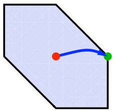
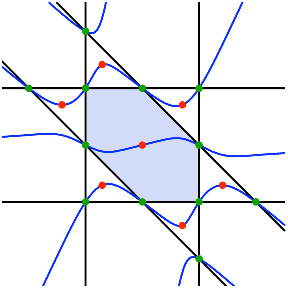
Recent Advances in the Geometry of Linear Programming

Jesús A. De Loera



$\longrightarrow$



Based on joint work with : (1) S. Klee (Math of OR), (2) B. Sturmfels, and C. Vinzant (Found. Comput. Math.).

The classical **Linear Programming Problem:**

$$\text{maximize } C_1x_1 + C_2x_2 + \cdots + C_dx_d$$

among all $x_1, x_2, \ldots, x_d,$ **satisfying:**

$$a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,d}x_d \leq b_1$$
$$a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,d}x_d \leq b_2$$
$$\vdots$$
$$a_{k,1}x_1 + a_{k,2}x_2 + \cdots + a_{k,d}x_d \leq b_k$$

Linear Equations are included in this model.

This reduces to the canonical problem

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

NOTE: Set of possible solutions is a **convex polyhedron**

This reduces to the canonical problem

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

NOTE: Set of possible solutions is a **convex polyhedron**

Mathematically speaking the solution of the **optimization problem** is equivalent to the solution of the **feasibility problem**:

Is there a vector $\mathbf{x}$, with $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$?

This reduces to the canonical problem

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

NOTE: Set of possible solutions is a **convex polyhedron**

Mathematically speaking the solution of the **optimization problem** is equivalent to the solution of the **feasibility problem**:

Is there a vector $\mathbf{x}$, with $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$?

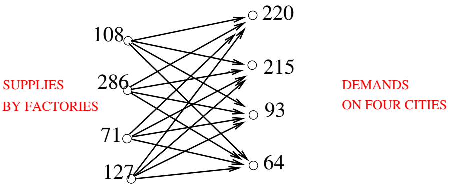
Linear Programming

$\Updownarrow$

Linear Algebra over NON-NEGATIVE reals

Transportation LPs

- My favorite LP problem: A company builds laptops in four factories, each with certain supply power. Four cities have laptop demands. There is a cost $c_{i,j}$ for transporting a laptop from factory i to city j. What is the best assignment of transport in order to minimize the cost?



- Linear programming is important for optimization:
  - Linear programs used in solution/approximation schemes for combinatorial and non-linear optimization.

- Linear programming is important for optimization:
  - Linear programs used in solution/approximation schemes for combinatorial and non-linear optimization.
  - Exciting new applications keep coming (e.g., *compressed sensing*!!).

- Linear programming is important for optimization:
  - Linear programs used in solution/approximation schemes for combinatorial and non-linear optimization.
  - Exciting new applications keep coming (e.g., *compressed sensing*!!).
- Impact of linear optimization goes well beyond optimization and reaches other areas of mathematics:
  - Combinatorics and graph theory.

- Linear programming is important for optimization:
  - Linear programs used in solution/approximation schemes for combinatorial and non-linear optimization.
  - Exciting new applications keep coming (e.g., *compressed sensing*!!).
- Impact of linear optimization goes well beyond optimization and reaches other areas of mathematics:
  - Combinatorics and graph theory.
  - Statistical Regression.

- Linear programming is important for optimization:
  - Linear programs used in solution/approximation schemes for combinatorial and non-linear optimization.
  - Exciting new applications keep coming (e.g., *compressed sensing*!!).
- Impact of linear optimization goes well beyond optimization and reaches other areas of mathematics:
  - Combinatorics and graph theory.
  - Statistical Regression.
  - Geometry (e.g., solution of Kepler's conjecture)

- Linear programming is important for optimization:
  - Linear programs used in solution/approximation schemes for combinatorial and non-linear optimization.
  - Exciting new applications keep coming (e.g., *compressed sensing*!!).
- Impact of linear optimization goes well beyond optimization and reaches other areas of mathematics:
  - Combinatorics and graph theory.
  - Statistical Regression.
  - Geometry (e.g., solution of Kepler's conjecture)
- Many different methods of solution known!!
  - Fourier-Motzkin elimination

- Linear programming is important for optimization:
  - Linear programs used in solution/approximation schemes for combinatorial and non-linear optimization.
  - Exciting new applications keep coming (e.g., *compressed sensing*!!).
- Impact of linear optimization goes well beyond optimization and reaches other areas of mathematics:
  - Combinatorics and graph theory.
  - Statistical Regression.
  - Geometry (e.g., solution of Kepler's conjecture)
- Many different methods of solution known!!
  - Fourier-Motzkin elimination
  - Simplex Method

- Linear programming is important for optimization:
  - Linear programs used in solution/approximation schemes for combinatorial and non-linear optimization.
  - Exciting new applications keep coming (e.g., *compressed sensing*!!).
- Impact of linear optimization goes well beyond optimization and reaches other areas of mathematics:
  - Combinatorics and graph theory.
  - Statistical Regression.
  - Geometry (e.g., solution of Kepler's conjecture)
- Many different methods of solution known!!
  - Fourier-Motzkin elimination
  - Simplex Method
  - Ellipsoid Method and its relatives

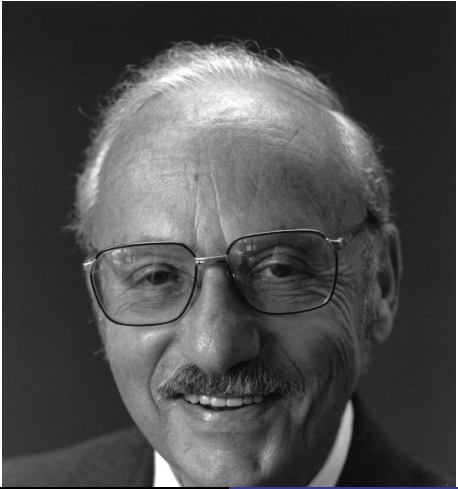
- Linear programming is important for optimization:
  - Linear programs used in solution/approximation schemes for combinatorial and non-linear optimization.
  - Exciting new applications keep coming (e.g., *compressed sensing*!!).
- Impact of linear optimization goes well beyond optimization and reaches other areas of mathematics:
  - Combinatorics and graph theory.
  - Statistical Regression.
  - Geometry (e.g., solution of Kepler's conjecture)
- Many different methods of solution known!!
  - Fourier-Motzkin elimination
  - Simplex Method
  - Ellipsoid Method and its relatives
  - Interior Point Methods

- Linear programming is important for optimization:
  - Linear programs used in solution/approximation schemes for combinatorial and non-linear optimization.
  - Exciting new applications keep coming (e.g., *compressed sensing*!!).
- Impact of linear optimization goes well beyond optimization and reaches other areas of mathematics:
  - Combinatorics and graph theory.
  - Statistical Regression.
  - Geometry (e.g., solution of Kepler's conjecture)
- Many different methods of solution known!!
  - Fourier-Motzkin elimination
  - Simplex Method
  - Ellipsoid Method and its relatives
  - Interior Point Methods
  - Others...

PART I: Advances in the Simplex Method

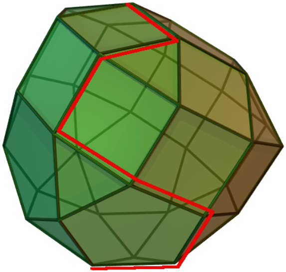
Selected one of *"10 top most-influential algorithms in the 20th century"* by SIAM news

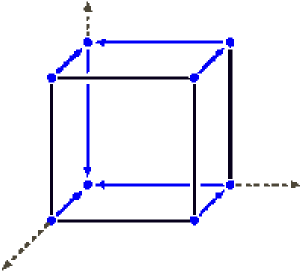
George Dantzig, inventor of the simplex algorithm



The simplex method

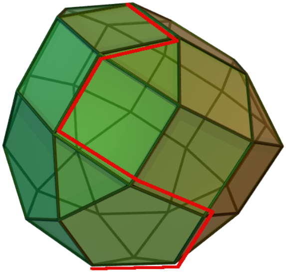
- **Lemma:** Feasible region is a polyhedron. An optimal solution for an LP is among the vertices of the polytope.
- The simplex method **walks** or searches along the graph of the polytope, each time moving to a better and better cost!

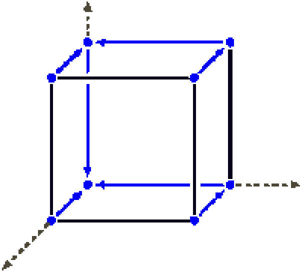




The simplex method

- **Lemma:** Feasible region is a polyhedron. An optimal solution for an LP is among the vertices of the polytope.
- The simplex method **walks** or searches along the graph of the polytope, each time moving to a better and better cost!





- Performance of the simplex method depends on the diameter of the graph of the polytope: largest distance between any pair of nodes.

Bounding the Diameter

QUESTION: Is there a polynomial bound of the diameter in terms of the number of facets and the dimension? The best general bounds are

Bounding the Diameter

QUESTION: Is there a polynomial bound of the diameter in terms of the number of facets and the dimension? The best general bounds are

Barnette-Larman: $\frac{2^{dim(P)-2}}{3}(\#facets(P) - dim(P) + 5/2)$.

Kalai-Kleitman: $(\#facets(P))^{log(dim(P))+1}$.

The Hirsch Conjecture on a Polytopes

The story so far:

- **1957:** Hirsch proposes that a d-dimensional polytope with n facets has diameter (maximal distance between any two vertices) no more than $n - d$.

The Hirsch Conjecture on a Polytopes

The story so far:

- **1957:** Hirsch proposes that a d-dimensional polytope with n facets has diameter (maximal distance between any two vertices) no more than $n - d$.
- **2010**: Santos constructs a counterexample (43 dimensions and 86 facets)

Some remarks

- **2012** B. Matschke, F. Santos, & C. Weibel improved the counterexample, now smaller in 20 dimensions and one can actually plug it into a computer!

Some remarks

- **2012** B. Matschke, F. Santos, & C. Weibel improved the counterexample, now smaller in 20 dimensions and one can actually plug it into a computer!
- The diameter gives a lower bound on the worst-case behavior of edge-following algorithms. If diameter is exponential then all edge-path algorithms will be exponential in the worst case.

Some remarks

- **2012** B. Matschke, F. Santos, & C. Weibel improved the counterexample, now smaller in 20 dimensions and one can actually plug it into a computer!
- The diameter gives a lower bound on the worst-case behavior of edge-following algorithms. If diameter is exponential then all edge-path algorithms will be exponential in the worst case.
- Hirsch conjecture known to be true in many instances, e.g. for polytopes with 0/1 vertices. More important today, bounds for special families are LINEAR!!!

Some remarks

- **2012** B. Matschke, F. Santos, & C. Weibel improved the counterexample, now smaller in 20 dimensions and one can actually plug it into a computer!
- The diameter gives a lower bound on the worst-case behavior of edge-following algorithms. If diameter is exponential then all edge-path algorithms will be exponential in the worst case.
- Hirsch conjecture known to be true in many instances, e.g. for polytopes with $0/1$ vertices. More important today, bounds for special families are LINEAR!!!
- But there are some very interesting cases where we do not know the tight bound
- Open: Is the Hirsch conjecture true for transportation LPs?
- Open: Is the Hirsch conjecture true for Network LPs?

Using the Coordinate Size

- **Theorem**[Onn & Kleinschmidt] The diameter of a d-polytope with all its vertices integer and contained in the box $[0, K]^d$ is no more than Kd
- **Theorem**[Bonifas, Bonifas, Di Summa, Eisenbrand, Hähnle, and Niemeier] Let $P = \{x \in R^n \colon Ax \leq b\}$ be a polytope where all sub-determinants of $A \in \mathbb{Z}^{n \times d}$ are bounded by Δ in absolute value. then the diameter of P is at most $O(\Delta^2 d^{3.5} \log(d\Delta))$.
- **Corollary** When A is a totally unimodular matrix, the bound of Bonifas et al. simplify to $O(d^{3.5} \log d)$.
- Kitahara and Mizuno Given a linear program of the form $\max\{c^T x : Ax = b, x \geq 0\}$ where A is a real $d \times n$ matrix, The number of different basic feasible solutions (BFSs) generated by Dantzig's simplex method is bounded by $n\lceil d\frac{\gamma}{\delta} \log(d\frac{\gamma}{\delta})\rceil$, where δ and γ are the minimum and the maximum values of all the positive elements of primal BFSs.

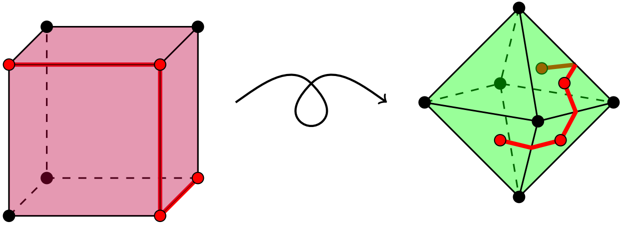
Could the diameter be still LINEAR ???

There is not a single case with diameter more than $2(n-d)$

Diameters of Simplicial Complexes

Definition

- The **distance** between two facets, F_1, F_2, is the length k of the shortest simplicial path $F_1 = f_0, f_1, \ldots, f_k = F_2$.
- The **diameter** of a simplicial complex is the maximum over all distances between all pairs of facets.



Why? We work with simplicial complexes, we can use topological arguments! Idea goes back to the 1980's.

Weak k-Decomposability

Definition

A d-dimensional complex Δ is **weakly k-decomposable** if

1. All the maximal-dimension pieces are of the same dimension, and
2. either Δ is a d-simplex, or there exists a face τ of Δ (called a **shedding face**) such that $\dim(\tau) \leq k$ and $\Delta \setminus \tau$ is d-dimensional and weakly k-decomposable.

Weak k-Decomposability

Definition

A d-dimensional complex Δ is **weakly k-decomposable** if

1. All the maximal-dimension pieces are of the same dimension, and
2. either Δ is a d-simplex, or there exists a face τ of Δ (called a **shedding face**) such that $\dim(\tau) \leq k$ and $\Delta \setminus \tau$ is d-dimensional and weakly k-decomposable.

Definition

When $k = 0$, weak k-decomposability is known as **weak vertex-decomposability**.

Weak k-Decomposability

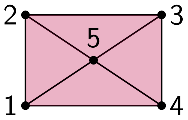
Definition

A d-dimensional complex Δ is **weakly k-decomposable** if

1. All the maximal-dimension pieces are of the same dimension, and
2. either Δ is a d-simplex, or there exists a face τ of Δ (called a **shedding face**) such that $\dim(\tau) \leq k$ and $\Delta \setminus \tau$ is d-dimensional and weakly k-decomposable.

Definition

When $k = 0$, weak k-decomposability is known as **weak vertex-decomposability**.



Weak k-Decomposability

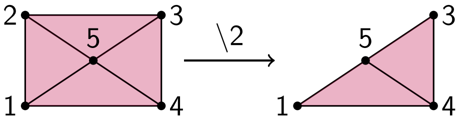
Definition

A d-dimensional complex Δ is **weakly k-decomposable** if

1. All the maximal-dimension pieces are of the same dimension, and
2. either Δ is a d-simplex, or there exists a face τ of Δ (called a **shedding face**) such that $\dim(\tau) \leq k$ and $\Delta \setminus \tau$ is d-dimensional and weakly k-decomposable.

Definition

When $k = 0$, weak k-decomposability is known as **weak vertex-decomposability**.



Weak k-Decomposability

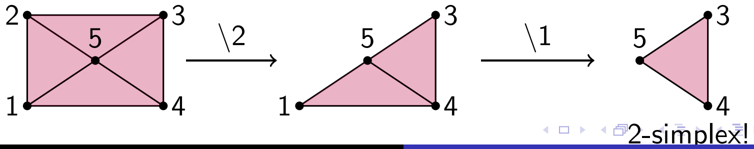
Definition

A d-dimensional complex Δ is **weakly k-decomposable** if

1. All the maximal-dimension pieces are of the same dimension, and
2. either Δ is a d-simplex, or there exists a face τ of Δ (called a **shedding face**) such that $\dim(\tau) \leq k$ and $\Delta \setminus \tau$ is d-dimensional and weakly k-decomposable.

Definition

When $k = 0$, weak k-decomposability is known as **weak vertex-decomposability**.



Why should we care??

Theorem (Billera,Provan, 1980)

If Δ is a weakly k-decomposable complex, $0 \leq k \leq d$, then

$$\mathrm{diam}(\Delta) \leq 2f_k(\Delta),$$

where $f_k(\Delta)$ is the number of k-faces Δ.
In the case of weakly 0-decomposable, we have the following linear bound ($n = f_0(\Delta)$):

$$\mathrm{diam}(\Delta) \leq 2f_0(\Delta) = 2n.$$

Why should we care??

Theorem (Billera,Provan, 1980)

If Δ *is a weakly* k*-decomposable complex,* $0 \leq k \leq d$*, then*

$$\mathrm{diam}(\Delta) \leq 2f_k(\Delta),$$

where $f_k(\Delta)$ *is the number of* k*-faces* Δ*.*
In the case of weakly 0*-decomposable, we have the following linear bound (*$n = f_0(\Delta)$*):*

$$\mathrm{diam}(\Delta) \leq 2f_0(\Delta) = 2n.$$

Theorem: All simplicial d-dimensional polytopes are weakly d-decomposable (because they are shellable!).

Why should we care??

Theorem (Billera,Provan, 1980)

If Δ is a weakly k-decomposable complex, $0 \leq k \leq d$, then

$$\mathrm{diam}(\Delta) \leq 2f_k(\Delta),$$

where $f_k(\Delta)$ is the number of k-faces Δ.
In the case of weakly 0-decomposable, we have the following linear bound ($n = f_0(\Delta)$):

$$\mathrm{diam}(\Delta) \leq 2f_0(\Delta) = 2n.$$

Theorem: All simplicial d-dimensional polytopes are weakly d-decomposable (because they are shellable!).
Theorem: [Provan] Some topological spheres are not weakly 0-decomposable

Why should we care??

Theorem (Billera,Provan, 1980)

If Δ is a weakly k-decomposable complex, $0 \leq k \leq d$, then

$$\mathrm{diam}(\Delta) \leq 2f_k(\Delta),$$

where $f_k(\Delta)$ is the number of k-faces Δ.
In the case of weakly 0-decomposable, we have the following linear bound ($n = f_0(\Delta)$):

$$\mathrm{diam}(\Delta) \leq 2f_0(\Delta) = 2n.$$

Theorem: All simplicial d-dimensional polytopes are weakly d-decomposable (because they are shellable!).
Theorem: [Provan] Some topological spheres are not weakly 0-decomposable

BIG OLD QUESTIONS from 1980's

- Are all simplicial polytopes weakly vertex-decomposable?

BIG OLD QUESTIONS from 1980's

- Are all simplicial polytopes weakly vertex-decomposable?
- Are all simplicial polytopes weakly k-decomposable?

BIG OLD QUESTIONS from 1980's

- Are all simplicial polytopes weakly vertex-decomposable?
- Are all simplicial polytopes weakly k-decomposable?
- Is there a constant k that works for all polytopes???

BIG OLD QUESTIONS from 1980's

- Are all simplicial polytopes weakly vertex-decomposable?
- Are all simplicial polytopes weakly k-decomposable?
- Is there a constant k that works for all polytopes???

All these questions have been finally answered

RESULTS

RESULTS

- **Theorem (2012, JDL and S. Klee)** Not all simplicial polytopes are weakly vertex decomposable!

RESULTS

- **Theorem (2012, JDL and S. Klee)** Not all simplicial polytopes are weakly vertex decomposable!
  We constructed explicit transportation problems whose polars are not weakly vertex-decomposable (from dimension 4 onward).

RESULTS

- **Theorem (2012, JDL and S. Klee)** Not all simplicial polytopes are weakly vertex decomposable!
  We constructed explicit transportation problems whose polars are not weakly vertex-decomposable (from dimension 4 onward).
- Question: Is there a different constant $0 < k < d$ such that every simplicial polytope is k-decomposable?

RESULTS

- **Theorem (2012, JDL and S. Klee)** Not all simplicial polytopes are weakly vertex decomposable!
  We constructed explicit transportation problems whose polars are not weakly vertex-decomposable (from dimension 4 onward).
- Question: Is there a different constant $0 < k < d$ such that every simplicial polytope is k-decomposable?
- **Theorem (2012, Hähnle, Klee, Pilaud)**: Our family of examples has in fact non-k-weakly decomposable members for any $k < d$.
- **Other contributions using "topological lenses":** work by Adler-Dantzig, Billera-Provan, Klee-Walkup, Klee-Kleinschmidt, Kim, Eisenbrand-Hähnle-Razborov-Rothvoss.
- **Theorem (2013) Santos** The diameter of d-manifold grows at least as $O(n^{\frac{2d}{3}})$.

Part II: Contributions to Interior-Point Methods

Narendra Karmarkar, inventor of interior point methods

The Central Path of a Linear Program: Optimizers view

Linear Program: $\text{Maximize}_{\mathbf{x}\in\mathbb{R}^n}$ $\mathbf{c}\cdot\mathbf{x}$ s.t. $A\cdot\mathbf{x}=\mathbf{b}$ and $\mathbf{x}\geq 0$.

The Central Path of a Linear Program: Optimizers view

Linear Program: $\text{Maximize}_{\mathbf{x}\in\mathbb{R}^n}$ $\mathbf{c}\cdot\mathbf{x}$ s.t. $A\cdot\mathbf{x}=\mathbf{b}$ and $\mathbf{x}\geq 0$.

Replace by : $\text{Maximize}_{\mathbf{x}\in\mathbb{R}^n}$ $f_\lambda(\mathbf{x})$ s.t. $A\cdot\mathbf{x}=\mathbf{b}$,

where $\lambda\in\mathbb{R}_+$ and $f_\lambda(\mathbf{x}) := \mathbf{c}^T\cdot\mathbf{x} + \lambda\sum_{i=1}^{n}\log|x_i|$.

The Central Path of a Linear Program: Optimizers view

Linear Program: Maximize$_{\mathbf{x}\in\mathbb{R}^n}$ $\mathbf{c}\cdot\mathbf{x}$ s.t. $A\cdot\mathbf{x}=\mathbf{b}$ and $\mathbf{x}\geq 0$.

Replace by : Maximize$_{\mathbf{x}\in\mathbb{R}^n}$ $f_\lambda(\mathbf{x})$ s.t. $A\cdot\mathbf{x}=\mathbf{b}$,

where $\lambda\in\mathbb{R}_+$ and $f_\lambda(\mathbf{x}) := \mathbf{c}^T\cdot\mathbf{x} + \lambda\sum_{i=1}^n \log|x_i|$.

The maximum of the function f_λ is attained by a unique point $\mathbf{x}^*(\lambda)$ in the the open polytope $\{\mathbf{x}\in(\mathbb{R}_{>0})^n : A\cdot\mathbf{x}=\mathbf{b}\}$.

The Central Path of a Linear Program: Optimizers view

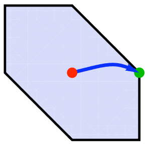
Linear Program: $\text{Maximize}_{\mathbf{x}\in\mathbb{R}^n}\ \mathbf{c}\cdot\mathbf{x}$ s.t. $A\cdot\mathbf{x}=\mathbf{b}$ and $\mathbf{x}\geq 0$.

Replace by : $\text{Maximize}_{\mathbf{x}\in\mathbb{R}^n}\ f_\lambda(\mathbf{x})$ s.t. $A\cdot\mathbf{x}=\mathbf{b}$,

where $\lambda\in\mathbb{R}_+$ and $f_\lambda(\mathbf{x}) := \mathbf{c}^T\cdot\mathbf{x} + \lambda\sum_{i=1}^n \log|x_i|$.

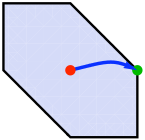
The maximum of the function f_λ is attained by a unique point $\mathbf{x}^*(\lambda)$ in the the open polytope $\{\mathbf{x}\in(\mathbb{R}_{>0})^n : A\cdot\mathbf{x}=\mathbf{b}\}$.

The central path is $\{\mathbf{x}^*(\lambda)\ :\ \lambda>0\}$.
As $\lambda\to 0$, the path leads from the analytic center of the polytope, $\mathbf{x}^*(\infty)$, to the optimal vertex, $\mathbf{x}^*(0)$.



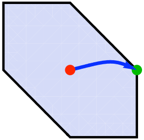
The Central Path of a Linear Program

The central path is $\{\mathbf{x}^*(\lambda) \ : \ \lambda > 0\}$.
As $\lambda \to 0$, the path leads from the analytic center of the polytope, $\mathbf{x}^*(\infty)$, to the optimal vertex, $\mathbf{x}^*(0)$.



The Central Path of a Linear Program

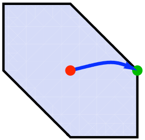
The central path is $\{\mathbf{x}^*(\lambda) \ : \ \lambda > 0\}$.
As $\lambda \to 0$, the path leads from the analytic center of the polytope, $\mathbf{x}^*(\infty)$, to the optimal vertex, $\mathbf{x}^*(0)$.



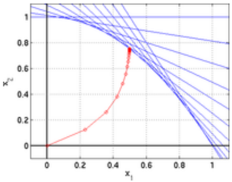
Interior point methods $=$ piecewise-linear approx. of this path

The Central Path of a Linear Program

The central path is $\{\mathbf{x}^*(\lambda) \ : \ \lambda > 0\}$.
As $\lambda \to 0$, the path leads from the analytic center of the polytope, $\mathbf{x}^*(\infty)$, to the optimal vertex, $\mathbf{x}^*(0)$.

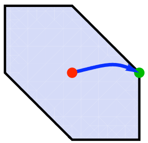


Interior point methods = piecewise-linear approx. of this path

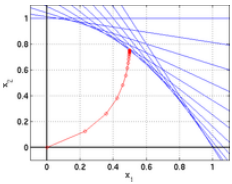


The Central Path of a Linear Program

The central path is $\{\mathbf{x}^*(\lambda) \ : \ \lambda > 0\}$.
As $\lambda \to 0$, the path leads from the analytic center of the polytope, $\mathbf{x}^*(\infty)$, to the optimal vertex, $\mathbf{x}^*(0)$.



Interior point methods = piecewise-linear approx. of this path



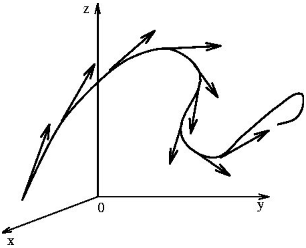
Key point: The convergence of Newton steps will depend on how "curvy" how "twisted" is the central path!!!

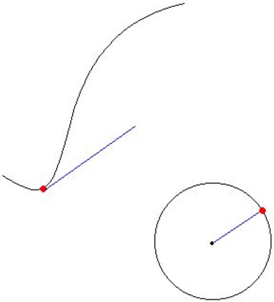
CURVATURE: measuring how “twisted” is the central path

The total curvature of a curve $\mathcal{C}$ in $\mathbb{R}^m$ is the arc length of its image under the Gauss map. $\gamma : \mathcal{C} \rightarrow \mathbb{S}^{m-1}$.

CURVATURE: measuring how "twisted" is the central path

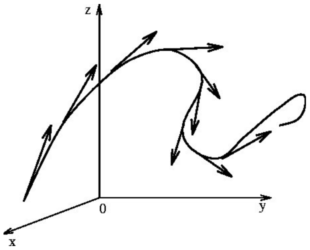
The total curvature of a curve $\mathcal{C}$ in $\mathbb{R}^m$ is the arc length of its image under the Gauss map. $\gamma : \mathcal{C} \to \mathbb{S}^{m-1}$.

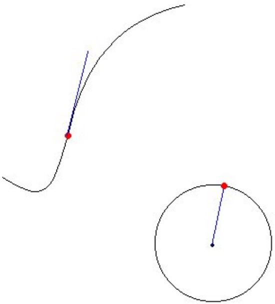




CURVATURE: measuring how “twisted” is the central path

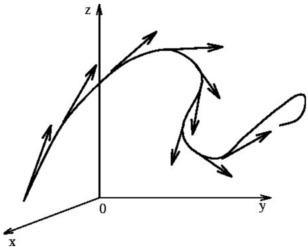
The total curvature of a curve $\mathcal{C}$ in $\mathbb{R}^m$ is the arc length of its image under the Gauss map. $\gamma : \mathcal{C} \to \mathbb{S}^{m-1}$.

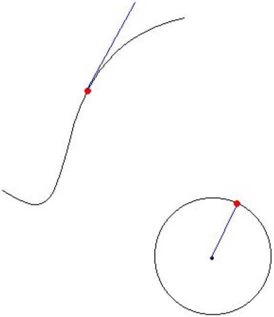




CURVATURE: measuring how "twisted" is the central path

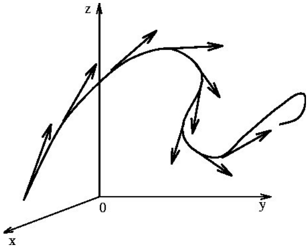
The total curvature of a curve $\mathcal{C}$ in $\mathbb{R}^m$ is the arc length of its image under the Gauss map. $\gamma : \mathcal{C} \to \mathbb{S}^{m-1}$.

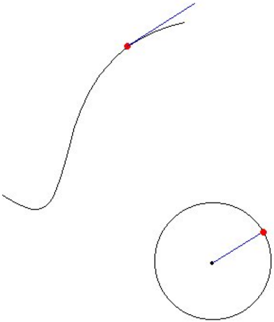




CURVATURE: measuring how “twisted” is the central path

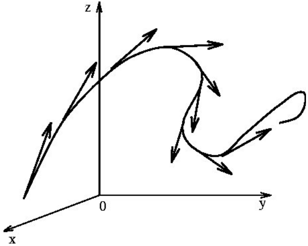
The total curvature of a curve $\mathcal{C}$ in $\mathbb{R}^m$ is the arc length of its image under the Gauss map. $\gamma : \mathcal{C} \to \mathbb{S}^{m-1}$.

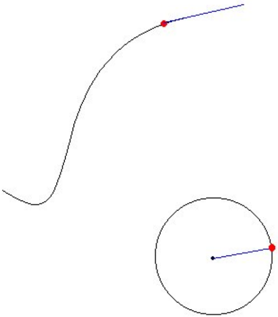




CURVATURE: measuring how “twisted” is the central path

The total curvature of a curve $\mathcal{C}$ in $\mathbb{R}^m$ is the arc length of its image under the Gauss map. $\gamma : \mathcal{C} \to \mathbb{S}^{m-1}$.

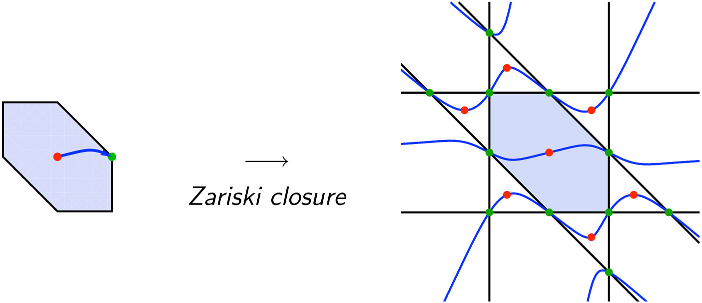




Bounds on total curvature $\to$ bounds on # Newton steps.

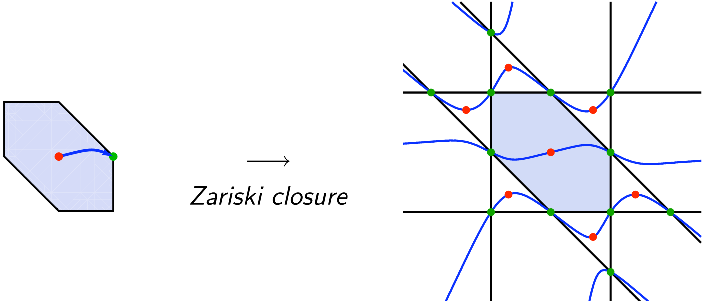
Megiddo-Shub (1989), Sonnevend-Stoer-Zhao (1991),Todd-Ye (1996), Vavasis-Ye (1996), Monteiro-Tsuchiya (2008), Dedieu-Malajovich-Shub (2005), Deza-Terlaky-Zinchenko (2008)....

The central curve $\mathcal{C}$ is the Zariski closure of the central path.
It contains the central paths of all polyhedra in the hyperplane arrangement $\{x_i = 0\}_{i=1,\ldots,n} \subset \{A \cdot \mathbf{x} = \mathbf{b}\}$.



The Central Path is part of an Algebraic Curve

The central curve $\mathcal{C}$ is the Zariski closure of the central path. It contains the central paths of all polyhedra in the hyperplane arrangement $\{x_i = 0\}_{i=1,\dots,n} \subset \{A \cdot \mathbf{x} = \mathbf{b}\}$.



Bayer-Lagarias (1989) first studied the central path as an algebraic curve and suggest the problem of identifying its defining equations.

Questions and Contributions

Motivating Question 1: Find the defining equations for the central curve, what is the degree of the curve?

Questions and Contributions

Motivating Question 1: Find the defining equations for the central curve, what is the degree of the curve?

Motivating Question 2: What is the degree and the maximum total curvature of the central path given a matrix A?

Questions and Contributions

Motivating Question 1: Find the defining equations for the central curve, what is the degree of the curve?

Motivating Question 2: What is the degree and the maximum total curvature of the central path given a matrix A?

Our contribution: We use algebraic geometry and matroid theory to find explicit defining equations of the central curve and refine bounds on its degree and total curvature.

Conditions defining the curve

Recall the function $f_\lambda(\mathbf{x}) \;=\; \mathbf{c}\cdot\mathbf{x} + \lambda\sum_{i=1}^{n}\log|x_i|$ in $\{A\cdot\mathbf{x}=\mathbf{b}\}$:

Lemma: A point $\mathbf{x}$ belongs to the central curve

Conditions defining the curve

Recall the function $f_\lambda(\mathbf{x}) \; = \; \mathbf{c}\cdot\mathbf{x} + \lambda \sum_{i=1}^{n} \log|x_i|$ in $\{A\cdot\mathbf{x} = \mathbf{b}\}$:

Lemma: A point $\mathbf{x}$ belongs to the central curve
$\Leftrightarrow \; \nabla f_\lambda(\mathbf{x}) \;\; = \;\; \mathbf{c} + \lambda\mathbf{x}^{-1} \in \mathrm{span}\{\mathrm{rows}(A)\}$

Conditions defining the curve

Recall the function $f_\lambda(\mathbf{x}) \;=\; \mathbf{c}\cdot\mathbf{x} + \lambda\sum_{i=1}^{n}\log|x_i|$ in $\{A\cdot\mathbf{x}=\mathbf{b}\}$:

Lemma: A point $\mathbf{x}$ belongs to the central curve

$$\Leftrightarrow\; \nabla f_\lambda(\mathbf{x}) \;=\; \mathbf{c} + \lambda\mathbf{x}^{-1} \in \operatorname{span}\{\text{rows}(A)\}$$

$$\Leftrightarrow\; \mathbf{x}^{-1} \in \operatorname{span}\{\text{rows}(A),\; \mathbf{c}\}$$

Conditions defining the curve

Recall the function $f_\lambda(\mathbf{x}) \;=\; \mathbf{c}\cdot\mathbf{x} + \lambda\sum_{i=1}^{n}\log|x_i|$ in $\{A\cdot\mathbf{x}=\mathbf{b}\}$:

Lemma: A point $\mathbf{x}$ belongs to the central curve

$$\Leftrightarrow \;\nabla f_\lambda(\mathbf{x}) \;=\; \mathbf{c}+\lambda\mathbf{x}^{-1} \in \text{span}\{\text{rows}(A)\}$$

$$\Leftrightarrow \quad \mathbf{x}^{-1} \in \text{span}\{\text{rows}(A),\; \mathbf{c}\}$$

Conditions defining the curve

Recall the function $f_\lambda(\mathbf{x}) \;=\; \mathbf{c}\cdot\mathbf{x} + \lambda\sum_{i=1}^{n}\log|x_i|$ in $\{A\cdot\mathbf{x} = \mathbf{b}\}$:

Lemma: A point $\mathbf{x}$ belongs to the central curve

$$\Leftrightarrow \;\nabla f_\lambda(\mathbf{x}) \;=\; \mathbf{c} + \lambda\mathbf{x}^{-1} \in \mathrm{span}\{\mathrm{rows}(A)\}$$

$$\Leftrightarrow \quad \mathbf{x}^{-1} \in \mathrm{span}\{\mathrm{rows}(A),\ \mathbf{c}\} =: \mathcal{L}_{A,\mathbf{c}}$$

Conditions defining the curve

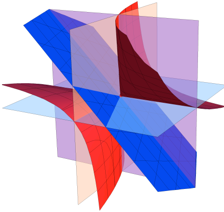
Recall the function $f_\lambda(\mathbf{x}) = \mathbf{c} \cdot \mathbf{x} + \lambda \sum_{i=1}^{n} \log |x_i|$ in $\{A \cdot \mathbf{x} = \mathbf{b}\}$:

Lemma: A point $\mathbf{x}$ belongs to the central curve

$$\Leftrightarrow \nabla f_\lambda(\mathbf{x}) = \mathbf{c} + \lambda \mathbf{x}^{-1} \in \text{span}\{\text{rows}(A)\}$$

$$\Leftrightarrow \quad \mathbf{x}^{-1} \in \text{span}\{\text{rows}(A), \ \mathbf{c}\} =: \mathcal{L}_{A,\mathbf{c}}$$

$$\Leftrightarrow \quad \mathbf{x} \in \mathcal{L}_{A,\mathbf{c}}^{-1}$$



where $\mathcal{L}_{A,\mathbf{c}}^{-1}$ is the coordinate-wise reciprocal $\mathcal{L}_{A,\mathbf{c}}$:

$$\mathcal{L}_{A,\mathbf{c}}^{-1} := \left\{ (u_1^{-1}, \ldots, u_n^{-1}) \quad \text{where} \quad (u_1, \ldots, u_n) \in \mathcal{L}_{A,\mathbf{c}} \right\}$$

Conditions defining the curve

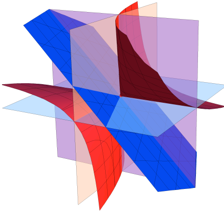
Recall the function $f_\lambda(\mathbf{x}) \;=\; \mathbf{c}\cdot\mathbf{x} + \lambda\sum_{i=1}^{n}\log|x_i|$ in $\{A\cdot\mathbf{x}=\mathbf{b}\}$:

Lemma: A point $\mathbf{x}$ belongs to the central curve

$$\Leftrightarrow \nabla f_\lambda(\mathbf{x}) \;=\; \mathbf{c}+\lambda\mathbf{x}^{-1} \in \operatorname{span}\{\text{rows}(A)\}$$

$$\Leftrightarrow \quad \mathbf{x}^{-1} \in \operatorname{span}\{\text{rows}(A),\ \mathbf{c}\} =: \mathcal{L}_{A,\mathbf{c}}$$

$$\Leftrightarrow \quad \mathbf{x} \in \mathcal{L}_{A,\mathbf{c}}^{-1}$$



where $\mathcal{L}_{A,\mathbf{c}}^{-1}$ is the coordinate-wise reciprocal $\mathcal{L}_{A,\mathbf{c}}$:

$$\mathcal{L}_{A,\mathbf{c}}^{-1} := \left\{ (u_1^{-1},\ldots,u_n^{-1}) \quad \text{where} \quad (u_1,\ldots,u_n)\in\mathcal{L}_{A,\mathbf{c}} \right\}$$

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

(Linear) Matroids

Given the matrix A the defines the LP:

$$A = \left[\ x^1 \ \middle| \ x^2 \ \middle| \ \cdots \ \middle| \ x^n \ \right]$$

$E := \{1, 2, \ldots, n\}$.
$\mathcal{C} := \{S \subset E \ : \ A_S$ has *linearly* dependent columns;
A_{S-e} has *linearly* independent columns$\}$.

(Linear) Matroids

Given the matrix A the defines the LP:

$$A = \left[\ x^1 \mid x^2 \mid \cdots \mid x^n\ \right]$$

$E := \{1, 2, \ldots, n\}$.
$\mathcal{C} := \{S \subset E\ :\ A_S$ has *linearly* dependent columns;
A_{S-e} has *linearly* independent columns$\}$.
These are called the CIRCUITS of the matroid of A.

(Linear) Matroids

Given the matrix A the defines the LP:

$$A = \left[\ x^1 \mid x^2 \mid \cdots \mid x^n\ \right]$$

$E := \{1, 2, \ldots, n\}$.
$\mathcal{C} := \{S \subset E\ :\ A_S$ has *linearly* dependent columns;
A_{S-e} has *linearly* independent columns$\}$.
These are called the CIRCUITS of the matroid of A.
They satisfy three axioms:

(C1) $\emptyset \notin \mathcal{C}$.

(C2) $X, Y \in \mathcal{C},\ X \subset Y\ \Rightarrow\ X = Y$.

(C3) $X, Y \in \mathcal{C},\ X \neq Y,\ e \in X \cap Y\ \Rightarrow$
$\exists\ Z \in \mathcal{C}$ with $Z \subset (X \cup Y) - e$.
(Elimination Property)

Example of circuits

"circuits 101":

$$E = \{1, 2, 3,\ 4,\ 5,\ 6\}.$$

$$A = \left[\begin{array}{c|c|c|r|r|c} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 \end{array}\right]$$

Q1: $\{2, 3, 4, 6\} \in \mathcal{C}$?

A1: Yes. $A_2 + A_3 + A_4 - A_6 = 0$.

$\det[A_2|A_3|A_4] = 1$, $\det[A_2|A_3|A_6] = 1$,

$\det[A_2|A_4|A_6] = -1$, $\det[A_3|A_4|A_6] = 1$.

Key Point: The curvature can be bounded for the curve using matroid properties of CIRCUITS!!!.

Equations

Applying results by Proudfoot and Speyer (2006) we obtained:

Equations

Applying results by Proudfoot and Speyer (2006) we obtained:
Lemma: The equations defining $\mathcal{L}_{A,\mathbf{c}}^{-1}$ are the homogeneous polynomials

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

where v is a vector in kernel$\binom{A}{\mathbf{c}}$ of minimal support circuits.

Equations

Applying results by Proudfoot and Speyer (2006) we obtained:

Lemma: The equations defining $\mathcal{L}_{A,\mathbf{c}}^{-1}$ are the homogeneous polynomials

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

where v is a vector in kernel$\binom{A}{\mathbf{c}}$ of minimal support circuits.

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

Equations

Applying results by Proudfoot and Speyer (2006) we obtained:
Lemma: The equations defining $\mathcal{L}_{A,\mathbf{c}}^{-1}$ are the homogeneous polynomials

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

where v is a vector in kernel$\binom{A}{\mathbf{c}}$ of minimal support circuits.

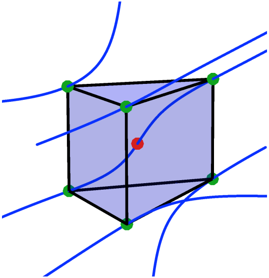
$$\begin{pmatrix} A \\ \mathbf{c} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 2 & 0 & 4 & 0 \end{pmatrix}$$

Circuit $v = \begin{pmatrix} -2 & 1 & 1 & 0 & 0 \end{pmatrix}$ produces $-2x_2x_3 + 1x_1x_3 + 1x_1x_2$.

Example

$(n = 5,\ d = 2)$

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



Example

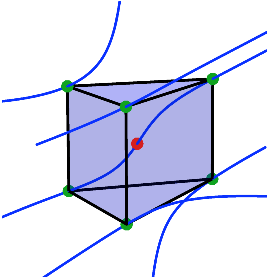
$(n = 5,\ d = 2)$

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

Equations for $\mathcal{C}$:

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

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



Degree, Genus, and the Broken Circuit Complexes

- How about the degree of the central curve?? Matroids again!

Degree, Genus, and the Broken Circuit Complexes

- How about the degree of the central curve?? Matroids again!
- Fix the standard ordering $1 < 2 < \cdots < n$ of $[n]$. A broken circuit of M is any subset of $[n]$ of the form $C \backslash \{\min(C)\}$ where C is a circuit.

Degree, Genus, and the Broken Circuit Complexes

- How about the degree of the central curve?? Matroids again!
- Fix the standard ordering $1 < 2 < \cdots < n$ of $[n]$. A broken circuit of M is any subset of $[n]$ of the form $C \setminus \{\min(C)\}$ where C is a circuit.
- The broken circuit complex of M is the simplicial complex $\mathrm{Br}(M)$ whose minimal non-faces are the broken circuits. A subset of $[n]$ is a face of $\mathrm{Br}(M)$ if it does not contain any broken circuit.

Degree, Genus, and the Broken Circuit Complexes

- How about the degree of the central curve?? Matroids again!
- Fix the standard ordering $1 < 2 < \cdots < n$ of $[n]$. A broken circuit of M is any subset of $[n]$ of the form $C \setminus \{\min(C)\}$ where C is a circuit.
- The broken circuit complex of M is the simplicial complex $\mathrm{Br}(M)$ whose minimal non-faces are the broken circuits. A subset of $[n]$ is a face of $\mathrm{Br}(M)$ if it does not contain any broken circuit.

Degree, Genus, and the Broken Circuit Complexes

- How about the degree of the central curve?? Matroids again!
- Fix the standard ordering $1 < 2 < \cdots < n$ of $[n]$. A broken circuit of M is any subset of $[n]$ of the form $C \backslash \{\min(C)\}$ where C is a circuit.
- The broken circuit complex of M is the simplicial complex $\mathrm{Br}(M)$ whose minimal non-faces are the broken circuits. A subset of $[n]$ is a face of $\mathrm{Br}(M)$ if it does not contain any broken circuit.
- Write $f_i = f_i(\mathrm{Br}(M))$ for the number of i-dimensional faces of the broken circuit complex $\mathrm{Br}(M)$.

Degree, Genus, and the Broken Circuit Complexes

- How about the degree of the central curve?? Matroids again!
- Fix the standard ordering $1 < 2 < \cdots < n$ of $[n]$. A broken circuit of M is any subset of $[n]$ of the form $C \backslash \{\min(C)\}$ where C is a circuit.
- The broken circuit complex of M is the simplicial complex $\mathrm{Br}(M)$ whose minimal non-faces are the broken circuits. A subset of $[n]$ is a face of $\mathrm{Br}(M)$ if it does not contain any broken circuit.
- Write $f_i = f_i(\mathrm{Br}(M))$ for the number of i-dimensional faces of the broken circuit complex $\mathrm{Br}(M)$.
- An easy linear transformation of the f_i gives a vector $(h_0, h_1, \ldots, h_d)$.

Degree from Matroid data

One can recover its Hilbert series from the Stanley-Reisner ring of the broken circuit complex

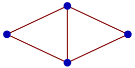
Degree from Matroid data

One can recover its Hilbert series from the Stanley-Reisner ring of the broken circuit complex

Using the matroid associated to $\mathcal{L}_{A,\mathbf{c}}^{-1}$, construct its broken circuit complex.

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 2 & 0 & 4 & 0 \end{pmatrix}$$

$\{123, 1245,$
$1345, 2345\}$



$h = (1, 2, 2)$

matrix $\binom{A}{\mathbf{c}}$ $\longrightarrow$ matroid $\longrightarrow$ "broken circuit" complex $\longrightarrow$ h-vector

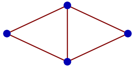
Degree from Matroid data

One can recover its Hilbert series from the Stanley-Reisner ring of the broken circuit complex

Using the matroid associated to $\mathcal{L}_{A,\mathbf{c}}^{-1}$, construct its broken circuit complex.

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 2 & 0 & 4 & 0 \end{pmatrix}$$

$\{123, 1245, 1345, 2345\}$



$h = (1, 2, 2)$

matrix $\binom{A}{\mathbf{c}}$ $\longrightarrow$ matroid $\longrightarrow$ "broken circuit" complex $\longrightarrow$ h-vector

$\Rightarrow$ $\deg(\mathcal{C}) = \sum_{i=0}^{d} h_i$ and $\text{genus}(\mathcal{C}) = 1 - \sum_{j=0}^{d}(1-j)h_j$.

Total Curvature

Total Curvature

Theorem: Dedieu-Malajovich-Shub (2005) The total curvature of $\mathcal{C}$ is bounded above by π times the degree $\deg(\gamma(\mathcal{C}))$ of the projective Gauss curve in $\mathbb{P}^{m-1}$.

Total Curvature

Theorem: Dedieu-Malajovich-Shub (2005) The total curvature of $\mathcal{C}$ is bounded above by π times the degree $\deg(\gamma(\mathcal{C}))$ of the projective Gauss curve in $\mathbb{P}^{m-1}$.

Classic algebraic geometry: $\deg(\gamma(\mathcal{C})) \leq 2 \cdot (\deg(\mathcal{C}) + \mathrm{genus}(\mathcal{C}) - 1)$

Total Curvature

Theorem: Dedieu-Malajovich-Shub (2005) The total curvature of $\mathcal{C}$ is bounded above by π times the degree $\deg(\gamma(\mathcal{C}))$ of the projective Gauss curve in $\mathbb{P}^{m-1}$.

Classic algebraic geometry: $\deg(\gamma(\mathcal{C})) \leq 2 \cdot (\deg(\mathcal{C}) + \mathrm{genus}(\mathcal{C}) - 1)$

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

$$\deg(\gamma(\mathcal{C})) \quad \leq \quad 2 \cdot \sum_{i=1}^{d} i \cdot h_i$$

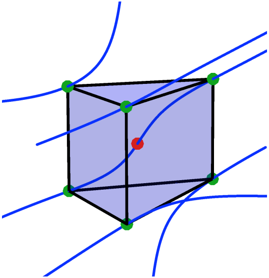
The **total curvature** of the central curve is no more than

$$(2 \cdot \pi) \cdot \sum_{i=1}^{d} i \cdot h_i$$

Example (continued!)

$(n = 5,\ d = 2)$

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



Example (continued!)

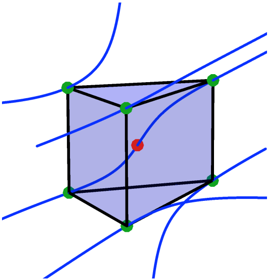
$(n = 5,\ d = 2)$

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

Equations for $\mathcal{C}$:

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

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



Example (continued!)

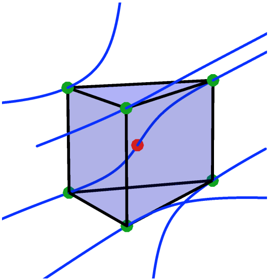
$(n = 5,\ d = 2)$

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

Equations for $\mathcal{C}$:

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

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



$h = (1, 2, 2) \ \Rightarrow\ \deg(\mathcal{C}) = 5, \quad \text{total curvature}(\mathcal{C}) \leq 12\pi$

Example (continued!)

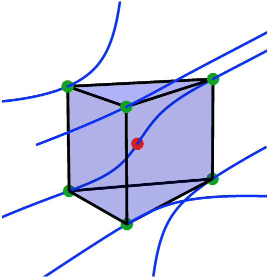
$(n = 5,\ d = 2)$

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

Equations for $\mathcal{C}$:

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

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



$$h = (1, 2, 2) \ \Rightarrow \ \deg(\mathcal{C}) = 5, \quad \text{total curvature}(\mathcal{C}) \leq 12\pi \ (\leq 16\pi)$$

Part III: Other Methods

Fourier & Motzkin

Linear Feasibility

We are interested in determining the feasibility of systems of the form

$$Ax = b, \quad Cx \leq d$$

with $x \in \mathbb{R}^n$, A an $m \times n$ matrix, and C an $l \times n$ matrix. The set of all $x \in \mathbb{R}^n$ that satisfy the above constraints is the feasible region.

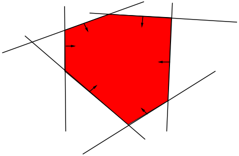


Figure: Example of a feasible region.

This is the Linear Feasibility Problem (LFP), it is computationally equivalent to the Linear Optimization Problem!!

The Relaxation Method

- Start at current best guess x_j.

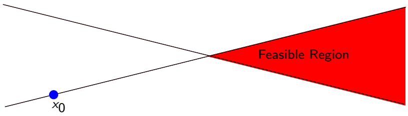


Figure: Projection onto violated constraints.

The Relaxation Method

- Start at current best guess x_j.
- Consider one constraint, i.e. $c_i x \leq d_i$, at a time.

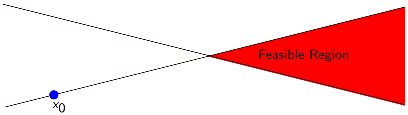


Figure: Projection onto violated constraints.

The Relaxation Method

- Start at current best guess x_j.
- Consider one constraint, i.e. $c_i x \leq d_i$, at a time.
- In the first algorithm proposed by Motzkin and Schoenberg [1954], project onto the most violated constraint $c_k x = d_k$, where $k = \mathrm{arg\ max}\{c_i x_j - d_i\}$.

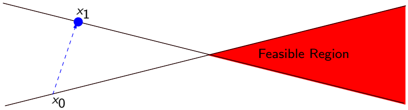


Figure: Projection onto violated constraints.

The Relaxation Method

- Start at current best guess x_j.
- Consider one constraint, i.e. $c_i x \leq d_i$, at a time.
- In the first algorithm proposed by Motzkin and Schoenberg [1954], project onto the most violated constraint $c_k x = d_k$, where $k = \arg\max\{c_i x_j - d_i\}$.
- Set projected point to x_{j+1} and repeat.

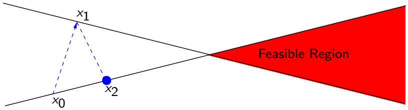


Figure: Projection onto violated constraints.

The Relaxation Method

- Start at current best guess x_j.
- Consider one constraint, i.e. $c_i x \leq d_i$, at a time.
- In the first algorithm proposed by Motzkin and Schoenberg [1954], project onto the most violated constraint $c_k x = d_k$, where $k = \mathrm{arg\ max}\{c_i x_j - d_i\}$.
- Set projected point to x_{j+1} and repeat.

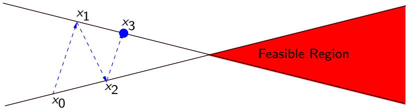


Figure: Projection onto violated constraints.

The Relaxation Method

- Start at current best guess x_j.
- Consider one constraint, i.e. $c_i x \leq d_i$, at a time.
- In the first algorithm proposed by Motzkin and Schoenberg [1954], project onto the most violated constraint $c_k x = d_k$, where $k = \arg\max\{c_i x_j - d_i\}$.
- Set projected point to x_{j+1} and repeat.
- Sequence of points converges to the feasible region.

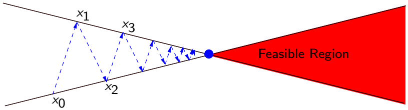


Figure: Projection onto violated constraints.

Agmon [1954], and Motzkin and Schoenberg [1954]

Pros of the relaxation method:

- Always terminates or converges to a point in the feasible region.
- Lends itself for parallelization.

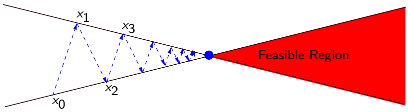


Figure: 2-dim example of the relaxation method.

Agmon [1954], and Motzkin and Schoenberg [1954]

Pros of the relaxation method:

- Always terminates or converges to a point in the feasible region.
- Lends itself for parallelization.

Cons of the relaxation method:

- Need to assume feasible region is nonempty.
- May take exponential time [Goffin 1982, Telgen 1982].

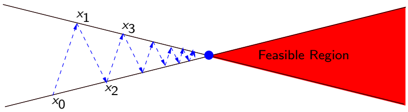


Figure: 2-dim example of the relaxation method.

What is the idea?

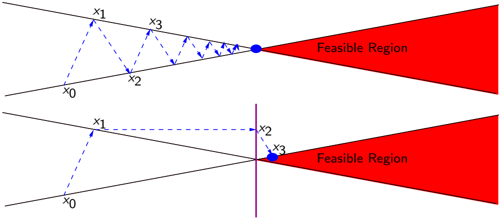


Figure: Projecting onto an induced hyperplane.

What is the idea? Induced Hyperplanes!

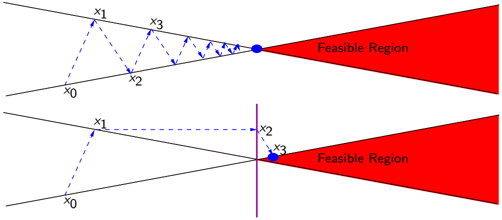


Figure: Projecting onto an induced hyperplane.

The Chubanov Relaxation Method [2011]

What is the idea? Induced Hyperplanes!

- These are new constraints derived as convex combinations of the original constraints.

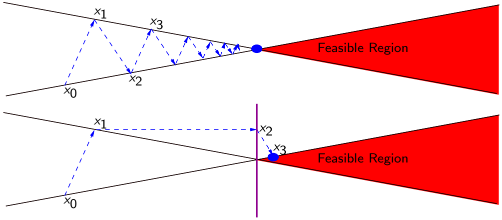


Figure: Projecting onto an induced hyperplane.

The Chubanov Relaxation Method [2011]

What is the idea? Induced Hyperplanes!

- These are new constraints derived as convex combinations of the original constraints.
- Such an advantage that when $Cx \leq d$ takes the form $\mathbf{0} \leq x \leq \mathbf{1}$, Chubanov's algorithm runs in strongly polynomial time.

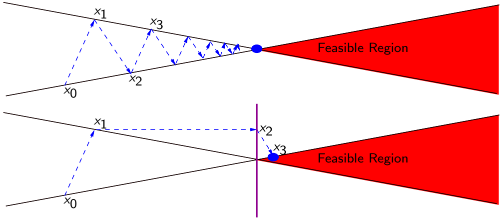


Figure: Projecting onto an induced hyperplane.

Theoretical Extensions and Numerical Results

Theory:

- Chubanov's algorithm either returns a feasible solution or determines no **integer** solutions exist.
- We use Chubanov's method to determine feasibility of strict LFPs.
- When the constraint matrix is totally unimodular, we get a strongly polynomial running time similar to that of Tardos [1986].

Numerics:

- Despite its theoretical advantages, Chubanov's relaxation method it is practically much slower than the original relaxation method.
- We are investigating how it influences the number of branches when solving 0/1 integer programs.

MANY OPEN PROBLEMS!!

- Is there a **strongly polynomial** time linear programming algorithm ??

MANY OPEN PROBLEMS!!

- Is there a **strongly polynomial** time linear programming algorithm ??
- Is there a polynomial-time pivot rule for the simplex method??

MANY OPEN PROBLEMS!!

- Is there a **strongly polynomial** time linear programming algorithm ??
- Is there a polynomial-time pivot rule for the simplex method??
- What are the best bounds for the diameter of convex polyhedra??

MANY OPEN PROBLEMS!!

- Is there a **strongly polynomial** time linear programming algorithm ??
- Is there a polynomial-time pivot rule for the simplex method??
- What are the best bounds for the diameter of convex polyhedra??

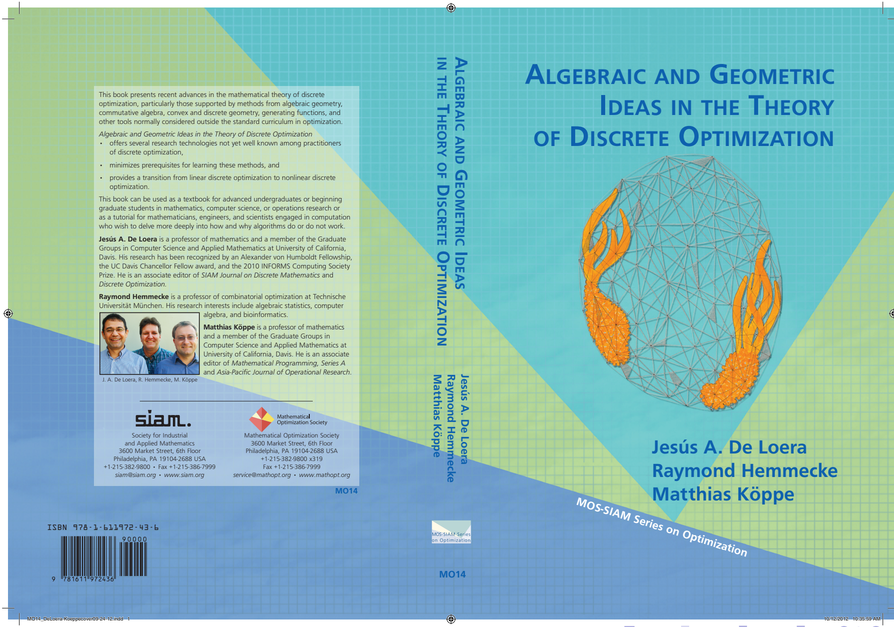
MANY OPEN PROBLEMS!!

- Is there a **strongly polynomial** time linear programming algorithm ??
- Is there a polynomial-time pivot rule for the simplex method??
- What are the best bounds for the diameter of convex polyhedra??
- What is total curvature of **just** the central path?
- Conjecture:(Deza, Terlaky, Zinchenko) **The total curvature of the central path in a polyhedron is** $\leq 2\pi$(#**number of facets).**

MANY OPEN PROBLEMS!!

- Is there a **strongly polynomial** time linear programming algorithm ??
- Is there a polynomial-time pivot rule for the simplex method??
- What are the best bounds for the diameter of convex polyhedra??

- Conjecture:(Deza, Terlaky, Zinchenko) **The total curvature of the central path in a polyhedron is** $\leq 2\pi$(#**number of facets**).


This book presents recent advances in the mathematical theory of discrete optimization, particularly those supported by methods from algebraic geometry, commutative algebra, convex and discrete geometry, generating functions, and other tools normally considered outside the standard curriculum in optimization.
Algebraic and Geometric Ideas in the Theory of Discrete Optimization
• offers several research technologies not yet well known among practitioners of discrete optimization,
• minimizes prerequisites for learning these methods, and
• provides a transition from linear discrete optimization to nonlinear discrete optimization.
This book can be used as a textbook for advanced undergraduates or beginning graduate students in mathematics, computer science, or operations research or as a tutorial for mathematicians, engineers, and scientists engaged in computation who wish to delve more deeply into how and why algorithms do or do not work.
Jesús A. De Loera is a professor of mathematics and a member of the Graduate Groups in Computer Science and Applied Mathematics at University of California, Davis. His research has been recognized by an Alexander von Humboldt Fellowship, the UC Davis Chancellor Fellow award, and the 2010 INFORMS Computing Society Prize. He is an associate editor of *SIAM Journal on Discrete Mathematics* and *Discrete Optimization*.
Raymond Hemmecke is a professor of combinatorial optimization at Technische Universität München. His research interests include algebraic statistics, computer algebra, and bioinformatics.
Matthias Köppe is a professor of mathematics and a member of the Graduate Groups in Computer Science and Applied Mathematics at University of California, Davis. He is an associate editor of *Mathematical Programming, Series A* and *Asia-Pacific Journal of Operational Research*.
J. A. De Loera, R. Hemmecke, M. Köppe
siam.
Society for Industrial and Applied Mathematics
3600 Market Street, 6th Floor
Philadelphia, PA 19104-2688 USA
+1-215-382-9800 • Fax +1-215-386-7999
siam@siam.org • www.siam.org
Mathematical Optimization Society
Mathematical Optimization Society
3600 Market Street, 6th Floor
Philadelphia, PA 19104-2688 USA
+1-215-382-9800 x319
Fax +1-215-386-7999
service@mathopt.org • www.mathopt.org
MO14
ISBN 978-1-611972-43-6
90000
9 781611 972436
ALGEBRAIC AND GEOMETRIC IDEAS IN THE THEORY OF DISCRETE OPTIMIZATION
Jesús A. De Loera
Raymond Hemmecke
Matthias Köppe
MOS-SIAM Series on Optimization
MO14
ALGEBRAIC AND GEOMETRIC IDEAS IN THE THEORY OF DISCRETE OPTIMIZATION
Jesús A. De Loera
Raymond Hemmecke
Matthias Köppe
MOS-SIAM Series on Optimization

Thank You!