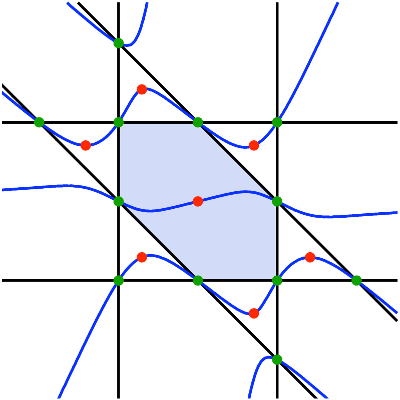
The Central Curve in Linear Programming

Bernd Sturmfels
UC Berkeley and MATHEON Berlin



joint work with Jesus De Loera and Cynthia Vinzant

Linear Programming

primal : Maximize $\mathbf{c}^T\mathbf{x}$ subject to $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq 0$

dual : Minimize $\mathbf{b}^T\mathbf{y}$ subject to $A^T\mathbf{y} - \mathbf{s} = \mathbf{c}$ and $\mathbf{s} \geq 0$

A is a fixed matrix of rank d having n columns,.
The vectors $\mathbf{c} \in \mathbb{R}^n$ and $\mathbf{b} \in \mathrm{image}(A)$ may vary.

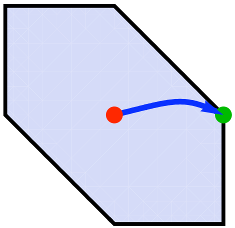
For any $\lambda > 0$, the *logarithmic barrier function* for the primal is

$$f_\lambda(\mathbf{x}) \;:=\; \mathbf{c}^T\mathbf{x} + \lambda \sum_{i=1}^{n} \log x_i,$$

This function is concave. Let $\mathbf{x}^*(\lambda)$ be the unique solution of

barrier : Maximize $f_\lambda(\mathbf{x})$ subject to $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq 0$

The Central Path



The set of all (primal) feasible solutions is a convex polytope

$$P \quad = \quad \left\{ \mathbf{x} \in \mathbb{R}^n_{\geq 0} : A\mathbf{x} = \mathbf{b} \right\}.$$

The logarithmic barrier function $f_\lambda(\mathbf{x})$ is defined on the relative interior of P. It tends to $-\infty$ when $\mathbf{x}$ approaches the boundary.

The *primal central path* is the curve $\left\{ \mathbf{x}^*(\lambda) \mid \lambda > 0 \right\}$ inside P. It connects the *analytic center* of P with the optimal solution:

$$\mathbf{x}^*(\infty) \longrightarrow \cdots \longrightarrow \mathbf{x}^*(\lambda) \longrightarrow \cdots \longrightarrow \mathbf{x}^*(0).$$

Complementary Slackness

Optimal primal and dual solutions are characterized by

$$\begin{array}{c} A\mathbf{x} = \mathbf{b}\,,\;\; A^T\mathbf{y} - \mathbf{s} = \mathbf{c}\,,\;\; \mathbf{x} \geq 0\,,\;\; \mathbf{s} \geq 0, \\ \text{and} \quad x_i \cdot s_i \;=\; 0 \;\; \text{for} \;\; i = 1, 2, \ldots, n. \end{array} \tag{1}$$

In textbooks on Linear Programming we find

Theorem

For all $\lambda > 0$, the system of polynomial equations

$$A\mathbf{x} = \mathbf{b}\,,\;\; A^T\mathbf{y} - \mathbf{s} = \mathbf{c},\;\; \textit{and}\;\; x_i s_i = \lambda \;\; \textit{for} \;\; i = 1, 2, \ldots, n,$$

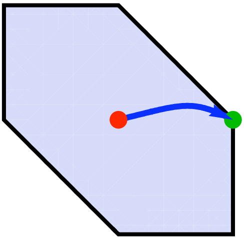
has a unique real solution $(\mathbf{x}^*(\lambda), \mathbf{y}^*(\lambda), \mathbf{s}^*(\lambda))$ with $\mathbf{x}^*(\lambda) > 0$ and $\mathbf{s}^*(\lambda) > 0$. The point $\mathbf{x}^*(\lambda)$ solves the barrier problem.

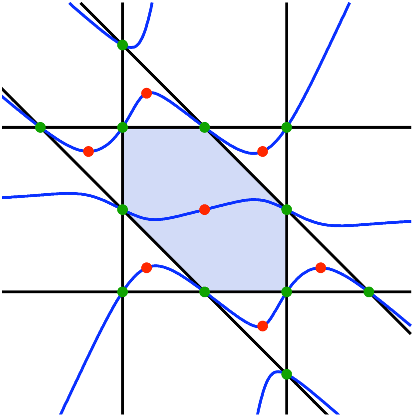
The limit point $(\mathbf{x}^*(0), \mathbf{y}^*(0), \mathbf{s}^*(0))$ uniquely solves (1).

Our Contributions

Bayer-Lagarias (1989) showed that the central path is an algebraic curve, and they suggested the problem of identifying its prime ideal.

We resolve this problem.

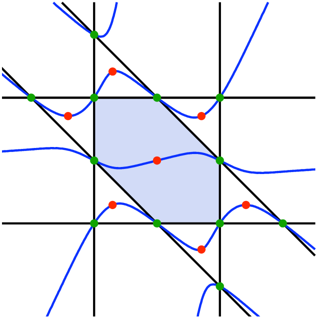




The central curve is the **Zariski closure** of the central path.

Our Contributions

The central curve is the union of the central paths over all polyhedra in the hyperplane arrangement:



Dedieu-Malajovich-Shub (2005) studied the global curvature of the central path, by bounding the degree of corresponding Gauss curve.

We offer a refined bound.

Curvature is important for numerical interior point methods.
Deza-Terlaky-Zinchenko (2008): continuous Hirsch conjecture.

Central Curves in the Plane

The dual problem for $d = 2$ is

$$\text{Minimize} \;\; \mathbf{b}^T\mathbf{y} \;\; \text{subject to} \;\; A^T\mathbf{y} \geq \mathbf{c}$$

The central curve has the parametric representation

$$\mathbf{y}^*(\lambda) \;=\; \mathrm{argmin}_{\mathbf{y}\in\mathbb{R}^2} \; b_1y_1 + b_2y_2 - \lambda \sum_{i=1}^{n} \log(a_{1i}y_1 + a_{2i}y_2 - c_i).$$

Its defining polynomial is

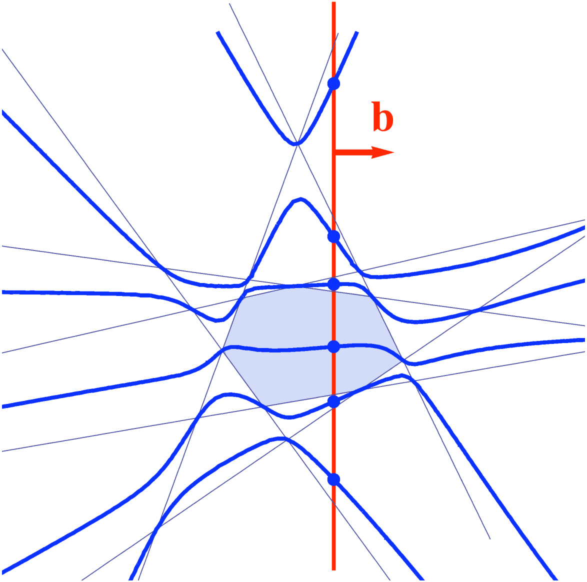
$$\mathcal{C}(y_1, y_2) \;=\; \sum_{i\in\mathcal{I}} (b_1a_{2i} - b_2a_{1i}) \prod_{j\in\mathcal{I}\backslash\{i\}} (a_{1j}y_1 + a_{2j}y_2 - c_j),$$

where $\mathcal{I} = \{i \;:\; b_1a_{2i} - b_2a_{1i} \neq 0\}$. The degree equals $|\mathcal{I}| - 1$.

Proposition

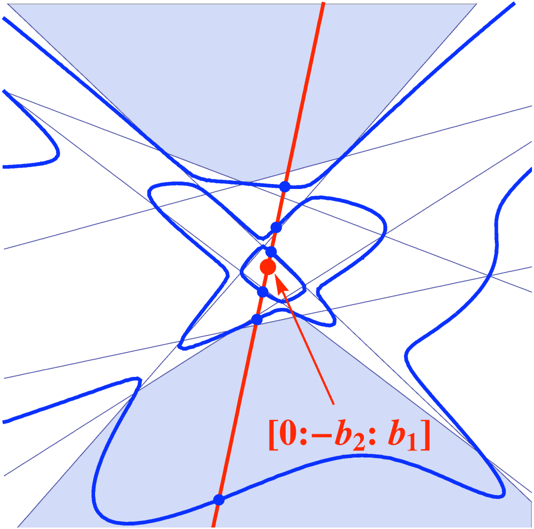
The central curve $\mathcal{C}$ is hyperbolic with respect to the point $[0 : -b_2 : b_1]$. This means that every line in $\mathbb{P}^2(\mathbb{R})$ passing through this special point meets $\mathcal{C}$ only in real points.

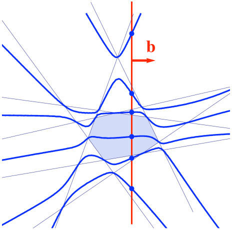
Sextic Central Curve



.... obtained as the *polar curve* of an arrangement of $n = 7$ lines.

After a Projective Transformation





hyperbolic curve = Vinnikov curve → Spectrahedron

Inflection Points

The number of inflection points of a plane curve of degree D in $\mathbb{P}^2_{\mathbb{C}}$ is at most $3D(D-2)$.

Felix Klein (1876) proved: *The number of real inflection points of a plane curve of degree D is at most $D(D-2)$.*

Theorem: The average total curvature of a central curve in the plane is at most 2π.

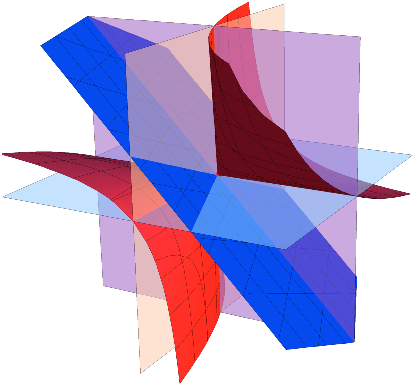
Dedieu (2005) et. al. had the bound 4π.

Question: What is the largest number of inflection points on a single oval of a hyperbolic curve of degree D in the real plane?

In particular, is this number *linear* in the degree D?

Central Sheet

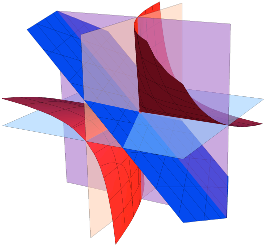
Let $K = \mathbb{Q}(A)(\mathbf{b}, \mathbf{c})$ and $\mathcal{L}_{A,\mathbf{c}}$ the subspace of K^n spanned by the rows of A and the vector $\mathbf{c}$.



Define the *central sheet* to be its coordinatewise reciprocal. Denoted $\mathcal{L}_{A,\mathbf{c}}^{-1}$, this is the Zariski closure of the set

$$\left\{ \left(\frac{1}{u_1}, \ldots, \frac{1}{u_n}\right) \in \mathbb{C}^n \ : \ (u_1, \ldots, u_n) \in \mathcal{L}_{A,\mathbf{c}} \ \text{and} \ u_i \neq 0 \ \text{for all} \ i \right\}$$

Equations



Lemma

The primal central curve $\mathcal{C}$ equals the intersection of the central sheet $\mathcal{L}_{A,\mathbf{c}}^{-1}$ with the affine space $\{A \cdot \mathbf{x} = \mathbf{b}\}$.

Theorem

The prime ideal of polynomials that vanish on the central curve $\mathcal{C}$ is $\langle A\mathbf{x} - \mathbf{b}\rangle + J_{A,\mathbf{c}}$, where $J_{A,\mathbf{c}}$ is the ideal of the central sheet $\mathcal{L}_{A,\mathbf{c}}^{-1}$. The common degree of $\mathcal{C}$ and $\mathcal{L}_{A,\mathbf{c}}^{-1}$ is the Möbius number $|\mu(A, \mathbf{c})|$.

Proudfoot and Speyer (2006) found a universal Gröbner basis for the prime ideal $J_{A,\mathbf{c}}$ of the central sheet $\mathcal{L}_{A,\mathbf{c}}^{-1}$.

We use this to answer the question of Bayer and Lagarias (1989).

Details

The universal Gröbner basis of $J_{A,\mathbf{c}}$ consists of the polynomials

$$\sum_{i \in \mathrm{supp}(v)} v_i \cdot \prod_{j \in \mathrm{supp}(v) \backslash \{i\}} x_j ,$$

where $(v_1, \ldots, v_n)$ runs over the cocircuits of the linear space $\mathcal{L}_{A,\mathbf{c}}$.

Cocircuits means non-zero vectors of minimal support.

The Möbius number $|\mu(A, \mathbf{c})|$ is an invariant from matroid theory. It gives the degree of the central curve. If A and $\mathbf{c}$ are generic then

$$|\mu(A, \mathbf{c})| \;=\; \binom{n-1}{d}.$$

Proudfoot and Speyer (2006) also determine the Hilbert series ....

Example: 2×3 Transportation Problem

Let $d = 4, n = 6$ and A the linear map taking a matrix $\binom{x_1\, x_2\, x_3}{x_4\, x_5\, x_6}$ to its row and column sums. The ideal of the affine subspace is

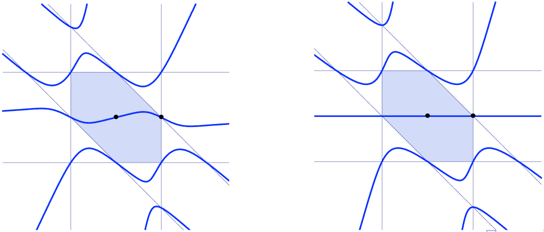
$$I_{A,\mathbf{b}} \;=\; \langle\, x_1+x_2+x_3-b_1\,,\; x_4+x_5+x_6-b_2\,,\; x_1+x_4-b_3\,,\; x_2+x_5-b_4\,\rangle.$$

The central sheet $\mathcal{L}^{-1}_{A,\mathbf{c}}$ is the quintic hypersurface

$$f_{A,\mathbf{c}}(\mathbf{x}) \;=\; \det\begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ c_1 & c_2 & c_3 & c_4 & c_5 & c_6 \\ x_1^{-1} & x_2^{-1} & x_3^{-1} & x_4^{-1} & x_5^{-1} & x_6^{-1} \end{pmatrix} \cdot x_1x_2x_3x_4x_5x_6.$$

The central curve is defined by $I_{A,\mathbf{b}} + \langle f_{A,\mathbf{c}}\rangle$.

It is irreducible for any **b** as long as **c** is generic:



Applying the Gauss Map

Dedieu-Malajovich-Shub (2005): The total curvature of any real algebraic curve $\mathcal{C}$ in $\mathbb{R}^m$ is the arc length of its image under the Gauss map $\gamma : \mathcal{C} \to \mathbb{S}^{m-1}$. This quantity is bounded above by π times the degree of the projective Gauss curve in $\mathbb{P}^{m-1}$. In symbols,

$$\int_a^b \left|\left|\frac{d\gamma(t)}{dt}\right|\right| dt \;\leq\; \pi \cdot \deg(\gamma(\mathcal{C})).$$

Our Theorem: *The degree of the projective Gauss curve of the central curve $\mathcal{C}$ satisfies a bound in terms of matroid invariants:*

$$\deg(\gamma(\mathcal{C})) \;\leq\; 2 \cdot \sum_{i=1}^{d} i \cdot h_i \;\leq\; 2 \cdot (n-d-1) \cdot \binom{n-1}{d-1}.$$

$(h_0, h_1, \ldots, h_d)$ = h-vector of the broken circuit complex of $\mathcal{L}_{A,\mathbf{c}}$

Example

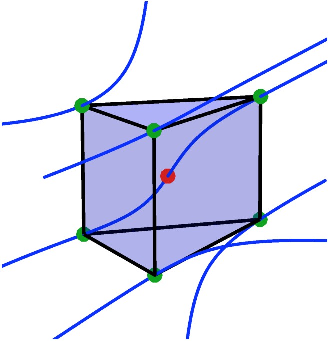
$n = 5,\ d = 2.$

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \qquad \mathbf{b} = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \qquad \mathbf{c} = \begin{pmatrix} 1 & 2 & 0 & 4 & 0 \end{pmatrix}$$

Equations for $\mathcal{C}$:

$2x_2x_3 - x_1x_3 - x_1x_2,$
$4x_2x_4x_5 - 4x_1x_4x_5 + x_1x_2x_5 - x_1x_2x_4,$
$4x_3x_4x_5 - 4x_1x_4x_5 - x_1x_3x_5 + x_1x_3x_4,$
$4x_3x_4x_5 - 4x_2x_4x_5 - 2x_2x_3x_5 + 2x_2x_3x_4,$

$x_1 + x_2 + x_3 = 3,\ x_4 + x_5 = 2$



$$h = (1,2,2) \ \Rightarrow\ deg(\mathcal{C}) = 5 \ \text{ and } \ \deg(\gamma(\mathcal{C})) \leq 12$$

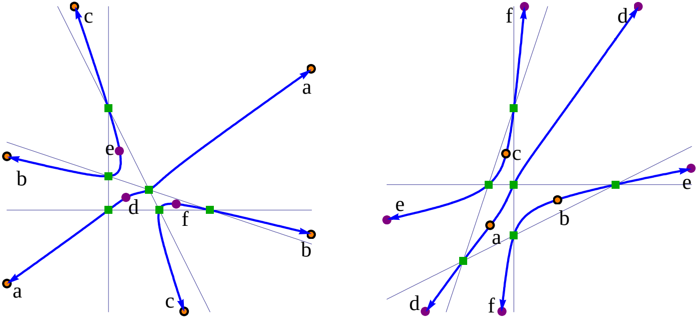
Primal-Dual Curve

Let $\mathcal{L}$ denote the row space of the matrix A and $\mathcal{L}^\perp$ its orthogonal complement in $\mathbb{R}^n$. Fix a vector $\mathbf{g} \in \mathbb{R}^n$ such that $A\mathbf{g} = \mathbf{b}$.

The primal-dual central path $(\mathbf{x}^*(\lambda), \mathbf{s}^*(\lambda))$ has the following description that is symmetric under duality:

$$\mathbf{x} \in \mathcal{L}^\perp + \mathbf{g}\,, \quad \mathbf{s} \in \mathcal{L} + \mathbf{c} \quad \text{and} \quad x_1 s_1 = x_2 s_2 = \cdots = x_n s_n = \lambda.$$

These equations define an irreducible curve in $\mathbb{P}^n \times \mathbb{P}^n$.



Analytic Centers

Consider the dual pair of hyperplane arrangements

$$\mathcal{H} \;=\; \{x_i = 0\}_{i\in[n]} \;\text{ in }\; \mathcal{L}^{\perp} + \mathbf{g} \;\subset\; \mathbb{P}^n \backslash \{x_0 = 0\}$$

$$\mathcal{H}^* \;=\; \{s_i = 0\}_{i\in[n]} \;\text{ in }\; \mathcal{L} + \mathbf{c} \;\subset\; \mathbb{P}^n \backslash \{s_0 = 0\}$$

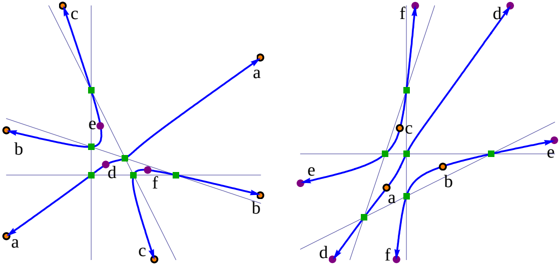
Proposition

The intersection $\mathcal{L}^{-1} \cap (\mathcal{L}^{\perp} + \mathbf{g})$ is a zero-dimensional variety. All its points are defined over $\mathbb{R}$. They are the analytic centers of the polytopes that form the bounded regions of the arrangement $\mathcal{H}$.

Proposition

The intersection $(\mathcal{L}^{\perp})^{-1} \cap (\mathcal{L} + \mathbf{c})$ is a zero-dimensional variety. All points are defined over $\mathbb{R}$. They are the analytic centers of the polytopes that form the bounded regions of the arrangement $\mathcal{H}^*$.

Global Geometry



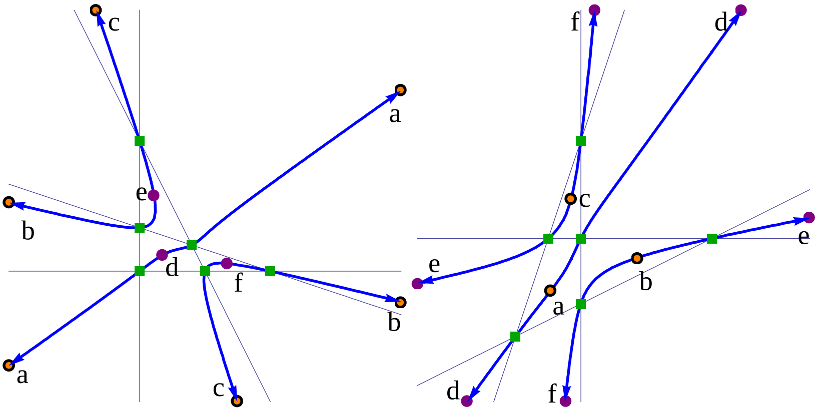
Theorem

The primal central curve in **x***-space* $\mathbb{R}^n$ *passes through all vertices of* $\mathcal{H}$*. In between these vertices, it passes through the analytic centers of the bounded regions. Similarly, the dual central curve in* **s***-space passes through all vertices and analytic centers of* $\mathcal{H}^*$*.*

Along the curve, vertices of $\mathcal{H}$ *correspond to vertices of* $\mathcal{H}^*$*.*

The analytic centers of bounded regions of $\mathcal{H}$ *correspond to points on the dual curve in* **s***-space at the hyperplane* $\{s_0 = 0\}$*, and the analytic centers of bounded regions of* $\mathcal{H}^*$ *correspond to points on the primal curve in* **x***-space at the hyperplane* $\{x_0 = 0\}$*.*

A Curve in $\mathbb{P}^2 \times \mathbb{P}^2$



Conclusion

... for Pure Mathematicians: Optimization is Beautiful.

... for Applied Mathematicians: Algebraic Geometry is Useful.

Is there a difference between “Pure” and “Applied” ?