The Algebraic Degree of Semidefinite Programming

BERND STURMFELS
UC BERKELEY AND TU BERLIN

joint work with

Jiawang Nie (Caltech and UCSD)
and Kristian Ranestad (Oslo, Norway)

with a decisive contribution by
Hans-Christian Graf von Bothmer (Hannover, Germany)

Our Question

Semidefinite programming is a numerical method in convex optimization.

SDP is very efficient, both in theory and in practice.

SDP is widely used in engineering and the sciences.

Input: Several symmetric $n \times n$-matrices
Output: One symmetric $n \times n$-matrix

What is the function from the input to the output?
How does the solution depend on the data?

Let's begin with two slides on context.....

Linear Programming

is semidefinite programming for diagonal matrices

The optimal solution of

$$\text{Maximize } c \cdot x \text{ subject to } A \cdot x = b \text{ and } x \geq 0$$

is a piecewise linear function of c and b.

It is a piecewise rational function in the entries of A.

To study the function *data* $\mapsto$ *solution*, one needs geometric combinatorics, namely matroids for the dependence on A and secondary polytopes for b, c.

Universality of Nash Equilibria

Consider the following problem in game theory:

Given payoff matrices, compute the Nash equilibria.

For two players, this is a piecewise linear problem.

For three and more players, it is non-linear.
Datta's Universality Theorem (2003):
Every real algebraic variety is isomorphic to the set of Nash equilibria of some three-person game.

Corollary: The coordinates of the Nash equilibria can be arbitrary algebraic functions of the payoff matrices.

Semidefinite Programming

Given: A matrix C and an m-dimensional affine subspace $\mathcal{U}$ of real symmetric $n \times n$-matrices

SDP Problem:

$$\text{Maximize} \quad \text{trace}(C \cdot X)$$
$$\text{subject to } X \in \mathcal{U} \text{ and } X \succeq 0.$$

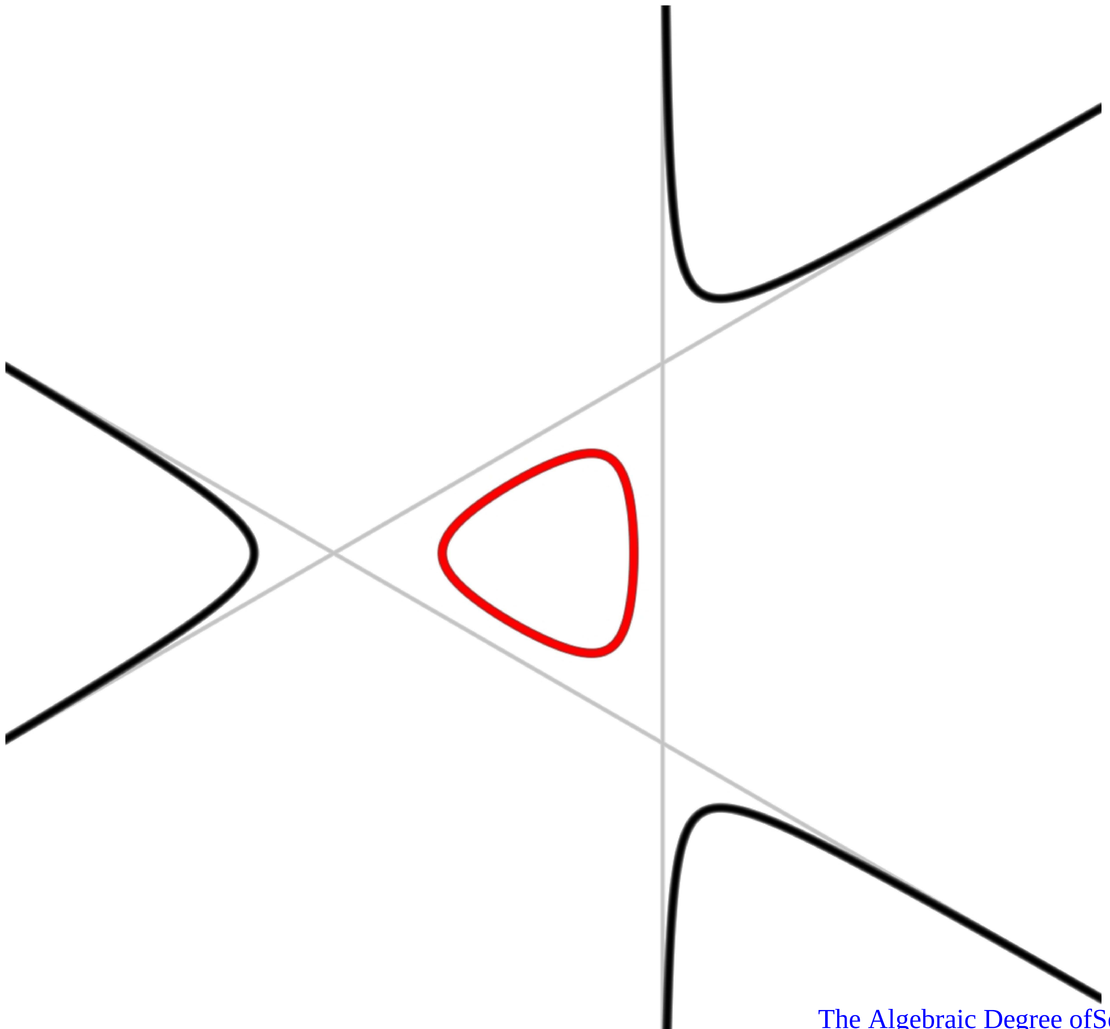
Here $X \succeq 0$ means that X is positive semidefinite.

The problem is feasible if and only if the subspace $\mathcal{U}$ intersects the cone of positive semidefinite matrices.

The optimal solution $\hat{X}$ is a (piecewise) algebraic function of the matrix C and of the subspace $\mathcal{U}$.

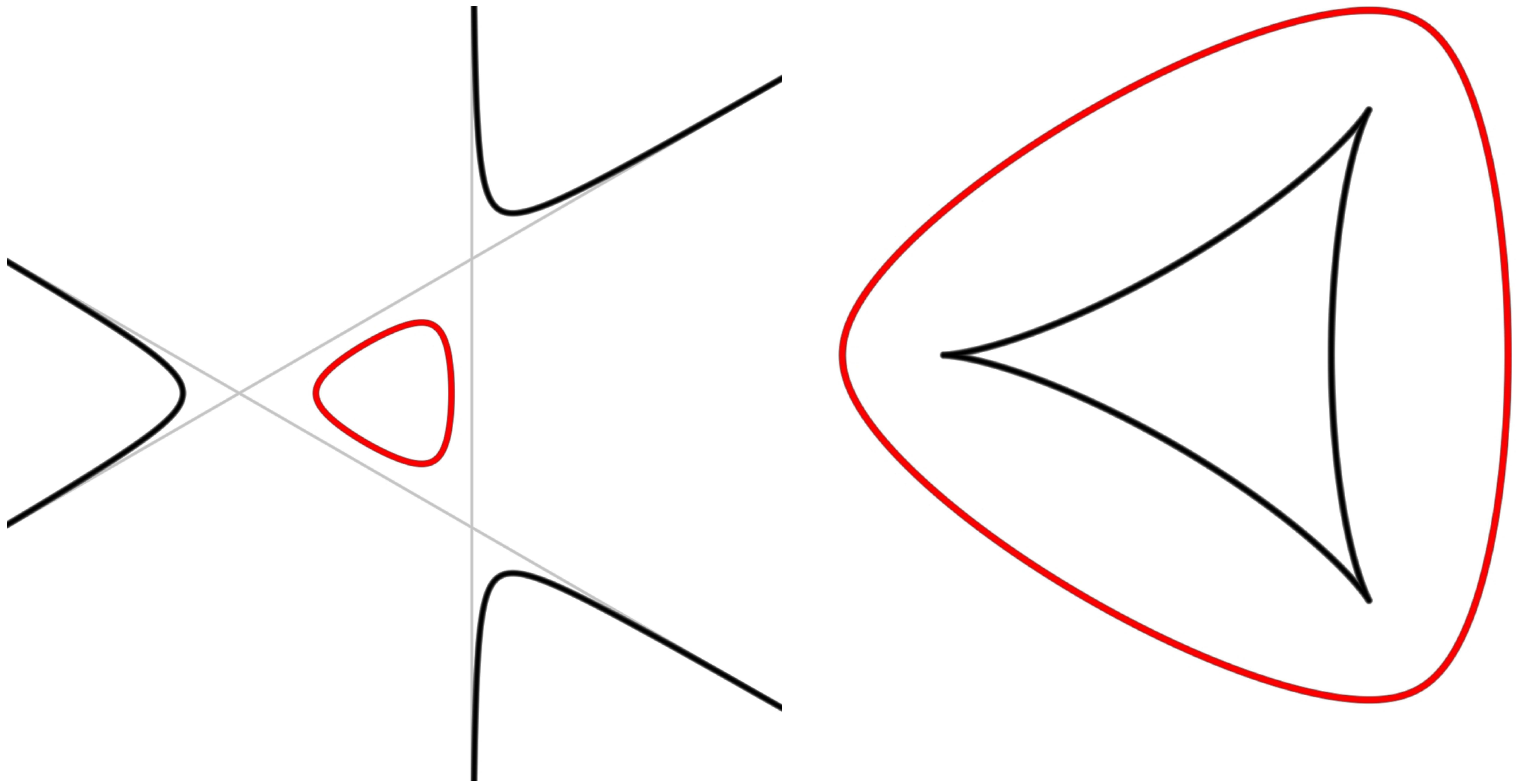
An Elliptic Curve

Let $m = 2$ and $n = 3$. Then $X \succeq 0$ defines a semialgebraic convex region in $\mathcal{U} \simeq \mathbb{R}^2$. It is bounded by the cubic curve $\{\det(X) = 0\}$.



Duality of Plane Curves

The dual to the cubic curve is a curve of degree six:



What does the number 6 mean for our SDP problem?

The Rank Inequalities

We always assume that the data C and $\mathcal{U}$ are generic.

Theorem 1. *(Alizadeh–Haeberly–Overton 1997; Pataki 2000) The rank r of the solution $\hat{X}$ to the SDP problem satisfies the two inequalities*

$$\binom{r+1}{2} \leq \binom{n+1}{2} - m$$

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

For fixed m and n, all ranks r in the specified range are attained for an open set of instances $(\mathcal{U}, C)$.

Distribution of the optimal rank

n	3		4		5		6	
m	rank	percent	rank	percent	rank	percent	rank	percent
3	2	24.00%	3	35.34%	4	79.18%	5	82.78%
	1	76.00%	2	64.66%	3	20.82%	4	17.22%
4			3	23.22%	4	16.96%	5	37.42%
	1	100 %	2	76.78%	3	83.04%	4	62.58%
5					4	5.90%	5	38.42%
	1	100 %	2	100 %	3	94.10%	4	61.58%
6							5	1.32%
			2	67.24%	3	93.50%	4	93.36%
			1	32.76%	2	6.50%	3	5.32%
7			2	52.94%	3	82.64%	4	78.82%
			1	47.06%	2	17.36%	3	21.18%
8					3	34.64%	4	45.62%
			1	100 %	2	65.36%	3	54.38%
9					3	7.60%	4	23.50%
			1	100 %	2	92.40%	3	76.50%

Algebraic Degree of SDP

Suppose that m, n and r satisfy the rank inequalities.

The degree $\delta(m,n,r)$ of the algebraic function $(C,\mathcal{U}) \mapsto \hat{X}$ is the algebraic degree of SDP.

Plane Curves: $\delta(2,n,n-1) \;=\; n(n-1)$

Bigger Example: $\delta(105,20,10) =$

$$1672239271455030620756919692689369762748\,80$$

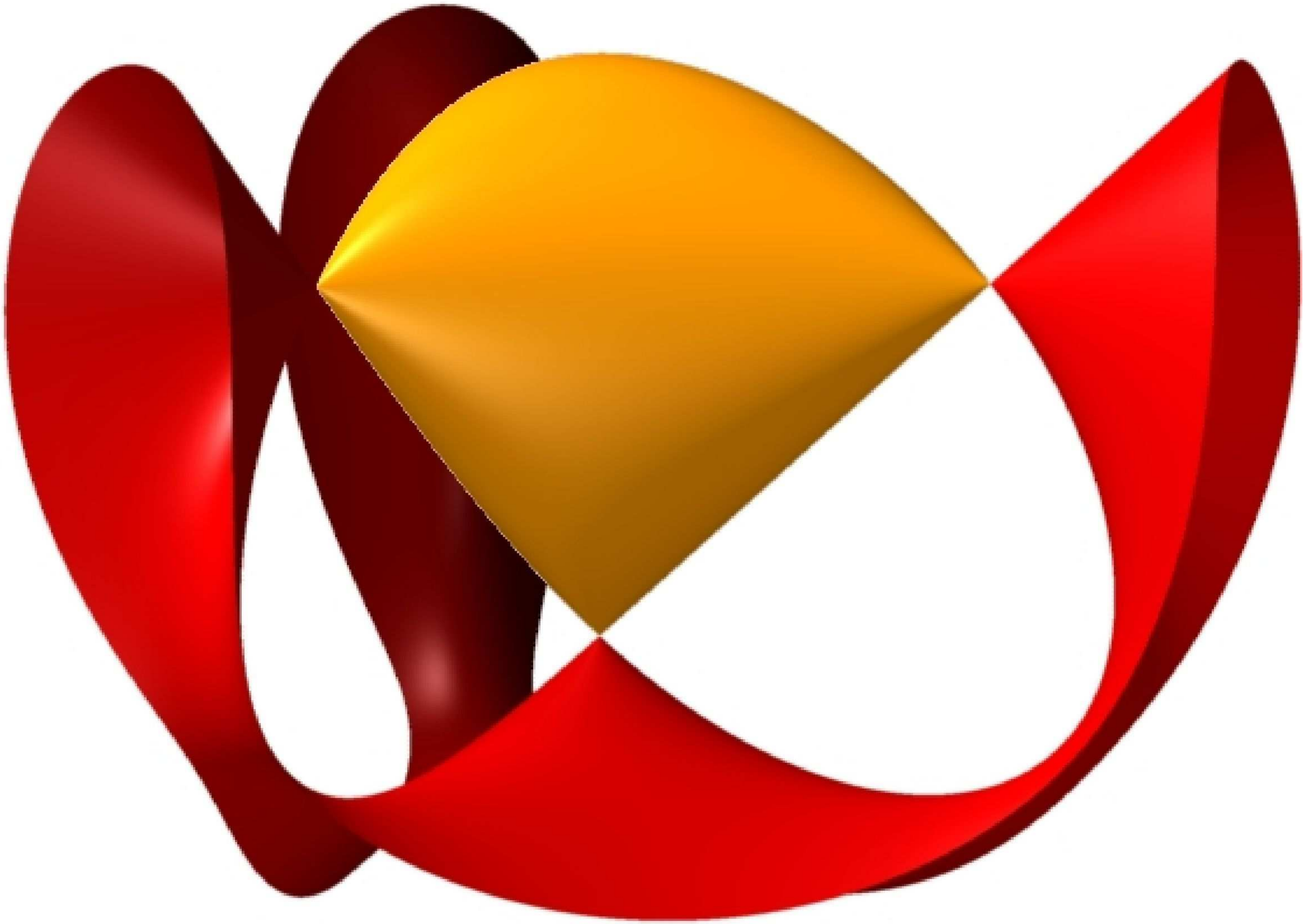
Duality: $\delta(m,n,r) \;=\; \delta(\binom{n+1}{2}-m,\, n\,,n-r).$

Cayley-Steiner: $\delta(3,3,1) \;=\; \delta(3,3,2) \;=\; 4$

m	r	degree	r	degree	r	degree	r	degree	r	degree
1	1	2	2	3	3	4	4	5	5	6
2	1	2	2	**6**	3	12	4	20	5	30
3			2	4	3	16	4	40	5	80
			1	4	2	10	3	20	4	35
4					3	8	4	40	5	120
			1	6	2	30	3	90	4	210
5							4	16	5	96
			1	3	2	42	3	207	4	672
6									5	32
					2	30	3	290	4	1400
					1	8	2	35	3	112
7					2	10	3	260	4	2040
					1	16	2	140	3	672
8							3	140	4	2100
					1	12	2	260	3	1992
9							3	35	4	1470
					1	4	2	290	3	3812

Cayley's Cubic Surface

Let $m = n = 3$. The cubic surface $\det(X) = 0$ is a Cayley cubic, with four singular points...



Analytic Solution

Let $m = n = 3$. The cubic surface $\det(X) = 0$ is a Cayley cubic. Its dual is a quartic Steiner surface.

SDP: Maximize a linear function over the convex region $X \succeq 0$ bounded by the Cayley cubic.

We can express the optimal solution $\hat{X}$ in terms of radicals $\sqrt{}$ using Cardano's formula:

Either $\hat{X}$ has rank one and is one of the four singular points of the Cayley cubic, or $\hat{X}$ has rank two and is found by intersecting the Steiner surface with a line.

Determinantal Varieties

Consider the complex projective space $\mathbb{P}\mathcal{U} \simeq \mathbb{P}^m$.

Let $D_{\mathcal{U}}^r$ denote the variety of all matrices of rank $\leq r$.

Theorem 2. *The codimension of* $D_{\mathcal{U}}^r$ *is* $\binom{n-r+1}{2}$.
If $m > \binom{n-r+1}{2}$ *then* $D_{\mathcal{U}}^r$ *is irreducible.*

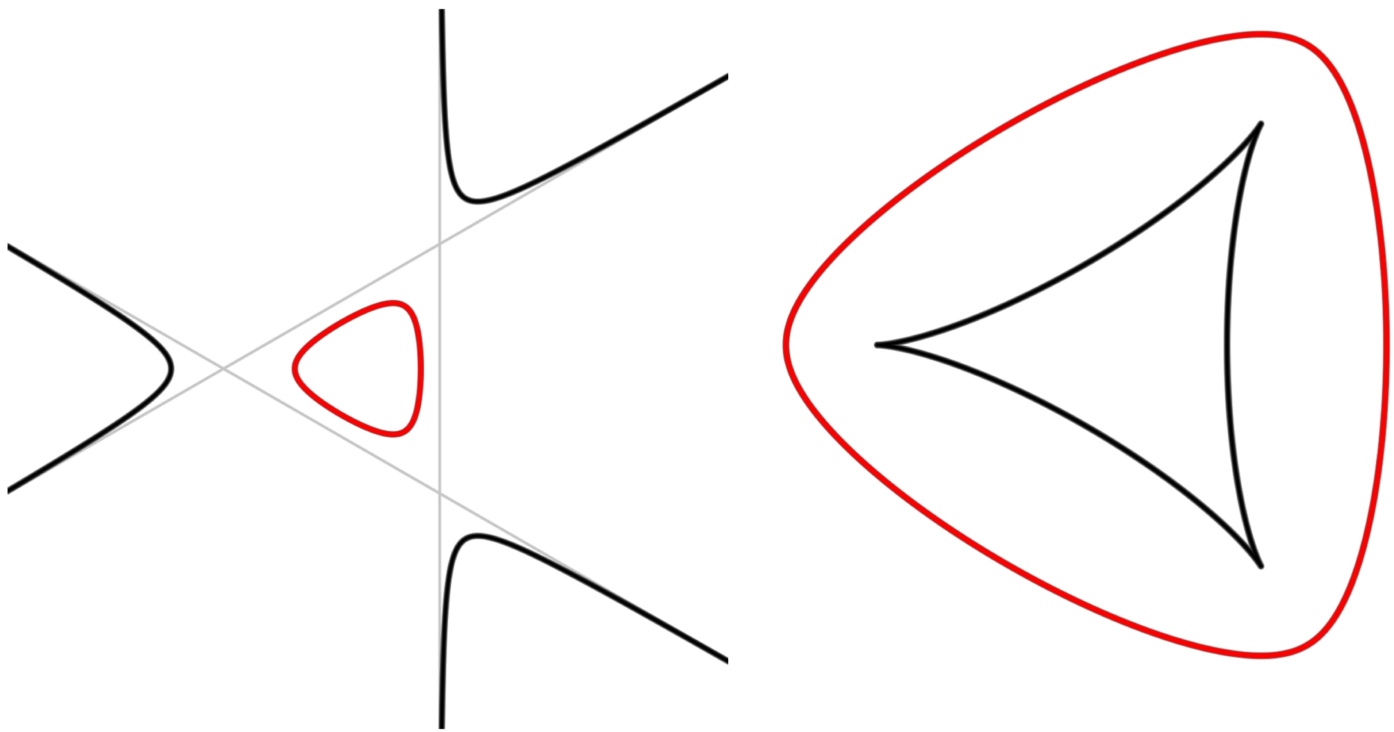
The singular locus of $D_{\mathcal{U}}^r$ *equals* $D_{\mathcal{U}}^{r-1}$, *and*

$$\mathrm{degree}(D_{\mathcal{U}}^r) \quad = \quad \prod_{j=0}^{n-r-1} \frac{\binom{n+j}{n-r-j}}{\binom{2j+1}{j}}$$

Projective Duality

Let $\mathbb{P}\mathcal{U}^*$ denote the dual projective space to $\mathbb{P}\mathcal{U}$.
The points in $\mathbb{P}\mathcal{U}^*$ correspond to hyperplanes in $\mathbb{P}\mathcal{U}$.

Any variety $\mathcal{V} \subset \mathbb{P}\mathcal{U}$ has a dual variety $\mathcal{V}^* \subset \mathbb{P}\mathcal{U}^*$.
$\mathcal{V}^*$ is the Zariski closure of the set of all hyperplanes in $\mathbb{P}\mathcal{U}$ that are tangent to $\mathcal{V}$ at a smooth point.



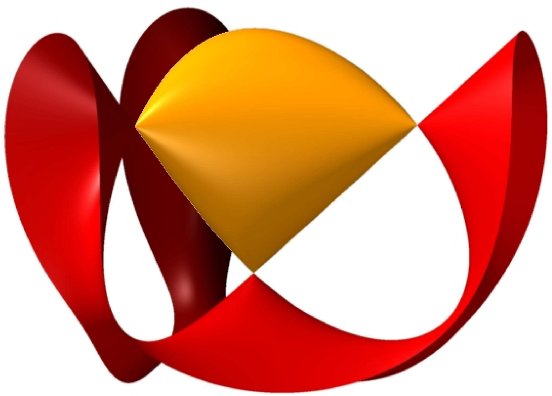
The Dual Hypersurface

Lemma 3. *If* $m = \binom{n+1}{2}$ *then the projective dual of* $D^r_{\mathcal{U}}$ *equals the complementary determinantal variety:*

$$(D^r_{\mathcal{U}})^* \quad = \quad D^{n-r}_{\mathcal{U}^*}$$

Theorem 4. *The variety* $D^r_{\mathcal{U}}$ *is non-degenerate if and only if the rank inequalities hold. the algebraic degree of SDP is the degree of the dual hypersurface:*

$$\delta(m,n,r) \quad = \quad \text{degree}\,(D^r_{\mathcal{U}})^*$$



Two Matrices with Product Zero

Theorem 5. *Let $\mathcal{Q}^{\{r\}}$ be the variety of pairs (X, Y) of symmetric $n \times n$-matrices with $X \cdot Y = 0$, $\operatorname{rank}(X) = r$ and $\operatorname{rank}(Y) = n - r$.*
The bidegree of $\mathcal{Q}^{\{r\}}$ equals the generating function for the algebraic degree of semidefinite programming:

$$\mathcal{C}(\mathcal{Q}^{\{r\}}; s, t) \quad = \quad \sum_{m=0}^{\binom{n+1}{2}} \delta(m, n, r) \cdot s^{\binom{n+1}{2} - m} \cdot t^m.$$

Setting $s = t = 1$ we get the scalar degree of $\mathcal{Q}^{\{r\}}$:

$$\begin{aligned} \mathcal{C}(\mathcal{Q}^{\{3\}}; 1, 1) &= 4 + 12 + 16 + 8 = 40 \\ \mathcal{C}(\mathcal{Q}^{\{2\}}; 1, 1) &= 10 + 30 + 42 + 30 + 10 = 122 \end{aligned}$$

Check with Macaulay 2

```
R = QQ[x11,x12,x13,x14,x22,x23,x24,x33,x34,x44,
       y11,y12,y13,y14,y22,y23,y24,y33,y34,y44];

X = matrix {{x11, x12, x13, x14},
            {x12, x22, x23, x24},
            {x13, x23, x33, x34},
            {x14, x24, x34, x44}};

Y = matrix {{y11, y12, y13, y14},
            {y12, y22, y23, y24},
            {y13, y23, y33, y34},
            {y14, y24, y34, y44}};

Q3 = minors(1,X*Y) + minors(4,X) + minors(2,Y);
codim Q3, degree Q3
```

(10, 40)

```
Q2 = minors(1,X*Y) + minors(3,X) + minors(3,Y);
codim Q2, degree Q2
```

(10, 122)

Bothmer-Ranestad Formula

Define a skew-symmetric matrix $(\psi_{ij})_{0 \le i < j \le n}$ by

$$\psi_{0j} \;=\; 2^{j-1} \quad \text{and} \quad \psi_{ij} \;=\; \sum_{k=i}^{j-1} \binom{i+j-2}{k}$$

For any subset $I = \{i_1, \ldots, i_r\}$ of $\{1, \ldots, n\}$ let ψ_I denote the sub-Pfaffian of (ψ_{ij}) indexed by I if $|I|$ is even and by $I \cup \{0\}$ if $|I|$ is odd.

Theorem 7. $$\delta(m, n, r) \quad = \quad \sum_I \psi_I \cdot \psi_{I^c}$$

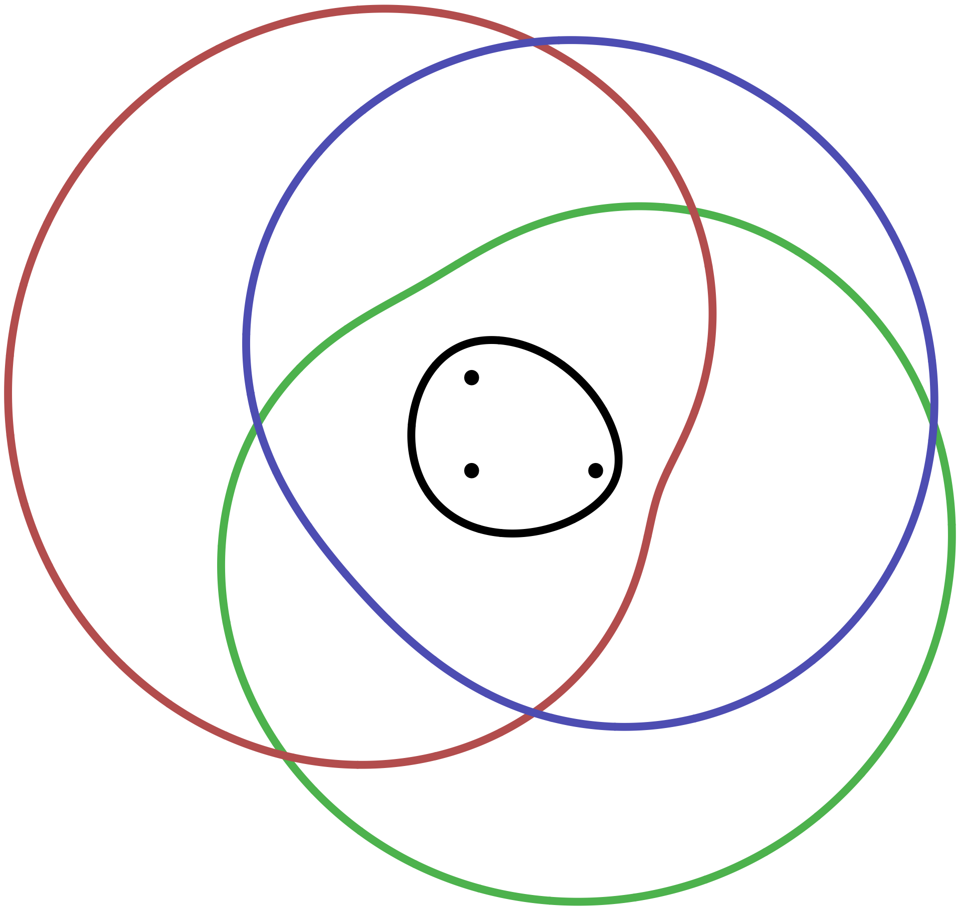
where $I = \{i_1, \ldots, i_r\}$ runs over all r-subsets of $\{1, 2, \ldots, n\}$ with $i_1 + \cdots + i_r \;=\; \binom{n+1}{2} - m$.

Reminder on Genericity

In this talk the subspace $\mathcal{U}$ was always generic.

For special instances, the algebraic degree is smaller.

Example: The 3-ellipse is a Helton-Vinnikov octic



What's the degree of its dual? Hint: $< \delta(2,8,7) = 56$.

Conclusion

Conclusion for Applied Mathematicians:

Algebraic Geometry might be useful.

Conclusion for Pure Mathematicians:

Optimization might be interesting.