomial as the determinant of a symmetric matrix of linear polynomials. Our representation extends to weighted m-ellipses and m-ellipsoids in arbitrary dimensions .....*

[J. Nie, P. Parrilo, B.St.: Semidefinite representation of the k-ellipse, in *Algorithms in Algebraic Geometry*, I.M.A. Volumes in Mathematics and its Applications, 146, Springer, New York, 2008, pp. 117-132]

In other words, *m*-ellipses and *m*-ellipsoids are spectrahedra. The problem of finding the Fermat-Weber point is an SDP.

2, 2, 8, 10, 32, 44, 128, ...

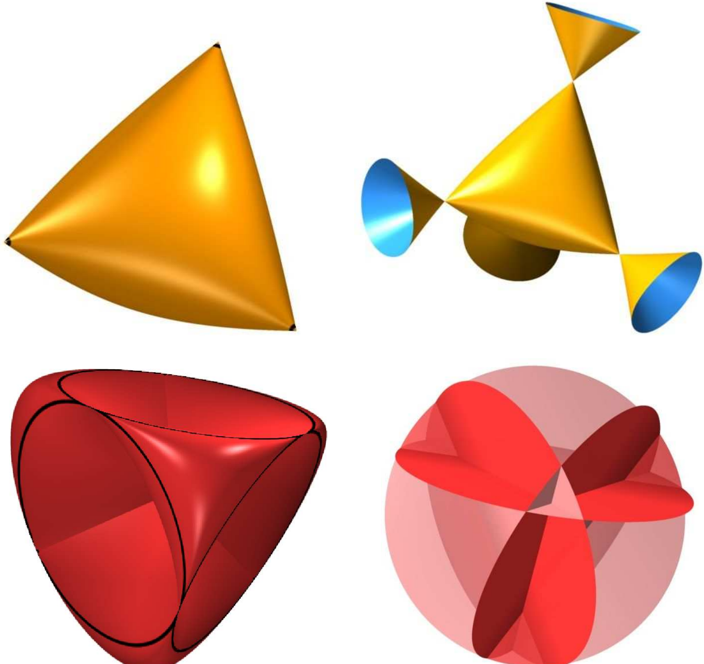
Theorem: *The polynomial equation defining the m-ellipse has degree* 2^m *if m is odd and degree* $2^m - \binom{m}{m/2}$ *if m is even. We express this polynomial as the determinant of a symmetric matrix of linear polynomials. Our representation extends to weighted m-ellipses and m-ellipsoids in arbitrary dimensions .....*

[J. Nie, P. Parrilo, B.St.: Semidefinite representation of the k-ellipse, in *Algorithms in Algebraic Geometry*, I.M.A. Volumes in Mathematics and its Applications, 146, Springer, New York, 2008, pp. 117-132]

In other words, m-ellipses and m-ellipsoids are spectrahedra.
The problem of finding the Fermat-Weber point is an SDP.

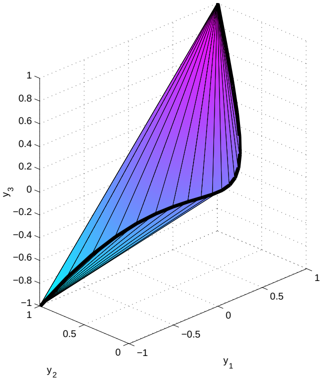
Let's now look at some spectrahedra in dimension three.
Our next picture shows the typical behavior for $m = 3$ and $n = 3$.

A Spectrahedron and its Dual



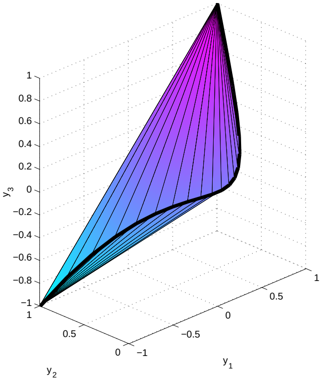
Non-Linear Convex Hull Computation

Input : $\left\{(t, t^2, t^3) \in \mathbb{R}^3 : -1 \leq t \leq 1\right\}$



Non-Linear Convex Hull Computation

Input : $\{(t, t^2, t^3) \in \mathbb{R}^3 : -1 \leq t \leq 1\}$



The convex hull of the moment curve is a spectrahedron.

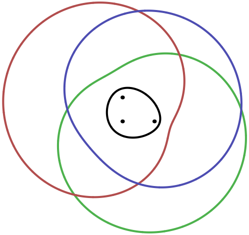
Output : $\begin{pmatrix} 1 & x \\ x & y \end{pmatrix} \pm \begin{pmatrix} x & y \\ y & z \end{pmatrix} \succeq 0$

Characterization of Spectrahedra

A convex hypersurface of degree d in $\mathbb{R}^n$ is *rigid convex* if every line passing through its interior meets (the Zariski closure of) that hypersurface in d real points.

Theorem (Helton–Vinnikov (2006))

Every spectrahedron is rigid convex. The converse is true for $n = 2$.

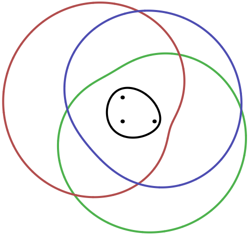


Characterization of Spectrahedra

A convex hypersurface of degree d in $\mathbb{R}^n$ is *rigid convex* if every line passing through its interior meets (the Zariski closure of) that hypersurface in d real points.

Theorem (Helton–Vinnikov (2006))

Every spectrahedron is rigid convex. The converse is true for $n = 2$.



Open problem: Is every compact convex basic semialgebraic set $\mathcal{S}$ the projection of a spectrahedron in higher dimensions?

Theorem (Helton–Nie (2008))

The answer is yes if the boundary of $\mathcal{S}$ is "sufficiently smooth".

Questions about 3-Dimensional Spectrahedra

What are the edge graphs of spectrahedra in $\mathbb{R}^3$?

How can one define their *combinatorial types*?

Is there an analogue to Steinitz' Theorem for polytopes in $\mathbb{R}^3$?

Consider 3-dimensional spectrahedra whose boundary is an irreducible surface of degree n. Can such a spectrahedron have $\binom{n+1}{3}$ isolated singularities in its boundary? How about $n = 4$?

A Pinch of Statistics

Every positive definite $n \times n$-matrix $\Sigma = (\sigma_{ij})$ is the covariance matrix of a multivariate normal distribution. A Gaussian graphical model is specified by requiring that some entries of Σ^{-1} are zero.

Maximum likelihood estimation is a *matrix completion problem*.

For example, under which condition on the visible entries σ_{ij} can we find x and y which make the following matrix positive definite?

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & x & \sigma_{14} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} & y \\ x & \sigma_{23} & \sigma_{33} & \sigma_{34} \\ \sigma_{14} & y & \sigma_{34} & \sigma_{44} \end{pmatrix}$$

A Pinch of Statistics

Every positive definite $n \times n$-matrix $\Sigma = (\sigma_{ij})$ is the covariance matrix of a multivariate normal distribution. A Gaussian graphical model is specified by requiring that some entries of Σ^{-1} are zero.

Maximum likelihood estimation is a *matrix completion problem*.

For example, under which condition on the visible entries σ_{ij} can we find x and y which make the following matrix positive definite?

$$\Sigma \quad = \quad \begin{pmatrix} \sigma_{11} & \sigma_{12} & x & \sigma_{14} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} & y \\ x & \sigma_{23} & \sigma_{33} & \sigma_{34} \\ \sigma_{14} & y & \sigma_{34} & \sigma_{44} \end{pmatrix}$$

The MLE is the point $(\hat{x}, \hat{y})$ which maximizes the determinant. In optimization, this is the *analytic center* of the spectrahedron.

[C. Uhler, B.St.: Multivariate Gaussians, Semidefinite Matrix Completion and Convex Algebraic Geometry, 2009]

Minimizing Polynomial Functions

Let $f(x_1, \ldots, x_m)$ be a polynomial of even degree $2d$.
We wish to compute the global minimum x^* of $f(x)$ on $\mathbb{R}^m$.

This optimization problem is equivalent to

Maximize λ such that $f(x) - \lambda$ is non-negative on $\mathbb{R}^m$.

This problem is very hard.

Minimizing Polynomial Functions

Let $f(x_1, \ldots, x_m)$ be a polynomial of even degree $2d$.
We wish to compute the global minimum x^* of $f(x)$ on $\mathbb{R}^m$.

This optimization problem is equivalent to

Maximize λ such that $f(x) - \lambda$ is non-negative on $\mathbb{R}^m$.

This problem is very hard.

The optimal value of the following relaxtion gives a lower bound.

Maximize λ such that $f(x) - \lambda$ is a sum of squares of polynomials.

The second problem is much easier. It is a semidefinite program.

Minimizing Polynomial Functions

Let $f(x_1, \ldots, x_m)$ be a polynomial of even degree $2d$.
We wish to compute the global minimum x^* of $f(x)$ on $\mathbb{R}^m$.

This optimization problem is equivalent to

Maximize λ such that $f(x) - \lambda$ is non-negative on $\mathbb{R}^m$.

This problem is very hard.

The optimal value of the following relaxtion gives a lower bound.

Maximize λ such that $f(x) - \lambda$ is a sum of squares of polynomials.

The second problem is much easier. It is a semidefinite program.

Empirically, the optimal value of the SDP almost always agrees with the global minimum. In that case, the optimal matrix of the dual SDP has rank one, and the optimal point x^* can be recovered from this. **How to reconcile this with Blekherman's results?**

SOS Programming: A Univariate Example

Let $m = 1$, $d = 2$ and $f(x) = 3x^4 + 4x^3 - 12x^2$. Then

$$f(x) - \lambda \quad = \quad \begin{pmatrix} x^2 & x & 1 \end{pmatrix} \begin{pmatrix} 3 & 2 & \mu - 6 \\ 2 & -2\mu & 0 \\ \mu - 6 & 0 & -\lambda \end{pmatrix} \begin{pmatrix} x^2 \\ x \\ 1 \end{pmatrix}$$

Our problem is to find (λ, μ) such that the 3×3-matrix is positive semidefinite and λ is maximal. The optimal solution of this SDP is

$$(\lambda^*, \mu^*) \;=\; (-32, -2).$$

Cholesky factorization reveals the SOS representation

$$f(x) - \lambda^* \quad = \quad \left((\sqrt{3}\,x - \frac{4}{\sqrt{3}}) \cdot (x+2)\right)^2 + \frac{8}{3}(x+2)^2.$$

We see that the global minimum is $x^* = -2$.
This approach works for many polynomial optimization problems.

Hankel Matrices

Consider the intersection of the cone of 6×6 PSD matrices with the 15-dimensional linear space consisting of all Hankel matrices

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

Dual to this intersection is the projection

$$\mathrm{Sym}_2(\mathrm{Sym}_2(\mathbb{R}^3)) \;\rightarrow\; \mathrm{Sym}_4(\mathbb{R}^3)$$

taking a 6×6-matrix to the ternary quartic it represents. Its image is a cone whose algebraic boundary is a *discriminant* of degree 27.

Problem: Determine the variety of all Bézout matrices H^{-1}.

Orbitopes

An *orbitope* is the convex hull of an orbit under a real algebraic representation of a compact Lie group. Primary examples are the groups $\mathrm{SO}(n)$ and their products. Orbitopes for their adjoint representations are continuous analogues of *permutohedra*.

Many of these special orbitopes are projections of spectrahedra.

A forthcoming paper with Raman Sanyal and Frank Sottile develops the basic theory of orbitopes and has many examples.

Orbitopes

An *orbitope* is the convex hull of an orbit under a real algebraic representation of a compact Lie group. Primary examples are the groups $\mathrm{SO}(n)$ and their products. Orbitopes for their adjoint representations are continuous analogues of *permutohedra*.

Many of these special orbitopes are projections of spectrahedra.

A forthcoming paper with Raman Sanyal and Frank Sottile develops the basic theory of orbitopes and has many examples.

Example: Consider the orbitope of $(x+y+z)^4$ under the $\mathrm{SO}(3)$-action on the space $\mathrm{Sym}_4(\mathbb{R}^3)$ of ternary quartics.

Quiz: Is this orbitope a spectrahedron?

Orbitopes

An *orbitope* is the convex hull of an orbit under a real algebraic representation of a compact Lie group. Primary examples are the groups $\mathrm{SO}(n)$ and their products. Orbitopes for their adjoint representations are continuous analogues of *permutohedra*.

Many of these special orbitopes are projections of spectrahedra.

A forthcoming paper with Raman Sanyal and Frank Sottile develops the basic theory of orbitopes and has many examples.

Example: Consider the orbitope of $(x+y+z)^4$ under the $\mathrm{SO}(3)$-action on the space $\mathrm{Sym}_4(\mathbb{R}^3)$ of ternary quartics.

Quiz: Is this orbitope a spectrahedron?

Answer: Yes, it is the set of psd Hankel matrices H that satisfy

$$\lambda_{400} + \lambda_{040} + \lambda_{004} + 2\lambda_{220} + 2\lambda_{202} + 2\lambda_{022} \;=\; 9.$$

Problem. Classify all $\mathrm{SO}(n)$-orbitopes that are spectrahedra.

Tautological Orbitopes

... are obtained by taking the convex hull of a matrix group.

Example (P. Parrilo): $\mathrm{conv}(SO(3))$ is the set of 3×3-matrices

$$\begin{pmatrix} u_{11}+u_{22}-u_{33}-u_{44} & 2u_{23}-2u_{14} & 2u_{13}+2u_{24} \\ 2u_{23}+2u_{14} & u_{11}-u_{22}+u_{33}-u_{44} & 2u_{34}-2u_{12} \\ 2u_{24}-2u_{13} & 2u_{12}+2u_{34} & u_{11}-u_{22}-u_{33}+u_{44} \end{pmatrix}$$

where $U = (u_{ij})$ runs over all 4×4 psd matrices having trace 1.

Tautological Orbitopes

... are obtained by taking the convex hull of a matrix group.

Example (P. Parrilo): $\mathrm{conv}(SO(3))$ is the set of 3×3-matrices

$$\begin{pmatrix} u_{11}+u_{22}-u_{33}-u_{44} & 2u_{23}-2u_{14} & 2u_{13}+2u_{24} \\ 2u_{23}+2u_{14} & u_{11}-u_{22}+u_{33}-u_{44} & 2u_{34}-2u_{12} \\ 2u_{24}-2u_{13} & 2u_{12}+2u_{34} & u_{11}-u_{22}-u_{33}+u_{44} \end{pmatrix}$$

where $U = (u_{ij})$ runs over all 4×4 psd matrices having trace 1.

Proof: Psd matrices having both trace 1 and rank 1 are of the form

$$U \;=\; \frac{1}{a^2+b^2+c^2+d^2} \begin{pmatrix} a^2 & ab & ac & ad \\ ab & b^2 & bc & bd \\ ac & bc & c^2 & cd \\ ad & bd & cd & d^2 \end{pmatrix}$$

Their images under the linear map parametrize the group $SO(3)$.

Barvinok-Novik Orbitopes

Consider the $\mathrm{SO}(2)\times\mathrm{SO}(2)$-orbitope BN_4 determined by the curve

$$\theta \;\mapsto\; \big(\cos(\theta), \sin(\theta), \cos(3\theta), \sin(3\theta)\big) \quad \in \; \mathbf{R}^4.$$

This is the projection of a 6-dimensional Hermitian spectrahedron:

Barvinok-Novik Orbitopes

Consider the $\mathrm{SO}(2)\times\mathrm{SO}(2)$-orbitope BN_4 determined by the curve

$$\theta \;\mapsto\; \big(\cos(\theta), \sin(\theta), \cos(3\theta), \sin(3\theta)\big) \quad \in \; \mathbf{R}^4.$$

This is the projection of a 6-dimensional Hermitian spectrahedron:

$$\begin{pmatrix} 1 & x_1 & x_2 & x_3 \\ y_1 & 1 & x_1 & x_2 \\ y_2 & y_1 & 1 & x_1 \\ y_3 & y_2 & y_1 & 1 \end{pmatrix} \qquad \text{where} \quad \begin{array}{l} x_j = c_j + \sqrt{-1}\cdot s_j, \\ y_j = c_j - \sqrt{-1}\cdot s_j, \end{array}$$

under the map $(c_1, c_2, c_3, s_1, s_2, s_3) \mapsto (c_1, c_3, s_1, s_3)$. Here the unknown c_j represents $\cos(j\theta)$, the unknown s_j represents $\sin(j\theta)$.

The curve is cut out by the 2×2-minors of the Toeplitz matrix.

Barvinok-Novik Orbitopes

Consider the $\mathrm{SO}(2)\times\mathrm{SO}(2)$-orbitope BN_4 determined by the curve

$$\theta \;\mapsto\; \big(\cos(\theta), \sin(\theta), \cos(3\theta), \sin(3\theta)\big) \quad \in \; \mathbf{R}^4.$$

This is the projection of a 6-dimensional Hermitian spectrahedron:

$$\begin{pmatrix} 1 & x_1 & x_2 & x_3 \\ y_1 & 1 & x_1 & x_2 \\ y_2 & y_1 & 1 & x_1 \\ y_3 & y_2 & y_1 & 1 \end{pmatrix} \qquad \text{where} \quad \begin{matrix} x_j = c_j + \sqrt{-1}\cdot s_j, \\ y_j = c_j - \sqrt{-1}\cdot s_j, \end{matrix}$$

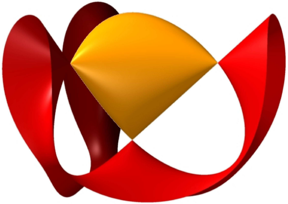
under the map $(c_1, c_2, c_3, s_1, s_2, s_3) \mapsto (c_1, c_3, s_1, s_3)$. Here the unknown c_j represents $\cos(j\theta)$, the unknown s_j represents $\sin(j\theta)$. The curve is cut out by the 2×2-minors of the Toeplitz matrix.

The faces of BN_4 are certain edges and triangles. Its algebraic boundary is the threefold defined by the degree 8 polynomial

$$x_3^2y_1^6 - 2x_1^3x_3y_1^3y_3 + x_1^6y_3^2 + 4x_1^3y_1^3 - 6x_1x_3y_1^4 - 6x_1^4y_1y_3 + 12x_1^2x_3y_1^2y_3 \\ - 2x_3^2y_1^3y_3 - 2x_1^3x_3y_3^2 - 3x_1^2y_1^2 + 4x_3y_1^3 + 4x_1^3y_3 - 6x_1x_3y_1y_3 + x_3^2y_3^2.$$

Conclusion

Spectrahedra and orbitopes deserve to be studied in their own right, independently of their important uses in applications.



A true understanding of these convex bodies will require the integration of three different areas of mathematics:

- Classical Convexity
- Algebraic Geometry
- Optimization Theory