

Computational Algebra: Big Ideas

Ioannis Z. Emiris

Dept. of Informatics & Telecoms



Outline

- 05. Idea: coefficients \equiv values (FFT)
- 17. Idea: matrices faster than Gauss
- 25. (Idea): real solving by remainders (Euclid)
- 41. intro to polynomial systems
- 48. Idea: algebra-geometry dictionary (Hilbert)
 - 52. Polynomial Degree
- 56. Idea: system solving by linear algebra
 - 68. Sylvester
 - 75. Macaulay
 - 80. Bilinear example
- 83: Idea: polynomials \equiv polytopes (Gelfand)
 - 92. Mixed subdivisions
 - 106. Sylvester-type sparse-resultant matrices
 - 122. Polynomial system solving
- 127: (Applications): geometric modeling, robotics, game theory

Big Questions for 2013

- Is multiplication of integers harder than adding?
Can we improve multiplication in the space of values (FFT)?
- What is the complexity of matrix multiplication?
- Can we efficiently solve (nonlinear) polynomial systems by linear algebra?
- Can combinatorics accelerate polynomial system solving?

Reading

coefficients \equiv values

matrices faster than Gauss

real solving by remainders (Euclid)

[Yap: Fundamental Problems in Algorithmic algebra]

varieties vs ideals (Hilbert) [Cox-L-O:Ideals,Varieties,Algorithms]

system solving by linear algebra [CLO:Using algebraic geometry,ch.3]

polynomials \equiv polytopes (Gelfand) [CLO:Using... ,ch.7]

[Sturmfels: Solving Systems of polynomial equations]

[Dickenstein-E: Solving polynomial equations: Foundations, Algorithms...]

Arithmetic operations [Yap, ch.1]

Computational model

Real RAM (Random Access Machine):

provides $O(1)$ storage/access time/space for reals,
requires $O(1)$ time for arithmetic operations on reals, performed exactly.
Hence counts arithmetic complexity, notation $O_A(\cdot)$.

Boolean RAM (or Turing machine):

provides $O(1)$ storage/access time/space for bits,
requires $O(1)$ time for operations on bits, performed exactly.
Hence counts bit/Boolean complexity, notation $O_B(\cdot)$.

Integers

Integers with n bits:

sum/difference with $\leq n + 1$ bits, in $\Theta_B(n)$.

Product with $\leq 2n$ bits, naive algorithm in $O_B(n^2)$.

Question: Is multiplication really harder? is it $O(n)$ additions?

Theorem. The asymptotic complexities of multiplication, division with remainder, inversion, and squaring are connected by constants.

Theorem [Karatsuba]

Divide+Conquer yields $O_B(n^{\lg 3}) = O_B(n^{1.585\dots})$.

Pf. $a = a_0 + 2^{n/2}a_1$, $ab = a_0b_0 + 2^{n/2}(a_0b_1 + b_0a_1) + 2^n a_1b_1$,
 $(a_0b_1 + b_0a_1) = (a_0 + a_1)(b_0 + b_1) - a_0b_0 - a_1b_1$.

$M(n) = 3M(n/2) + 4A(n/2) + 2A(n) = 3M(n/2) + O(n) = O(n^{\lg 3})$,
where complexities $M(n)$ of multiplication, $A(n)$ of addition.

Theorem. Fast Fourier Transform yields $O_B(n \log n \log \log n)$.

Univariate polynomials

$p_1(x), p_2(x) \in \mathbb{Z}[x]$, degrees d_1, d_2 , and t_1, t_2 terms. Let $d = \max\{d_1, d_2\}$.

The **sum** has degree $\leq d$, $\leq t_1 + t_2$ terms, cost $\Theta(d)$.

Product of degree $d_1 + d_2$, $\leq t_1 t_2$ terms, cost depending on the algorithm:

$$O_A(d_1 d_2), O_A(d^{\lg 3}), O_A(d \log^2 d), O_A(d \log d),$$

by school, D+C, evaluate/interpolate, FFT (no carry needed) algorithms.

In sparse representation: $O_A(t_1 t_2)$, $O_A(t^{\lg 3})$.

The arithmetic complexities of multiplication, squaring, division with remainder are connected with constants.

Integers to polynomials: Given binary integer $[c_{n-1} \ c_{n-2} \ \cdots \ c_0]$,

$\exists!$ polynomial $c_{n-1}x^{n-1} + c_{n-2}x^{n-2} + \cdots + c_0 \in \mathbb{Z}_2[x]$.

Evaluation

Horner's rule $p(a) = (\cdots (c_n a + c_{n-1})a + \cdots) + c_0$.

Requires n additions, n products, which is optimal.

Equivalent: $p(a) = p(x) \bmod (x - a)$,

since $p(x) = q(x)(x - a) + r(x)$, $\deg(r(x)) = 0$.

General Problem: Given k points/values x_0, \dots, x_{k-1} , and $n+1$ coefficients of $p(x)$, i.e. $\deg(p) = n$, compute k values $p(x_0), \dots, p(x_{k-1})$.

Horner yields $O_A(kn)$, we'll see a quasi-linear algorithm.

Quasi-linear multi-evaluation

[Note: this can be avoided if you go directly to FFT.]

Theorem. D+C algorithm $= O_A(n \lg^2 n)$, for $k = \Theta(n)$.

Lem. $a, b, c \geq 0 \Rightarrow (a \bmod (bc)) \bmod b = a \bmod b$.

Lem.

$$p(x) \bmod (x - x_i) = [p(x) \bmod \prod_{j \in J} (x - x_j)] \bmod (x - x_i), \quad i \in J \subset \mathbb{N}.$$

Quasi-linear algorithm: fan-in

Assume we have $k = n$ points. We compute $\prod_j (x - x_j)$, $j = 2^i - 1, \dots, 2^{i+1} - 1$, using fan-in, for appropriate i (see next page).

Leaves: Compute $n/2$ products of degree=2:

$$(x - x_{2i})(x - x_{2i+1}), i = 0, \dots, \frac{n}{2} - 1.$$

Then $n/4$ products of degree 4, then $n/2^j$ products of degree 2^j in

$$O_A((n/2^j)M(2^{j-1})) = O_A(nj), \quad j = 1, \dots, \lg n,$$

$M(t) = O(t \log t)$ corresponds to FFT multiplication.

Total $O(n(1 + \dots + \lg n)) = O(n \lg^2 n)$.

Quasi-linear algorithm: Fan-out

Given $q(x) = p(x) \bmod \prod_{i=0}^{n-1} (x - x_i)$, compute

$$p(x) \bmod \prod_{i=0}^{n/2-1} (x - x_i) = q(x) \bmod \prod_{i=0}^{n/2-1} (x - x_i),$$

and $q(x) \bmod \prod_{i=n/2}^{n-1} (x - x_i)$, i.e. 2 polynomials of degree $n/2 - 1$, in $O(n \log n)$ by FFT. Then, 4 mod operations in $4(n - 2)/2 \cdot O(\log n)$.

Stage k : compute 2^k remainders with divisor

$$\prod_i (x - x_i), \quad i = \frac{mn}{2^k}, \dots, \frac{(m+1)n}{2^k} - 1, \quad m = 0, \dots, 2^k - 1.$$

Divisor degree = $n/2^k$, remainder degree = $n/2^k - 1$, $k = 0, 1, \dots, \lg n$.

Cost per level = $2^k n / 2^k \cdot O(\lg n)$.

Total $T(n) = 2T(n/2) + 2O(n \lg n) = 2nkO(\lg n) = O(n \lg^2 n)$.

Example

$$p(x) = 5x^3 + x^2 + 3x - 2, x_i = -1, 0, 3, 9.$$

$$q_0(x) = p(x) \bmod (x+1)x = 7x - 2,$$

$$q_1(x) = p(x) \bmod (x-3)(x-9) = 600x - 1649.$$

$$p(x) \bmod (x+1) = q_0(x) \bmod (x+1) = -9,$$

$$p(x) \bmod x = q_0(x) \bmod x = -2,$$

$$p(x) \bmod (x-3) = q_1(x) \bmod (x-3) = 151,$$

$$p(x) \bmod (x-9) = q_1(x) \bmod (x-9) = 3751.$$

Interpolation

Def.: compute $n + 1$ coefficients of $p(x)$ given $n + 1$ values $r_i = p(x_i)$, $i = 0, \dots, n$ for distinct x_i 's, assuming the degree n is known.

Lagrange: $L(x) := \prod_{i=0, \dots, n} (x - x_i)$, $L'(x) = \sum_{i=0}^n \prod_{j \neq i} (x - x_j)$.
Then $L'(x_k) = \prod_{j \neq k} (x_k - x_j)$. Now define:

$$L_i(x) := \prod_{j=0, \dots, n, j \neq i} \frac{x - x_j}{x_i - x_j}.$$

Hence the solution is:

$$p(x) = L(x) \sum_{i=0}^n \frac{r_i}{L'(x_i)(x - x_i)} = \sum_{i=0}^n \frac{r_i}{L'(x_i)} \prod_{j \neq i} (x - x_j) = \sum_{i=0}^n r_i L_i(x).$$

Clearly p satisfies the data; it is also unique with degree $\leq n$.

Fan-in computes $L(x)$, $L'(x)$, $L'(x_0), \dots, L'(x_n)$, and $p(x)$ in $O_A(n \lg^2 n)$

FFT

Given polynomial

$$p(x) = c_{n-1}x^{n-1} + \dots + c_0,$$

compute values at the complex n -th roots of unity:

$$\{1, \omega = e^{2\pi i/n}, \omega^2 = e^{4\pi i/n}, \dots, \omega^{n-1} = e^{2\pi i(n-1)/n}\}.$$

Assume n is a power of 2:

$$\begin{aligned} p(x) &= (c_0 + c_2x^2 + \dots + c_{n-2}x^{n-2}) + x(c_1 + c_3x^2 + \dots + c_{n-1}x^{n-2}) = \\ &= q(x^2) + xs(x^2), \end{aligned}$$

and set $y = x^2$, where $q(y), s(y)$ of degree $(n-2)/2$.

Property 1. $x = \omega^j$, $j = 0, \dots, n-1$, then $y = \omega^{2j}$ takes only $n/2$ values.

Property 2. $\omega^j = -\omega^{j+n/2}$ reduces half of $q(y) + \dots$ to $q(y) - \dots$.

Complexity:

$$T(n) = 1.5n + 2T\left(\frac{n}{2}\right) = 1.5kn + 2^k T\left(\frac{n}{2^k}\right) = 1.5n \lg n + O(n) = O_A(n \lg n)$$

Inverse Fourier Transform

Def. Interpolate $(n - 1)$ -degree polynomial from values at n -th roots of 1

Let $n \times n$ Vandermonde matrix Ω with $\Omega_{ij} = [\omega^{ij} / \sqrt{n}]$, $0 \leq i, j < n$.

Fourier Transform computes

$$\sqrt{n} \Omega \begin{bmatrix} c_0 \\ \vdots \\ c_{n-1} \end{bmatrix} = [\omega^{ij}]_{i,j} \begin{bmatrix} c_0 \\ \vdots \\ c_{n-1} \end{bmatrix} = \begin{bmatrix} p(\omega^0) \\ \vdots \\ p(\omega^{n-1}) \end{bmatrix} =: p^T.$$

Inverse Transform: solve for c , given p : $c = \frac{1}{\sqrt{n}} \Omega^{-1} p^T$.

Lem. $\Omega^{-1} = [\omega^{-ij} / \sqrt{n}]$.

Pf. $\sum_k \omega^{-ik} \omega^{kj} = \sum_k \omega^{k(j-i)} = n$, if $i = j$; otherwise 0.

Cor. Since ω^{-1} is n -th root of 1, c is obtained by FFT.

Idea: Matrices faster than Gauss [Aho-Hopcroft-Ullman]

Matrices

Dense matrices $n \times m$: add/subtract in $\Theta_A(nm)$
(as opposed to sparse or structured matrices)

Square matrices $n \times n$: **Multiplication** $= \Omega_A(n^2)$.
Question: Is this tight?

Algorithms: school $= O_A(n^3)$.

D+C [Strassen'69] $O_A(n^{\lg 7}) = O_A(n^{2.81})$.

[Coppersmith-Winograd'90] $O_A(n^{2.376})$.

Record bound still holds, also achieved (2010) by other approach.

Strassen's algorithm

Given 2×2 matrices $[a_{ij}]$, $[b_{ij}]$, $i, j = 1, 2$, let the product be $[c_{ij}]$.

Set: $m_1 = (a_{12} - a_{22})(b_{21} + b_{22})$, $m_2 = (a_{11} + a_{22})(b_{11} + b_{22})$,
 $m_3 = (a_{11} + a_{12})b_{22}$, $m_4 = a_{22}(b_{21} - b_{11})$, $m_5 = a_{11}(b_{12} - b_{22})$,
 $m_6 = (a_{21} + a_{22})b_{11}$, $m_7 = (a_{11} - a_{21})(b_{11} + b_{12})$

$$\Rightarrow (c_{ij}) = \begin{bmatrix} m_1 + m_2 - m_3 + m_4 & m_3 + m_5 \\ m_4 + m_6 & m_2 + m_5 - m_6 - m_7 \end{bmatrix}$$

General dimension: replace a_{ij}, b_{ij}, c_{ij} by $\frac{n}{2} \times \frac{n}{2}$ submatrices A_{ij}, B_{ij}, C_{ij} .
Then,

$$M(n) = 7M\left(\frac{n}{2}\right) + O(n^2) \leq \dots \leq 7^k M(n/2^k) + kn^2 = O(n^{\lg 7}).$$

Matrix operations

Let $T(n)$ be the asymptotic arithmetic complexity of multiplication.

Inversion, determinant, solving $Mx = b$, factoring $M = LU$, and factoring with permutation $M = LUP$ (Gaussian elimination), all lie in $\Theta(T(n))$.

Compute the kernel $\{x : Mx = 0\}$ and the rank: both in $O(T(n))$.

Compute the characteristic polynomial in $O(T(n) \log^2 n)$.

Numeric approximation of eigen-vectors/values in $25n^3$.

Integer Determinant

Given is integer matrix $[a_{ij}]$, max entry length $L = \max_{ij} \{\lg |a_{ij}|\}$:
 Worst-case optimal bound on value [Hadamard]:

$$|\det A| \leq \prod_{i=1}^n \|a_i\|_2 \leq n^{n/2} \max\{|a_{ij}|\}^n.$$

1. Chinese remaindering avoids intermediate swell: $O^*(nL)$ evaluations modulo constant-length primes, each in $O^*(n^{2.38})$; Lagrange in $O_B^*(n^2 L^2)$.

Total: $O_B^*(n^{3.38} L)$.

2. Avoid rationals [Bareiss'68] in $\sum_{i=1}^n n^2 i L = O_B^*(n^4 L)$.

Let $[12k] = |a_{ij} : i = 1, 2, 3, j = 1, 2, k|$: Multiply by a_{11} rows 2 \dots , n , eliminate:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots \\ 0 & a_{11}a_{22} - a_{12}a_{21} & a_{11}a_{23} - a_{12}a_{21} \\ 0 & a_{11}a_{32} - a_{12}a_{31} & a_{11}a_{33} - a_{12}a_{31} \end{bmatrix} \rightarrow \begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots \\ 0 & a_{11}a_{22} - a_{12}a_{21} & \cdots & \cdots \\ 0 & 0 & a_{11}[123] & a_{11}[124] \end{bmatrix}$$

3. Baby steps / giant steps $O_B(n^{3.2} L)$ [Kaltofen-Villard'01]

$n \times n$ linear system

$\text{rank}(M) = r \leq n$:

- $r = n \Rightarrow \exists!$ solution.
- $r < n \Rightarrow$ system defined by r equations.

remaining equations trivial ($0=0$) implies ∞ roots.

existence of incompatible equation ($0=b$) implies no roots.

$\text{rank}(M)$ also defined for rectangular M .

Structured matrices

Defined by $O(n)$ elements, matrix-vector product is quasi-linear.

Two important examples:

- **Vandermonde**: matrix-vector multiply and solving in $O_A(n \log^2 n)$.
- Rectangular matrix is **Toeplitz** iff $M(a+i, b+i) = M(a, b)$, $i > 0$, when defined, i.e. constant diagonals. Lower triangular * vector is polynomial multiplication, hence in $O_A^*(n)$; same for vector * upper triangular.
 - More types: Hankel (constant anti-diagonals), Cauchy, Hilbert.

Thm [Wiedemann (Lanszos)]. Matrix determinant reduced to $O^*(n)$ matrix-vector products.

Proof. Krylov sequence $M^i v$ computed as $M(M^{i-1} v)$,
charpoly $\chi(\lambda) = \det(M - \lambda I) = (-1)^{\pm 1} \lambda^n \pm \text{tr}(M) \lambda^{n-1} + \dots \pm \det M$.
Caley-Hamilton thm: $\chi(M) = 0$, so $\chi(M)v = 0$.

Berlekamp-Massey: finds k -recurrence from $2k$ (vector) elements.

Toeplitz example

$$P_1(x) = x^4 - 2x^3 + 3x + 5, \quad P_2(x) = 5x^3 + 2x - 11.$$

Upper triangular Toeplitz T has rows corresponding to P_2 multiples:

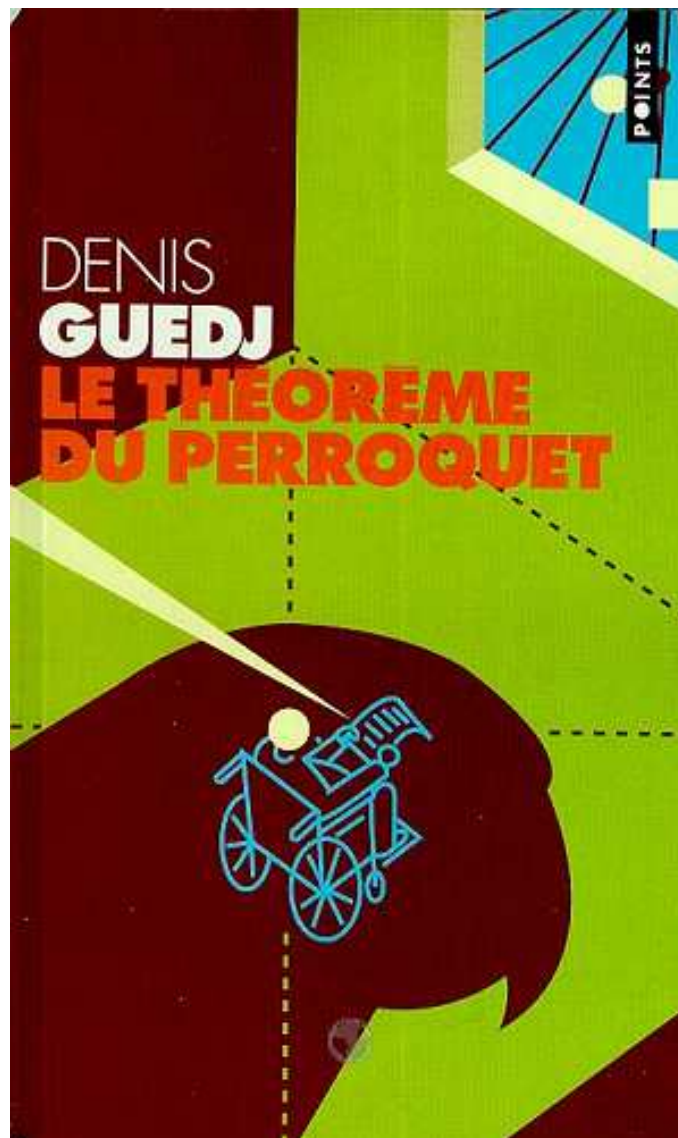
$$\begin{bmatrix} 5 & 0 & 2 & -11 & & & & 0 \\ & 5 & 0 & 2 & -11 & & & \\ & & 5 & 0 & 2 & -11 & & \\ & & & 5 & 0 & 2 & -11 & \\ 0 & & & & 5 & 0 & 2 & -11 \end{bmatrix} \begin{matrix} x^4 P_2 \\ x^3 P_2 \\ x^2 P_2 \\ x P_2 \\ P_2 \end{matrix}$$

Row vector $v = [1, -2, 0, 3, 5]$ expresses P_1 , then Vector-matrix multiplication vT is equivalent to polynomial multiplication

$$(P_1 P_2)(x) = 5x^7 - 10x^6 + 2x^5 + 47x^3 + 6x^2 - 23x - 55.$$

If multiplying polynomials of degree d costs $F(d)$, then multiplying $d \times d$ Toeplitz matrix by vector is in $O(F(d))$.

Real numbers



Univariate real solving

Univariate solving

- **Counting / Exclusion**

- Interval arithmetic (cf. Matlab)
- Descartes' rule, Bernstein basis (fast)
- **Sturm sequences**
- Thom's encoding (good asymptotics)

- **Approximation**

- Numeric solvers $O(d^3L)$
- Continued Fractions [E-Tsigaridas] (fast)

Polynomial in $\mathbb{Z}[x]$ of degree d and bitsize L .
Input size in $O(dL)$, output in $\Omega(dL)$.

Bit complexity of exact solvers

Cont.Frac.	Sturm	Descartes	Bernstein
$O^*(2^L)$ [Uspensky48] $O(d^5 L^3)$ [Akritis'80]	$O^*(d^7 L^3)$ [Heidel'71] $O^*(d^6 L^3)$ [Davenport'88]	$O^*(d^6 L^2)$ [Collins,Akritis'76] $O^*(d^5 L^2)$ [Krandick'95] [Johnson'98]	[LaneReisenfeld81] $O^*(d^6 L^3)$ [MourrainVrahatis] [-Yakoubson'04]
$O^*(d^8 L^3)$ [Sharma07]	$O^*(d^4 L^2)$ [DuSharmaYap05] [EigenwilligSharmaYap06] [E,Mourrain,T'06] + square-free + multiplicities [E,Mourrain,Tsigaridas'06]		
$O^*(d^4 L^2)$ [ET'06]	$O^*(r d^2 L^2)$	$O^*(d^3 L^2)$ [E,Tsigaridas]	

Polynomial in $\mathbb{Z}[x]$ of degree d and bitsize L .

Best numerical algorithm in $O(d^3 L)$, input = $O(dL)$.

Worst-case vs. [average-case](#) complexities, $r = \# \text{real-roots}$.

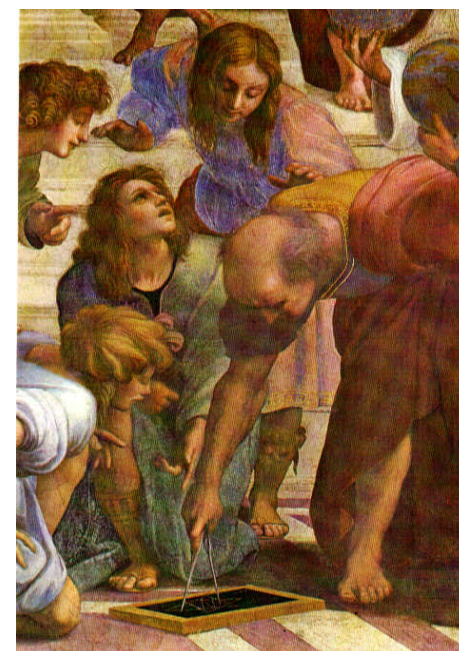
Sturm theory

Sturm sequences

Definition. Given univariate polynomials $P_0, P_1 \in \mathbb{R}[x]$, their **Sturm sequence** is any (pseudo-remainder) sequence of polynomials $P_0, P_1, \dots, P_n \in \mathbb{R}[x]$, $n \geq 1$ such that

$$\alpha_i P_{i-1} = Q P_i + \beta_i P_{i+1}, \quad i = 1, \dots, n-1,$$

for some $Q \in \mathbb{R}[x]$, $\alpha_i, \beta_i \in \mathbb{R}$, and $\alpha_i \beta_i < 0$.



Remember *Ευκλείδης*

Example of Sturm sequence

Input: $f_i = \alpha_i x^2 - 2\beta_i x + \gamma_i$, $i = 1, 2$.

Hypothesis: the f_i are relatively prime, $\alpha_i, \Delta_i \neq 0$.

The **Sturm sequence** (P_i) of $f_1, f_1' f_2$:

$$\begin{aligned}P_0(x) &= f_1(x) \\P_1(x) &= f_1'(x) f_2(x) \\P_2(x) &= -f_1(x) \\P_3(x) &= 2\alpha_1 [-(\alpha_1 K + 2\alpha_2 \Delta_1)x + (\gamma_1 J - \alpha_1 J')] \\P_4(x) &= -\alpha_1 \Delta_1 (\alpha_1 K + 2\alpha_2 \Delta_1)^2 (G^2 - 4JJ') \\&= -\alpha_1 \Delta_1 (\alpha_1 G - 2\beta_1 J)^2 (G^2 - 4JJ')\end{aligned}$$

Root counting

Theorem [Tarski]. Suppose that

- $f_0, f_1 \in \mathbb{R}[x]$ are relatively prime,
- f_0 is square-free, and
- $p < q$ are not roots of f_0 .

Then, for any Sturm sequence $P = (f_0, f'_0 f_1, \dots)$,

$$V_P(p) - V_P(q) = \sum_{f_0(\rho)=0, p < \rho < q} \text{sign}(f_1(\rho)),$$

where $V_P(p) := \# \text{sign variations in } P_0(p), \dots, P_n(p)$.

The **Sturm sequence** here may be $(f_0, f'_0 f_1, -f_0, \dots)$.

More uses of Sturm sequences

Corollary. For $p < q$ non-roots of $f \in \mathbb{R}[x]$, the number of distinct real roots of f in (p, q) equals $V_{f,f'}(p) - V_{f,f'}(q)$.

Proof. Let $f_0 = f, f_1 = 1$ in Tarski's theorem.

Theorem [Schwartz-Sharir]. For square-free $f_0, f_1 \in \mathbb{R}[x]$ and $p < q$ non-roots of f_0 ,

$$V_{f_0, f_1}(p) - V_{f_0, f_1}(q) = \sum_{f_0(\rho)=0, p < \rho < q} \text{sign}(f_0'(\rho)f_1(\rho)).$$

- Yields previous theorem by using $f_0, f_0'f_1$.

[Yap: Fundamental Problems of Algorithmic Algebra, 2000]

Generalizations of Sturm theory

Systems of univariate polynomials

Recall [Tarski]. For $f_0, f_1 \in \mathbb{R}[x]$ relatively prime, f_0 square-free and $p < q$ not roots of f_0 , consider the Sturm sequence $(f_0, f'_0 f_1, \dots)$. Then

$$V(p) - V(q) = \sum_{f_0(\rho)=0, p < \rho < q} \text{sign}(f_1(\rho)).$$

This equals

$$\# \{ \rho \in (p, q) : f_0(\rho) = 0, f_1(\rho) > 0 \} - \# \{ \rho \in (p, q) : f_0(\rho) = 0, f_1(\rho) < 0 \}.$$

Algorithm [Ben-Or, Kozen, Reif], [Canny]. Compute

$$\sum_{i=1}^n \# \{ \rho \in (p, q) : P_0(\rho) = 0, P_i(\rho) \otimes_i 0 \}, \quad \otimes_i \in \{<, >\}.$$

Generalized Sturm sequences

Definition. Given univariate polynomials $P_0, P_1 \in \mathbb{R}[x]$, where P_0 is square-free, their generalized Sturm sequence over an interval $[a, b] \subset \mathbb{R} \cup \{-\infty, +\infty\}$ is any sequence $P_0, P_1, \dots, P_n \in \mathbb{R}[x], n \geq 1$ s.t.

1. $P_0(a)P_0(b) \neq 0$,
2. $\forall c \in [a, b], P_n(c) \neq 0$,
3. $\forall c \in [a, b], P_j(c) = 0 \Rightarrow P_{j-1}(c)P_{j+1}(c) < 0$,
4. $\forall c \in [a, b] : P_0(c) = 0 \Rightarrow \exists [c_1, c), (c, c_2]$ s.t. $u \in [c_1, c) \Rightarrow P_0(u)P_1(u) < 0$ and $u \in (c, c_2] \Rightarrow P_0(u)P_1(u) > 0$.

Corollary (Existence). For any $P_0, P_1 \in \mathbb{R}[x]$, the previously-defined Sturm sequence, using the pseudo-remainders and starting with $P_0 / \gcd(P_0, P_0')$ and P_1 is “generalized” over an interval $[a, b]$ such that $P_0(a)P_0(b) \neq 0$.

Further generalization

Corollary. It is possible to omit [1. $P_0(a)P_0(b) \neq 0$] provided that, (4) is stated only in the appropriate subinterval of $[a, b]$, when $c = a$ or $c = b$.

Corollary. Relax (4) to require that the number of roots of $P_0(x)$ is odd between any two roots of $P_1(x)$.

Real Closed fields generalize \mathbb{R}

Definition. An **ordered field** K contains a positive subset $P \subset K$, ie.
 $a \in K - \{0\} \Rightarrow a \in P \text{ xor } -a \in P$.

Examples: $\mathbb{Q}, \mathbb{R}, \mathbb{Q}(\epsilon), \mathbb{R}(x), \mathbb{Q}(\sqrt[3]{2}) \equiv \mathbb{Q}[x]/\langle x^3 - 2 \rangle$.

Counter-example: \mathbb{C} .

Definition. A **real closed field** K is

- ordered (hence contains positive $P \subset K$),
- $a \in P \Rightarrow \sqrt{a} \in P$ (ie. $x^2 = a$ has a root in P),
- equations of odd degree have a root in P .

Examples: $\mathbb{R}, \mathbb{R}(\epsilon), \mathbb{R}(\epsilon_1, \epsilon_2)$.

Counter-example: \mathbb{Q} , algebraic closure $\overline{\mathbb{Q}}$, $\mathbb{Q}(\sqrt[3]{2})$.

Sturm sequences are defined, and all stated properties hold, for polynomials over **real closed fields**.

Descartes' rule

Descartes' rule of sign

Theorem. The number of **sign variations** in the coefficients of a univariate polynomial exceeds the number of positive **real roots** by an even non-negative integer.

Proof by induction, using Sturm sequences.

Step: $V[(x - a)f] = V[f] + \text{odd natural number}$.

Corollary. If **all roots** of the univariate polynomial are nonzero and real, then the number of sign variations in the coefficient sequence gives **precisely** the number of positive roots.

Proof by induction on the degree: the number of variations in the coefficients of $f(-x)$ bounds the number of negative roots.

Notions of Algebraic geometry

Introduction

Single polynomial

$$f(x) = c_0 + c_1x + c_2x^2 + \cdots + c_dx^d \in K[x].$$

- Fundamental theorem of **algebra**: There are d roots in \overline{K} .
E.g. $\overline{\mathbb{Q}} = \mathbb{C}$.
- Fundamental problem of **real algebra**: How many roots are real?
- Fundamental problem of **computational real algebra**: Isolate all real roots of a given polynomial equation.
- Fundamental problem of **computational algebraic geometry**: Isolate/approximate all complex roots of a given polynomial system.
- Fundamental problem of **computational real algebraic geometry**: Isolate all real roots of a given polynomial system.

Algebraic varieties

$$f_1, \dots, f_m \in \mathbb{Q}[x_1, \dots, x_n].$$

Defn. The polynomial system's **variety (or zero-set)** is

$$V(f_1, \dots, f_m) := \{x \in \mathbb{C}^n : f_1(x) = \dots = f_m(x) = 0\}.$$

Examples.

- $V(x^2 + 1) = \{\pm\sqrt{-1}\},$
- $V(\mathbb{Q}[x_1, \dots, x_n]) = \emptyset,$
- $V(\emptyset) = \mathbb{C}^n.$

Properties.

- $S \subset T \Rightarrow V(T) \subset V(S)$

Dimension of a variety

Def. $\dim(V) = \# \text{degrees of freedom of } V = \# \text{parameters for covering } V$

- $\dim(\text{point}) = 0$, $\dim(\text{line}) = 1$, $\dim(\text{surface}) = 2$.
- $\dim(V) = n \Leftrightarrow V = \mathbb{C}^n$.
- $\dim V(f_i) = n - 1$ generically.

Def. Dimension $\dim(V) := \max_C \{ \dim(C) : \text{component } C \subset V \}$.

- $\dim(V) = 0 \Leftrightarrow V = \text{point set (iff finite cardinality)}$.
- $\dim(V) = 1 \Leftrightarrow V$ contains a curve (possibly straight line), may contain points, but no component of $\dim \geq 2$.
- $\dim(V) = 2 \Leftrightarrow V$ contains a surface (possibly planar), may contain 0-dim or 1-dim components, but no higher-dim component.

Algebraic varieties (cont'd)

System $f_1, \dots, f_m \in K[x_1, \dots, x_n]$.

- Well-constrained: $m = n$, generically 0-dim variety.
- Over-constrained: $m > n$, generically no roots (empty).
- Under-constrained: $m < n$, generically ∞ roots.

Lemma.

- $V(f_1, \dots, f_m) = V(f_1) \cap \dots \cap V(f_m) \subset \mathbb{C}^n$.
- $\dim(V \cap W) = \dim(V) - \text{codim}(W)$,

where $\text{codim}(W) = n - \dim(W)$;

clearly, $\dim(V \cap W) = \dim(W) - \text{codim}(V)$.

E.g. \mathbb{C}^2 : $V, W = \text{curves}$, $\dim(W \cap V) = 0$ (points).

E.g. \mathbb{C}^3 : $V, W = \text{surfaces}$, $\dim(W \cap V) = 1$ (curve).

E.g. \mathbb{C}^3 : $V = \text{surface}$, $W = \text{curve}$, $\dim(W \cap V) = 0$.

$n \times n$ linear system

$\text{rank}(M) = r \leq n$:

- $r = n \Rightarrow \exists!$ solution.
- $r < n \Rightarrow$ system defined by r equations.

remaining equations trivial ($0=0$) implies ∞ roots.

existence of incompatible equation ($0=b$) implies no roots.

Hilbert's Nullstellensatz

Algebraic ideals

Given a polynomial ring $R = K[x_1, \dots, x_n]$,

- a subring $S \subset R$ is closed under addition and multiplication: $a, b \in S \Rightarrow a + b, ab \in S$;
- an (algebraic) ideal $I \subset R$ is closed under addition and multiplication by any ring element: $a, b \in I, p \in R \Rightarrow a + b, ap \in I$.

E.g. $\langle x^2, x^5 \rangle = \langle x^2 \rangle$, $\langle x, x + y \rangle = \langle x, y \rangle$.

Fact. Given a set of polynomials, all elements in the generated (algebraic) ideal vanish at the set's variety.

Corollary. The ideal is the largest set of polynomials vanishing precisely at this variety.

Varieties vs Ideals

Definition. Given set $X \subset \mathbb{C}^n$, $J(X) := \{f \in \mathbb{Q}[x] : f(x) = 0, \forall x \in X\}$.

Fact. $J(X)$ is an ideal.

Properties.

- $J(\mathbb{C}^n) = \emptyset, J(\emptyset) = \mathbb{Q}[x],$
- $X \subset Y \Rightarrow J(Y) \subset J(X),$
- $X = V(J(X))$
- $S \subset J(V(S))$: when is it tight?
- Counter-example: $\langle x^2 \rangle \neq J(\{0\}) = \langle x \rangle$:

How do the roots of x and x^2 differ?

Hilbert's Nullstellensatz

Recall definition $J(X) := \{f \in \mathbb{Q}[x] : f(x) = 0, \forall x \in X \subset \mathbb{C}^n\}$.

Defn. Given an ideal I in a commutative ring R , its **radical ideal** is

$$\sqrt{I} := \{r \in R \mid r^n \in I, \exists n > 0\}.$$

Property. $I \subset \sqrt{I}$.

Intuition: taking the radical removes the multiplicities.

Eg. In ring \mathbb{Z} : $\sqrt{\langle 8 \rangle} = \langle 2 \rangle$, $\sqrt{\langle 12 \rangle} = \langle 6 \rangle$,

In a polynomial ring: $\sqrt{\langle x^3 \rangle} = \langle x \rangle$, $\sqrt{\langle x^2, x - 2y, y^3 \rangle} = \langle x, y \rangle$.

Hilbert's zeroes theorem. $J(V(I)) = \sqrt{I}$.

Specifies the algebra-geometry dictionary.

Polynomial Degree

Degree

Defn: (total) degree of polynomial $F(x_1, \dots, x_n)$ is the maximum sum of exponents in any monomial (term).

E.g. $\deg(x^2 - xy^2 + z) = 3$.

We also talk of degree in some variable(s).

E.g.: $\deg_x(F) = 2$, $\deg_y(F) = 2$, $\deg_z(F) = 1$.

The polynomial is **homogeneous** (wrt to all n variables) if all monomials have the same degree.

E.g. $x^2w - xy^2 + zw^2$.

Here $w \neq 0$ is the homogenizing variable. So, for every (affine) root $(x, y, z) \in \mathbb{C}^3$ there is now a (projective) root $(x : y : z : 1) \in \mathbb{P}^3$.

Intersection theory

Geometrically, $\deg f(x_1, \dots, x_n)$ equals the number of intersection points of $f(x_1, \dots, x_n) = 0$ with a generic line in \mathbb{C}^n .

Defn. The degree of a variety V is $\#$ points in the intersection of V with a generic **linear subspace** L of dimension $= \operatorname{codim}(V)$:

$$\deg V = \#(V \cap L) : \dim L = \operatorname{codim} V.$$

E.g. curve $V \subset \mathbb{C}^3$ defined by $f(x, y, z) = g(x, y, z)$. L is a generic plane.

Number of roots

Defn. The complex **projective** space $\mathbb{P}_{\mathbb{C}}^n$ or \mathbb{P}^n or $\mathbb{P}(\mathbb{C})^n$ is the following set of equivalence classes:

$$\begin{aligned} & \left\{ (\alpha_0 : \cdots : \alpha_n) \in \mathbb{C}^{n+1} - \{0^{n+1}\} \mid \alpha \sim \lambda\alpha, \lambda \in \mathbb{C}^* \right\} = \\ & = \{(1 : \beta) \mid \beta \in \mathbb{C}^n\} \cup \{(0 : \beta) \mid \beta \in \mathbb{C}^n - \{0^n\}, \beta \sim \lambda\beta\}. \end{aligned}$$

E.g. $n = 1$: $\mathbb{P}^1 \simeq \mathbb{C} \cup \{(0 : 1)\}$.

Theorem [Bézout,1790]. Given (homogeneous) $f_1, \dots, f_n \in K[x_1, \dots, x_n]$, the number of its common roots (counting multiplicities) in $\mathbb{P}(\overline{K})^n$ is bounded by

$$\prod_{i=1}^n \deg f_i,$$

where $\deg(\cdot)$ is the polynomial's total degree.

The bound is exact for generic coefficients.

Note: The theorem considers dense polynomials.

Polynomial system solving

A perspective. . .



on La Boca

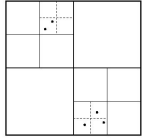
A perspective...



on system solving

Input: n polynomial equations in n variables, coefficients in a ring (e.g. \mathbb{Z} , \mathbb{R} , \mathbb{C}).

Output: All n -vectors of values s.t. all polynomials evaluate to 0.

Type	Algebraic	Analytic
Approach	Combine constraints	Use values (or signs)
Computation	Exact (+ possibly numerical)	Numerical mostly
Methods	<p>Matrix-based: resultant</p> <p>symbolic-numeric computation</p> <ul style="list-style-type: none"> + exploit structure + continuity w.r.t. coefficients – high-dimensional components <p>$O_b^*(d^n)$</p> <p>Gröbner bases</p> <ul style="list-style-type: none"> + complete information – discontinuity w.r.t. coefficients <p>dimension=0: $O_b^*(d^{n^2})$, else $O_b^*(d^{2^n})$</p> <p>Characteristic sets</p> <p>dimension=0: $O_b^*(d^n)$, else $O_b^*(d^{n^2})$</p> <p>Normal forms, boundary bases</p> <p>Straight-line programs</p> <p>express evaluation</p>	<p>Newton-like, optimization, discretization</p> <ul style="list-style-type: none"> + simple, fast – local, may need initial point <p>Exclusion, interval, topological degree</p> <ul style="list-style-type: none"> + simple, flexible, robust + focuses on given domain – costly for large n  <p>$O_b^*(\log \frac{D}{\epsilon})$</p> <p>Homotopy continuation</p> <ul style="list-style-type: none"> + exploit structure – divergent paths

Resultants

Resultant definition

Given $n + 1$ **Laurent** polynomials $f_0, \dots, f_n \in K[x_1, \dots, x_n, x_1^{-1}, \dots, x_n^{-1}]$ with indeterminate coefficients \vec{c} , their **projective**, resp. **toric / sparse**, *resultant* is the unique (up to sign) irreducible polynomial $R(\vec{c}) \in \mathbb{Z}[\vec{c}]$ such that

$$R(\vec{c}) = 0 \Leftrightarrow \exists \xi = (\xi_1, \dots, \xi_n) \in X : f_0(\xi) = \dots = f_n(\xi) = 0$$

where the variety X equals:

- the projective space \mathbb{P}^n over the algebraic closure \overline{K} ,
- resp. the **toric variety** X , $(\overline{K}^*)^n \subset X \subset \mathbb{P}^N$.

[van der Waerden, Gelfand-Kapranov-Zelevinsky, Cox-Little-O'Shea]

Resultant degree

The **projective**, resp. **toric**, resultant polynomial $R \in \mathbb{Z}[\vec{c}]$ is separately homogeneous in the coefficients of each f_i , with *degree* equal to $\prod_{j \neq i} \deg f_j$ (**Bézout's number**), resp. the n -fold **mixed volume**:

$$\text{MV}_{-i} := \text{MV}(f_0, \dots, f_{i-1}, f_{i+1}, \dots, f_n),$$

provided the supports of the f_i generate \mathbb{Z}^n .

Generalizations

The **toric** resultant reduces to:

- the determinant of the coefficient matrix of a *linear* system,
- the Sylvester or Bézout determinant of 2 *univariate* polynomials,
- the **projective** resultant for $n + 1$ *dense* polynomials, where the toric variety equals \mathbb{P}^n and $\text{MV}_{-i} = \prod_{j \neq i} \deg f_j$.

Resultants

- **Projective** (classical): over \mathbb{P}^n .
- **Toric** (sparse): over toric variety $X : (\overline{K}^*)^n \subset X \subset \mathbb{P}^N$
- **Over unirational variety** X s.t. $(\overline{K}^*)^n \subset X \subset \mathbb{P}^N$ [Busé-Elkadi-Mourrain].
- **Residual**: over $V(F : G)$ $F : G = \{f : fg \in F, \forall g \in G\}$ [Bus-Elk-Mour].

Resultant matrices

- **Sylvester**, Macaulay, Dixon, [Canny-E'93], rational [D'Andrea'02, E-Konaxis'09].
- **Bézout**, [Kapur et.al], [Elkadi-Mourrain].
- **Hybrid**: Morley, [Jouanolou], [D'Andrea-Dickenstein'01], Tate resolution [Khetan'02], multigraded [Dickenstein-E'03, E-Mantzaflaris'09] etc

Resultant formulae

How to compute and represent the resultant polynomial?

- The resultant divides a resultant matrix determinant.
- The resultant equals the ratio of two matrix determinants (rational, or Macaulay-type formula)
- The resultant equals a matrix determinant (optimal determinantal formula)
- Poisson formula.

Poisson

Given $f_0, \dots, f_n \in K[x_1, \dots, x_n]$, with coefficients $c = (c_0, \dots, c_n)$ in K .

Poisson formula:

$$R = T \cdot \prod_{\alpha \in V(f_1, \dots, f_n)} f_0(\alpha)$$

where V is (generically) a 0-dimensional variety $\subset \mathbb{C}^n$, and T is a polynomial in c_1, \dots, c_n such that R is a polynomial in c .

Corollary. By Bézout's bound on $\#V$:

$$\deg_{c_0} R = \prod_{i=1}^n \deg f_i.$$

In the toric setting, $\deg_{c_0} R = \text{MV}(f_1, \dots, f_n)$ by BKK.

Toy example

$$f_0 = c_{01}x + c_{00}$$

$$f_1 = c_{11}x + c_{10}$$

$$R = \det \begin{bmatrix} c_{01} & c_{00} \\ c_{11} & c_{10} \end{bmatrix} = c_{01}c_{10} - c_{00}c_{11}$$

Solve f_0 yields $x_0 = -c_{00}/c_{01}$. Substitute, then

$$R \sim f_1(x_0) = c_{11}(-c_{00}/c_{01}) + c_{10}.$$

Compare to the Poisson formula.

Linear system

$$f_0 = c_{01}x + c_{02}y + c_{00}$$

$$f_1 = c_{11}x + c_{12}y + c_{10}$$

$$f_2 = c_{21}x + c_{22}y + c_{20}$$

$$R = \det \underbrace{\begin{bmatrix} x & y & 1 \\ c_{01} & c_{02} & c_{00} \\ c_{11} & c_{12} & c_{10} \\ c_{21} & c_{22} & c_{20} \end{bmatrix}}_M \begin{matrix} f_0 \\ f_1 \\ f_2 \end{matrix}$$

For indeterminates c_{ij} : $R \neq 0$ iff there is **no common solution**.

$R = 0$ iff there is a (unique) solution of $f_i = 0 \Leftrightarrow \exists \vec{v} \neq \vec{0} : M\vec{v} = \vec{0}$.

Note $M \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} = \begin{bmatrix} f_0(x_0, y_0) \\ f_1(x_0, y_0) \\ f_2(x_0, y_0) \end{bmatrix}$ vectors indexed by the **column monomials**

Linear system (cont'd)

Develop $\det M$ along the f_0 row:

$$\det M = c_{01} \begin{vmatrix} c_{12} & c_{10} \\ c_{22} & c_{20} \end{vmatrix} + c_{02} \begin{vmatrix} c_{11} & c_{10} \\ c_{21} & c_{20} \end{vmatrix} + c_{00} \begin{vmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{vmatrix}$$

equals $f_0(x_0 : y_0 : 1)$, where $\alpha = (x_0, y_0) \in \mathbb{C}^2$ is the root of $f_1 = f_2 = 0$.

Poisson formula: $R = \begin{vmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{vmatrix} f_0(\alpha), \quad \alpha \in \mathbb{C}^2 : f_1(\alpha) = f_2(\alpha) = 0.$

Sylvester matrix

Overconstrained system

$$\begin{aligned} f_0 &= a_{d_0}x^{d_0} + \cdots + a_0, & a_{d_0} &\neq 0, \\ f_1 &= b_{d_1}x^{d_1} + \cdots + b_0, & b_{d_1} &\neq 0. \end{aligned}$$

Define S :

$$R = \det \begin{matrix} & x^{d_0+d_1-1} & \cdots & x & 1 \\ \left[\begin{array}{ccccc} a_{d_0} & \cdots & a_0 & & 0 \\ & \ddots & & \ddots & \\ 0 & & a_{d_0} & \cdots & a_0 \\ b_{d_1} & \cdots & \cdots & b_0 & 0 \\ 0 & b_{d_1} & \cdots & \cdots & b_0 \end{array} \right] & \begin{matrix} f_0^* \\ \\ \\ f_1^* \end{matrix} & \begin{matrix} x^{d_1-1} \\ \vdots \\ 1 \\ x^{d_0-1} \\ \vdots \\ 1 \end{matrix} \end{matrix} \left. \begin{matrix} \vphantom{\begin{matrix} x^{d_1-1} \\ \vdots \\ 1 \\ x^{d_0-1} \\ \vdots \\ 1 \end{matrix}} \right\} \begin{matrix} B_0 \\ B_1 \end{matrix} \right\}$$

Lem. S is a square matrix, and $R = \det S$ (see below).

Poisson formula:
$$R = b_{d_1}^{d_0} \prod_{\alpha: f_1(\alpha)=0} f_0(\alpha).$$

Sylvester matrix: vector multiplication

Lemma. M is a Sylvester, or Sylvester-type matrix, of f_0, \dots, f_n , and row vector v contains the coefficients of g_0, \dots, g_n . Then, computing $v^T M$ is equivalent to computing $\sum_{i=0}^n f_i g_i$.

Example. $f_1 := c_0 + c_1x + c_2xy$, $B_1 := \{1, x\}$, $g_1 := s_0 + s_1x$:

$$\begin{bmatrix} 1 & x & \cdots \end{bmatrix} \begin{bmatrix} 1 & x & xy & x^2 & x^2y \\ c_0 & c_1 & c_2 & 0 & 0 \\ 0 & c_0 & 0 & c_1 & c_2 \\ \vdots & & & & \end{bmatrix} \begin{bmatrix} f_1 \\ xf_1 \\ \vdots \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & x & xy & x^2 & x^2y \\ s_0c_0 & s_0c_1 + s_1c_0 & s_0c_2 & s_1c_1 & s_1c_2 \end{bmatrix} + \cdots.$$

□

Sylvester matrix (cont'd)

Lemma. If S is the Sylvester matrix of f_0, f_1 , then

$$\det S = 0 \Leftrightarrow \deg \gcd(f_0, f_1) \geq 1.$$

Proof. $[\Leftarrow]$ $\deg \gcd(f_0, f_1) \geq 1 \Rightarrow \exists r \in \mathbb{C}$: root of the (univariate) gcd, hence $f_0(r) = f_1(r) = 0$.

Now nonzero column vector $[r^{d_0+d_1-1}, \dots, r^2, r, 1]$ lies in the right kernel of S , hence $\det S = 0$.

$[\Rightarrow]$ $\det S = 0 \Rightarrow \exists w \neq 0$ vector s.t. $wS = 0$. Consider w contains the coefficients of polynomials q_0, q_1 of degrees $d_1 - 1, d_0 - 1$, hence

$$f_0 q_0 + f_1 q_1 = 0 \Rightarrow f_0 q_0 = -f_1 q_1,$$

which has degree $< d_0 + d_1$, hence $\deg \text{lcm}(f_0, f_1) < d_0 + d_1$. Then,

$$\gcd = \frac{f_0 f_1}{\text{lcm}} \Rightarrow \deg \gcd(f_0, f_1) = d_0 + d_1 - \deg \text{lcm}(f_0, f_1) > 0.$$

Cor. $R = \det S$.

Sylvester matrix (for solving)

Well-constrained system:

$$\begin{aligned} f_0 &= (2y)x + (y - 3), \\ f_1 &= yx^2 + 4x + (-y + 5). \end{aligned}$$

$$\begin{array}{c} x^2 \quad x \quad 1 \\ \begin{array}{c} xf_0 \\ f_0 \\ f_1 \end{array} \end{array} \begin{bmatrix} 2y & y-3 & 0 \\ 0 & 2y & y-3 \\ y & 4 & -y+5 \end{bmatrix} \begin{bmatrix} x^2 \\ x \\ 1 \end{bmatrix} = \begin{bmatrix} xf_0(x,y) \\ f_0(x,y) \\ f_1(x,y) \end{bmatrix}$$

System solving reduced to an eigenproblem:

$$\left(\underbrace{\begin{bmatrix} 0 & -3 & 0 \\ 0 & 0 & -3 \\ 0 & 4 & 5 \end{bmatrix}}_{M_0} + \beta \underbrace{\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & -1 \end{bmatrix}}_{M_1} \right) \underbrace{\begin{bmatrix} \alpha^2 \\ \alpha \\ 1 \end{bmatrix}}_v = \vec{0}$$

$$\Rightarrow \exists v : (M - \beta I) v = 0 \text{ for } |M_1| \neq 0, \quad M := -M_1^{-1} M_0.$$

The u -resultant: example

$$\begin{aligned} f_0 &= u_1 x_1 + u_2 x_2 + u_0 = 0, \\ f_1 &= x_1^2 + x_1 x_2 + 2x_1 + x_2 - 1 = 0, \\ f_2 &= x_1^2 + 3x_1 - x_2^2 + 2x_2 - 1 = 0. \end{aligned}$$

The Macaulay matrix =

$$\begin{bmatrix} 1 & 1 & 2 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 2 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 2 & 0 & 0 & 1 & -1 \\ 1 & 0 & 3 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 3 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 3 & 0 & -1 & 2 & -1 \\ 0 & u_1 & 0 & u_2 & u_0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & u_1 & 0 & u_2 & u_0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & u_1 & 0 & 0 & u_2 & u_0 & 0 \\ 0 & 0 & 0 & 0 & 0 & u_1 & 0 & 0 & u_2 & u_0 \end{bmatrix}$$

The u -resultant is $|M| = (u_1 - u_2 + u_0)(-3u_1 + u_2 + u_0)(u_2 + u_0)(u_1 - u_2)$
 \Rightarrow the roots are $(1, -1), (-3, 1), (0, 1), (0 : 1 : -1)$.

Deleting row 5, column 7 yields a toric resultant matrix with:

number of rows per polynomial = 4,3,2, whereas the optimal would be $MV_{-i} = 4, 2, 2$.

The u -resultant (cont'd)

$$\begin{aligned} f_0 &= u_1 x_1 + u_2 x_2 + u_0, \\ f_1 &= c_{12} x_1^2 + c_{11} x_1 + c_{10} \\ f_2 &= c_{22} x_2^2 + c_{21} x_2 + c_{20} \end{aligned}$$

A toric resultant matrix =

$$\begin{bmatrix} c_{10} & c_{12} & c_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{10} & c_{11} & c_{12} & 0 \\ c_{20} & 0 & 0 & c_{22} & c_{21} & 0 & 0 & 0 \\ 0 & 0 & c_2 & 0 & 0 & c_{22} & 0 & c_{22} \\ u_0 & 0 & u_1 & 0 & u_2 & 0 & 0 & 0 \\ 0 & u_1 & u_0 & 0 & 0 & u_2 & 0 & 0 \\ 0 & 0 & 0 & u_2 & u_0 & u_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & u_0 & u_1 & u_2 \end{bmatrix} \begin{matrix} f_1 \\ x_2 f_1 \\ f_2 \\ x_1 f_2 \\ f_0 \\ x_1 f_0 \\ x_2 f_0 \\ x_1 x_2 f_0 \end{matrix}$$

defined by “row” monomials s.t. the matrix be **square** (dialytic elimination).
There also exist algorithms [Canny-E'92,'93,'00].

The number of rows per polynomial = 4,2,2 = MV_{-i} i.e. optimal, hence:

$$R(u) = \det M(u) = C \cdot \prod_{f_i(\alpha)=0, i=1,2} (u_1 \alpha_1 + u_2 \alpha_2 + u_0).$$

Matrix construction

$$n \geq 2$$

Problem: construct M , equivalently define the row monomials, s.t.:

- M is square
- $\det M \neq 0$ for generic coefficients
- $\det M = 0$ if $R = 0$ (ideally iff $R = 0$)
- Hopefully $\deg_{f_0} \det M = \deg_{f_0} R$.
- Sylvester-type: rows contain the coefficient vector of f_i

Algorithms for Sylvester-type matrices:

- Dialytic elimination (heuristic)
- Macaulay's for the projective resultant
- Canny-E for the toric/sparse resultant
- Further methods for special cases: Khetan, D'Andrea . . .

Macaulay's construction

Given $f_0, \dots, f_n \in K[x_1, \dots, x_n]$ **dense**, of total degree d_i .

Let T be the set of monomials $t \in K[x]$ of degree $\leq \nu = \sum_{i=0}^n d_i - n$.

Let $B_n := \{t \in T \mid \deg_{x_n} t \geq d_n\}$, $B_{n-1} := \{t \in T - B_n \mid \deg_{x_{n-1}} t \geq d_{n-1}\}, \dots$

Generally $B_i := \left\{ t \in T - \bigcup_{j=i+1}^n B_j \mid \deg_{x_i} t \geq d_i \right\}$, $i = 1, \dots, n$,

and $B_0 := T - \bigcup_{i=1}^n B_i$.

Macaulay's construction (cont'd)

Lemma. The B_i 's partition T , such that

$$|B_i| \geq \prod_{j \neq i} d_j, \text{ and } |B_0| = \prod_{j \neq 0} d_j.$$

Proof.

$|T| = \binom{n+\nu}{n}$ correspond to lattice points in a n -dim simplex of "size" ν .

B_0 contains no point of total degree ν .

Also, the $B_i, i \geq 1$ cut out pieces from the ν -simplex, hence B_0 corresponds to a hyper-rectangle with x_i -edge length $= d_i$.

B_n contains points in an n -dim simplex of size $\nu - d_n = \sum_{i < n} d_i - n$. So $|B_n| = \binom{\sum_{i < n} d_i}{n} \geq \prod_{i < n} d_i$ by a combinatorial argument.

Exercise: prove the lemma for B_1, \dots, B_{n-1} .

Macaulay matrix

Define the Macaulay matrix with rows expressing,

$$tf_0, \text{ for } t \in B_0, \quad \frac{t}{x_i^{d_i}} f_i, \text{ for } t \in B_i, i = 1, \dots, n.$$

Thm. [Macaulay'1902] The Macaulay matrix M is:

- (a) square, (b) generically nonsingular, (c) $R \mid \det M$,
- (d) \exists submatrix $M' : R = \det M / \det M'$.

Pf. Next slide.

Cor. The above construction implies $M = M_0$ is s.t.

$$\deg_{f_0} R = \prod_{i=1}^n \deg f_i = \deg_{f_0} |M_0|.$$

Analogously can define M_1, \dots, M_n . Then $R = \gcd(|M_0|, \dots, |M_n|)$.

Proof of theorem (a-c)

Thm. The Macaulay matrix M is: (a) square,
(b) generically nonsingular, (c) $R \mid \det M$.

Proof.

(a) Columns and rows are (essentially) indexed by T .

(b) Take a typical row of f_i :

$$f_i = \cdots + c_i x^{d_i} + \cdots \Rightarrow \text{row contains } \frac{t}{x_i^{d_i}} f_i = \cdots + c_i t + \cdots,$$

hence c_i appears on the diagonal. For a specialization where all f_i coefficients $\rightarrow 0$, $c_i \rightarrow 1$, we have $f_i \rightarrow x_i^{d_i}$, $f_0 \rightarrow 1$, thus $M \rightarrow I$.

(c) We showed $\deg_{f_i} |M| \geq \deg_{f_i} R$. Now, $R = 0 \Rightarrow \exists v$: contains values of T at root; $v \neq 0$ because $1 \in T$, and $Mv = 0 \Rightarrow \det M = 0$
(holds for every **Sylvester-type** resultant matrix).

Proof of theorem (d)

Thm. For the Macaulay matrix M :

(d) \exists submatrix $M' : R = \det M / \det M'$.

Proof. (d) Define the **reduced** monomials in T :

- (i) of degree ν , divisible by $x_i^{d_i}$, for exactly one $i \in \{1, \dots, n\}$, or
- (ii) of degree $\leq \nu - d_0$, not divisible by any $x_i^{d_i}$, $i \in \{1, \dots, n\}$,
i.e. divisible by $x_0^{d_0}$ only, for homogenizing variable x_0 .

M' has rows/columns indexed by the **non-reduced** monomials in T .

Properties:

- $\det M' \not\equiv 0$ by similar proof as for $\det M$.
- $\deg_{f_i} |M'| = \deg_{f_i} R - \deg_{f_i} |M|$.
- $\det M' = 0 \Rightarrow \det M = 0$: for specialization in (b) $\det M \mid \det M'$.

A bilinear example

Example: Bilinear surface

A bilinear surface in \mathbb{R}^3 is the set of **values** (x_1, x_2, x_3) :

$$x_i = c_{i0} + c_{i1}s + c_{i2}t + c_{i3}st, \quad i = 1, 2, 3, \quad \text{for } s, t \in [0, 1],$$

as well as the set of **roots** of some polynomial equation $H(x_1, x_2, x_3) = 0$.



Modeling/CAD use **parametric** AND **implicit/algebraic** representations
 \Rightarrow need to implicitize a (hyper)surface given a (rational) parameterization.

Bilinear system: Resultant matrix

$$f_i = (c_{i0} - x_i) + c_{i1}s + c_{i2}t + c_{i3}st, \quad i = 1, 2, 3.$$

The classical **projective** resultant vanishes identically.

The **toric (sparse)** resultant has $\deg R = 3 \cdot \deg_{f_i} R = 6$.

A **determinantal** Sylvester-type formula for the toric resultant is:

$$R = \det \begin{array}{cccccc} & 1 & s & t & st & s^2 & s^2t \\ \left[\begin{array}{cccccc} c_{10} - x_1 & c_{11} & c_{12} & c_{13} & 0 & 0 \\ c_{20} - x_2 & c_{21} & c_{22} & c_{23} & 0 & 0 \\ c_{30} - x_3 & c_{31} & c_{32} & c_{33} & 0 & 0 \\ 0 & c_{10} - x_1 & 0 & c_{12} & c_{11} & c_{13} \\ 0 & c_{20} - x_2 & 0 & c_{22} & c_{21} & c_{23} \\ 0 & c_{30} - x_3 & 0 & c_{32} & c_{31} & c_{33} \end{array} \right] & \begin{array}{l} f_1 \\ f_2 \\ f_3 \\ sf_1 \\ sf_2 \\ sf_3 \end{array} \end{array}$$

Sparse elimination theory

Newton polytopes

The **support** A_i of a polynomial $f_i \in K[x_1^{\pm 1}, \dots, x_n^{\pm 1}]$, s.t.

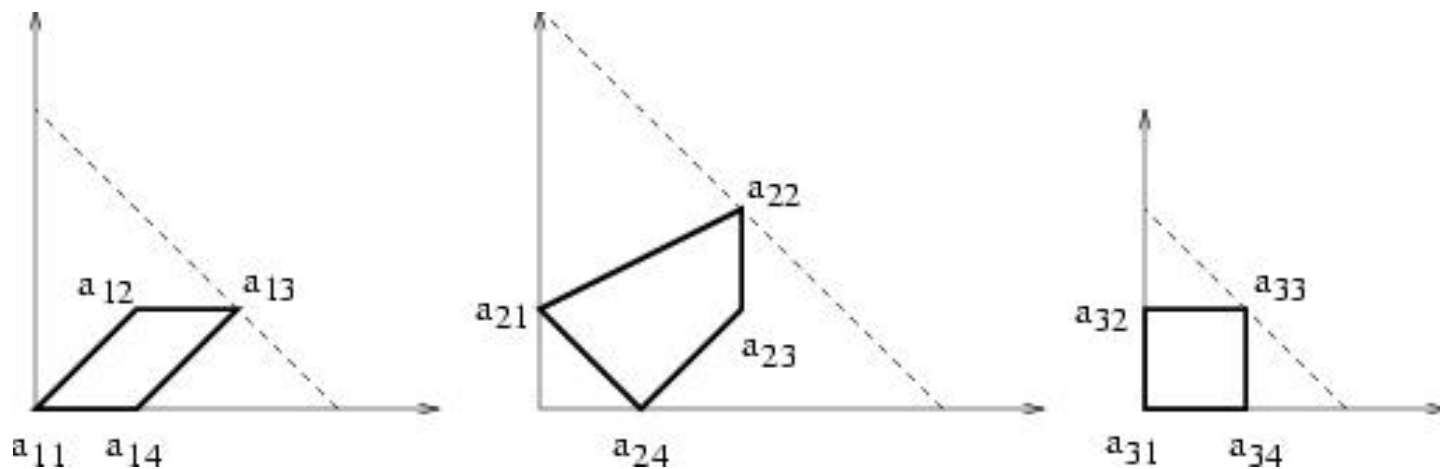
$$f_i = \sum_j c_{ij} x^{a_{ij}}, \quad c_{ij} \neq 0,$$

is defined as the set $A_i := \{a_{ij} \in \mathbb{Z}^n : c_{ij} \neq 0\}$.

The **Newton polytope** $Q_i \subset \mathbb{R}^n$ of f_i is the **Convex Hull** of all $a_{ij} \in A_i$.

Example:

$$\begin{aligned} f_1 &= c_{11} + c_{12}xy + c_{13}x^2y + c_{14}x \\ f_2 &= c_{21}y + c_{22}x^2y^2 + c_{23}x^2y + c_{24}x + c_{25}xy \\ f_3 &= c_{31} + c_{32}y + c_{33}xy + c_{34}x \end{aligned}$$



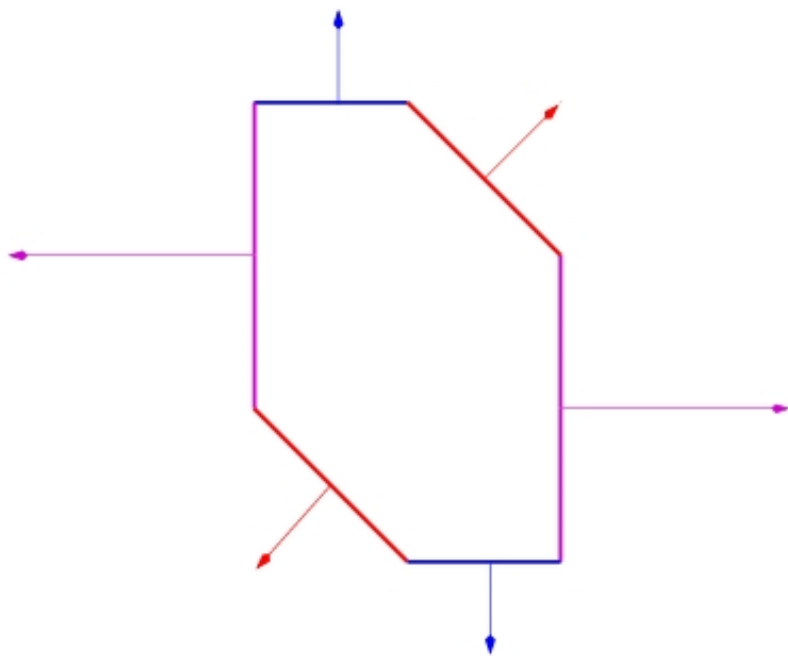
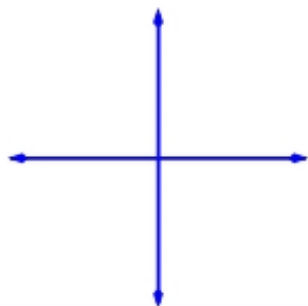
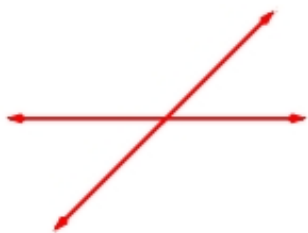
Minkowski addition

- The **Minkowski sum** of **convex** polytopes $P_1, P_2 \subset \mathbb{R}^n$ is **convex** polytope $P_1 + P_2 = \{p_1 + p_2 \mid p_i \in P_i\} \subset \mathbb{R}^n$.
If P_1, P_2 have integral vertices, then so does $P_1 + P_2$.

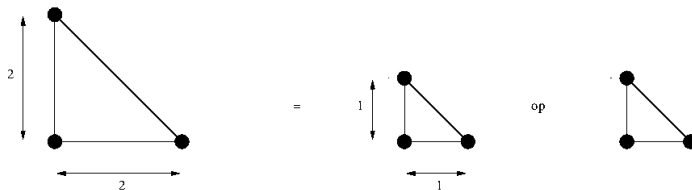
- **Minkowski addition** of polytopes $P_i \subset \mathbb{R}^n$, $i \in I$ is a **many-to-one** map

$$(P_i)_{i \in I} \rightarrow P := \sum_{i \in I} P_i \subset \mathbb{R}^n : (p_i \in P_i)_{i \in I} \mapsto \sum_{i \in I} p_i.$$

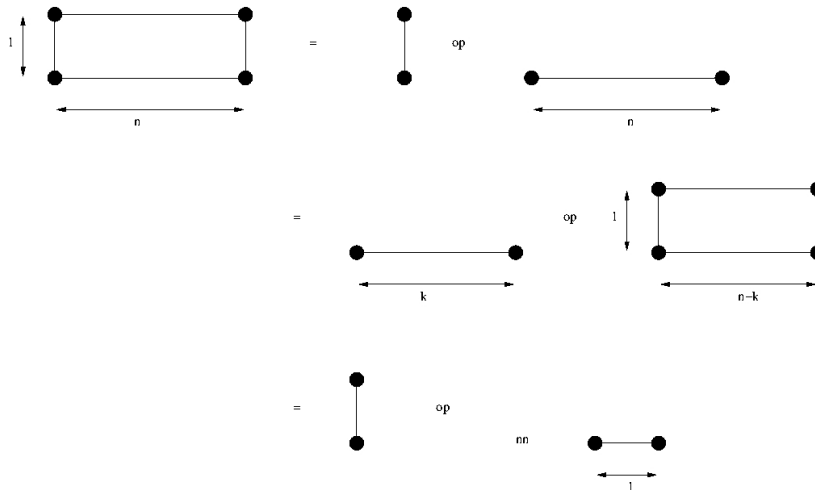
- Complexity in \mathbb{R}^2 : Minkowski addition is linear, but Minkowski decomposition is NP-hard.



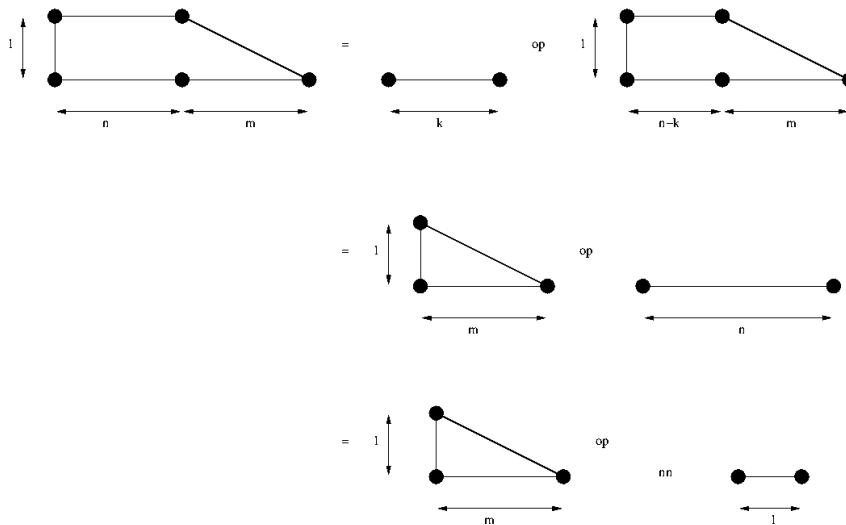
A



B



C



Mixed volume

1. The **mixed volume** $MV(P_1, \dots, P_n) \in \mathbb{R}$ of **convex** polytopes $P_i \subset \mathbb{R}^n$
 - is **multilinear** wrt Minkowski addition and scalar multiplication:
$$MV(P_1, \dots, \lambda P_i + \mu P'_i, \dots, P_n) = \\ = \lambda MV(P_1, \dots, P_i, \dots, P_n) + \mu MV(P_1, \dots, P'_i, \dots, P_n), \quad \lambda, \mu \in \mathbb{R},$$
 - st. $MV(P_1, \dots, P_1) = n! \operatorname{vol}(P_1)$.
2. Equivalently, $\operatorname{vol}(\lambda_1 P_1 + \dots + \lambda_n P_n)$ is a **polynomial** in scalar variables $\lambda_1, \dots, \lambda_n$, with **multilinear term** $MV(P_1, \dots, P_n) \lambda_1 \dots \lambda_n$.
3. **Exclusion-Inclusion** definition:
$$MV := \sum_{I \subset \{1, \dots, n\}} (-1)^{n-|I|} \operatorname{vol} \left(\sum_{i \in I} P_i \right).$$

Mixed Volume characterization

Property	MV: $\text{vtx}(Q_i) \subset \mathbb{Z}^n$	Generic number of isolated solutions
$\in \mathbb{Z}_{\geq 0}$	$\text{MV}(\dots, Q_i, \dots)$	$\#\{x \in (\overline{K}^*)^n \mid \dots = f_i(x) = \dots = 0\}$
Invariance by permutation	$\text{MV}(\dots, Q_j, \dots, Q_i, \dots) = \text{MV}(\dots, Q_i, \dots, Q_j, \dots)$	$\#\{x \mid \dots = f_j(x) = \dots = f_i(x) = \dots = 0\} = \#\{x \mid \dots = f_i(x) = \dots = f_j(x) = \dots = 0\}$
Linearity wrt Minkowski addition	$\text{MV}(\dots, Q_i + Q'_i, \dots) = \text{MV}(\dots, Q_i, \dots) + \text{MV}(\dots, Q'_i, \dots)$	$\#\{x \mid \dots = (f_i f'_i)(x) = \dots = 0\} = \#\{x \mid \dots = f_i(x) = \dots = 0\} + \#\{x \mid \dots = f'_i(x) = \dots = 0\}$
Linearity wrt scalar product	$\text{MV}(\dots, \lambda Q_i, \dots) = \lambda \text{MV}(\dots, Q_i, \dots)$	$\#\{x \mid \dots = (f_i(x))^\lambda = \dots = 0\} = \lambda \#\{x \mid \dots = f_i(x) = \dots = 0\}$
Monotone wrt volume	$\text{MV}(\dots, Q_i \cup \{a\}, \dots) \geq \text{MV}(\dots, Q_i, \dots)$	$\#\{x \mid \dots = f_i(x) + cx^a = \dots = 0\} \geq \#\{x \mid \dots = f_i(x) = \dots = 0\}$
[Kushnirenko]	$\text{MV}(Q_1, \dots, Q_1) = n!V(Q_1)$	$\#\{x \mid f_1(x) = \dots = f_n(x) = 0\} = n!V(Q_1)$

Bernstein (BKK) bound

Theorem [Bernstein'75,Kushnirenko'75,Khovanskii'78] [Danilov'78]:

Given polynomials $f_1, \dots, f_n \in K[x_1^{\pm 1}, \dots, x_n^{\pm 1}]$, for any field K , the number of **common isolated zeros** in $(\overline{K} - \{0\})^n$, counting multiplicities, is bounded by the **mixed volume** of the Newton polytopes $MV(Q_1, \dots, Q_n)$ (irrespective of the variety's dimension).

Dense homogeneous: $MV(Q_1, \dots, Q_n) = \prod_{i=1}^n d_i = \text{Bézout's bound}$, where $d_i = \deg(f_i)$ and $Q_i = \text{simplex}\{0, (d_i, 0, \dots, 0), \dots, (0, \dots, 0, d_i)\}$.

Dense multi-homogeneous: $MV(Q_1, \dots, Q_n) = \text{m-Bézout's bound}$:

the coefficient of $\prod_{j=1}^r y_j^{l_j}$ in $\prod_{i=1}^n (d_{i1}y_1 + \dots + d_{ir}y_r)$,

where $\deg_{X_j} f_i = d_{ij}$, $j = 1, \dots, r$, and X_j groups l_j variables.

Exactness of BKK

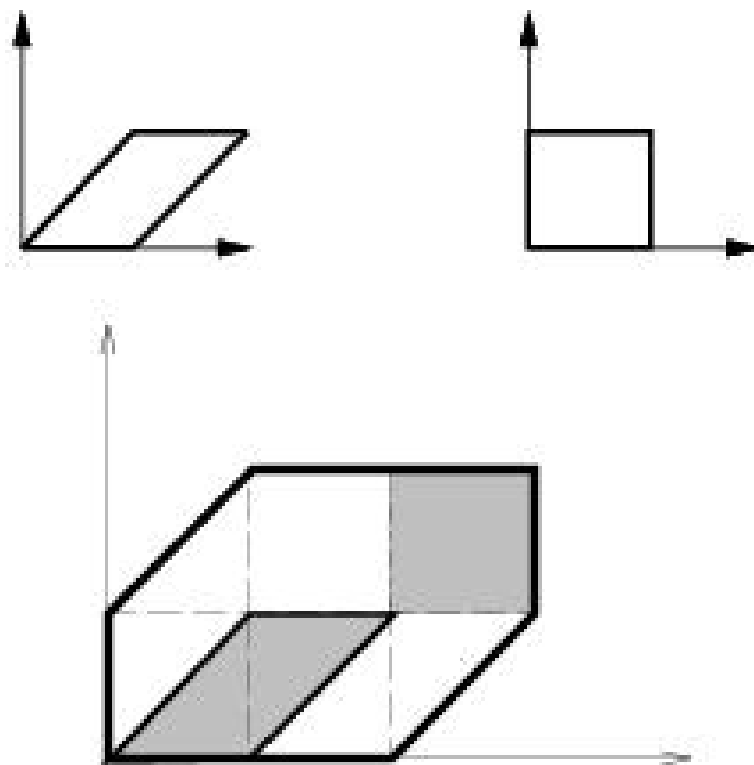
Theorem 2 [Bernstein'75] BKK is exact if, $\forall v \in \mathbb{R}^n$, the face system $\partial_v f_1 = \dots = \partial_v f_n = 0$ has **no solution** in $(\overline{K}^*)^n$.

[Canny,Rojas'91]: BKK is exact if the extremal coefficients are **generic**.

[Huber,Sturmfels'95]: BKK is exact if all **facet** systems $\partial_v f_1 = \dots = \partial_v f_n = 0$, the sparse resultant equals a constant.

Extension [Huber,Sturmfels'95]: Let $A'_i = \text{supp}(f_i) \cup \{0\}$. The number of isolated zeros in \overline{K}^n , counting multiplicities, is bounded by the **stable mixed volume** of A'_1, \dots, A'_n (irrespective of the variety's dimension). Equality holds for generic extremal coefficients.

Example: mixed subdivision for well-constrained problem



Given $f_1 = c_{11} + c_{12}xy + c_{13}x^2y + c_{14}x$, $f_3 = c_{31} + c_{32}y + c_{33}xy + c_{34}x$,

- construct their **Newton polytopes** in \mathbb{R}^2
- compute a **mixed subdivision** of the Minkowski Sum (3 mixed cells)
- compute the Mixed Volume using the formula $MV = \sum_{\sigma} V(\sigma)$, over all **mixed cells** σ of the mixed subdivision (here $MV=3$).

Sparse / Toric Resultant

Given $n + 1$ **Laurent** polynomials $f_0, \dots, f_n \in K[x_1, \dots, x_n, x_1^{-1}, \dots, x_n^{-1}]$ with indeterminate coefficients \vec{c} , their **projective**, resp. **toric/sparse**, *resultant* is the unique (up to sign) irreducible polynomial $R(\vec{c}) \in \mathbb{Z}[\vec{c}]$ s.t.

$$R(\vec{c}) = 0 \Leftrightarrow \exists \xi = (\xi_1, \dots, \xi_n) \in X : f_0(\xi) = \dots = f_n(\xi) = 0$$

where the variety X equals:

- the projective space \mathbb{P}^n over the algebraic closure \overline{K} ,
- resp. the **toric variety** X , $(\overline{K}^*)^n \subset X \subset \mathbb{P}^N$.

The **projective**, resp. **toric**, resultant polynomial $R \in \mathbb{Z}[\vec{c}]$ is separately homogeneous in the coefficients of each f_i , with *degree* equal to $\prod_{j \neq i} \deg f_j$ (**Bézout's number**), resp. the n -fold **mixed volume**:

$$\text{MV}_{-i} := \text{MV}(f_0, \dots, f_{i-1}, f_{i+1}, \dots, f_n),$$

provided the supports of the f_i generate \mathbb{Z}^n .

Mixed subdivisions

Regular (induced) subdivisions

For $Q_i \subset \mathbb{R}^n$, $(Q_i)_{i \in I} \rightarrow Q = \sum_{i \in I} Q_i : (q_i)_{i \in I} \mapsto \sum_{i \in I} q_i$.

Consider (affine) **lifting** functions $l_i : \mathbb{R}^n \rightarrow \mathbb{R}$, which define

$$\hat{Q}_i := \text{CH}\{(p_i, l_i(p_i)) : p_i \in Q_i\} \subset \mathbb{R}^{n+1}.$$

Let \hat{Q} be the Minkowski sum $\sum_i \hat{Q}_i$.

Lemma: If the l_i are sufficiently generic, then every face in the **lower-hull** of \hat{Q} is written uniquely as $\sum_i \hat{F}_i$, for faces $\hat{F}_i \subset \hat{Q}_i$.

Then, \hat{Q} projects onto Q , so the lower-hull faces induce a **regular** subdivision of Q , with faces (cells) $\sum_i F_i$, where \hat{F}_i is the lifted face $F_i \subset Q_i$. In particular, facets on the lower-hull project to maximal cells ($\dim = n$).

Tight coherent mixed subdivisions

In general: $\dim(\sum_i F_i) \leq \sum_i \dim F_i$.

Defn. A **tight/exact/fine** subdivision occurs when equality holds.

In particular, for a cell of maximum dimension, $n = \sum_i \dim F_i$.

Thus, the lower-hull of \hat{Q} corresponds bijectively to Q .

Eg: **Not** tight subdivision: 2 segments lifted in parallel:

$$\dim(F_0 + F_1) = 1 < \dim F_0 + \dim F_1 = 1 + 1.$$

Lemma. A regular subdivision by a **generic** lifting is tight and **coherent**.

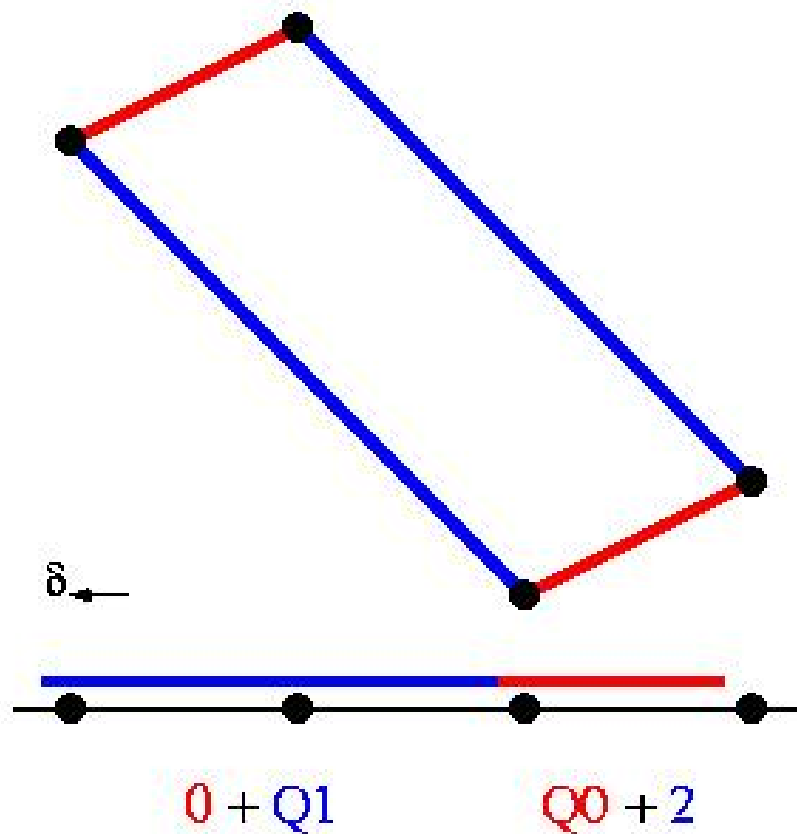
The latter captures continuity of the (unique) expressions of cells as Minkowski sums.

We call tight coherent mixed subdivisions simply **mixed subdivisions**.

Generalized mixed subdivisions are coherent but *not* always tight.

Lifting in the Sylvester case

$$f_0 = c_{00} + c_{01}x, \quad f_1 = c_{10} + c_{11}x + c_{12}x^2$$



Point $2 = 0 + 2$ from both maximal cells.

Example for the over- constrained problem

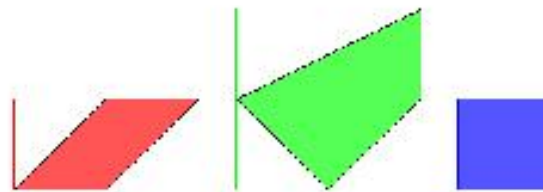


Figure 1: The given polytopes.

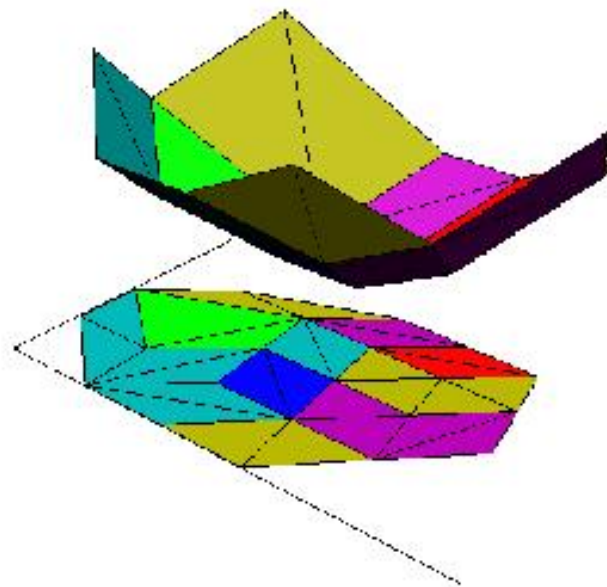


Figure 2: The lower hull of the lifted Minkowski Sum and its planar projection.

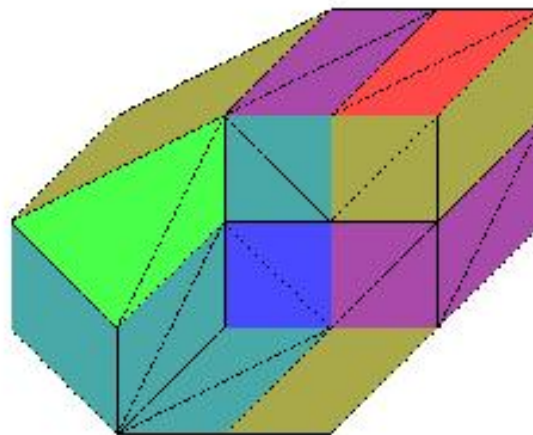


Figure 3: The mixed subdivision.

Cells in mixed subdivisions

Tightness of mixed subdivision implies for all maximal cells $\sigma = \sum_i F_i$:

$$\sum_i \dim F_i = n, \quad F_i \subset Q_i.$$

Corollary: when $n + 1$ summands, $\exists i : F_i = \text{vertex}$.

Example: For linear/affine liftings, certain cells are copies of the original Q_i . Then, all other summands are vertices in the Q_j , for $j \neq i$.

Resultant case of $n + 1$ polytopes: $Q = Q_0 + Q_1 + \cdots + Q_n$.

Then, every cell has at least one vertex summand.

Example of all possible summand dimensions (up to permutation):

$n = 2$, $Q_1 + Q_2$: 0,2 (Q_i) and 1,1 (mixed).

$n = 2$, $Q_0 + Q_1 + Q_2$: 0,0,2 (Q_i) and 0,1,1 (mixed).

$n = 3$, $Q_1 + \cdots + Q_3$: 0,0,3 (Q_i), 0,1,2 (unmixed), 1,1,1 (mixed).

Mixed cells

Defn. A maximal cell σ , in a mixed subdivision Δ , is **mixed** iff it has precisely n linear summands, i.e. n edge summands F_i : $\dim F_i = 1$.

- n polytopes: $Q = Q_1 + \cdots + Q_n$, mixed cells are sums of edges.

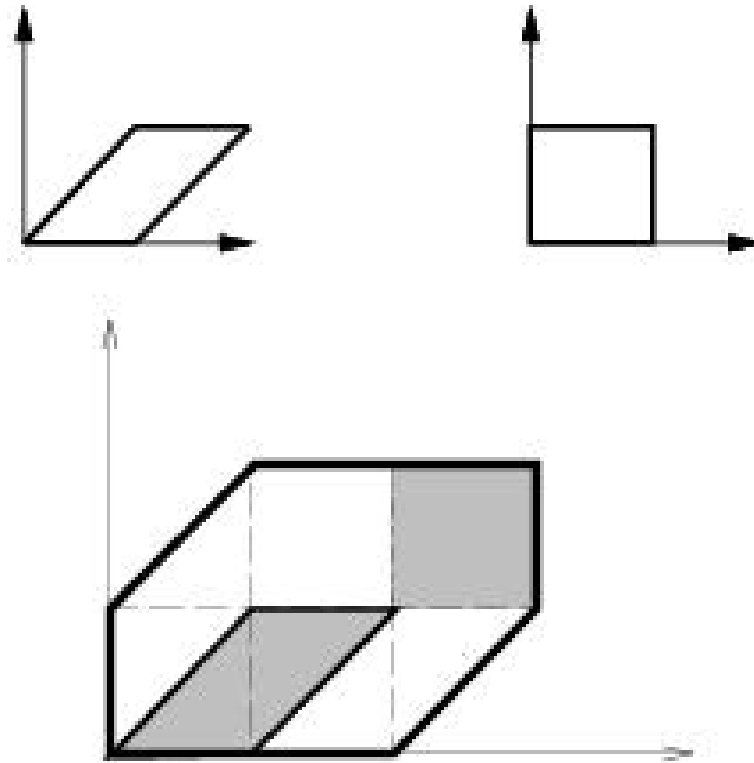
Thm: $MV(Q_1, \dots, Q_n) = \sum_{\sigma} \text{vol}(\sigma)$, over all **mixed cells** $\sigma \in \Delta$.

- $n + 1$ polytopes: $Q = Q_0 + Q_1 + \cdots + Q_n$, i -mixed cells are sums of edges plus vertex $a_i \in Q_i$.

Thm: $MV(Q_0, \dots, Q_{i-1}, Q_{i+1}, \dots, Q_n) = \sum_{\sigma} \text{vol}(\sigma)$,

over all **i -mixed cells** $\sigma \in \Delta$.

Algorithm for planar mixed subdivisions



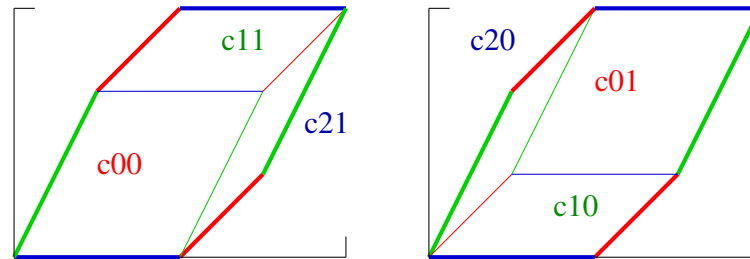
1. Construct the Minkowski sum $Q = \sum_i Q_i$.
2. Place the Q_i 's appropriately: no intersection of dimension ≥ 1 .
3. Move edges to the boundary ∂Q : paths intersect at mixed cells.

Mixed subdivision: example

The system from [Buchberger'88]

$$f_0 = c_{00} - c_{01}st, \quad f_1 = c_{10} - c_{11}st^2, \quad f_2 = c_{20} - c_{21}s^2,$$

has 2 possible mixed subdivisions, depending on the lifting:



Each subdivision contains exactly 3 maximal cells, all of which are mixed (vertex summands shown).

Preview of Matrix construction

Consider Minkowski sum $Q = Q_0 + \cdots + Q_n \subset \mathbb{R}^n$,
and infinitesimal perturbation $\delta \in \mathbb{R}^n$ in generic direction.

For every point $p \in \mathcal{E} = (Q + \delta) \cap \mathbb{Z}^n$, \exists **unique cell** $\sigma + \delta \ni p$.
Define $\text{RC}(p) := (i, F_i)$: unique if σ is i -mixed, else pick max i .

Construct (sparse) resultant **matrix** M with rows/columns **indexed by \mathcal{E}** :
for $p, q \in \mathcal{E}$, assume $p - \delta \in \sigma = F_0 + \cdots + a_i + \cdots + F_n$ (max i), i.e.
 $\text{RC}(p) = (i, a_i)$. Then, the matrix row indexed by p contains

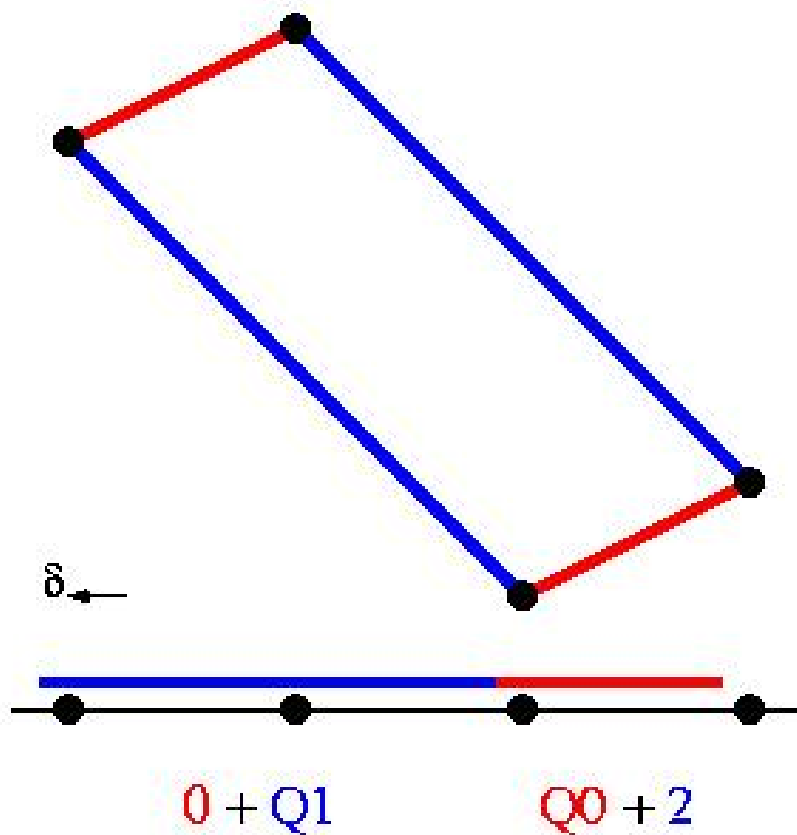
$$\text{polynomial } x^{p-a_i} f_i,$$

hence the matrix element (p, q) is

$$\text{the coefficient of } x^q \text{ in } x^{p-a_i} f_i.$$

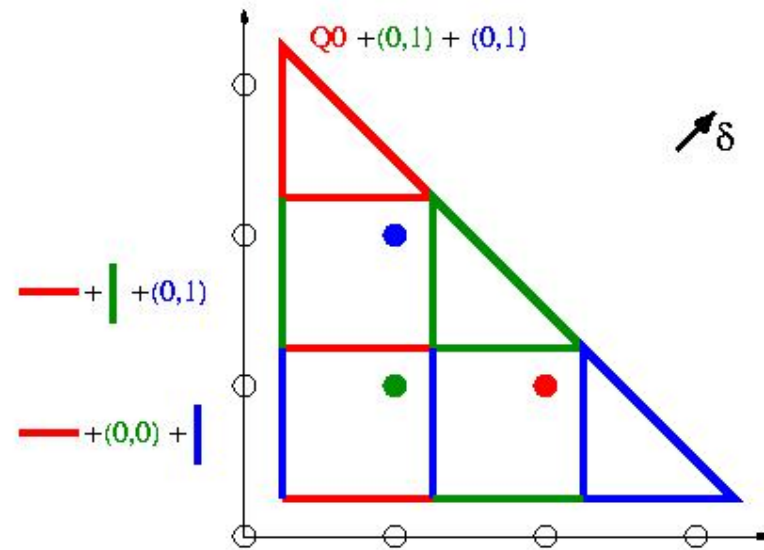
Lifting in the Sylvester case

$$f_0 = c_{00} + c_{01}x, \quad f_1 = c_{10} + c_{11}x + c_{12}x^2$$



$$\text{RC}(2) = (1; 2) \text{ ie. } x^2 \mapsto x^{2-2}f_1.$$

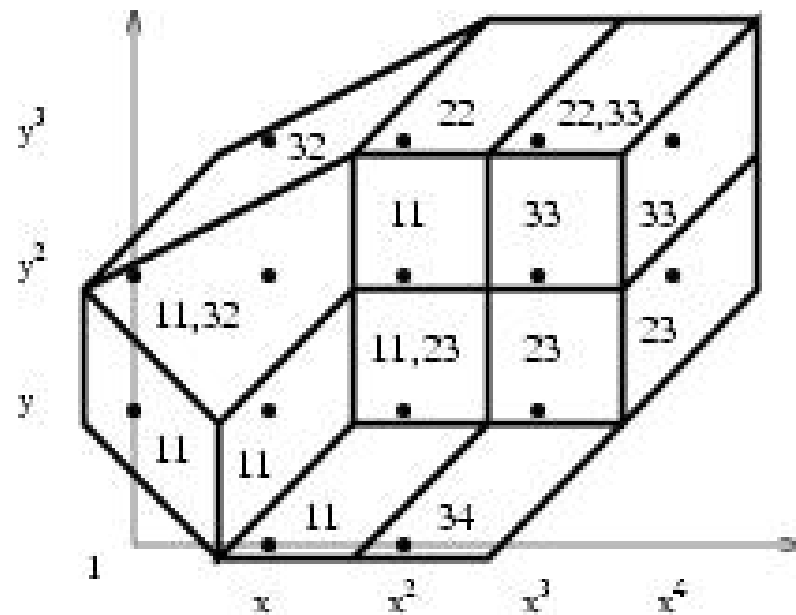
Mixed subdivision of a linear system



$$\begin{aligned} \text{RC}(1, 2) &= [2, (0, 1)] \text{ ie. } x_1 x_2^2 \mapsto x^{(1,2)-(0,1)} f_2 = x^{(1,1)} f_2 \\ \text{RC}(1, 1) &= [1, (0, 0)] \text{ ie. } x_1 x_2 \mapsto x^{(1,1)-(0,0)} f_1 = x^{(1,1)} f_1 \\ \text{RC}(2, 1) &= [0, (1, 0)] \text{ ie. } x_1^2 x_2 \mapsto x^{(2,1)-(1,0)} f_0 = x^{(1,1)} f_0 \end{aligned}$$

$$M = \begin{bmatrix} x_1^2 x_2 & x_1 x_2^2 & x_1 x_2 \\ c_{01} & c_{02} & c_{03} \\ c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix} \quad \begin{aligned} & x_1 x_2 f_0 \\ & x_1 x_2 f_1 \\ & x_1 x_2 f_2 \end{aligned}$$

Example: mixed subdivision for the over-constrained problem



Eg: $x \mapsto (x, y)^{(1,0)-(0,0)} f_1$, $x^2 y \mapsto (x, y)^{(2,1)-(2,1)} f_2$.

Example: subdivision-based matrix

$$f_1 = c_{11} + c_{12}xy + c_{13}x^2y + c_{14}x,$$

$$f_2 = c_{21}y + c_{22}x^2y^2 + c_{23}x^2y + c_{24}x,$$

$$f_3 = c_{31} + c_{32}y + c_{33}xy + c_{34}x.$$

	1, 0	2, 0	0, 1	1, 1	2, 1	3, 1	0, 2	1, 2	2, 2	3, 2	4, 2	1, 3	2, 3	3, 3	4, 3
1, 0)x	c_{11}	c_{14}	0	0	c_{12}	c_{13}	0	0	0	0	0	0	0	0	0
2, 0)x	c_{31}	c_{34}	0	c_{32}	c_{33}	0	0	0	0	0	0	0	0	0	0
0, 1)y	0	0	c_{11}	c_{14}	0	0	0	c_{12}	c_{13}	0	0	0	0	0	0
1, 1)xy	0	0	0	c_{11}	c_{14}	0	0	0	c_{12}	c_{13}	0	0	0	0	0
2, 1)	c_{24}	0	c_{21}	0	c_{23}	0	0	0	c_{22}	0	0	0	0	0	0
3, 1)x	0	c_{24}	0	c_{21}	0	c_{23}	0	0	0	c_{22}	0	0	0	0	0
0, 2)y	0	0	c_{31}	c_{34}	0	0	c_{32}	c_{33}	0	0	0	0	0	0	0
1, 2)xy	0	0	0	c_{31}	c_{34}	0	0	c_{32}	c_{33}	0	0	0	0	0	0
2, 2)x ² y ²	0	0	0	0	0	0	0	0	c_{11}	c_{14}	0	0	0	c_{12}	c_{13}
3, 2)x ² y	0	0	0	0	c_{31}	c_{34}	0	0	c_{32}	c_{33}	0	0	0	0	0
4, 2)x ² y	0	0	0	0	0	c_{24}	0	0	c_{21}	0	c_{23}	0	0	0	c_{22}
1, 3)xy ²	0	0	0	0	0	0	0	c_{31}	c_{34}	0	0	c_{32}	c_{33}	0	0
2, 3)y	0	0	0	c_{24}	0	0	c_{21}	0	c_{23}	0	0	0	c_{22}	0	0
3, 3)x ² y ²	0	0	0	0	0	0	0	0	c_{31}	c_{34}	0	0	c_{32}	c_{33}	0
4, 3)x ³ y ²	0	0	0	0	0	0	0	0	0	c_{31}	c_{34}	0	0	c_{32}	c_{33}

$\dim M = 15$, greedy [CanPed]: 14, incremental [ECan]: 12.

Mixed volumes = 4, 3, 4 $\Rightarrow \deg R_{tor} = 11$ while $\deg(\text{classical resultant}) = 26$.

Matrices of Sylvester-type

Limitation of subdivision-based algorithm

Bilinear system: $f_i = c_{i0} + c_{i1}x_1 + c_{i2}x_2 + c_{i3}x_1x_2$, $i = 1, 2, 3$.

The toric resultant has $\deg R = 3 \cdot \deg_{f_i} R = 6$.

$|\mathcal{E}| = 9 \Rightarrow$ the **subdivision-based** algorithm cannot yield an optimal matrix.

The **greedy** variant [Canny-Pedersen'93] *may* obtain an optimal matrix.

The **incremental** algorithm gets the following optimal matrix:

$$R = \det \begin{bmatrix} 1 & x_1 & x_2 & x_1x_2 & x_1^2 & x_1^2x_2 \\ c_{10} & c_{11} & c_{12} & c_{13} & 0 & 0 \\ c_{20} & c_{21} & c_{22} & c_{23} & 0 & 0 \\ c_{30} & c_{31} & c_{32} & c_{33} & 0 & 0 \\ 0 & c_{10} & 0 & c_{12} & c_{11} & c_{13} \\ 0 & c_{20} & 0 & c_{22} & c_{21} & c_{23} \\ 0 & c_{30} & 0 & c_{32} & c_{31} & c_{33} \end{bmatrix} \begin{matrix} f_1 \\ f_2 \\ f_3 \\ x_1f_1 \\ x_1f_2 \\ x_1f_3 \end{matrix}$$

Sylvester-type matrices

Given: $f_0, \dots, f_m \in K[x]$ (resp. $K[x^{\pm 1}]$) where $x = (x_1, \dots, x_n)$, $m \geq n$; let $A_i = \text{supp}(f_i)$. The supports $B_0, \dots, B_m \subset \mathbb{Z}^n$ **define** map

$$M^T : P(B_0) \times \dots \times P(B_m) \rightarrow P\left(\bigcup_{i=0}^m A_i + B_i\right),$$

$$(g_0, \dots, g_m) \mapsto \sum_{i=0}^m f_i g_i,$$

s.t. $P(B) = \{g \in K[x^{\pm 1}] : \text{supp}(g) \subset B\}$.

Necessary condition. For any specialization c of the f_i coefficients such that the f_i have a common root, the matrix $M^T(c)$ is **not surjective**.

Algorithmic Problem:

Find B_i such that $\sum_{i=0}^m |B_i| \geq \left| \bigcup_{i=0}^m A_i + B_i \right| = \text{rank}(M)$ generically.

Matrices of Sylvester-type

Algorithms: subdivision-based [Canny-E'93,'00], incremental [E-Canny'95] yield a square “Newton” matrix M of the toric resultant, such that:

$$\begin{aligned}\det(M) &\neq 0, \\ R &\mid \det(M), \\ \deg_{f_0} \det(M) &= \deg_{f_0} R,\end{aligned}$$

where R is the toric resultant. Same properties for the Macaulay matrix of the projective resultant [Mac'02].

Complexity [E'96] $O(e^n \deg R (\text{vtx} Q_i)^3)$, when n -fold Mixed Volumes > 0 , and the Newton polytopes do not differ “too much” (bounded scaling).

Rational form [D'Andrea'01] : $R = \det(M) / \det(M')$,
where M' is a submatrix of M , generalizing Macaulay's construction.

Open: single lifting.

Estimate $\text{vol}(Q_1 + \cdots + Q_n) / \text{MV}(Q_1, \dots, Q_n)$

The Aleksandrov-Fenchel inequality [1935-6]:

$\text{MV}^2(Q_1, \dots, Q_n) \geq \text{MV}(Q_1, Q_1, Q_3, \dots) \text{MV}(Q_2, Q_2, Q_3, \dots)$, implies:

$$\text{MV}(Q_1, \dots, Q_n) \geq n! \sqrt[n]{\prod_{i=1}^n \text{vol}(Q_i)}.$$

If $\text{vol}(Q_\mu)$ is minimal, the **system's scalar factor** is set to be the minimum real $s \geq 1$ s.t. $Q_i \subset sQ_\mu, \forall i$ (mod translations).

Thus, $s < \infty \Leftrightarrow$ all Q_i of the same dimension, $s = 1 \Leftrightarrow Q_1 = \cdots = Q_n$.

Corollary [E'94].

$$\frac{\text{vol}(Q_1 + \cdots + Q_n)}{\text{MV}(Q_1, \dots, Q_n)} < e^n s^n / \sqrt{2\pi n}.$$

Corollary [E'94]. For $\deg R = \sum_{i=0}^n \text{MV}(Q_0, \dots, Q_{i-1}, Q_{i+1}, \dots, Q_n)$,

$$\text{vol}(Q_0 + \cdots + Q_n) = O\left(\frac{e^n s^n}{n^{3/2}} \deg R\right) = O^*(s^n \deg R).$$

Matrix construction [Canny,E'93,00]

1. Pick (affine) **liftings** $\omega_i : \mathbb{Z}^n \rightarrow \mathbb{R} : \text{supp}(f_i) \rightarrow \mathbb{Q}$.

2. Define (tight coherent polyhedral) **mixed subdivision** of the Minkowski sum $Q = Q_0 + \cdots + Q_n$ of the Newton polytopes.

Maximal cells are **uniquely** expressed as

$$\sigma = F_0 + \cdots + F_n, \quad \text{with } \dim F_0 + \cdots + \dim F_n = n,$$

where F_i is a face of Q_i . σ is **i -mixed** $\iff \exists! i : \dim F_i = 0$.

3. For every point $p \in \mathcal{E} = (Q + \delta) \cap \mathbb{Z}^n$, \exists **unique** $\sigma + \delta \ni p$.

Define function $\text{RC}(p) = (i, F_i) : \text{unique if } \sigma \text{ } i\text{-mixed, else pick max } i$.

4. Construct resultant **matrix** M with rows/columns **indexed by** \mathcal{E} :

for $p, q \in \mathcal{E}$, element (p, q) is the coefficient of x^q in $x^{p-a_i} f_i$:

$p - \delta \in \sigma = F_0 + \cdots + a_i + \cdots + F_n$ (max i), i.e. $\text{RC}(p) = (i, a_i)$.

Correctness

Lemma. $\text{RC}(p) = (i, a_i) \Rightarrow \text{support}(x^{p-a_i} f_i) \subset \mathcal{E}$.

Proof. $p \in \sigma + \delta \subset Q_0 + \cdots + Q_{i-1} + a_i + Q_{i+1} + \cdots + Q_n + \delta$ implies $p - a_i \in \sum_{i \neq j} Q_i + \delta$, hence $p - a_i + q \in \mathcal{E}$ for all $q \in \text{supp}(f_i)$.

Corollary. The diagonal entry at the row indexed by p contains the f_i coefficient of x^{a_i} .

Proof. Consider the row indexed by p , s.t. $\text{RC}(p) = (i, a_i)$.

Then, the f_i coefficient of x^{a_i} is the coefficient of x^p in $x^{p-a_i} f_i$, hence it appears at the column indexed by p .

Incremental algorithm [E-Canny'95]

Idea: The **rows** express $x^b f_i : b \in Q_{-i} \cap \mathbb{Z}^n$, where $Q_{-i} = Q_0 + \dots + Q_{i-1} + Q_{i+1} + \dots + Q_n$ so that column monomials $\subset \sum_i Q_i$.

1. **Sort** $Q_{-i} \cap \mathbb{Z}^n$ on their distance $\text{dist}_v(\cdot)$ from the boundary of Q_{-i} along some **vector** $v \in \mathbb{Q}^n$.
2. Define the **rows** of M by points $B_i = \{b : \text{dist}_v(b) > \beta\}$, for bound $\beta \in \mathbb{R}$. The **columns** are indexed by $\cup_i \cup_{b \in B_i} \text{supp}(x^b f_i)$.
3. Enlarge M by decreasing β until M (i) has at least **as many rows as columns** and (ii) is **generically of full rank**.

For **multihomogeneous** systems: Deterministic vector v yields:

- exact matrices if possible [Sturmfels-Zelevinsky'94],
- otherwise minimum matrices [Dickenstein-E'02].

Complexity in $\sim e^{2n}(\deg R)^2$ (by quasi-Toeplitz structure)

Unmixed multihomogeneous systems

Partition the variables to r subsets: every polynomial is **homogeneous in each subset**. The i -th subset has $l_i + 1$ homogeneous variables, of total degree d_i . Then the polynomial is of **type** $(l_1, \dots, l_r; d_1, \dots, d_r)$.

Type $(2, 1; 2, 1) : (x_1, x_2, y_1) \in \mathbb{P}^2 \times \mathbb{P}^1 : c_0 + c_1x_1 + c_2x_2 + c_3x_1x_2 + c_4x_1^2 + c_5x_2^2 + c_6y_1 + c_7x_1y_1 + c_8x_2y_1 + c_9x_1x_2y_1 + c_{10}x_1^2y_1 + c_{11}x_2^2y_1$.

A system is of **type** (l, d) iff all polynomials are of type (l, d) .

[Sturmfels, Zelevinsky'94]. If $l_i = 1$ or $d_i = 1$, $\forall i$, then \exists **determinantal** resultant formula i.e. $\det M = R$.

Type $(2, 1; 1, 1) : c_0 + c_1x_1 + c_2x_2 + c_3y_1 + c_4x_1y_1 + c_5x_2y_1$.

[Dickenstein, E'02] find minimum (non-optimal) Sylvester-type matrix; extended by [E-Mantzaflaris]

The **incremental algorithm** [E, Canny'95] constructs all these matrices.

Rational form

Recursive lifting on n , using the subdivision algorithm [D'Andrea'01].

Bilinear: $f_i = a_i + b_i x_1 + c_i x_2 + d_i x_1 x_2$, $i = 0, 1, 2$.

Linear lift $(-\infty, \dots), (0, 1, 1, 2), (0, 0, 7, 7)$, $\delta = (\frac{2}{3}, \frac{1}{2}) \Rightarrow \dim M = 16$ (numerator):

$$M = \begin{pmatrix} a_1 & b_1 & 0 & c_1 & d_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_0 & b_0 & 0 & c_0 & d_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_1 & b_1 & 0 & c_1 & d_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_2 & b_2 & 0 & c_2 & d_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_2 & b_2 & 0 & c_2 & d_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_0 & 0 & 0 & c_0 & d_0 & b_0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_2 & b_2 & 0 & c_2 & d_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & a_2 & b_2 & 0 & c_2 & d_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_1 & b_1 & 0 & c_1 & d_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_1 & 0 & 0 & c_1 & d_1 & b_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_2 & 0 & 0 & c_2 & d_2 & b_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_2 & 0 & 0 & c_2 & 0 & 0 & 0 & 0 & d_2 & b_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & a_2 & b_2 & 0 & 0 & 0 & 0 & c_2 & d_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_2 & b_2 & 0 & 0 & 0 & 0 & c_2 & d_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_2 & b_2 & 0 & 0 & 0 & 0 & c_2 & d_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_1 & b_1 & 0 & 0 & 0 & 0 & c_1 & d_1 \end{pmatrix}$$

Rational form: denominator

$$M' = \begin{pmatrix} a_1 & 0 & c_1 & d_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b_1 & 0 & c_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_2 & 0 & c_2 & d_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b_2 & 0 & c_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_2 & b_2 & c_2 & d_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_2 & 0 & c_2 & d_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_1 & 0 & c_1 & d_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & c_2 & b_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_2 & b_2 & 0 & c_2 & d_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & a_2 & 0 & 0 & c_2 \end{pmatrix}$$

$\det(M) = \pm R \cdot \det(M')$: M' is a submatrix of M ,

$$|M'| = -c_2^3(-c_1a_2 + a_1c_2)b_2(c_1d_2 - d_1c_2)(-b_2c_1 + b_1c_2)$$

Main step: lifting of some $b \in Q_0$ is very negative.

The mixed subdivision provides all info.

Open: \exists single lifting yielding both numerator and denominator?

YES if $n = 2$, or systems with sufficiently different Newton polytopes, or unmixed systems [E-Konaxis'11].

Single-lifting rational formula

$$M = \begin{array}{c} \begin{array}{ccccccccc} 00 & 10 & 01 & 11 & 21 & 12 & 20 & 02 & 22 \end{array} \\ \left[\begin{array}{ccccccccc} c_{10} & c_{11} & c_{12} & c_{13} & 0 & 0 & 0 & 0 & 0 \\ c_{20} & c_{21} & c_{22} & c_{23} & 0 & 0 & 0 & 0 & 0 \\ c_{30} & c_{31} & c_{32} & c_{33} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{10} & c_{11} & c_{12} & 0 & 0 & c_{13} \\ 0 & c_{20} & 0 & c_{22} & c_{23} & 0 & c_{21} & 0 & 0 \\ 0 & 0 & c_{30} & c_{31} & 0 & c_{33} & 0 & c_{32} & 0 \\ 0 & c_{30} & 0 & c_{32} & c_{33} & 0 & c_{31} & 0 & 0 \\ 0 & 0 & c_{20} & c_{21} & 0 & c_{23} & 0 & c_{22} & 0 \\ 0 & 0 & 0 & c_{20} & c_{21} & c_{22} & 0 & 0 & c_{23} \end{array} \right] \end{array} \begin{array}{l} f_1 \\ f_2 \\ f_3 \\ x_1 x_2 f_1 \\ x_1 f_2 \\ x_2 f_3 \\ x_1 f_3 \\ x_2 f_2 \\ x_1 x_2 f_2 \end{array}$$

Same linear lifting $(-\infty, \dots), (0, 1, 1, 2), (0, 0, 7, 7); \delta = (\frac{2}{3}, \frac{1}{2})$.

Denominator = submatrix of points in non-mixed cells:

$$M' = \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_1 x_2 \end{array} \\ \left[\begin{array}{ccc} c_{21} & c_{22} & 0 \\ c_{31} & c_{32} & 0 \\ 0 & 0 & c_{23} \end{array} \right] \end{array} \begin{array}{l} f_2 \\ f_3 \\ x_1 x_2 f_2 \end{array} \Rightarrow R = \det M / \det M'.$$

Bézout matrices

The Bezoutian

Definition. For $f_0, \dots, f_n \in K[t_1, \dots, t_n]$, the **Bezoutian** is

$$\Theta_{f_i}(t, z) = \begin{vmatrix} f_0(t) & \theta_1(f_0)(t, z) & \cdots & \theta_n(f_0)(t, z) \\ \vdots & \vdots & \ddots & \vdots \\ f_n(t) & \theta_1(f_n)(t, z) & \cdots & \theta_n(f_n)(t, z) \end{vmatrix}$$

where $\theta_i(f_j)(t, z) :=$

$$\frac{f_j(z_1, \dots, z_{i-1}, t_i, \dots, t_n) - f_j(z_1, \dots, z_i, t_{i+1}, \dots, t_n)}{t_i - z_i}.$$

Let $\Theta_{f_0, \dots, f_n}(t, z) = \sum_{a, b} \theta_{ab} t^a z^b, \theta_{a, b} \in K, a, b \in \mathbb{N}^n$

Then the **Bezoutian matrix** of f_0, \dots, f_n is the matrix $B_{f_0, \dots, f_n} = (\theta_{ab})_{a, b}$.

Theorem [Cardinal-Mourrain'96]

The resultant divides all maximal nonzero minors of the Bezoutian matrix.

The dimension of the matrix is $O(e^n d^n)$, $d = \max\{\deg f_i\}$.

Examples of Bezoutian matrices

$n = 1$ [Béz1779]

$$\begin{aligned} f_0 &= x_0^2 + x_0 x_1 + 2x_0 + x_1 - 1, \\ f_1 &= x_0^2 + 3x_0 - x_1^2 + 2x_1 - 1. \end{aligned}$$

$$R = -x_0^3 - 2x_0^2 + 3x_0 =$$

$$\det \begin{bmatrix} x_0 + 1 & x_0^2 + 2x_0 - 1 \\ -x_0^2 - 4x_0 - 1 & -(x_0 + 1)(x_0^2 + 3x_0 - 1) \end{bmatrix}$$

$n = 2$: Hide t_3 in cyclohexane system:

$$\begin{aligned} f_i &= (13 + t_3^2) - 24t_j t_3 + (1 + t_3^2)t_j^2 = 0, \{i, j\} = \{1, 2\} \\ f_3 &= 13 + t_2^2 - 24t_1 t_2 + t_1^2 + t_1^2 t_2^2 = 0 \end{aligned}$$

B is 8×8 ,

$$|B| = 186624 (t_3^4 - 118t_3^2 + 13) (t_3^4 - 22t_3^2 + 13)^3 (t_3^2 + 1)^8.$$

Bezoutians on Maple with multires

```
Theta([u[0]+u[1]*x[1]+u[2]*x[2],f1,f2],[x[1],x[2]]);
```

$$\begin{aligned}
 & -4u_0y_1y_2 + 5u_0y_1^2 + \left(4u_0 - \frac{2}{3}u_1\right)y_2 + \left(8u_0 - \frac{10}{3}u_1 + \frac{25}{6}u_2\right)y_1 \\
 & + \left(\frac{4}{3}u_1 - 8u_0 - \frac{10}{3}u_2\right) - 4(u_2y_2 + u_1y_1 + u_0)x_1x_2 - 4(u_2y_2 + u_1y_1 + u_0)x_2^2 \\
 & + \left(-4u_1y_1y_2 + 5u_1y_1^2 - 4u_0y_2 + (8u_1 + 5u_2 + 5u_0)y_1 + \left(8u_0 + \frac{25}{6}u_2\right)\right)x_1 \\
 & - \left(4u_2y_1y_2 - 5u_2y_1^2 - 4(u_2 - u_0)y_2 - (8u_1 - 4u_0)y_1 - \left(\frac{10}{3}u_2 + 12u_0 - \frac{2}{3}u_1\right)\right)x_2
 \end{aligned}$$

```
mbezout([u[0]+u[1]*x[1]+u[2]*x[2],f1,f2],
[x[1],x[2]]);
```

$$\begin{bmatrix}
 -4u_0 & 5u_0 & 4u_0 - \frac{2}{3}u_1 & 8u_0 - \frac{10}{3}u_1 + \frac{25}{6}u_2 & \frac{4}{3}u_1 - 8u_0 - \frac{10}{3}u_2 \\
 -4u_1 & 5u_1 & -4u_0 & 8u_1 + 5u_2 + 5u_0 & 8u_0 + \frac{25}{6}u_2 \\
 -4u_2 & 5u_2 & -4u_0 + 4u_2 & 8u_1 - 4u_0 & -\frac{2}{3}u_1 + \frac{10}{3}u_2 + 12u_0 \\
 0 & 0 & -4u_2 & -4u_1 & -4u_0 \\
 0 & 0 & -4u_2 & -4u_1 & -4u_0
 \end{bmatrix}$$

```
factor(det(submatrix(%,1..4,2..5)));
```

$$\frac{5}{11664} \left(u_0 + \frac{1}{3}u_1 + \frac{7}{6}u_2\right)^2 \left(u_0 - \frac{1}{3}u_1 + \frac{5}{6}u_2\right)^2$$

Polynomial system solving

Polynomial System Solving I

Given $f_1, \dots, f_n \in K[x_1^{\pm 1}, \dots, x_n^{\pm 1}]$ defining a 0-dimensional radical ideal.
 Add polynomial $f_0 = u + r_1x_1 + \dots + r_nx_n$, random r_i , indeterminate u .

Construct resultant matrix $M(u)$ for f_0, f_1, \dots, f_n . At root α , $u = -\sum r_i\alpha_i$,

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22}(u) \end{bmatrix} \begin{bmatrix} \vdots \\ \alpha^p \\ \vdots \\ \alpha^q \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ \alpha^a f_i(\alpha) \\ \vdots \\ \alpha^b f_0(u, \alpha) \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \end{bmatrix}.$$

If $\det M_{11} \neq 0$, let $M'(u) = M_{22}(u) - M_{21}M_{11}^{-1}M_{12}$,

$$(M' + uI)v'_\alpha = 0, \quad \dim M' = \text{MV}(f_1, \dots, f_n).$$

- Ratios of the entries of eigenvectors v'_α yield α , if the q span \mathbb{Z}^n .
- Otherwise, use some entries of $v_\alpha = -M_{11}^{-1}M_{12}v'_\alpha$, where $(v_\alpha, v'_\alpha)^T$ is the respective eigenvector of M .

Polynomial System Solving I (factoring)

For $f_0 = u_0 + u_1x_1 + \cdots + u_nx_n$, with indeterminates u_i , the Poisson formula implies

$$R(u_0, \dots, u_n) = C \prod_{\alpha \in V(f_1, \dots, f_n)} (u_0 + \alpha_1 u_1 + \cdots + \alpha_n u_n)^{m_\alpha},$$

over all roots α with multiplicity m_α , where C depends on the coefficients of f_1, \dots, f_n .

Setting $u_i = r_i$, $i = 1, \dots, n$, for random r_i , we have

$$R(u_0) = C \prod_{\alpha} (u_0 + r_1 \alpha_1 + \cdots + r_n \alpha_n)^{m_\alpha}.$$

Solving $R(u_0)$ for u_0 yields $u_0 = -\sum_i r_i \alpha_i$ for all α .

$R(u_0)$ is used in the method of Rational Univariate Representation (primitive element) [Canny,Rouillier] for isolating all real α .

Polynomial System Solving II

“Hide” a variable in the coefficient field: $f_0, f_1, \dots, f_n \in (K[x_0])[x_1, x_1^{-1}, \dots, x_n, x_n^{-1}]$

Hypothesis: x_0 -coordinates of roots distinct, $|M(x_0)| \neq 0$.

$$\det M(x_0) = \begin{vmatrix} M_{11} & M_{12}(x_0) \\ M_{21} & M_{22}(x_0) \end{vmatrix} = \begin{vmatrix} M_{11} & M_{12}(x_0) \\ 0 & M'(x_0) \end{vmatrix},$$

$$|M'(x_0)| = |A_d x_0^d + \dots + A_1 x_0 + A_0| = \det A_d \det(x_0^d + \dots + A_d^{-1} A_1 x_0 + A_d^{-1} A_0).$$

- If $\det A_d \neq 0$, define companion matrix C :

$$C = \begin{bmatrix} 0 & I & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & I \\ -A_d^{-1} A_0 & -A_d^{-1} A_1 & \dots & -A_d^{-1} A_{d-1} \end{bmatrix}$$

The eigenvalues of C are the x_0 -coordinates of the solutions and its eigenvectors contain the values of the monomials indexing M' at the roots.

- Rank balancing improves the conditioning (of A_d) by $x \mapsto (t_1 y + t_2)/(t_3 y + t_4)$, $t_i \in_R \mathbb{Z}$.
- If A_d remains ill-conditioned, solve the generalized eigenproblem

$$\begin{bmatrix} I & & & \\ & \ddots & & \\ & & I & \\ & & & A_d \end{bmatrix} x + \begin{bmatrix} 0 & -I & & \\ & & \ddots & \\ & & & -I \\ A_0 & A_1 & \dots & A_{d-1} \end{bmatrix}.$$

Matrix-based methods for system solving

Theorem. Let $\{z_k\}_k \subset \mathbb{C}^n$ be the isolated zeros of $f_1, \dots, f_n \in \mathbb{Q}[x_1, \dots, x_n]$. There exists **matrix** M_a expressing multiplication by $a \bmod \langle f_i \rangle$ s.t.

- the **eigenvalues** of M_a are $a(z_k)$, and
- the **eigenvectors** of M_a^t are, up to a scalar, $\mathbf{1}_{z_k} : p(x) \mapsto p(z_k)$.

Construct multiplication matrices by means of

- resultant matrices, e.g. Sylvester, Bézout, sparse, or
- normal forms, boundary bases (generalize Gröbner bases).

Stable with respect to input perturbations.

Handles multiplicities and zero sets at infinity.

Extends to over-constrained systems and 1-dimensional zero sets.

Complexity: single exponential in n .

Synaps/Mathemagix library: C++, fast univariate solvers (e.g. [E-Tsigaridas]), connections (GMP, MPFR, LAPACK, SparseLU etc).

Application: Geometric modeling

Implicitization of parametric surfaces

Example: sphere

The sphere in \mathbb{R}^3 is the set of **values** (x, y, z) :

$$x = \frac{t_1^2 - t_2^2 - 1}{t_1^2 + t_2^2 + 1}, y = \frac{2t_1}{t_1^2 + t_2^2 + 1}, z = \frac{2t_1 t_2}{t_1^2 + t_2^2 + 1}, \quad t_1, t_2 \in [0, 1],$$

as well as the set of **roots** of $H(x, y, z) := x^2 + y^2 + z^2 - 1 = 0$.



Modeling/CAD use **parametric** and **implicit/algebraic** representations due to their complementary advantages. This is crucial in operations such as intersecting two surfaces. \Rightarrow must implicitize a (hyper)surface given a (rational) parameterization

Implicitization of rational parametric surfaces

Given is a parametrization of a rational surface:

$$x_1 = \frac{p_1(t_1, t_2)}{p_0(t_1, t_2)}, \quad x_2 = \frac{p_2(t_1, t_2)}{p_0(t_1, t_2)}, \quad x_3 = \frac{p_3(t_1, t_2)}{p_0(t_1, t_2)}.$$

Homogenize the p_i $\theta : \mathbb{P}^2 \rightarrow \mathbb{P}^3 : (t_0 : t_1 : t_2) \mapsto (p_0 : p_1 : p_2 : p_3)$.

Problem: compute the smallest algebraic surface $H(x_1, x_2, x_3, x_0)$ containing $\overline{\text{Im}(\theta)}$, including the case of **base points** $t \in \mathbb{P}^2 : p_i(t) = 0$.

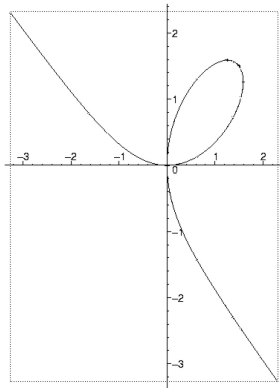
Methods: Gröbner bases, moving surfaces, resultant (perturbation, residual, Bezoutian), residue, Newton sums, numerical methods...

Implicitization examples

[Descartes' folium]

[1596-1650]

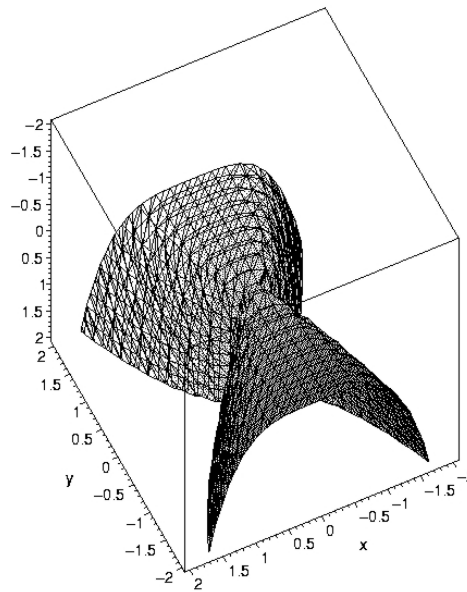
$$(x, y) = \left(\frac{3 t^2}{t^3 + 1}, \frac{3 t}{t^3 + 1} \right)$$



$$H = x^3 + y^3 - 3 x y$$

[Buchberger'88]

$$(x, y, z) = (st, st^2, s^2)$$



$$H = x^4 - y^2 z$$

[Busé'01]

$$x = \frac{s^2}{s^3 + t^3},$$

$$y = \frac{s^3}{s^3 + t^3},$$

$$z = \frac{t^2}{s^3 + t^3}$$

$$H = x^3 - 2x^3y + x^3y^2 - y^2z^3$$

Implicitization by linear algebra

S = monomials forming (a superset of) the **implicit support**.

C = unknown **coefficients** of implicit equation wrt S , $|C| = |S|$.

- $MC = \vec{0}$, where matrix M is $|S| \times |S|$, and contains values of S at points $(s_i, t_i), i = 1, \dots, |S|$. Try roots of unity [Sturmfels-Tevelev-Yu'07].
- $(SS^T)C = \vec{0}$, substitute x, y, z by parametric expressions in $K[s, t]$, integrate over s, t ; solve for C [Corless-Galligo-Kotsireas-Watt'01].

Example: $\text{supp}(H) \subset \{x^3y, x^3, x^3y^2, y^2z^3\}$, then

$$SS^T = \begin{bmatrix} x^6y^2 & x^6y & x^6y^3 & x^3y^3z^3 \\ x^6y & x^6 & x^6y^2 & x^3y^2z^3 \\ x^6y^3 & x^6y^2 & x^6y^4 & x^3y^4z^3 \\ x^3y^3z^3 & x^3y^2z^3 & x^3y^4z^3 & y^4z^6 \end{bmatrix} \Rightarrow C = \begin{bmatrix} -2 \\ 1 \\ 1 \\ -1 \end{bmatrix}.$$

- Approximate implicitization [Dokken].

Implicit Newton polytope

Consider parameterizations with fixed supports.

- **Generic** coefficients:

- Compute the resultant's Newton polytope, then specialize:
[E-Kotsireas'03] developed Maple code calling Topcom [Rambau];
[E-Konaxis-Palios'07] specify implicit Newton polygon for curves.
[E-Konaxis-Fysikopoulos-Penaranda'11] fast algorithm for projecting resultant polytope in high-dim.
- Tropical geometry for varieties of $\text{codim} > 1$.
For curves, specified implicit polygon [Sturmfels-Tevelev-Yu'07].

- **Arbitrary** coefficients:

- Implicit Newton polygon for curves:
[Dickenstein-Feichtner-Sturmfels'07] study tropical discriminants;
[D'Andrea-Sombra'07] use mixed fiber polytopes [Esterov-Khovanskii'07].

Voronoi / Apollonius diagrams

Apollonius diagrams

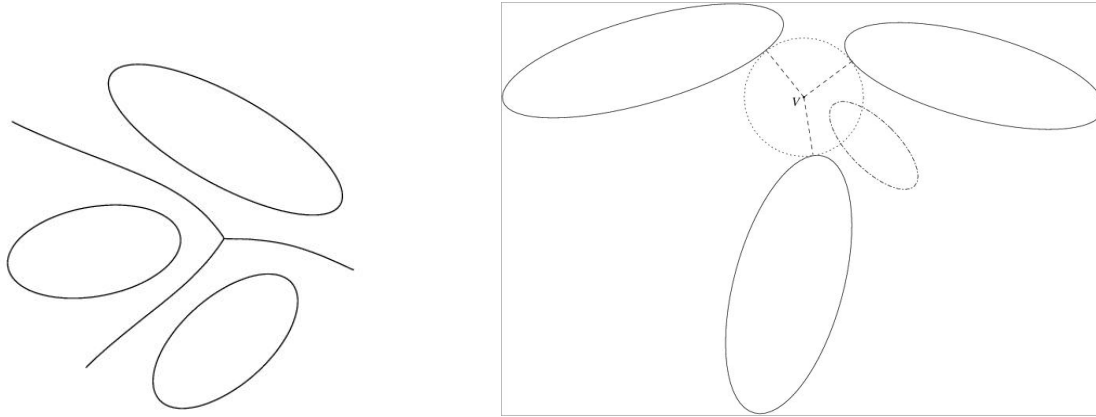
Def. Given n objects in \mathbb{R}^2 , their **Voronoi diagram** is a subdivision into n cells, each comprising the points closer to one object.

Nonlinear computational geometry considers circles, spheres, and ellipses. So, we refer to **Apollonius diagrams**.



Apollonius diagram of **green circles** [Karavelas-E'03], code in CGAL.

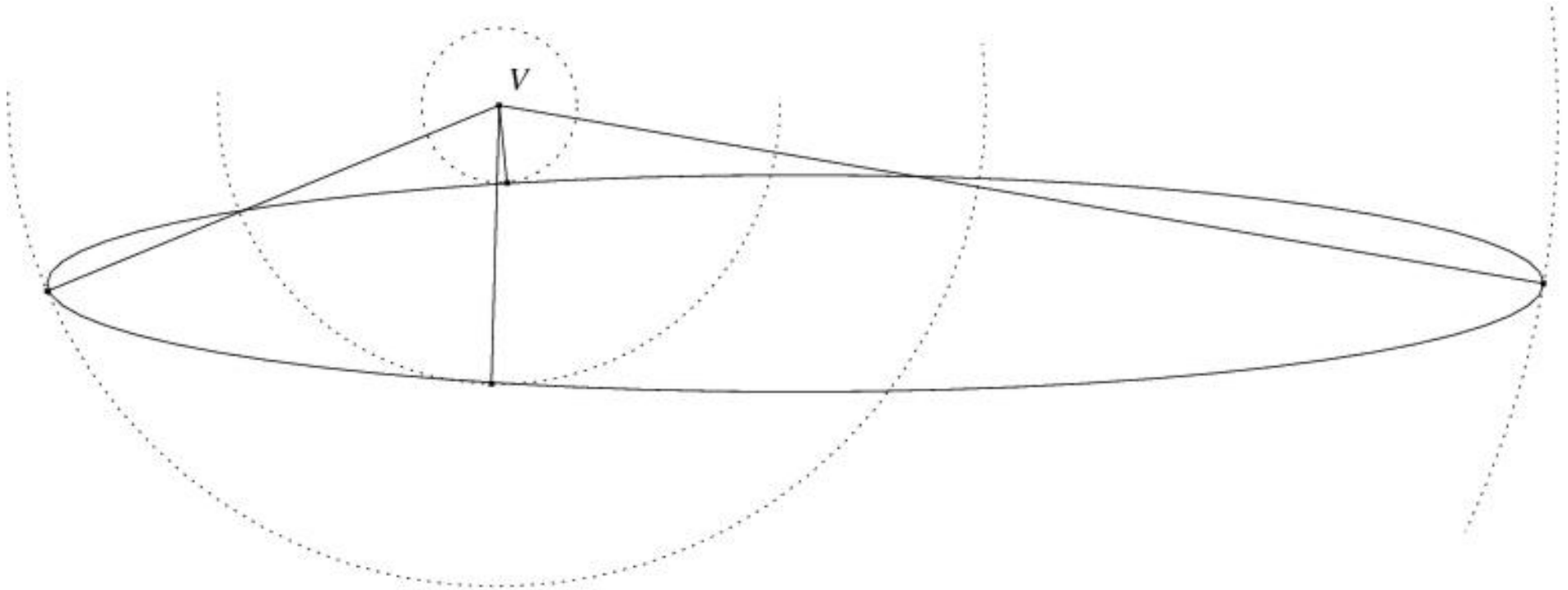
Apollonius diagram of ellipses



- Standard incremental **algorithm**.
- Problem: **predicates**, under Euclidean distance.
- For now: n **disjoint** ellipses.
- **Predicate 1**. Given 2 ellipses and an external point, decide which ellipse is closer to the point.
- **Main predicate**: 3 ellipses define one **Apollonius circle** externally tri-tangent to all: decide relative position of 4th ellipse wrt circle.

Point-ellipse distance

For a point outside an ellipse, there are **2-4 normals** onto the ellipse, depending on the point's position wrt the evolute curve.



Pencil of conics

General conic, M symmetric:

$$[x, y, 1]M[x, y, 1]^T = 0$$

Given ellipse, and circle centered at (v_1, v_2) with parametric radius:

$$E = \begin{pmatrix} a & b & d \\ b & c & e \\ d & e & f \end{pmatrix}, \quad C(s) = \begin{pmatrix} 1 & 0 & -v_1 \\ 0 & 1 & -v_2 \\ -v_1 & -v_2 & v_1^2 + v_2^2 - s \end{pmatrix}.$$

- Their pencil is $\lambda E + C(s)$,
- the characteristic polynomial is $\phi(s, \lambda) = |\lambda E + C(s)|$,
- and $\Delta(s)$ is ϕ 's discriminant (wrt λ).

Comparing point-ellipse distances

Thm. $\Delta(s) = 0 \Leftrightarrow E, C(s)$ have a multiple intersection

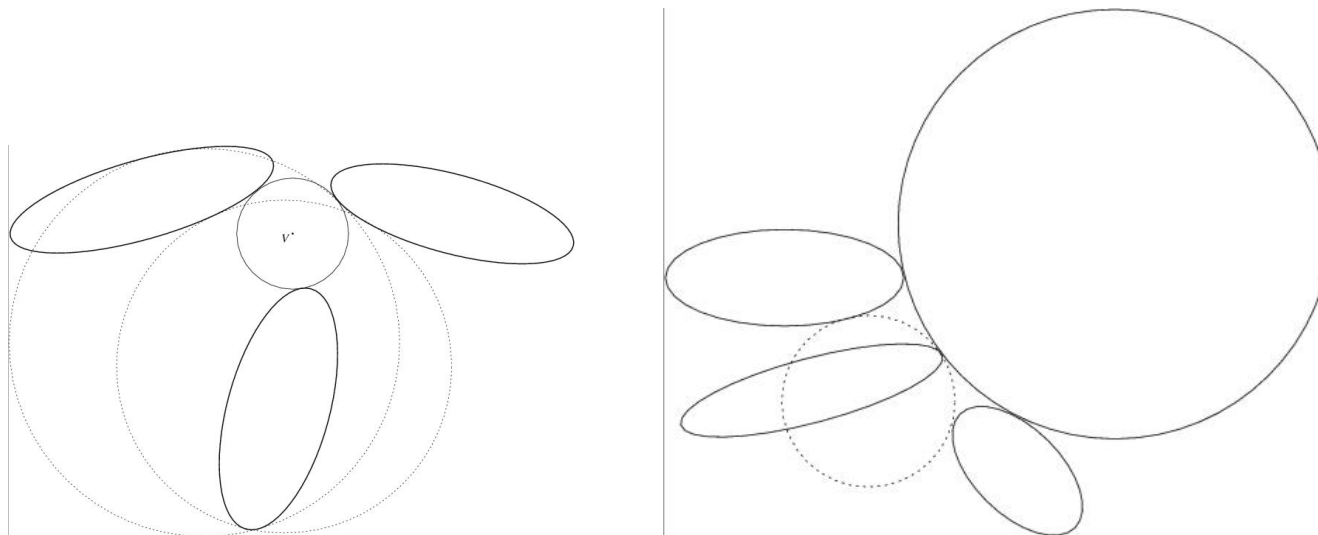
Given ellipse E and point v outside E , their **distance** is the square-root of the smallest positive root of the discriminant $\Delta(s)$.

Deciding which ellipse is closest to an external point reduces to comparing two **algebraic numbers** of degree 4. This degree is optimal.

Implemented in SYNAPS [E-Tsigaridas'04].

Apollonius circles

Given 3 ellipses, **how many** (real) tritangent circles are defined?



$$\text{MV} [\Delta_1(v_1, v_2, s), \Delta_2(v_1, v_2, s), \Delta_3(v_1, v_2, s)] = 256.$$

$$q := v_1^2 + v_2^2 - s \quad \Rightarrow \quad C(s) = \begin{pmatrix} 1 & 0 & -v_1 \\ 0 & 1 & -v_2 \\ -v_1 & -v_2 & q \end{pmatrix} \quad \Rightarrow \quad \text{MV} = 184.$$

Arguments from real algebraic geometry yield same [Sottile].

Unmixed bivariate systems

Given: unmixed system of 3 bivariate polynomials (identical supports).

\exists hybrid determinantal formula [Khetan'02]: $M = \begin{bmatrix} B & S \\ S^T & 0 \end{bmatrix}$

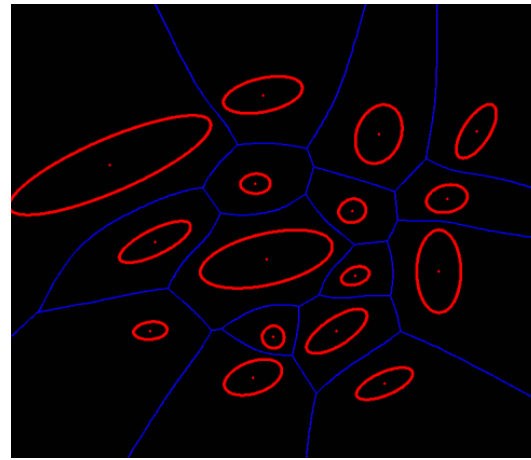
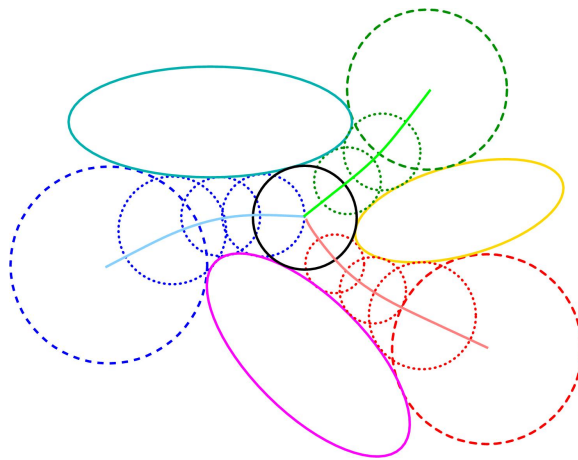
Eliminate $(v_1, v_2) \rightarrow 58 \times 58$ matrix with Sylvester and Bézout blocks:
sparse resultant $= \det(M)$, of degree 184 in q .

Open: How many real tritangent circles, in general?

Random example yields 8 real roots.

Voronoi diagram of ellipses

- **Sparse elimination**, Mixed Volume: 184 complex tritangent circles
- **Resultants**, factoring: sparse, successive Sylvester
- Adapted Newton's: quadratic convergence, **certified**
- **Real solving**: Complexity and software [E-Tsigaridas]
- Switch **representation**: implicit, parametric



- Geometric CGAL C++ software relying on algebra (Synaps, NTL).
- About 1sec per non-intersecting ellipse
- Faster than Voronoi of k -gons, $k \geq 15$ edges or $k \geq 200$ points.

[E-Tsigaridas-Tzoumas, SoCG'06] [E-Tz, CAD'08] [E-Ts-Tz, ACM/SIAM-GPM'09]

Parallel robots

Robot kinematics

Forward Kinematics: Compute all **displacements** for given configuration.
Easy/hard for serial/parallel robots respectively.

Inverse Kinematics: Compute all **configurations** that result to given translation – rotation (displacement).
Hard/easy for serial/parallel robots resp.

Parallel robots

Advantages: precision, rigidity, manipulation, force.

Examples: micro-surgery, flight simulation, heavy-duty objects etc.

Forward kinematics of **Stewart platform**: “The major outstanding problem in all of manipulator direct and inverse kinematics” [Roth93].

Configuration defined by the lengths of 6 articulations, system of 6 to 10 equations, ≤ 40 real solutions.

Stewart platform

Two rigid bodies connected with 6 sliding joints rotating freely at attachments: parallel mechanism.

Forward kinematics: Given joint lengths, compute pose of platform.

Rotation/translation/attachment quaternions $\dot{q}, \dot{t}, \dot{a}_i, \dot{a}'_i$

$$(-\dot{a}_i + \dot{t} + \dot{q}\dot{a}'_i\dot{q}^*)^T(-\dot{a}_i + \dot{t} + \dot{q}\dot{a}'_i\dot{q}^*) = L_i^2, \quad i = 1, \dots, 6.$$

Bézout bound = 256, m -Bézout = 144.

Exact bound = 40 [Ronga-Vust'92] [Mourrain'93] [Husty'94].

Can have 40 real solutions) [Dietmaier'98].

6×6 original system has $MV = 160$.

7×7 system with $\dot{x} = \dot{q}^* \dot{t}$ has $MV = 84$, $\deg R_{tor} = 214$, $\dim M = 405$.

10×10 system with $y_0 = \|\dot{q}\|^2, \dot{z} = \dot{q}^* \dot{t} \dot{q}$, has $MV = 54$.