Solving Polynomial Systems by Polyhedral Homotopies

Jan Verschelde

Department of Math, Stat & CS
University of Illinois at Chicago
Chicago, IL 60607-7045, USA

email: jan@math.uic.edu

URL: http://www.math.uic.edu/~jan

The Zero Problem: Theory and Applications

Korea Institute for Advanced Study, 19-21 July 2007

Summary

- many polynomial systems in practical applications are sparse; the sparser a system is, the fewer its number of solutions
- the number of solutions paths in a polyhedral homotopy is determined by the mixed volume of the Newton polytopes of the polynomials in the system
- mixed volumes give a sharp root count for generic systems; with stable mixed volumes we count roots in affine space
- to solve for large applications we use parallel implementations

Solving Polynomial Systems

- algebraic geometry studies solutions of polynomial systems
- applications in science and engineering lead to systems with approximate input coefficients
- most systems occur with parameters, often easier to solve first for a generic choice of the parameters
- although the complexity of the problem is intractable, homotopy continuation methods are pleasingly parallel

goal: turn applications into benchmark problems

Some Introductions and Surveys

I.M. Gel'fand, M.M. Kapranov, and A.V. Zelevinsky: **Discriminants, Resultants and Multidimensional Determinants**. Birkhäuser, 1994.

L. Blum, F. Cucker, M. Shub, and S. Smale: **Complexity and Real Computation**. Springer-Verlag, 1998.

B. Sturmfels: **Solving Systems of Polynomial Equations**. AMS, 2002.

T.Y. Li.: **Numerical solution of polynomial systems by homotopy continuation methods**. In F. Cucker, editor, *Handbook of Numerical Analysis. Volume XI. Special Volume: Foundations of Computational Mathematics*, pages 209–304. North-Holland, 2003.

A.J. Sommese and C.W. Wampler: **The Numerical Solution of Systems of Polynomials Arising in Engineering and Science**. World Scientific, 2005.

Homotopy Continuation Methods

Solve $f(\mathbf{x}) = \mathbf{0}$ in two stages:

1. The homotopy $h(\mathbf{x}, t) = \gamma(1-t)g(\mathbf{x}) + tf(\mathbf{x}) = \mathbf{0}$, $\gamma \in \mathbb{C}$, defines solution paths $\mathbf{x}(t)$, for t going from 0 to 1.

   $g(\mathbf{x}) = \mathbf{0}$ is a start system with the same structure as f.
   All solutions of $g(\mathbf{x}) = \mathbf{0}$ are isolated and regular.

2. Continuation methods apply predictor-corrector techniques to track the solution paths defined by the homotopy $h(\mathbf{x}, t) = \mathbf{0}$.
   Singularities do not occur for $t < 1$ for a generic choice of γ.

Knowing the right #paths is critical to the performance!

The Gamma Trick

Consider the homotopy $h(\mathbf{x}, t) = \gamma(1-t)g(\mathbf{x}) + tf(\mathbf{x}) = \mathbf{0}$, $\gamma \in \mathbb{C}$.
All solutions of $g(\mathbf{x}) = \mathbf{0}$ are isolated and regular.

1. **Singular** solutions of $h(\mathbf{x}, t) = \mathbf{0}$ satisfy

$$H(\mathbf{x}, t) = \left\{ \begin{array}{r} h(\mathbf{x}, t) = \mathbf{0} \\ \det(J_h(\mathbf{x}, t)) = 0 \end{array} \right. \qquad J_h \text{ is the Jacobian of } h.$$

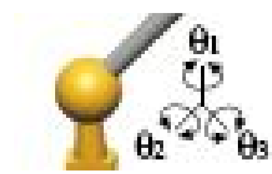
2. Embed $(\mathbf{x}, t) \in \mathbb{C}^n \times \mathbb{C}$ into projective space: $(\mathbf{z}, t) \in \mathbb{P}^n \times \mathbb{C}$.
   Apply the **main theorem of elimination theory** to $H^{-1}(\mathbf{0})$,
   eliminating $\mathbf{z}$, i.e.: apply $\pi : \mathbb{P}^n \times \mathbb{C} \rightarrow \mathbb{C} : (\mathbf{z}, t) \mapsto t$.
   Then $\pi(H^{-1}(\mathbf{0}))$ is an algebraic set, defined by $p(t) = 0$.

3. Because all solutions of $g(\mathbf{x}) = \mathbf{0}$ are isolated and regular,
   $p(0) \neq 0$. So there are **only finitely many** singularities.

For a generic choice of γ, $H(\mathbf{x}, t) = \mathbf{0}$ has no solutions for $t \in [0, 1)$.

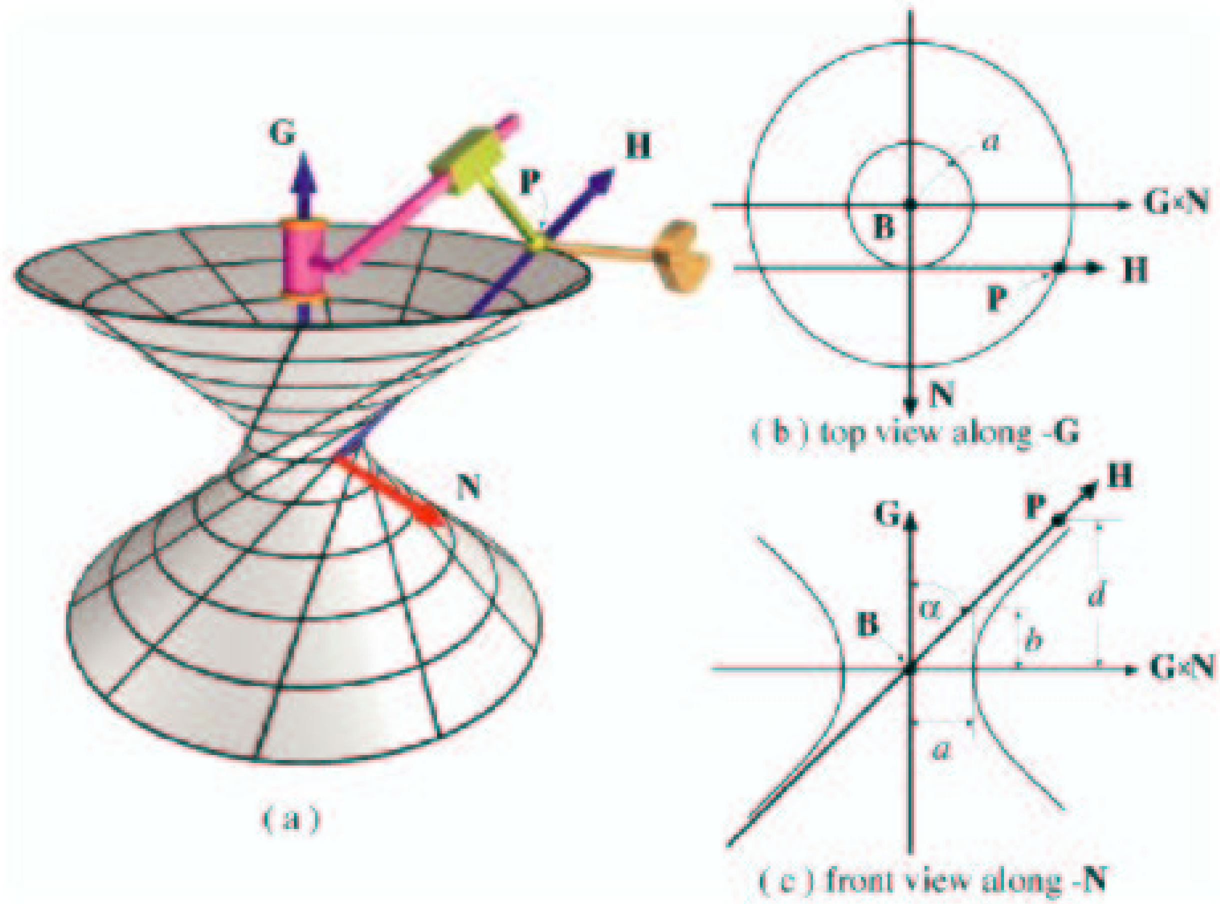
Design of an RPS robot

robot = serial chain in which each joint is actuated

type of joint	diagram	symbol	DOF
Revolute		R	1
Prismatic		P	1
Spherical		S	3

H.J. Su. *Computer-Aided Constrained Robot Design Using Mechanism Synthesis Theory.* PhD thesis, University of California, Irvine, 2004.

Reachable Surfaces: Circular Hyperboloid



H.J. Su. *Computer-Aided Constrained Robot Design Using Mechanism Synthesis Theory*. PhD thesis, University of California, Irvine, 2004.

The Equations for RPS10

Ten unknowns: $\mathbf{g} = (g_1, g_2, g_3)$ is the axis of the R-joint, $\mathbf{p} = (p_1, p_2, p_3)$ is the center of the S-joint, and k_0, $\mathbf{k} = (k_1, k_2, k_3)$ are derived from $\mathbf{g}$ and the center of the hyperboloid.

Ten parameters: $\mathbf{p}_i = [T_i]\mathbf{p}$, $i = 1, 2, \ldots, 10$, goal positions defined by transformations $[T_i]$.

$$\left\{ \begin{array}{c} k_0(\mathbf{p}_2 \cdot \mathbf{p}_2 - \mathbf{p}_1 \cdot \mathbf{p}_1) + 2\mathbf{k} \cdot (\mathbf{p}_2 - \mathbf{p}_1) - (\mathbf{p}_2 \cdot \mathbf{g})^2 + (\mathbf{p}_1 \cdot \mathbf{g})^2 = 0 \\ k_0(\mathbf{p}_3 \cdot \mathbf{p}_3 - \mathbf{p}_1 \cdot \mathbf{p}_1) + 2\mathbf{k} \cdot (\mathbf{p}_3 - \mathbf{p}_1) - (\mathbf{p}_3 \cdot \mathbf{g})^2 + (\mathbf{p}_1 \cdot \mathbf{g})^2 = 0 \\ \vdots \\ k_0(\mathbf{p}_{10} \cdot \mathbf{p}_{10} - \mathbf{p}_1 \cdot \mathbf{p}_1) + 2\mathbf{k} \cdot (\mathbf{p}_{10} - \mathbf{p}_1) - (\mathbf{p}_{10} \cdot \mathbf{g})^2 + (\mathbf{p}_1 \cdot \mathbf{g})^2 = 0 \\ c_1 g_1 + c_2 g_2 + c_3 g_3 - 1 = 0 \end{array} \right.$$

The coefficients c_1, c_2, c_3 are random numbers to scale $\mathbf{g}$.

Expected Number of Solutions

A system of nine quartics and one linear equation ...

total degree: $4^9 \times 1 = 262,144$

linear-product Bézout bound: 9,216

H.J. Su and J.M. McCarthy: **Kinematic synthesis of RPS serial chains for a given Set of task positions**. *Mechanism and Machine Theory*, 40(7):757-775, 2005.

mixed volume: 1,204 is sharp in this case.

For generic choices of the parameters, there will always be 1,024 complex solutions.

For a specific choice of the parameters, we find 128 real solutions.

On a (slow) Mac OS X laptop, solving takes about 40 minutes, less than two seconds are spent on computing the mixed volume.

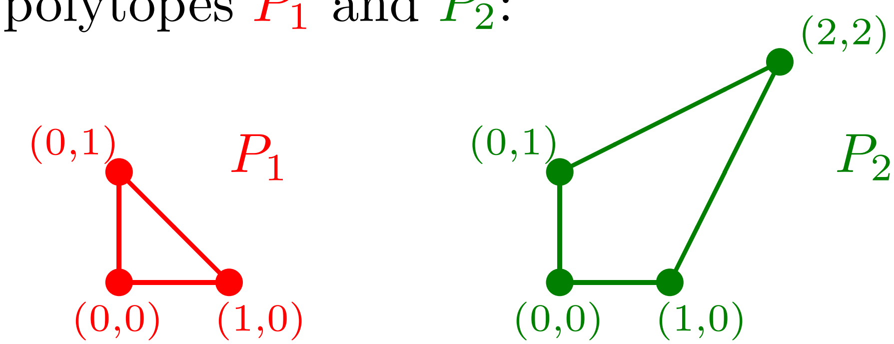
Newton Polytopes

$$f(x_1, x_2) = \left\{ \begin{array}{lcl} f_1 & : & 0 = c_{1,(0,0)} + c_{1,(0,1)} x_2 + c_{1,(1,0)} x_1 \\ f_2 & : & 0 = c_{2,(0,0)} + c_{2,(0,1)} x_2 + c_{2,(1,0)} x_1 + c_{2,(2,2)} x_1^2 x_2^2 \end{array} \right.$$

coefficients $c_{i,\mathbf{a}} \in \mathbb{C}^* = \mathbb{C} \setminus \{0\}$ exponents $\mathbf{a} = (a_1, a_2) \in \mathbb{Z}^2$

supports $A_1 = \{(0,0), (1,0), (0,1)\}$ and $A_2 = \{(0,0), (1,0), (0,1), (2,2)\}$

span Newton polytopes P_1 and P_2:



notation:

$$f(\mathbf{x}) = \sum_{\mathbf{a} \in A_i} c_{i,\mathbf{a}} \mathbf{x}^{\mathbf{a}}, \quad \mathbf{x}^{\mathbf{a}} = x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$$

Puiseux Series

Let $g(\mathbf{x}) = \mathbf{0}$ be a generic system with same supports as f.

Consider the homotopy $h(\mathbf{x}, t) = \gamma(1-t)g(\mathbf{x}) + tf(\mathbf{x}) = \mathbf{0}$.

Then, for $t \to 1$: $$\left\{ \begin{array}{ll} x_i(s) = b_i s^{v_i}(1 + O(s)) & \qquad b_i \in \mathbb{C}^*, v_i \in \mathbb{Z}, \\ \;\; t(s) = 1 - s^m & \qquad s \to 0, m \in \mathbb{N}^+, \end{array} \right.$$

where m is the common denominator in the fractional power series.

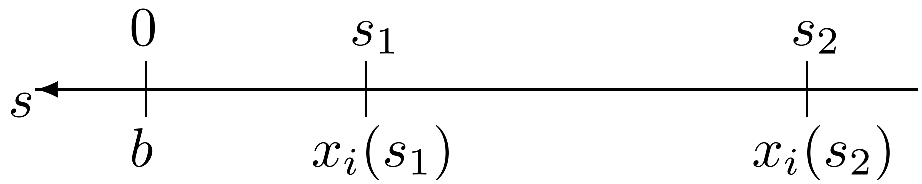
As $t \to 1$, observe the sign of v_i:

$$v_i \left\{ \begin{array}{lcl} > 0 & \Rightarrow & x_i \to 0 \notin \mathbb{C}^*, \\ = 0 & \Rightarrow & x_i \to b_i \in \mathbb{C}^*, \\ < 0 & \Rightarrow & x_i \to \infty \notin \mathbb{C}^*. \end{array} \right.$$

Whether x_i converges to $b_i \in \mathbb{C}^*$ depends on $v_i = 0$, $v_i \in \mathbb{Z}$.

Extrapolation

Assume $m = 1$:



To compute v_i in

$$
\begin{aligned}
x_i(s_1) &= b_i s_1^{v_i}(1 + O(s_1)) \\
x_i(s_2) &= b_i s_2^{v_i}(1 + O(s_2))
\end{aligned}
$$

we take logarithms:

$$
\begin{array}{rcl}
\log|x_i(s_1)| & = & \log|b_i| + v_i \log|s_1| + \log|1 + O(s_1)| \\
-(\ \log|x_i(s_2)| & = & \log|b_i| + v_i \log|s_2| + \log|1 + O(s_2)|\ ) \\
\hline
\end{array}
$$

and eliminate $\log|b_i|$: $\quad v_i = \dfrac{\log|x_i(s_1)| - \log|x_i(s_2)|}{\log|s_1| - \log|s_2|} \in \mathbb{Z}.$

For $m > 1$, higher order extrapolation is needed.

If $v_i = 0$, $x_i(s) \to b$. If $v_i < 0$, $x_i(s) \to \infty$. If $v_i > 0$, $x_i(s) \to 0$.

Diverging Paths

Substitute $\mathbf{x}(s) = \mathbf{b}s^{\mathbf{v}}(1 + O(s))$ and $t(s) = 1 - s^m$, $s \longrightarrow 0$,
in the homotopy $h(\mathbf{x}, t) = \gamma(1-t)g(\mathbf{x}) + tf(\mathbf{x}) = \mathbf{0}$, $t \longrightarrow 1$:

Only f matters: $h(\mathbf{x}(s), s) = f(\mathbf{x}(s)) + s^m(f(\mathbf{x}(s)) - \gamma g(\mathbf{x}(s)))$.

$$\text{Then,} \quad f_i(\mathbf{x}(s)) = \sum_{\mathbf{a} \in A_i} c_{i,\mathbf{a}} \mathbf{x}(s)^{\mathbf{a}} = \sum_{\mathbf{a} \in A_i} c_{i,\mathbf{a}} (\mathbf{b}s^{\mathbf{v}}(1 + O(s)))^{\mathbf{a}}.$$

$$\text{Ignore } O(s) : (\mathbf{b}s^{\mathbf{v}})^{\mathbf{a}} = b_1 s^{v_1 a_1} b_2 s^{v_2 a_2} \cdots b_n s^{v_n a_n} = \left(\prod_{j=1}^{n} b_j \right) s^{\langle \mathbf{v}, \mathbf{a} \rangle}.$$

Collect exponents of dominant terms as $s \longrightarrow 0$ in

$$\partial_{\mathbf{v}} A_i = \{ \ \mathbf{a} \in A \mid \langle \mathbf{v}, \mathbf{a} \rangle = \min_{\mathbf{a}' \in A} \langle \mathbf{v}, \mathbf{a}' \rangle \ \}.$$

The set $\partial_{\mathbf{v}} A_i$ spans the face $\partial_{\mathbf{v}} P_i$ of the Newton polytope P_i of f_i.

Bernshteĭn's second theorem

- Face $\partial_{\mathbf{v}} f = (\partial_{\mathbf{v}} f_1, \partial_{\mathbf{v}} f_2, \ldots, \partial_{\mathbf{v}} f_n)$ of system $f = (f_1, f_2, \ldots, f_n)$ with Newton polytopes $\mathcal{P} = (P_1, P_2, \ldots, P_n)$ and mixed volume $V(\mathcal{P})$.

$$\partial_{\mathbf{v}} f_i(\mathbf{x}) = \sum_{\mathbf{a} \in \partial_{\mathbf{v}} A_i} c_{i,\mathbf{a}} \mathbf{x}^{\mathbf{a}} \qquad \underset{\text{face of Newton polytope}}{\partial_{\mathbf{v}} P_i = \mathrm{conv}(\partial_{\mathbf{v}} A_i)}$$

Theorem: **If $\forall \mathbf{v} \neq \mathbf{0}$, $\partial_{\mathbf{v}} f(\mathbf{x}) = \mathbf{0}$ has no solutions in $(\mathbb{C}^*)^n$, then $V(\mathcal{P})$ is exact and all solutions are isolated.**

Otherwise, for $V(\mathcal{P}) \neq 0$: $V(\mathcal{P}) > \#$isolated solutions.

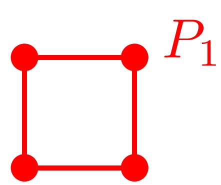
- Newton polytopes ***in general position***: $V(\mathcal{P})$ is **exact for every nonzero choice of the coefficients**.

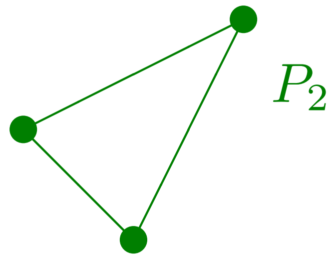
Newton polytopes in general position

Consider

$$f(\mathbf{x}) = \begin{cases} c_{1,(1,1)}x_1x_2 + c_{1,(1,0)}x_1 + c_{1,(0,1)}x_2 + c_{1,(0,0)} = 0 \\ c_{2,(2,2)}x_1^2x_2^2 + c_{2,(1,0)}x_1 + c_{2,(0,1)}x_2 = 0 \end{cases}$$

The Newton polytopes:





$A_1 = \{(1,1),(1,0),(0,1),(0,0)\}$ $\qquad$ $A_2 = \{(2,2),(1,0),(0,1)\}$

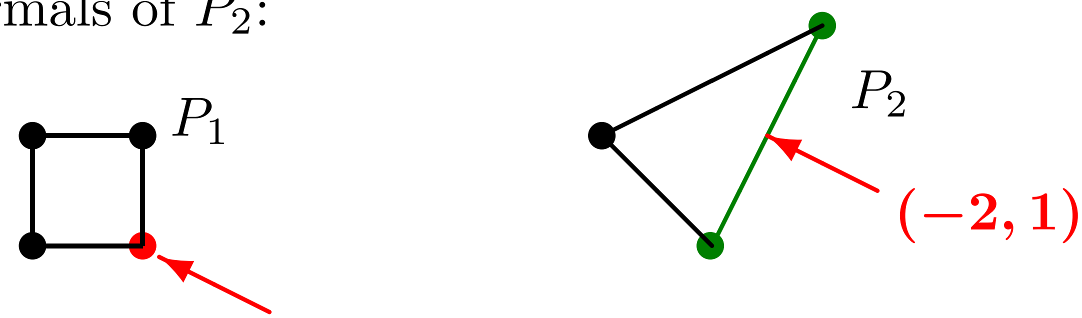
$P_1 = \mathrm{conv}(A_1)$ $\qquad$ $P_2 = \mathrm{conv}(A_2)$

Newton polytopes in general position

Consider

$$f(\mathbf{x}) = \begin{cases} c_{1,(1,1)}x_1x_2 + c_{1,(1,0)}x_1 + c_{1,(0,1)}x_2 + c_{1,(0,0)} = 0 \\ c_{2,(2,2)}x_1^2x_2^2 + c_{2,(1,0)}x_1 + c_{2,(0,1)}x_2 = 0 \end{cases}$$

Look at the inner normals of P_2:



$\rightarrow$ the corresponding face system $\begin{cases} c_{1,(1,0)}x_1 = 0 \\ c_{2,(2,2)}x_1^2x_2^2 + c_{2,(1,0)}x_1 = 0 \end{cases}$

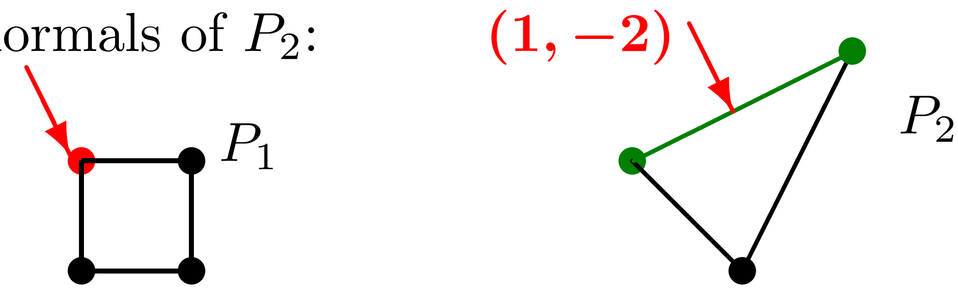
does not have a solution in $(\mathbb{C}^*)^2$.

Newton polytopes in general position

Consider

$$f(\mathbf{x}) = \begin{cases} c_{1,(1,1)}x_1x_2 + c_{1,(1,0)}x_1 + c_{1,(0,1)}x_2 + c_{1,(0,0)} = 0 \\ c_{2,(2,2)}x_1^2x_2^2 + c_{2,(1,0)}x_1 + c_{2,(0,1)}x_2 = 0 \end{cases}$$

Look at the inner normals of P_2:



$\longrightarrow$ the corresponding face system $\begin{cases} c_{1,(1,0)}x_2 = 0 \\ c_{2,(2,2)}x_1^2x_2^2 + c_{2,(1,0)}x_2 = 0 \end{cases}$

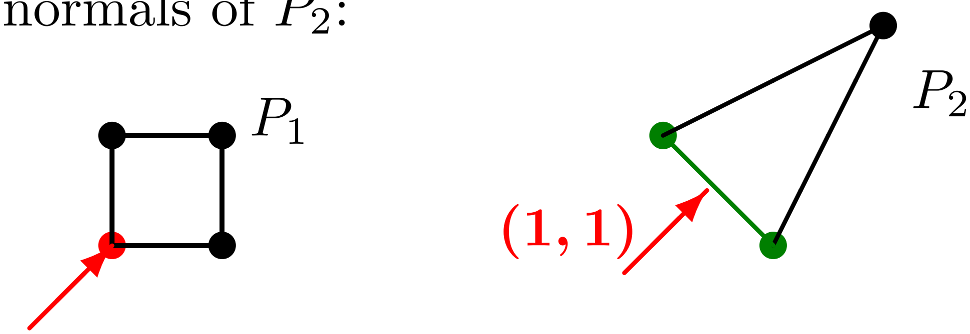
does not have a solution in $(\mathbb{C}^*)^2$.

Newton polytopes in general position

Consider

$$f(\mathbf{x}) = \begin{cases} c_{1,(1,1)}x_1x_2 + c_{1,(1,0)}x_1 + c_{1,(0,1)}x_2 + c_{1,(0,0)} = 0 \\ c_{2,(2,2)}x_1^2x_2^2 + c_{2,(1,0)}x_1 + c_{2,(0,1)}x_2 = 0 \end{cases}$$

Look at the inner normals of P_2:



$\rightarrow$ the corresponding face system $\begin{cases} c_{1,(0,0)} = 0 \\ c_{2,(1,0)}x_1 + c_{2,(0,1)}x_2 = 0 \end{cases}$

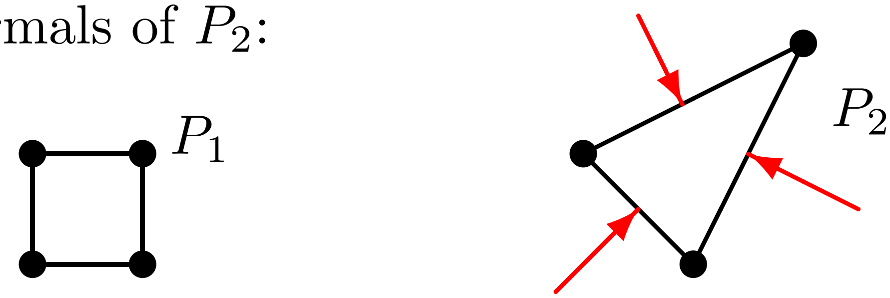
does not have a solution in $(\mathbb{C}^*)^2$.

Newton polytopes in general position

Consider

$$f(\mathbf{x}) = \begin{cases} c_{1,(1,1)}x_1x_2 + c_{1,(1,0)}x_1 + c_{1,(0,1)}x_2 + c_{1,(0,0)} = 0 \\ c_{2,(2,2)}x_1^2x_2^2 + c_{2,(1,0)}x_1 + c_{2,(0,1)}x_2 = 0 \end{cases}$$

Look at the inner normals of P_2:



$$\forall \mathbf{v} \neq \mathbf{0} : \partial_{\mathbf{v}} A_1 + \partial_{\mathbf{v}} A_2 \leq 3 \quad \Rightarrow \quad V(P_1, P_2) = 4$$ **always exact**

for all nonzero coefficients

Richardson Extrapolation for $\mathbf{v}$ and m

$$\left\{ \begin{array}{rcl} x_i(s) & = & b_i s^{v_i}(1 + O(s)) \\ t(s) & = & 1 - s^m \end{array} \right.$$

$$x_i(s_k) = b_i h^{k v_i/m} s_0 (1 + O(h^{k/m} s_0))$$

Geometric sampling $0 < h < 1$

$$1-t_k = h(1-t_k) = \cdots = h^k(1-t_0)$$

$$s_k = h^{1/m} s_{k-1} = \cdots = h^{k/m} s_0$$

Input: $(\mathbf{x}(s), t(s))$ solutions along a path, $h(\mathbf{x}(s), t(s)) = \mathbf{0}$.

Output: approximations for $\mathbf{v}$ and m.

Richardson Extrapolation for $\mathbf{v}$ and m

$$\left\{\begin{array}{rcl} x_i(s) & = & b_i s^{v_i}(1+O(s)) \\ t(s) & = & 1-s^m \end{array}\right.$$

$$x_i(s_k) = b_i h^{k v_i/m} s_0 (1 + O(h^{k/m} s_0))$$

Geometric sampling $0 < h < 1$

$$1-t_k = h(1-t_k) = \cdots = h^k(1-t_0)$$

$$s_k = h^{1/m} s_{k-1} = \cdots = h^{k/m} s_0$$

Take logarithms to find exponents of power series:

- $\log|x_i(s_k)| = \log|b_i| + \frac{k v_i}{m}\log(h) + v_i \log(s_0)$
  $+ \log(1 + \sum_{j=0}^{\infty} b'_j (h^{k/m} s_0)^j)$

  $w_{kk+1} := \log|x_i(s_k+1)| - \log|x_i(s_k)|$

Extrapolation on samples

$$w_{k..l} = w_{k..l-1} + \frac{w_{k+1..l} - w_{k..l-1}}{1-h}$$

$$v_i = m\frac{w_{0..r}}{\log(h)} + O(s_0^r)$$

$\rightarrow$ first-order approximation for $\mathbf{v}$

... is okay for $m = 1$

Richardson Extrapolation for v and m

$$\begin{cases} x_i(s) &= b_i s^{v_i}(1+O(s)) \\ t(s) &= 1-s^m \end{cases}$$

$$x_i(s_k) = b_i h^{kv_i/m} s_0(1+O(h^{k/m}s_0))$$

Geometric sampling $0 < h < 1$

$$1-t_k=h(1-t_k)=\cdots=h^k(1-t_0)$$

$$s_k=h^{1/m}s_{k-1}=\cdots=h^{k/m}s_0$$

- $\log|x_i(s_k)| = \log|b_i| + \frac{kv_i}{m}\log(h) + v_i\log(s_0)$

  $+\log(1+\sum_{j=0}^{\infty} b'_j(h^{k/m}s_0)^j)$

  $w_{kk+1} := \log|x_i(s_k+1)| - \log|x_i(s_k)|$

Extrapolation on samples

$$w_{k..l}=w_{k..l-1}+\frac{w_{k+1..l}-w_{k..l-1}}{1-h}$$

$$v_i = m\frac{w_{0..r}}{\log(h)} + O(s_0^r)$$

- $e_i^{(k)} = (\log|x_i(s_k)| - \log|x_i(s_{k+1})|)$

  $-(\log|x_i(s_{k+1})| - \log|x_i(s_{k+2})|)$

  $= c_1 h^{k/m} s_0(1+0(h^{k/m}))$

  $e_i^{(kk+1)} := \log(e_i^{(k+1)}) - \log(e_i^{(k)})$

Extrapolation on errors

$$e_i^{(k..l)}=e_i^{(k+1..l)}+\frac{e_i^{(k..l-1)}-e_i^{(k+1..l)}}{1-h_{k..l}}$$

$$h_{k..l} = h^{(l-k-1)/m_{k..l}}$$

$$m_{k..l} = \frac{\log(h)}{e_i^{(k..l)}} + O(h^{(l-k)k/m})$$

the system of Cassou-Noguès

$$f(b,c,d,e) =$$
$$\left\{\begin{array}{rcl}
15b^4cd^2 + 6b^4c^3 + 21b^4c^2d - 144b^2c - 8b^2c^2e & & \\
-28b^2cde - 648b^2d + 36b^2d^2e + 9b^4d^3 - 120 & = & 0 \\
30c^3b^4d - 32de^2c - 720db^2c - 24c^3b^2e - 432c^2b^2 + 576ec & & \\
-576de + 16cb^2d^2e + 16d^2e^2 + 16e^2c^2 + 9c^4b^4 + 5184 & & \\
+39d^2b^4c^2 + 18d^3b^4c - 432d^2b^2 + 24d^3b^2e - 16c^2b^2de - 240c & = & 0 \\
216db^2c - 162d^2b^2 - 81c^2b^2 + 5184 + 1008ec - 1008de & & \\
+15c^2b^2de - 15c^3b^2e - 80de^2c + 40d^2e^2 + 40e^2c^2 & = & 0 \\
261 + 4db^2c - 3d^2b^2 - 4c^2b^2 + 22ec - 22de & = & 0
\end{array}\right.$$

Root counts: $D = 1344$, $B = 312$, $V(\mathcal{P}) = 24 > 16$ finite roots.

$$\begin{array}{l} \partial_{(0,0,0,-1)} f(b,c,d,e) = \\ m = 2 \end{array}
\left\{\begin{array}{rcccl}
-8b^2c^2e - 28b^2cde + 36b^2d^2e & = & 0 & = & -2c^2 - 7cd + 9d^2 \\
-32de^2c + 16d^2e^2 + 16e^2c^2 & = & 0 & = & -2dc + d^2 + c^2 \\
-80de^2c + 40d^2e^2 + 40e^2c^2 & = & 0 & = & -2dc + d^2 + c^2 \\
22ec - 22de & = & 0 & = & c - d
\end{array}\right.$$

Geometric Root Counting

$$f_i(\mathbf{x}) = \sum_{\mathbf{a} \in A_i} c_{i\mathbf{a}} \mathbf{x}^{\mathbf{a}}$$

$$c_{i,\mathbf{a}} \in \mathbb{C}^* = \mathbb{C} \setminus \{0\}$$

$$f = (f_1, f_2, \ldots, f_n)$$

$$P_i = \mathrm{conv}(A_i)$$

Newton polytope

$$\mathcal{P} = (P_1, P_2, \ldots, P_n)$$

$L(f)$ root count in $(\mathbb{C}^*)^n$	desired properties
$L(f) = L(f_2, f_1, \ldots, f_n)$	invariant under permutations
$L(f) = L(f_1 \mathbf{x}^{\mathbf{a}}, \ldots, f_n)$	shift invariant
$L(f) \leq L(f_1 + \mathbf{x}^{\mathbf{a}}, \ldots, f_n)$	monotone increasing
$L(f) = L(f_1(\mathbf{x}^{U\mathbf{a}}), \ldots, f_n(\mathbf{x}^{U\mathbf{a}}))$	unimodular invariant
$L(f_{11} f_{12}, \ldots, f_n)$ $= L(f_{11}, \ldots, f_n) + L(f_{12}, \ldots, f_n)$	root count of product is sum of root counts

Geometric Root Counting

$$f_i(\mathbf{x}) = \sum_{\mathbf{a} \in A_i} c_{i\mathbf{a}} \mathbf{x}^{\mathbf{a}}$$

$$c_{i,\mathbf{a}} \in \mathbb{C}^* = \mathbb{C} \setminus \{0\}$$

$$f = (f_1, f_2, \ldots, f_n)$$

$$P_i = \mathrm{conv}(A_i)$$

Newton polytope

$$\mathcal{P} = (P_1, P_2, \ldots, P_n)$$

properties of $L(f)$	$V(\mathcal{P})$ mixed volume
invariant under permutations	$V(P_2, P_1, \ldots, P_n) = V(\mathcal{P})$
shift invariant	$V(P_1 + \mathbf{a}, \ldots, P_n) = V(\mathcal{P})$
monotone increasing	$V(\mathrm{conv}(P_1 + \mathbf{a}), \ldots, P_n) \geq V(\mathcal{P})$
unimodular invariant	$V(UP_1, \ldots, UP_n) = V(\mathcal{P})$
root count of product is sum of root counts	$V(P_{11} + P_{12}, \ldots, P_n) = V(P_{11}, \ldots, P_n) + V(P_{12}, \ldots, P_n)$

Geometric Root Counting

$$f_i(\mathbf{x}) = \sum_{\mathbf{a} \in A_i} c_{i\mathbf{a}} \mathbf{x}^{\mathbf{a}}$$

$$c_{i,\mathbf{a}} \in \mathbb{C}^* = \mathbb{C} \setminus \{0\}$$

$$f = (f_1, f_2, \ldots, f_n)$$

$$P_i = \mathrm{conv}(A_i)$$

Newton polytope

$$\mathcal{P} = (P_1, P_2, \ldots, P_n)$$

$L(f)$ root count in $(\mathbb{C}^*)^n$	$V(\mathcal{P})$ mixed volume
$L(f) = L(f_2, f_1, \ldots, f_n)$	$V(P_2, P_1, \ldots, P_n) = V(\mathcal{P})$
$L(f) = L(f_1 \mathbf{x}^{\mathbf{a}}, \ldots, f_n)$	$V(P_1 + \mathbf{a}, \ldots, P_n) = V(\mathcal{P})$
$L(f) \leq L(f_1 + \mathbf{x}^{\mathbf{a}}, \ldots, f_n)$	$V(\mathrm{conv}(P_1 + \mathbf{a}), \ldots, P_n) \geq V(\mathcal{P})$
$L(f) = L(f_1(\mathbf{x}^{U\mathbf{a}}), \ldots, f_n(\mathbf{x}^{U\mathbf{a}}))$	$V(UP_1, \ldots, UP_n) = V(\mathcal{P})$
$L(f_{11} f_{12}, \ldots, f_n)$ $= L(f_{11}, \ldots, f_n) + L(f_{12}, \ldots, f_n)$	$V(P_{11} + P_{12}, \ldots, P_n)$ $= V(P_{11}, \ldots, P_n) + V(P_{12}, \ldots, P_n)$

exploit sparsity $\boxed{L(f) = V(\mathcal{P})}$ 1st theorem of Bernshteĭn

The Theorems of Bernshteĭn

Theorem A: The number of roots of a generic system equals the mixed volume of its Newton polytopes.

Theorem B: Solutions at infinity are solutions of systems supported on faces of the Newton polytopes.

D.N. Bernshteĭn: **The number of roots of a system of equations.** *Functional Anal. Appl.*, 9(3):183–185, 1975.

Structure of proofs: First show Theorem B, looking at power series expansions of diverging paths defined by a linear homotopy starting at a generic system. Then show Theorem A, using Theorem B with a homotopy defined by *lifting* the polytopes.

Some References

J. Canny and J.M. Rojas: **An optimal condition for determining the exact number of roots of a polynomial system.** In ***Proceedings of ISSAC 1991***, pages 96–101. ACM, 1991.

J. Verschelde, P. Verlinden, and R. Cools: **Homotopies exploiting Newton polytopes for solving sparse polynomial systems.** ***SIAM J. Numer. Anal.*** 31(3):915–930, 1994.

B. Huber and B. Sturmfels: **A polyhedral method for solving sparse polynomial systems**. ***Math. Comp.*** 64(212):1541–1555, 1995.

I.Z. Emiris and J.F. Canny: **Efficient incremental algorithms for the sparse resultant and the mixed volume**. ***J. Symbolic Computation*** 20(2):117–149, 1995.

Bernshteĭn's first theorem

Let $g(\mathbf{x}) = \mathbf{0}$ have the same Newton polytopes $\mathcal{P}$ as $f(\mathbf{x}) = \mathbf{0}$, but with randomly choosen complex coefficients.

I. Compute $V_n(\mathcal{P})$:			II. Solve $g(\mathbf{x}) = \mathbf{0}$:	
I.1	lift polytopes	$\Leftrightarrow$	II.1	introduce parameter t
I.2	mixed cells	$\Leftrightarrow$	II.2	start systems
I.3	volume of mixed cell	$\Leftrightarrow$	II.3	path following

III. Coefficient-parameter continuation to solve $f(\mathbf{x}) = \mathbf{0}$:

$$h(\mathbf{x}, t) = \gamma(1-t)g(\mathbf{x}) + tf(\mathbf{x}) = \mathbf{0}, \quad \text{for } t \text{ from 0 to 1.}$$

#isolated solutions in $(\mathbb{C}^*)^n$, $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$, of $f(\mathbf{x}) = \mathbf{0}$ is bounded by the mixed volume of the Newton polytopes of f.

Finding Mixed Cells

$$g(x_1, x_2, t) = \begin{cases} c_{1,(1,1)} x_1 x_2 t^2 + c_{1,(1,0)} x_1 t^7 + c_{1,(0,1)} x_2 t^3 + c_{1,(0,0)} t^3 = 0 \\ c_{2,(2,2)} x_1^2 x_2^2 t^5 + c_{2,(1,0)} x_1 t^3 + c_{2,(0,1)} x_2 t^2 = 0 \end{cases}$$

- At $t = 1$: the system $g(\mathbf{x}, 1) = g(\mathbf{x}) = \mathbf{0}$ we want to solve.
- Where to start?

  $\rightarrow$ look for inner normals $\mathbf{v} \in \mathbb{Z}^3$, $v_3 > 0$, such that after

  $$x_1 = y_1 s^{v_1}, \quad x_2 = y_2 s^{v_2}, \quad t = s^{v_3},$$

  the system $g(\mathbf{y}, s) = \mathbf{0}$ has solutions in $(\mathbb{C}^*)^2$ at $s = 0$.

Coordinate Transformations give Homotopies

$$g(x_1 = y_1, x_2 = y_2 s, t = s)$$

$$= \begin{cases} c_{1,(1,1)} y_1 (y_2 s) s^2 + c_{1,(1,0)} y_1 s^7 + c_{1,(0,1)} (y_2 s) s^3 + c_{1,(0,0)} s^3 = 0 \\ c_{2,(2,2)} y_1^2 (y_2 s)^2 s^5 + c_{2,(1,0)} y_1 s^3 + c_{2,(0,1)} (y_2 s) s^2 = 0 \end{cases}$$

$$= \begin{cases} c_{1,(1,1)} y_1 y_2 s^3 + c_{1,(1,0)} y_1 s^7 + c_{1,(0,1)} y_2 s^4 + c_{1,(0,0)} s^3 = 0 \\ c_{2,(2,2)} y_1^2 y_2^2 s^7 + c_{2,(1,0)} y_1 s^3 + c_{2,(0,1)} y_2 s^3 = 0 \end{cases}$$

$$= \begin{cases} c_{1,(1,1)} y_1 y_2 + c_{1,(1,0)} y_1 s^4 + c_{1,(0,1)} y_2 s + c_{1,(0,0)} = 0 \\ c_{2,(2,2)} y_1^2 y_2^2 s^4 + c_{2,(1,0)} y_1 + c_{2,(0,1)} y_2 = 0 \end{cases}$$

At $s = 0$ we find a binomial system which has two solutions.
The two solutions extend to solutions of $g(\mathbf{x}) = g(\mathbf{x}, s = 1) = \mathbf{0}$.

Coordinate Transformations and Inner Normals

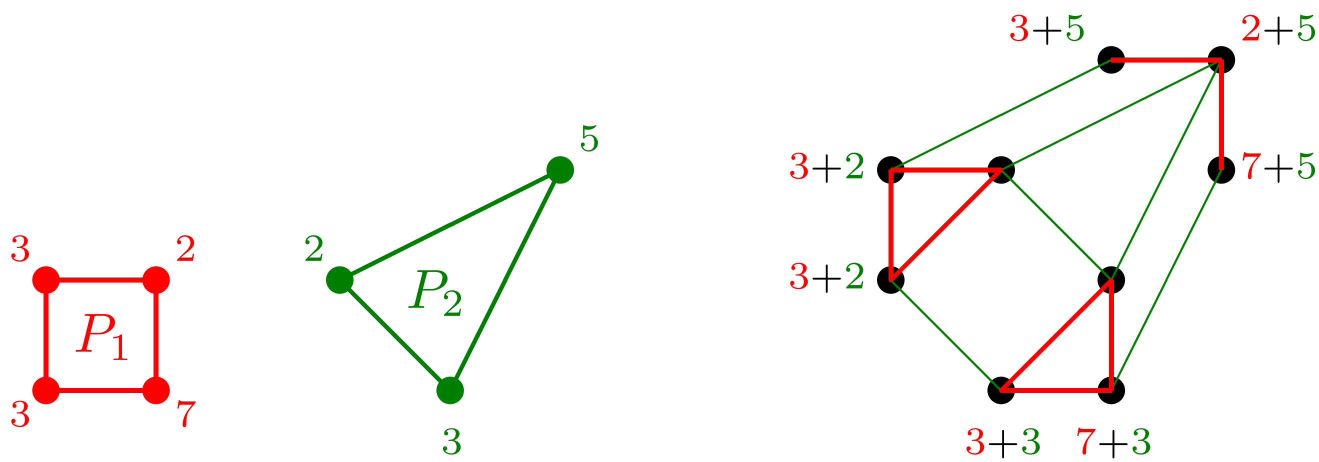
Applying the transformation $(x_1 = y_1 s^{v_1}, x_2 = y_2 s^{v_2}, t = s^{v_3})$, to a lifted monomial $\mathbf{x}^{\mathbf{a}} t^{l(\mathbf{a})}$ yields

$$\begin{aligned} x_1^{a_1} x_2^{a_2} t^{l(\mathbf{a})} &= (y_1 s^{v_1})^{a_1} (y_2 s^{v_2})^{a_2} (s^{v_3})^{l(\mathbf{a})} \\ &= y_1^{a_1} y_2^{a_2} s^{a_1 v_1 + a_2 v_2 + l(\mathbf{a}) v_3} \\ &= y_1^{a_1} y_2^{a_2} s^{\langle (\mathbf{a}, l(\mathbf{a})), \mathbf{v} \rangle}. \end{aligned}$$

A binomial system contains $\mathbf{x}^{\mathbf{a}} t^{l(\mathbf{a})}$ and $\mathbf{x}^{\mathbf{b}} t^{l(\mathbf{b})}$ if there exists an inner normal $\mathbf{v} \in \mathbb{Z}^3$, $v_3 > 0$, such that

$$\left\{ \begin{array}{l} \langle (\mathbf{a}, l(\mathbf{a})), \mathbf{v} \rangle = \langle (\mathbf{b}, l(\mathbf{b})), \mathbf{v} \rangle \\ \langle (\mathbf{a}, l(\mathbf{a})), \mathbf{v} \rangle < \langle (\mathbf{e}, l(\mathbf{e})), \mathbf{v} \rangle, \quad \forall \mathbf{e} \in A \setminus \{\mathbf{a}, \mathbf{b}\}. \end{array} \right.$$

A Regular Mixed Subdivision



Three mixed cells:

$$\begin{array}{lcc}
(\{(1,1,2),(1,0,7)\},\{(2,2,5),(1,0,3)\}) & \mathbf{v}=(-12,5,1) & V=1 \\
(\{(1,1,2),(0,1,3)\},\{(2,2,5),(0,1,2)\}) & \mathbf{v}=(1,-5,1) & V=1 \\
(\{(1,1,2),(0,0,3)\},\{(1,0,3),(0,1,2)\}) & \mathbf{v}=(0,1,1) & V=2 \\
\hline
\end{array}$$

$$\text{mixed volume} \;=\; \text{sum of volumes of mixed cells}: \quad V=4$$

More References on Polyhedral Methods

J.M. Rojas: **Toric Laminations, Sparse Generalized Characteristic Polynomials, and a Refinement of Hilbert's Tenth Problem**. In F. Cucker and M. Shub, editors, ***Foundations of Computational Mathematics***, pages 369–381, Springer-Verlag 1997.

B. Huber and B. Sturmfels: **Bernstein's theorem in affine space**. ***Discrete Comput. Geom.*** 17(2):137-141, 1997.

B. Huber and J. Verschelde: **Polyhedral end games for polynomial continuation**. ***Numerical Algorithms*** 18(1):91–108, 1998.

J.M. Rojas: **Toric intersection theory for affine root counting**. ***Journal of Pure and Applied Algebra*** 136(1):67–100, 1999.

T. Gao, T.Y. Li., and X. Wang: **Finding isolated zeros of polynomial systems in C^n with stable mixed volumes**. ***J. of Symbolic Computation*** 28(1-2):187–211, 1999.

Affine Roots

Mixed volumes count roots in $(\mathbb{C}^*)^n$, what about $\mathbb{C}^n$?

Consider

$$f(x,y) = \begin{cases} x + x^2 + x^3 y = 0 \\ y + y^2 + y^3 x = 0 \end{cases} \qquad V = 3 \text{ ignores} \qquad (0,a), (b,0), (0,0).$$

An idea: add artificial origins to the polytopes and first solve

$$g(x,y) = \begin{cases} c_1 + x + x^2 + x^3 y = 0 \\ c_2 + y + y^2 + y^3 x = 0 \end{cases} \qquad \text{where } c_1, c_2 \text{ are random.}$$

Then remove added constants with a homotopy.

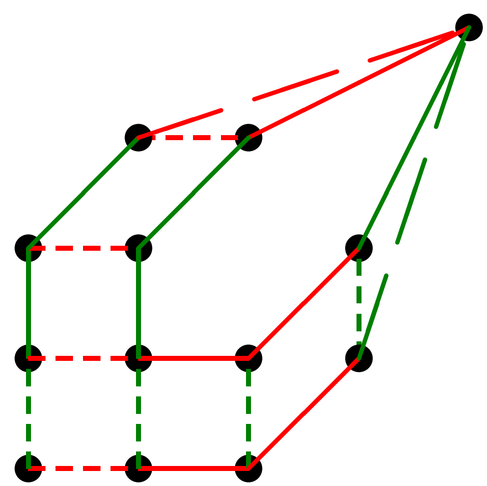
Drawback: too many diverging paths; $V = 8$ for $g(x,y) = \mathbf{0}$.

Stable Mixed Volumes

Lift artificial origins to height one, the rest remains at zero.

$$g(x,y) = \begin{cases} c_1 + x + x^2 + x^3 y = 0 & \quad A_1 = \{(0,0),(1,0),(2,0),(3,1)\} \\ c_2 + y + y^2 + y^3 x = 0 & \quad A_2 = \{(0,0),(0,1),(0,2),(1,3)\} \end{cases}$$

Observe a mixed subdivision:



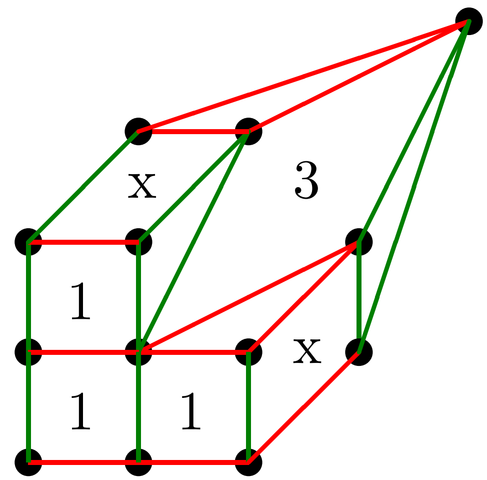
artificial origins introduced five new mixed cells, but only three are stable i.e.: converge to affine roots

A mixed cell is stable if all components of its inner normal $\mathbf{v}$ are ≥ 0. Recall: $x_i = y_i s^{v_i}$. For $v_i < 0$: $x_i \rightarrow \infty$ as $s \rightarrow 0$.

One Single Lifting

Lift the original points in the supports to heights in $[0,1]$ and the artificial origins sufficiently heigh.

One single lifting computes a stable mixed volume:



spurious cells marked with x

one cell with volume 3

3 extra stable mixed cells

each contributes one root

with zero components

Height of artificial origins: $n(n+1)d^n$, $d = \max_{i=1}^{n} \deg(f_i)$.

Software for Polyhedral Homotopies

J. Verschelde: **Algorithm 795: PHCpack: A general-purpose solver for polynomial systems by homotopy continuation.** ***ACM Trans. Math. Softw.***, 25(2):251–276, 1999.
Version 2.3.31 is available via `http://www.math.uic.edu/~jan`.

T. Gao, T.Y. Li, and M. Wu: **Algorithm 846: MixedVol: a software package for mixed-volume computation.**
ACM Trans. Math. Softw., 31(4):555–560, 2005.
HOM4PS is available via `http://www.math.msu.edu/~li`.

T. Gunji, S. Kim, M. Kojima, A. Takeda, K. Fujisawa, and T. Mizutani: **PHoM – a polyhedral homotopy continuation method for polynomial systems.** ***Computing***, 73(4):55–77, 2004.
Available via `http://www.is.titech.ac.jp/~kojima`.

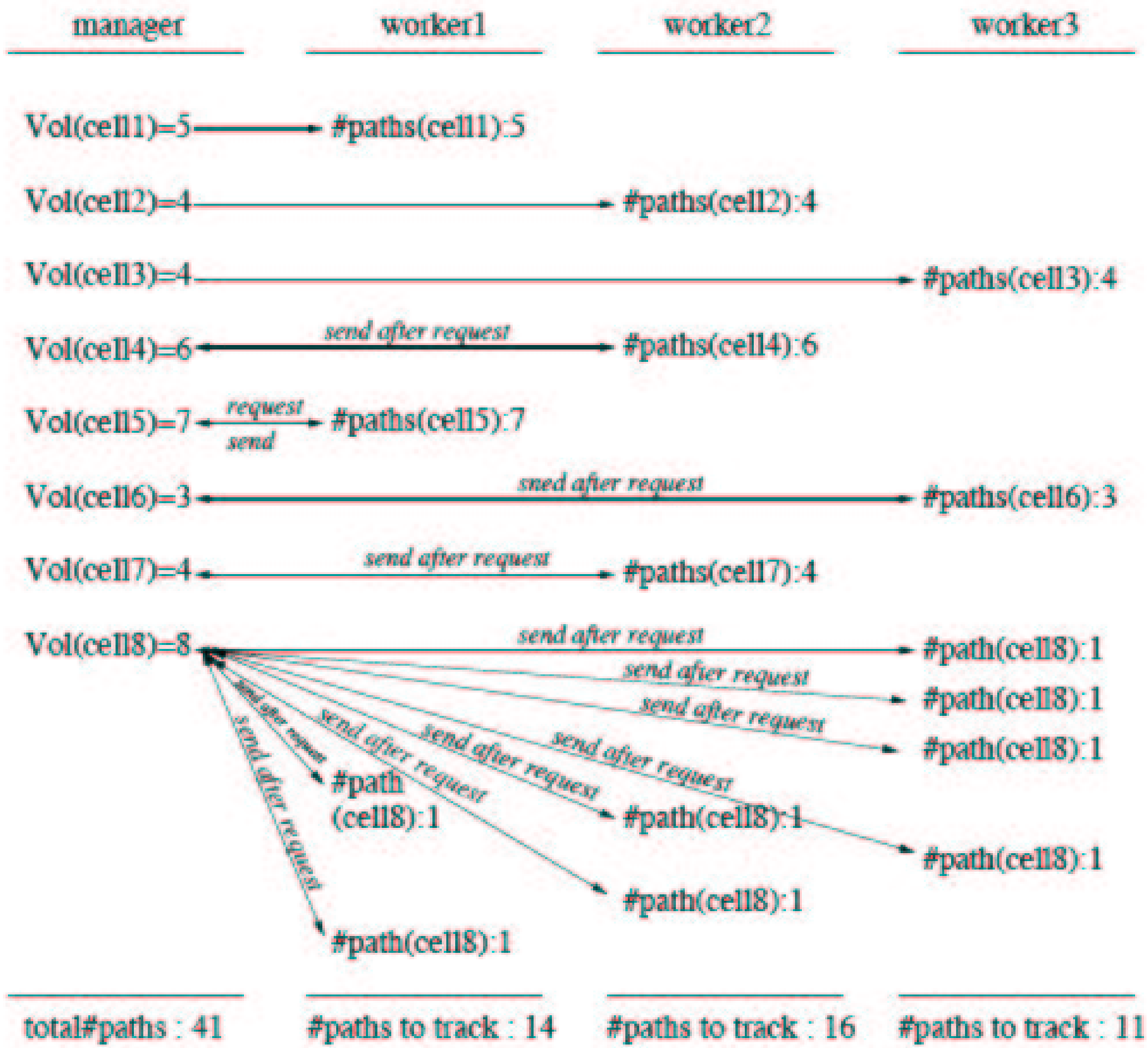
T. Gunji, S. Kim, K. Fujisawa, and M. Kojima: **PHoMpara – parallel implementation of the Polyhedral Homotopy continuation Method for polynomial systems.**
Computing 77(4):387–411, 2006.

Parallel Polyhedral Methods

joint with Yan Zhuang

- large systems from mechanical design
- dynamic load balancing achieves good speedup
- quality up still an issue...

Dynamic Load Balancing



An academic Benchmark: cyclic n-roots

The system

$$f(\mathbf{x}) = \left\{ \begin{array}{l} f_i = \displaystyle\sum_{j=0}^{n=1} \prod_{k=1}^{i} x_{(k+j) \bmod n} = 0, \quad i = 1, 2, \ldots, n-1 \\ f_n = x_0 x_1 x_2 \cdots x_{n-1} - 1 = 0 \end{array} \right.$$

appeared in

G. Björck: **Functions of modulus one on Z_p whose Fourier transforms have constant modulus** In *Proceedings of the Alfred Haar Memorial Conference, Budapest*, pages 193–197, 1985.

very sparse, well suited for polyhedral methods

Results on the cyclic n-roots problem

Problem	#Paths	CPU Time
cyclic 5-roots	70	0.13m
cyclic 6-roots	156	0.19m
cyclic 7-roots	924	0.30m
cyclic 8-roots	2,560	0.78m
cyclic 9-roots	11,016	3.64m
cyclic 10-roots	35,940	21.33m
cyclic 11-roots	184,756	2h 39m
cyclic 12-roots	500,352	24h 36m

Wall time for start systems to solve the cyclic n-roots problems, using 13 workers at 2.4Ghz, with static load distribution.

Dynamic versus Static Workload Distribution

#workers	Static versus Dynamic on our cluster				Dynamic on argo	
	Static	Speedup	Dynamic	Speedup	Dynamic	Speedup
1	50.7021	–	53.0707	–	29.2389	–
2	24.5172	2.1	25.3852	2.1	15.5455	1.9
3	18.3850	2.8	17.6367	3.0	10.8063	2.7
4	14.6994	3.4	12.4157	4.2	7.9660	3.7
5	11.6913	4.3	10.3054	5.1	6.2054	4.7
6	10.3779	4.9	9.3411	5.7	5.0996	5.7
7	9.6877	5.2	8.4180	6.3	4.2603	6.9
8	7.8157	6.5	7.4337	7.1	3.8528	7.6
9	7.5133	6.8	6.8029	7.8	3.6010	8.1
10	6.9154	7.3	5.7883	9.2	3.2075	9.1
11	6.5668	7.7	5.3014	10.0	2.8427	10.3
12	6.4407	7.9	4.8232	11.0	2.5873	11.3
13	5.1462	9.8	4.6894	11.3	2.3224	12.6

Wall time in seconds to solve a start system for the cyclic 7-roots problem.

Design of Serial Chains I

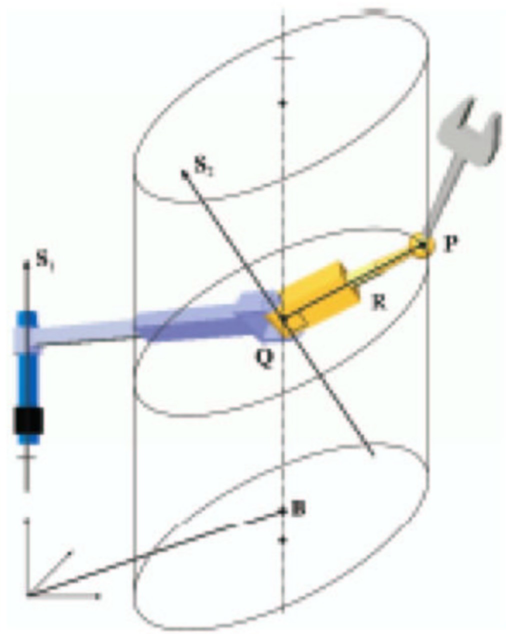


Figure 4.4: The elliptic cylinder reachable by a PRS serial chain.

H.J. Su. *Computer-Aided Constrained Robot Design Using Mechanism Synthesis Theory.* PhD thesis, University of California, Irvine, 2004.

Design of Serial Chains II

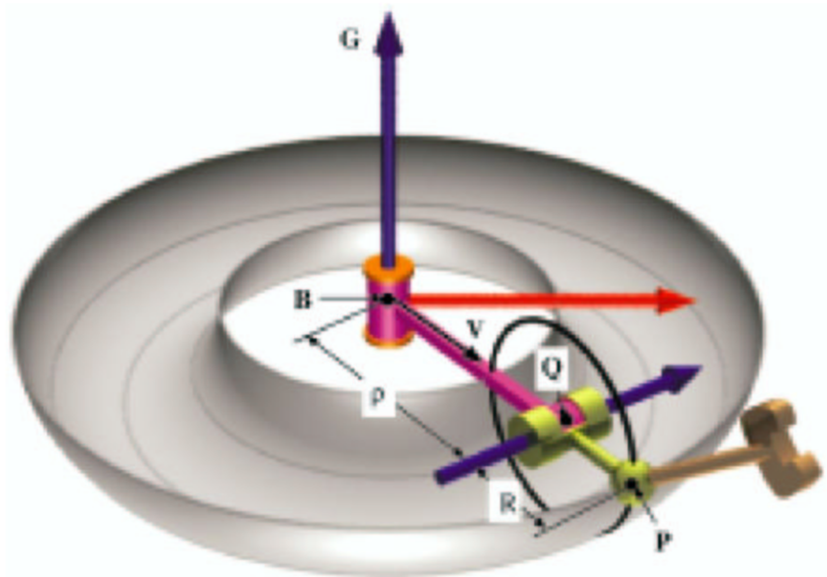


Figure 4.7: The circular torus traced by the wrist center of a “right” RRS serial chain.

H.J. Su. *Computer-Aided Constrained Robot Design Using Mechanism Synthesis Theory.* PhD thesis, University of California, Irvine, 2004.

Design of Serial Chains III

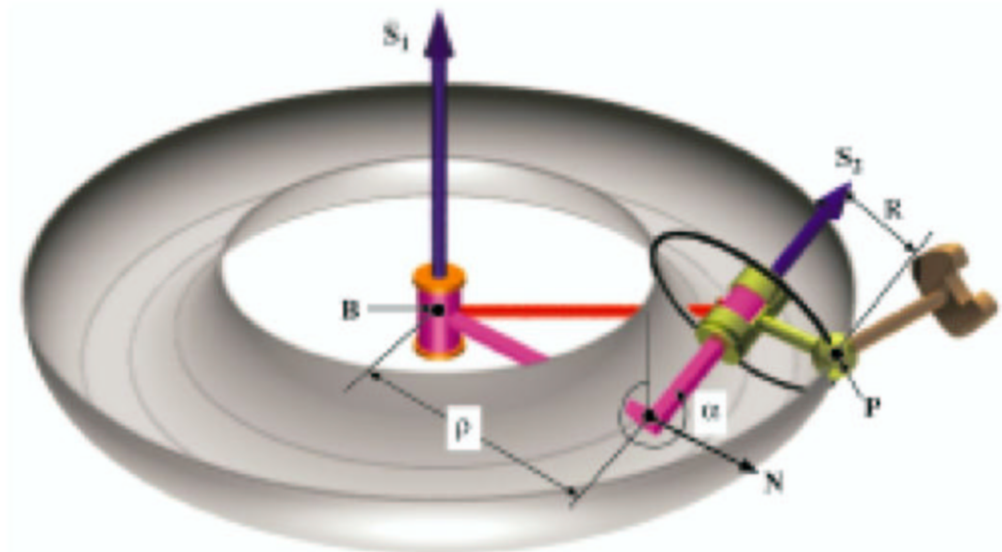


Figure 4.8: The general torus reachable by the wrist center of an RRS serial chain.

H.J. Su. *Computer-Aided Constrained Robot Design Using Mechanism Synthesis Theory*. PhD thesis, University of California, Irvine, 2004.

For more about these problems:

H.-J. Su and J.M. McCarthy: **Kinematic synthesis of RPS serial chains**. In the ***Proceedings of the ASME Design Engineering Technical Conferences*** (CDROM), Chicago, IL, Sep 2-6, 2003.

H.-J. Su, C.W. Wampler, and J.M. McCarthy: **Geometric design of cylindric PRS serial chains**. ***ASME Journal of Mechanical Design*** 126(2):269–277, 2004.

H.-J. Su, J.M. McCarthy, and L.T. Watson: **Generalized linear product homotopy algorithms and the computation of reachable surfaces**. ***ASME Journal of Information and Computer Sciences in Engineering*** 4(3):226–234, 2004.

H.-J. Su, J.M. McCarthy, M. Sosonkina, and L.T. Watson: **POLSYS_GLP: A parallel general linear product homotopy code for solving polynomial systems of equations**. ***ACM Trans. Math. Softw.*** 32(4):561-597, 2006.

Results on Mechanical Design Problems

Bézout vs Bernshteĭn

Surface	Bounds on #Solutions			Wall Time	
	D	B	V	our cluster	on argo
elliptic cylinder	2,097,152	247,968	125,888	11h 33m	6h 12m
circular torus	2,097,152	868,352	474,112	7h 17m	4h 3m
general torus	4,194,304	448,702	226,512	14h 15m	6h 36m

D = total degree; B = generalized Bézout bound; V = mixed volume

Wall time for mechanism design problems on our cluster and argo.

- Compared to the linear-product bound, polyhedral homotopies cut the #paths about in half.
- The second example is easier (despite the larger #paths) because of increased sparsity, and thus lower evaluation cost.