$\boldsymbol{A}$-discriminants

Alicia Dickenstein

KTH, May 9, 2011

◇ Projective toric varieties and their duals

A toric variety is an irreducible algebraic variety equipped with an action of an algebraic torus $\boldsymbol{T} = (\mathbf{k}^*)^n$ having an open dense orbit, i.e. it is the closure of a torus orbit

Normality is not required, char$(\mathbf{k}) = \mathbf{0}$ and algebraically closed.

◇ Projective toric varieties and their duals

Any projective toric variety in an equivariant embedding is of the following form:

- A an integer $d \times n$-matrix of maximal rank d, columns of A (weights of the torus action) span $\mathbb{Z}^d$, $(1, 1, \ldots, 1)$ is in the row span of A.
- This defines a *projective toric variety* X_A in $\mathbb{P}^{n-1}(\mathbf{k})$ as the closure of the image of the map $\phi_A : T^d \to \mathbb{P}^{n-1}(\mathbf{k})$
$$\phi_A(t) \; = \; (t^{a_1} : t^{a_2} : \cdots : t^{a_n}).$$
- We identify the matrix A with the point configuration $\{a_1, a_2, \ldots, a_n\}$. The convex hull of A is a $(d{-}1)$-dimensional polytope with $\leq n$ vertices. The toric variety X_A is an affine invariant of A.
- X_A is the set of all points $x \in \mathbb{P}^{n-1}(\mathbf{k})$ such that $x^u = x^v$ for all $u, v \in \mathbb{N}^n$ with $Au = Av$ (i.e. it is cut out by binomials)

◇ Projective toric varieties and their duals

- Any projective toric variety X_A in an equivariant embedding is the closure in $\mathbb{P}^{n-1}(\mathbf{k})$ of the image of the map $\phi_A : T^d \to \mathbb{P}^{n-1}(\mathbf{k})$

$$\phi_A(t) \;=\; (t^{a_1} : t^{a_2} : \cdots : t^{a_n}).$$

with $A \in \mathbb{Z}^{d\times n}$.

That is

- The torus action of T^d on $\mathbb{P}^{n-1}(\mathbf{k})$ prescribed by A is the coordinatewise product

$$t * x = \; (t^{a_1}x_1 : t^{a_2}x_2 : \cdots : t^{a_n}x_n).$$

- X_A is the closure of the orbit $\mathcal{O} \cdot \mathbf{1}$ of the unit point $\mathbf{1} = (1 : \cdots : 1)$.

◇ Projective toric varieties and their duals

- Let $\mathbb{P}^{n-1}(\mathbf{k})^*$ denote the projective space dual to $\mathbb{P}^{n-1}(\mathbf{k})$.
- The point $y = (y_1 : \cdots : y_n)$ in $\mathbb{P}^{n-1}(\mathbf{k})^*$ corresponds to the hyperplane

$$H_y = \{x \in \mathbb{P}^{n-1}(\mathrm{k}) \ : \sum_{i=1}^{n} y_i \, x_i = 0\}.$$

- The *dual variety* X_A^* is defined as the closure in $\mathbb{P}^{n-1}(\mathbf{k})^*$ of the set of points y such that the hyperplane H_y intersects the toric variety X_A at a regular point p and contains the tangent space $T_{X_A}(p)$ of X_A at p.

- The expected codimension of X_A^* is one and in this case its defining equation Δ_A (up to sign) is called the A-discriminant.
- When X_A^* is not a hypersurface, X is called *defective* and its defect equals $\mathrm{def}(X_A) = \mathrm{codim}(X_A^*) - 1$.

◇ Projective toric varieties and their duals

- X_A^* is also the closure in $\mathbb{P}^{n-1}(\mathbf{k})^*$ of the set of points y such that the polynomial

$$f_A(y;t) = \sum_{i=1}^{n} y_i t^{a_i}$$

with support in A and coefficients y defines a non smooth hypersurface of the torus

$$(t \in T^d \,/\, f_A(y;t) = 0),$$

that is, such that there exists a point $t \in T^d$ with

$$f_A(y;t) = \frac{\partial}{\partial t_1}(f_A)(y;t) = \cdots = \frac{\partial}{\partial t_d}(f_A)(y;t) = 0.$$

Therefore, A-discriminants generalize classical discriminants

◇ Homogeneous version of the Horn-Kapranov parametrization

- The dual variety $\boldsymbol{X_A^*}$ of the toric variety $\boldsymbol{X_A}$ is the closure of the image of the map $\boldsymbol{\varphi_A : \mathbb{P}(\mathbf{ker}(A)) \times T^d \longrightarrow \mathbb{P}^{n-1}(\mathbf{k})^*}$ which is given by

$$\boldsymbol{\varphi_A(u,t) \ = \ (u_1 t^{-a_1} \ : \ u_2 t^{-a_2} \ : \ \cdots \ : \ u_n t^{-a_n}).}$$

- So, the dual variety $\boldsymbol{X_A^*}$ is the closure of the union of the torus orbits of the points in the linear subspace $\boldsymbol{\mathbb{P}(\mathbf{ker}(A))}$ of $\boldsymbol{\mathbb{P}^{n-1}(\mathbf{k})^*}$.

The discriminant of a cubic polynomial in 1 variable

$$A := \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix} \qquad X_A \text{ is the twisted cubic.}$$

$$f_A(x;t) = x_1 t^0 + x_2 t^1 + x_3 t^2 + x_4 t^3$$

$$\Delta_A = 27\, x_1^2 x_4^2 - 18\, x_1 x_2 x_3 x_4 + 4\, x_1 x_3^3 + 4\, x_2^3 x_4 - x_2^2 x_3^2$$

$$\mathrm{in}_{(-1,-1,-1,0)}(\Delta_A) = 4x_1 x_3^3 - x_2^2 x_3^2 \qquad (-1,-1,-1,0) \in \tau(X_A^*)$$

$$\mathrm{in}_{(1,0,1,0)}(\Delta_{}) = 4x_2^3 x_4 \qquad (1,0,1,0) \notin \tau(X_A^*)$$

The discriminant of a cubic polynomial in 1 variable

$$\Delta_A = 27\, x_1^2 x_4^2 - 18\, x_1 x_2 x_3 x_4 + 4\, x_1 x_3^3 + 4\, x_2^3 x_4 - x_2^2 x_3^2$$

The Newton polytope $N(\Delta_A)$ is the convex hull of the exponent vectors $\alpha_1 = (2,0,0,2), \alpha_2 = (1,0,3,0), \alpha_3 = (0,3,0,1)$, $\alpha_4 = (0,2,2,0), \alpha_5 = (1,1,1,1)$, in $\mathbb{R}^4$.

As the discriminant has two homogeneities read in the rows of the matrix A: the linear functions

$$\langle (1,1,1,1), \alpha_i \rangle, \quad \langle (0,1,2,3), \alpha_i \rangle$$

take the same values (4 and 6 respectively) for any $i = 1, \ldots, 5$, $N(\Delta_A)$ is a polygon lying in a two dimensional plane in $\mathbb{R}^4$. It is straightforward to check that $\alpha_1, \ldots, \alpha_4$ are vertices and α_5 is an interior point.

The results in [D. - Feichtner - Sturmfels, '07] allow to predict the monomials in Δ_A, even if we couldn't compute it (as it happens in the general case).

$$\tau(\Delta_A) = \left(\cup_{i=1}^4 \mathbb{R}_{\geq 0} e_i \right) + \langle (1,1,1,1), (0,1,2,3) \rangle,$$

(and all the multiplicities are equal to 1).

To visualize it, we mod out by the row space of A to have a two-dimensional representation.

Figure 1: Newton polygon, tropicalization and extreme monomials of the discriminant of a degree $\mathbf{3}$ polynomial: $\boldsymbol{x_2^3x_4}$, $\boldsymbol{x_2^2x_3^2}$, $\boldsymbol{x_1x_3^3}$, $\boldsymbol{x_1^2x_4^2}$

The discriminant of a cubic polynomial in 1 variable

So we choose a basis $\{(1,-2,1,0),\ (0,1,-2,1)\}$ of the kernel of A and we project $\pi : \tau(\Delta_A) \to \mathbb{R}^2$, so that $b_1 := \pi(e_1) = (1,0), b_2 := \pi(e_2) = (-2,1)$, $b_3 := \pi(e_3) = (1,-2), b_4 := \pi(e_4) = (0,1)$, and the image if the tropicalization is the union of the four positive rays generated by these vectors.

To compute a monomial in Δ_A we pick a point $w \notin \pi(\tau(\Delta_A))$, for instance a point in the interior of the positive cone $\mathcal{C}$ generated by b_2 and b_3. We now “place” the projection $\pi(\tau(\Delta_A))$at w and we see which rays emanating from there intersect $\pi(\tau(\Delta_A))$. The intersections are given by the point $(w + \mathbb{R}_{\geq 0}b_1) \cap \mathbb{R}_{\geq 0}b_3$, with $|\det(b_1, b_3)| = 2$ and by the point $(w+\mathbb{R}_{\geq 0}b_4)\cap\mathbb{R}_{\geq 0}b_2$, with $|\det(b_2, b_4)| = 2$. Therefore, the vertex of $N(\Delta_A)$ dual to $\mathcal{C}$ is the point $(2, 0, 0, 2)$.

The Wilkinson polynomial

Consider the Wilkinson polynomial

$$W_{20} = \prod_{i=1}^{20}(x+i) = \sum_{j=0}^{20} c_j x^j,$$

which is well known by its numerical unstability.

For instance, it clearly has 20 real roots, but the polynomial $W_{20}(x) + 10^{-9}x^{19}$ –obtained by adding an apparently small perturbation in the coefficient of x^{19}– has only only 12 real roots and 4 pairs of complex roots, which do not "seem" to have small imaginary part.

For instance, one of these pairs is approximately equal to $-16.57173899 \pm 0.8833156071i$.

On the other side, if we substract $10^{-9}x^{19}$ from W_{20} we get a polynomial with 14 real zeros.

We think that this so unstable behaviour could be explained by the fact that the vector of coefficients $\mathbf{w} = (20!, \ldots, 210, 1)$ of W_{20} is very close not only to the variety $\Delta_A = 0$ if ill-posed polynomials, but also very close to a *singular point* of the discriminant variety $\Delta_A = 0$, where $A = \{0, 1, \ldots, 20\}$.

The Wilkinson polynomial

The discriminantal variety is a hypersurface in an affine space of dimension $\mathbf{21}$, and its singularities have codimension one, that is, they define a $\mathbf{19}$-dimensional variety.

We have experimented with the following $\mathbf{2}$-dimensional family of polynomials of degree $\mathbf{20}$

$$\boldsymbol{W(a,b,x) := W_{20}(x) + ax^{19} + bx^{18}.}$$

The corresponding discriminant $\boldsymbol{D(a,b)}$ (which is a specialization of $\boldsymbol{\Delta_A}$) defines a singular curve traced inside the discriminant locus. The singularities of this curve $\boldsymbol{D(a,b) = 0}$ are close to the point $\boldsymbol{a = b = 0}$, that is, to the vector of coefficients $\mathbf{w}$ of the Wilkinson polynomial.

Figure 2 features sample points of $\boldsymbol{D(a,b) = 0}$ inside a small box around the origin, which is the point lying in the intersection of the two coordinate arrows.

The marks near the ends of these arrows indicate distance $\mathbf{10^{-9}}$ from $\mathbf{(0,0)}$ and we see several branches of $\boldsymbol{D(a,b) = 0}$ very close to the origin: $\mathbf{4}$ to the right –which are crossed when moving $\boldsymbol{a}$ from $\mathbf{0}$ to $\mathbf{10^{-9}}$, causing the drop in $\mathbf{8}$ of the number of real roots–, another $\mathbf{3}$ close to the left –which cause the drop in $\mathbf{6}$ of the number of real roots when we decrease $\boldsymbol{a}$–, plus two other branches still close more to the left.

These drawings were done by Bernard Mourrain using the subivision solver of Mathemagix – by means of the package *subdivix* developed by Elias Tsigaridas and Bernard

Mourrain – and visualized with the software Axel developed by Julien Wintz.

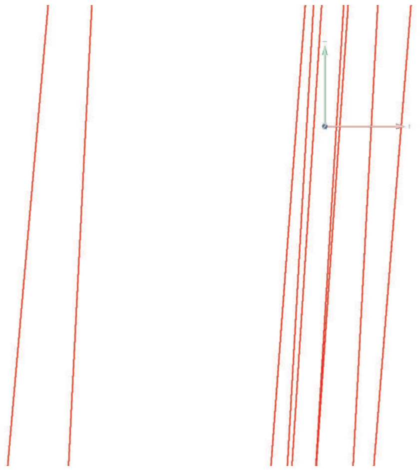


Figure 2: Branches of $D(a, b) = 0$ close to the origin

As the figure suggests, once we increment the parameter a from 0 to $2 \cdot 10^{-10}$, we already get a polynomial with only $16 = 20 - 2.2$ real roots (as we cross 2 of the branches). Also, it suggests that there are intersections of these branches close to $(0, 0)$, thus giving singularities of $D(a, b) = 0$ near the origin.

Considering not just the distance to the variety of ill posed problems $\Delta_A = 0$ but also to its singular locus would correspond in the case of conditioning of square $m \times m$ matrices in linear algebra, to consider not only the smallest and greatest absolute values of the singular values (or the distance to matrices of rank at most $m - 1$) but also the behaviour of the intermediate ones (or the distance to matrices of different ranks strictly smaller than $m - 1$).

Example: a mixed discriminant

- Consider the matrix

$$A := \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 6 & 0 & 0 & 0 & 3 & 1 \\ 0 & 3 & 1 & 6 & 0 & 0 \end{pmatrix}.$$

- A is the Cayley matrix associated to 2 planar configurations, and the A-discriminant $\Delta_A(y_1, \ldots, y_6)$ is the *mixed discriminant* of the family of polynomials

$$\begin{cases} h_1(y;t,s) := y_1 t^6 + y_2 s^3 + y_3 s^1 \\ h_2(y;t,s) := y_4 s^6 + y_5 t^3 + y_6 t^1 \end{cases}$$

- $\Delta_A(y) = 0$ whenever there exists a common zero $(s,t) \in (\mathrm{k}^*)^2$ which is not simple.

- The Horn-Kapranov parametrization of $X_A^* = (\Delta_A(y) = 0)$ is given by

$$y_1 = (-2u_1 + u_2)\, t_1 t^6, \quad y_2 = (u_1 - 6u_2)\, t_1 s^3,$$

$$y_3 = (-3u_1 + 6u_2)\, t_1 s, \quad y_4 = 2u_2\, t_2 s^6,$$

$$y_5 = (-6u_1 + u_2)\, t_2 t^3, \quad y_6 = (6u_1 - 3u_2)\, t_2 t.$$

..

and $\Delta_A(1, a, -1, 1, b, -1)$ equals

$$827540249418686807788221390646682295944670 72 * a^{47} * b^{33} +$$
$$24519711093887016527058411574716512472434688 * a^{46} * b^{39} -$$
$$24519711093887016527058411574716512472434688 * b^{46} * a^{39} +$$
$$23662740309026457547478521970718496800134567046336 0 * a^{28} * b^{7} +$$
$$17631004810327637966335552676449435712814331054687500 * a^{4} * b^{11} +$$

53 additional monomial terms of comparable size

It is a polynomial of degree 90 with 58 monomials and huge integer coefficients!

◇ Tropical $\boldsymbol{A}$-discriminants

- $\boldsymbol{A}$-discriminants are in general complicated polynomials which carry a lot of combinatorial information.
- In principle, we can compute $\boldsymbol{\Delta_A}$ by standard methods in elimination . . . but in practice we reach the limits of the current computations very easily.
- So, instead, we can try to get a first combinatorial approximation, which can nonetheless give us the information about discrete invariants as dimension and degree (and asymptotics), by computing its Newton polytope $\boldsymbol{N(\Delta_A)}$ or its tropicalization $\boldsymbol{\tau(X_A^*)}$.
- Tropicalization is an operation that turns complex projective varieties into polyhedral fans.

◇ Tropical $\boldsymbol{A}$-discriminants

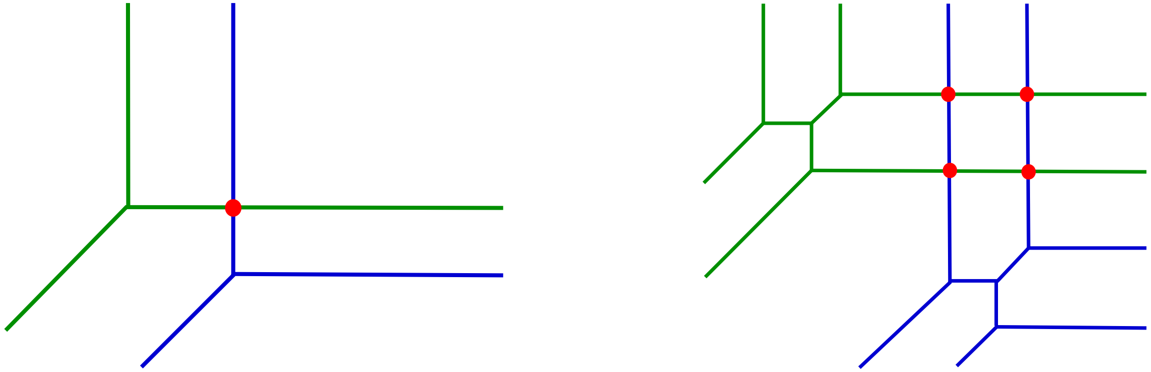
- The *tropicalization* $\boldsymbol{\tau(Y)}$ of a variety $\boldsymbol{Y}$ is the set

$$\boldsymbol{\tau(Y) \;=\; \{w \in \mathbb{R}^n \;:\; \mathbf{in}_w(I_Y)} \text{ does not contain a monomial} \boldsymbol{\}.}$$

- ... plus intersection theory information attached to each of the cones in the polyhedral fan $\boldsymbol{\tau(Y)}$ (Sturmfels-Tevelev '07)
- $\boldsymbol{\tau(Y)}$ is a pure dimensional fan of dimension $\mathbf{dim}\boldsymbol{(Y)}$ (Bieri-Groves)
- $\boldsymbol{\tau(Y)}$ can also be defined via valuations of the Puiseux series points of $\boldsymbol{Y}$ (Kapranov, etc.)

Tropical geometry is algebraic geometry over the tropical semiring

- $(\mathbb{R} \cup \{\infty\}, \oplus, \otimes)$, $\quad x \oplus y \; := \; \min\{x, y\}$, $\; x \otimes y \; := \; x+y$.
- algebraic varieties over valuated K ($\mathbb{C}$) $\xrightarrow{\tau}$ tropical varieties, (polyhedral fans)



- $f \; \in \mathbb{C}[x_1, \ldots, x_n]$ irreducible polynomial defining a hypersurface Y
- $\mathrm{New}(f)$ Newton polytope, $\mathcal{N}_{\mathrm{New}(f)}$ its normal fan. As a set

$$\tau(Y) \quad = \quad \text{codim } \mathbf{1}\text{-skeleton of } \mathcal{N}_{\mathrm{New}(f)}$$

- $\mathcal{L}(A)$ = geometric lattice whose elements are the sets of zero-entries of the vectors in $\mathbf{kernel}(A)$, ordered by inclusion.
- $\mathcal{C}(A)$ = set of proper maximal chains in $\mathcal{L}(A)$.
- We represent these chains as $(n-d-1)$-element subsets $\sigma = \{\sigma_1, \ldots, \sigma_{n-d-1}\}$ of $\{0,1\}^n$.
- The tropicalization of the kernel of A equals

$$\tau(\mathbf{kernel}(A)) \quad = \quad \bigcup_{\sigma \in \mathcal{C}(A)} \mathbb{R}_{\geq 0}\, \sigma \, .$$

- This tropical linear space is a subset of $\mathbb{R}^n$.
- Theorem: The tropical A-discriminant is the Minkowski sum of this tropical linear space and the (classical) row space of the $d \times n$-matrix A.

This is the tropical version of the Horn-Kapranov parametrization.

◇ Real polynomial systems with many real solutions

- Descartes' theorem (1637) for univariate polynomials allows to bound the number of real solutions in terms of the number of monomials independently of the degree.
- e.g. $\boldsymbol{x^d - a,\ 0 \neq a} \in \mathbb{R}$ has $\boldsymbol{d}$ complex solutions but at most $\mathbf{2}$ real solutions (and only one positive).
- A generalization to the multivariate setting is currently an open problem.

◇ Real polynomial systems with many real solutions

- Khovanskii (1980): There exists a (huge, non sharp) bound for the number of real solutions of a system of multivariate real polynomials in terms of the number of monomials which are present.
- Better bounds: only a few partial results (Li, Rojas & Wang (2002); Bihan, Sottile (2006))
- There exists a (false) conjecture by Koushnirenko, which in particular would imply that the number of positive simple real roots of a system of two trinomials in two variables is at most $\mathbf{4}$.
- There exists a counterexample by Haas (2002), with polynomials of degree $\mathbf{106}$ and $\mathbf{5}$ positive simple real solutions. In fact, $\mathbf{5}$ is the correct bound.
- "It is hard to find real sparse polynomials systems with many real solutions".

◇ Real polynomial systems with many real solutions

Understanding discriminants

we could prove that the two parameter family of real bivariate trinomials

$$H_{(a,b)} := \begin{cases} h_1(x,y) := x^6 + a\, y^3 - y \\ h_2(x,y) := y^6 + b\, x^3 - x \end{cases}$$

gives a far simpler family of counter-examples to Kushnirenko's Conjecture for $a = b = \frac{44}{31}$.

In fact, the area of the set of points $(a,b) \in \mathbb{R}^2$ such that the system has 5 positive real simple roots is smaller than 5.701×10^{-7}.

This is a dehomogenization of the generic family associated to the configuration

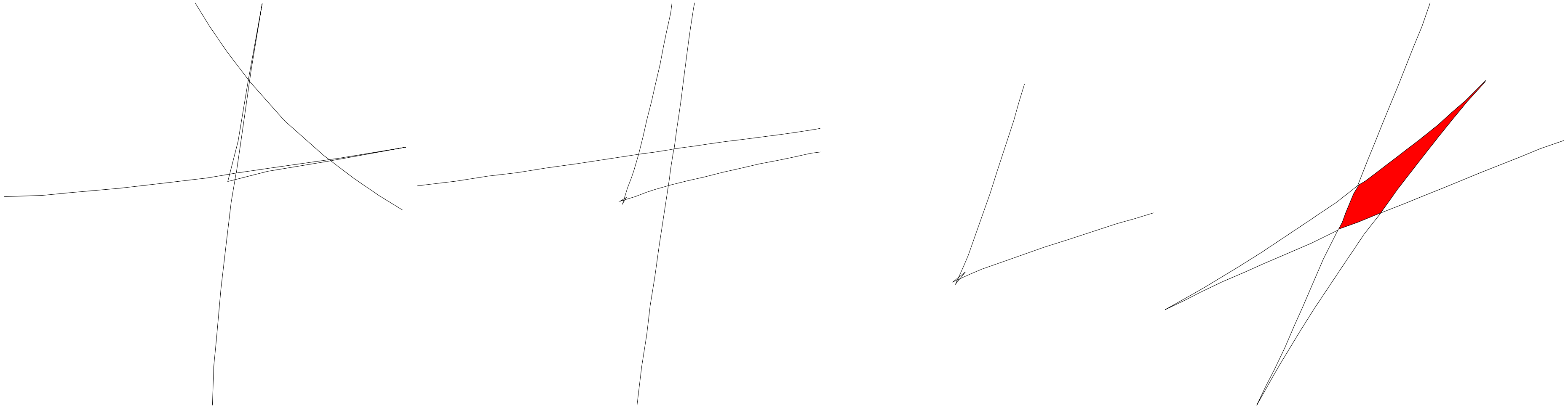
$$A = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 6 & 0 & 0 & 0 & 3 & 1 \\ 0 & 3 & 1 & 6 & 0 & 0 \end{pmatrix}.$$

◇ Real polynomial systems with many real solutions

- Two points in the same chamber (connected component) of the complement in $\mathbb{R}^2$ of the zero set of the dehomogenized discriminant $\nabla(a,b) = \Delta(1,a,-1,1,b,-1)$ have the same number of real roots (also the same number of positive real roots in this case).
- In particular, $\nabla_A(\frac{44}{31},\frac{44}{31}) \neq 0$, and this implies that $H_{(44/31,44/31)}$ has no degenerate roots.
- An implicit plot of $(\nabla_A = 0)$ has very poor quality, but instead we can draw it efficiently using the dehomogenized version of the Horn-Kapranov parametrization!

◇ Real polynomial systems with many real solutions

Below is a sequence of 4 plots, drawn on a logarithmic scale and successively magnified up to a factor of about $\mathbf{1700}$, of the real part of the discriminant variety $(\boldsymbol{\nabla_A = 0})$



♢ Recipe to cook the tropicalization of a rational planar curve from its parametrization

- DATA: $f = (f_1, f_2)$ rational functions of degree 0 in $t = (t_0, t_1)$, non cst.
- OUTPUT: The tropicalization of the closure of the image of f = the tropicalization of the principal ideal generated by an equation h of the closure of the image, or equivalently, its Newton poltope.
- PICTURES: On the blackboard...
- Step 1: Write $f_i = c_i \prod_j l_j^{v_{ij}}$, $l_j = \alpha_{0j} t_0 + \alpha_{1j} t_1$ linear forms, two by two linearly independent, $c_i \neq 0$
- Step 2: Call $b_j = (v_{1j}, v_{2j}), j = 1, \ldots, r$ lattice planar vectors. For each half line in the plane, add all b_j in that line. Call $b'_j, j = 1, \ldots, r'$ the resulting configuration and let d_j be the lattice length of b'_j.
- Step 3: Concatenate cyclically the vectors b'_j. As their sum is 0, we get a convex polygon. Rotate it 90 degrees and shift it until it intersects both positive coordinate axes.
- Step 4: Bring it to the table and enjoy it.
- It is not necessary to factorize f_i to get the vectors b_j [D'A-S]
- Note that the actual coefficients of the linear forms l_j are not needed, as soon as they are two by two linearly independent (i.e., as soon as we don't change the associated matroid).

- Two other approaches: D'Andrea-Sombra'07, Emiris-Konaxis-Palios'07.