Polyhedral Homotopy Methods
to Solve Polynomial Systems

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

IMA postdoc seminar
31 October 2006
Polyhedral Homotopies

Theorems A & B of Bernshteĭn
  constructive proofs via deformations
  solutions at $\infty$ are solutions of face systems

Numerical Tools
  extrapolation to compute certificates of divergence
  separate $z$ from $\log(|z|)$ in a numerically stable simplicial solver

Applications from Mechanism Design
  design of serial chains, systems of H.-J. Su and J.M. McCarthy
Solving Systems with Homotopies

Concerns *(of anyone who tries to use numerical homotopies)*

1. efficiency: \( \# \text{paths} = \text{bound on } \# \text{solutions} \);
   
   how can we find good bounds on \( \# \text{solutions} \)?

2. validation: how can we be sure to have \textbf{all} solutions?
Solving Systems with Homotopies

Concerns  
(of anyone who tries to use numerical homotopies)

1. efficiency: \#paths = bound on \#solutions;  
   how can we find good bounds on \#solutions?

2. validation: how can we be sure to have \textbf{all} solutions?

Answers  
(why we should consider polyhedral methods)

1. generically sharp root counts,  
   which can be computed by fully automatic blackboxes

2. certificates for diverging paths,  
   which are cheap by-products of continuation
Geometric Root Counting

\[ f_i(x) = \sum_{a \in A_i} c_{i a} x^a \]
\[ c_{i a} \in \mathbb{C}^* = \mathbb{C} \setminus \{0\} \]
\[ f = (f_1, f_2, \ldots, f_n) \]
\[ P_i = \text{conv}(A_i) \]
Newton polytope
\[ \mathcal{P} = (P_1, P_2, \ldots, P_n) \]

<table>
<thead>
<tr>
<th>( L(f) ) root count in ((\mathbb{C}^*)^n)</th>
<th>desired properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>( L(f) = L(f_2, f_1, \ldots, f_n) )</td>
<td>invariant under permutations</td>
</tr>
<tr>
<td>( L(f) = L(f_1 x^a, \ldots, f_n) )</td>
<td>shift invariant</td>
</tr>
<tr>
<td>( L(f) \leq L(f_1 + x^a, \ldots, f_n) )</td>
<td>monotone increasing</td>
</tr>
<tr>
<td>( L(f) = L(f_1(x^{U a}), \ldots, f_n(x^{U a})) )</td>
<td>unimodular invariant</td>
</tr>
<tr>
<td>( L(f_{11} f_{12}, \ldots, f_n) )</td>
<td>root count of product</td>
</tr>
<tr>
<td>( = L(f_{11}, \ldots, f_n) + L(f_{12}, \ldots, f_n) )</td>
<td>is sum of root counts</td>
</tr>
</tbody>
</table>
Geometric Root Counting

\[ f_i(x) = \sum_{a \in A_i} c_{ia} x^a \]
\[ c_{ia} \in \mathbb{C}^* = \mathbb{C} \setminus \{0\} \]
\[ f = (f_1, f_2, \ldots, f_n) \]

\[ P_i = \text{conv}(A_i) \]

Newton polytope

\[ P = (P_1, P_2, \ldots, P_n) \]

<table>
<thead>
<tr>
<th>properties of ( L(f) )</th>
<th>( V(P) ) mixed volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>invariant under permutations</td>
<td>( V(P_2, P_1, \ldots, P_n) = V(P) )</td>
</tr>
<tr>
<td>shift invariant</td>
<td>( V(P_1 + a, \ldots, P_n) = V(P) )</td>
</tr>
<tr>
<td>monotone increasing</td>
<td>( V(\text{conv}(P_1 + a), \ldots, P_n) \geq V(P) )</td>
</tr>
<tr>
<td>unimodular invariant</td>
<td>( V(UP_1, \ldots, UP_n) = V(P) )</td>
</tr>
</tbody>
</table>
| 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

\[ f_i(x) = \sum_{a \in A_i} c_{ia}x^a \]
\[ c_{ia} \in \mathbb{C}^* = \mathbb{C} \setminus \{0\} \]
\[ f = (f_1, f_2, \ldots, f_n) \]
\[ P_i = \text{conv}(A_i) \]
\[ \mathcal{P} = (P_1, P_2, \ldots, P_n) \]

<table>
<thead>
<tr>
<th>( L(f) ) root count in ((\mathbb{C}^*)^n)</th>
<th>( V(\mathcal{P}) ) mixed volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>( L(f) = L(f_2, f_1, \ldots, f_n) )</td>
<td>( V(P_2, P_1, \ldots, P_n) = V(\mathcal{P}) )</td>
</tr>
<tr>
<td>( L(f) = L(f_1x^a, \ldots, f_n) )</td>
<td>( V(P_1 + a, \ldots, P_n) = V(\mathcal{P}) )</td>
</tr>
<tr>
<td>( L(f) \leq L(f_1 + x^a, \ldots, f_n) )</td>
<td>( V(\text{conv}(P_1 + a), \ldots, P_n) \leq V(\mathcal{P}) )</td>
</tr>
<tr>
<td>( L(f) = L(f_1(x^Ua), \ldots, f_n(x^Ua)) )</td>
<td>( V(UP_1, \ldots, UP_n) = V(\mathcal{P}) )</td>
</tr>
<tr>
<td>( L(f_{11}f_{12}, \ldots, f_n) )</td>
<td>( V(P_{11} + P_{12}, \ldots, P_n) )</td>
</tr>
<tr>
<td>= ( L(f_{11}, \ldots, f_n) + L(f_{12}, \ldots, f_n) )</td>
<td>( = V(P_{11}, \ldots, P_n) + V(P_{12}, \ldots, P_n) )</td>
</tr>
</tbody>
</table>

exploit sparsity \( L(f) = V(\mathcal{P}) \) 1st theorem of Bernshtein
The Theorems of Bernshteǐn

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.


*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 on Polyhedral Methods


**Bernshteĭn’s second theorem**

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

$$\partial_\omega f_i(x) = \sum_{a \in \partial_\omega A_i} c_{ia} x^a \quad \partial_\omega P_i = \text{conv}(\partial_\omega A_i)$$

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

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

- Newton polytopes *in general position*: $V(P)$ is exact for every nonzero choice of the coefficients.
Consider $f(x) = \begin{cases} c_{111}x_1x_2 + c_{110}x_1 + c_{101}x_2 + c_{100} = 0 \\ c_{222}x_1^2x_2^2 + c_{210}x_1 + c_{201}x_2 = 0 \end{cases}$

The Newton polytopes:

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

$P_1 = \text{conv}(A_1)$  \hspace{1cm}  $P_2 = \text{conv}(A_2)$
Consider \( f(x) = \begin{cases} c_{111}x_1x_2 + c_{110}x_1 + c_{101}x_2 + c_{100} = 0 \\ c_{222}x_1^2x_2^2 + c_{210}x_1 + c_{201}x_2 = 0 \end{cases} \)

Look at the inner normals of \( P_2 \):

\[ P_1 \]

\[ P_2 \]

\((-2, 1)\)

→ the corresponding face system \( \begin{cases} c_{110}x_1 = 0 \\ c_{222}x_1^2x_2^2 + c_{210}x_1 = 0 \end{cases} \) does not have a solution in \((\mathbb{C}^*)^2\).
Consider \( f(x) = \begin{cases} c_{111}x_1x_2 + c_{110}x_1 + c_{101}x_2 + c_{100} = 0 \\ c_{222}x_1^2x_2^2 + c_{210}x_1 + c_{201}x_2 = 0 \end{cases} \)

Look at the inner normals of \( P_2: \) \((1, -2)\)

\( \rightarrow \) the corresponding face system

\( \rightarrow \) does not have a solution in \((\mathbb{C}^*)^2\).
Consider \( f(x) = \begin{cases} c_{111}x_1x_2 + c_{110}x_1 + c_{101}x_2 + c_{100} = 0 \\ c_{222}x_1^2x_2^2 + c_{210}x_1 + c_{201}x_2 = 0 \end{cases} \)

Look at the inner normals of \( P_2 \):

\[ \text{the corresponding face system} \begin{cases} c_{100} = 0 \\ c_{210}x_1 + c_{201}x_2 = 0 \end{cases} \]

does not have a solution in \((\mathbb{C}^*)^2\).
Newton polytopes in general position

Consider $f(x) = \begin{cases} c_{111}x_1x_2 + c_{110}x_1 + c_{101}x_2 + c_{100} = 0 \\ c_{222}x_1^2x_2^2 + c_{210}x_1 + c_{201}x_2 = 0 \end{cases}$

Look at the inner normals of $P_2$:

$\forall \omega \neq 0: \partial_\omega A_1 + \partial_\omega A_2 \leq 3 \implies V(P_1, P_2) = 4$ always exact for all nonzero coefficients
B. the 2nd theorem

**Power Series**

**Theorem:** \( \forall x(t), H(x(t), t) = (1 - t)g(x(t)) + tf(x(t)) = 0, \)

\[ \exists s > 0, m \in \mathbb{N} \setminus \{0\}, \omega \in \mathbb{Z}^n: \]

\[ \begin{cases} 
    x_i(s) = b_i s^\omega_i (1 + O(s)), & i = 1, 2, \ldots, n \\
    t(s) = 1 - s^m & \text{for } t \approx 1, s \approx 0
\end{cases} \]

\[ \lim_{t \to 1} x_i(t) \in \mathbb{C}^*? \quad x_i(t) \begin{cases} 
    \to \infty & < 0 \\
    \in \mathbb{C}^* & \iff \omega_i = 0 \\
    \to 0 & > 0
\end{cases} \]

\( m \) is the *winding number*, i.e.: the smallest number so that

\[ z(2\pi m) = z(0), \quad H(z(\theta), t(\theta)) = 0, \quad t = 1 + (t_0 - 1)e^{i\theta}, \quad t_0 \approx 1. \]
B. the 2nd theorem

**Face Systems and Power Series**

assume $\lim_{t \to 1} x_i(t) \not\in \mathbb{C}^*$, thus $\omega_i \neq 0$, a **diverging** path

- $H(x, t) = (1 - t)g(x) + tf(x) = 0$
  
  substitute power series

$$H(x(s), t(s)) = f(x(s)) + s^m(g(x(s)) - f(x(s))) = 0$$

  dominant as $s \to 0$

  only $f$ matters
Face Systems and Power Series

Assume \( \lim_{t \to 1} x_i(t) \notin \mathbb{C}^*, \) thus \( \omega_i \neq 0, \) a **diverging** path.

- \( H(x, t) = (1 - t)g(x) + tf(x) = 0 \) substitute power series
  \[
  H(x(s), t(s)) = \underbrace{f(x(s))}_{\text{dominant as } s \to 0} + s^m(g(x(s)) - f(x(s))) = 0
  \]

- \( f_i(x) = \sum_{a \in A_i} c_i a x^a \to f_i(x(s)) = \sum_{a \in A_i} c_i a \prod_{i=1}^{n} b_i^{a_i} s^{\iang{a}{\omega}} (1 + O(s)) \)

  face \( \partial_\omega A_i := \{ a \in A_i \mid \iang{a}{\omega} = \min_{a' \in A_i} \iang{a'}{\omega} \} \)

  \( \Rightarrow \partial_\omega f(b) = 0, b \in (\mathbb{C}^*)^n \)

  solution at \( \infty \) is a solution of a face system.
Richardson Extrapolation for $\omega$ and $m$

\[
\begin{aligned}
  x_i(s) &= b_is^{\omega_i}(1 + O(s)) \\
  t(s) &= 1 - s^m \\
  x_i(s_k) &= b_ih^{k\omega_i/m}s_0(1 + O(h^{k/m}s_0))
\end{aligned}
\]

Geometric sampling $0 < h < 1$

\[
\begin{aligned}
  1 - t_k &= h(1 - t_k) = \ldots = h^k(1 - t_0) \\
  s_k &= h^{1/m}s_{k-1} = \ldots = h^{k/m}s_0
\end{aligned}
\]

Input: $(x(s), t(s))$ solutions along a path, $H(x(s), t(s)) = 0$.

Output: approximations for $\omega$ and $m$. 
Richardson Extrapolation for $\omega$ and $m$

\[
\begin{align*}
\begin{cases}
    x_i(s) &= b_is^{\omega_i}(1 + O(s)) \\
    t(s) &= 1 - s^m \\
    x_i(s_k) &= b_i h^{k\omega_i/m} s_0(1 + O(h^{k/m} s_0))
\end{cases}
\end{align*}
\]

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:

- \[
\log |x_i(s_k)| = \log |b_i| + \frac{k\omega_i}{m} \log(h) + \omega_i \log(s_0)
\]
- \[
+ \log(1 + \sum_{j=0}^{\infty} b_j^i (h^{k/m} s_0)^j)
\]

\[
v_{kk+1} := \log |x_i(s_k + 1)| - \log |x_i(s_k)|
\]

\[
\omega_i = m \frac{v_{0..r}}{\log(h)} + O(s_0^r)
\]

→ first-order approximation for $\omega$

... is okay for $m = 1$
B. the 2nd theorem

Richardson Extrapolation for $\omega$ and $m$

\[
\begin{align*}
\left\{ \begin{array}{l}
x_i(s) = b_i s^{\omega_i} (1 + O(s)) \\
t(s) = 1 - s^m
\end{array} \right.
\]

\[x_i(s_k) = b_i h^{k\omega_i/m} s_0 (1 + O(h^{k/m} s_0))\]

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

- \[\nu_{kk+1} := \log |x_i(s_k + 1)| - \log |x_i(s_k)|\]

- \[e_i^{(k)} = (\log |x_i(s_k)| - \log |x_i(s_{k+1})|)\]
  \[-(\log |x_i(s_{k+1})| - \log |x_i(s_{k+2})|)\]
  \[= c_1 h^{k/m} s_0 (1 + 0(h^{k/m}))\]

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

Geometric sampling $0 < h < 1$

- \[1 - t_k = h(1 - t_k) = \cdots = h^k (1 - t_0)\]
  \[s_k = h^{1/m} s_{k-1} = \cdots = h^{k/m} s_0\]

Extrapolation on samples

- \[v_{k..l} = v_{k..l-1} + \frac{v_{k+1..l} - v_{k..l-1}}{1-h}\]
  \[\omega_i = m \frac{v_0 - r}{\log(h)} + O(s_0^r)\]

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

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

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

\[
\partial_{(0,0,0,-1)} f(b, c, d, e) = \begin{cases} 
-8b^2c^2e - 28b^2cde + 36b^2d^2e &= 0 &= -2c^2 - 7cd + 9d^2 \\
-32de^2c + 16d^2e^2 + 16e^2c^2 &= 0 &= -2dc + d^2 + c^2 \\
-80de^2c + 40d^2e^2 + 40e^2c^2 &= 0 &= -2dc + d^2 + c^2 \\
22ec - 22de &= 0 &= c - d
\end{cases}
\]
More References on Polyhedral Methods


Bernshtein’s first theorem

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

I. Compute $V_n(\mathcal{P})$:
   I.1 lift polytopes $\Leftrightarrow$ I.2 mixed cells $\Leftrightarrow$ I.3 volume of mixed cell

II. Solve $g(x) = 0$:
   II.1 introduce parameter $t$ $\Leftrightarrow$ II.2 start systems $\Leftrightarrow$ II.3 path following

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

$$H(x, t) = (1 - t)g(x) + tf(x) = 0,$$ for $t$ from 0 to 1.

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

C. the 1st theorem
Finding Mixed Cells

\[ g(x_1, x_2, t) = \begin{cases} 
  c_{111}x_1x_2t^2 + c_{110}x_1t^7 + c_{101}x_2t^3 + c_{100}t^3 = 0 \\
  c_{222}x_1^2x_2^2t^5 + c_{210}x_1t^3 + c_{201}x_2t^2 = 0 
\end{cases} \]

- At \( t = 1 \): the system \( g(x, 1) = g(x) = 0 \) we want to solve.
- Where to start?
  → look for inner normals \( \omega \in \mathbb{Z}^3, \omega_3 > 0 \), such that after

\[
x_1 = y_1 s^{\omega_1}, \quad x_2 = y_2 s^{\omega_2}, \quad t = s^{\omega_3},
\]

the system \( g(y, s) = 0 \) has solutions in \((\mathbb{C}^*)^2\) at \( s = 0 \).
C. the 1st theorem

Coordinate Transformations give Homotopies

\[ g(x_1 = y_1, x_2 = y_2 s, t = s) \]

\[ = \begin{cases} 
  c_{111}y_1(y_2 s)^2 + c_{110}y_1 s^7 + c_{101}(y_2 s)^3 + c_{100}s^3 = 0 \\
  c_{222}y_1^2(y_2 s)^2 s^5 + c_{210}y_1 s^3 + c_{201}(y_2 s)s^2 = 0 
\end{cases} \]

\[ = \begin{cases} 
  c_{111}y_1 y_2 s^3 + c_{110}y_1 s^7 + c_{101}y_2 s^4 + c_{100}s^3 = 0 \\
  c_{222}y_1^2 y_2^2 s^7 + c_{210}y_1 s^3 + c_{201}y_2 s^3 = 0 
\end{cases} \]

\[ = \begin{cases} 
  c_{111}y_1 y_2 + c_{110}y_1 s^4 + c_{101}y_2 s + c_{100} = 0 \\
  c_{222}y_1^2 y_2^2 s^4 + c_{210}y_1 + c_{201}y_2 = 0 
\end{cases} \]

At \( s = 0 \) we find a binomial system which has two solutions. The two solutions extend to solutions of \( g(x) = g(x, s = 1) = 0 \).
Coordinate Transformations and Inner Normals

Applying the transformation \( (x_1 = y_1 s^{\omega_1}, x_2 = y_2 s^{\omega_2}, t = s^{\omega_3}) \), to a lifted monomial \( x^a t^l(a) \) yields

\[
x_1^{a_1} x_2^{a_2} t^l(a) = (y_1 s^{\omega_1})^{a_1} (y_2 s^{\omega_2})^{a_2} (s^{\omega_3})^{l(a)} = y_1^{a_1} y_2^{a_2} s^{a_1 \omega_1 + a_2 \omega_2 + a_3 l(a)} = y_1^{a_1} y_2^{a_2} s^{\langle (a, l(a)), \omega \rangle}.
\]

A binomial system contains \( x^a t^l(a) \) and \( x^b t^l(b) \) if there exists an inner normal \( \omega \in \mathbb{Z}^3, \omega_3 > 0 \), such that

\[
\begin{align*}
\langle (a, l(a)), \omega \rangle &= \langle (b, l(b)), \omega \rangle \\
\langle (a, l(a)), \omega \rangle &< \langle (e, l(e)), \omega \rangle, \quad \forall e \in A \setminus \{a, b\}.
\end{align*}
\]
A Regular Mixed Subdivision

Three mixed cells:

1. \(((1, 1, 2), (1, 0, 7)), \{(2, 2, 5), (1, 0, 3)\}\) \(\omega = (-12, 5, 1)\) \(V = 1\)
2. \(((1, 1, 2), (0, 1, 3)), \{(2, 2, 5), (0, 1, 2)\}\) \(\omega = (1, -5, 1)\) \(V = 1\)
3. \(((1, 1, 2), (0, 0, 3)), \{(1, 0, 3), (0, 1, 2)\}\) \(\omega = (0, 1, 1)\) \(V = 2\)
Algorithms and Software for Mixed Volumes


A well conditioned polynomial system

just one of the 11,417 start systems generated by polyhedral homotopies

12 equations, 13 distinct monomials (after division):

\[
\begin{align*}
\theta \\
b_1x_5x_8 + b_2x_6x_9 &= 0 \\

b_3x_2^2 + b_4 &= 0 \\
b_5x_1x_4 + b_6x_2x_5 &= 0 \\
c_1^{(k)}x_1x_4x_7x_{12} + c_2^{(k)}x_1x_6x_{10} + c_3^{(k)}x_2x_4x_8x_{10} + c_4^{(k)}x_2x_4x_{11}^2 + c_5^{(k)}x_2x_6x_8x_{11} + c_6^{(k)}x_3x_4x_9x_{10} + c_7^{(k)}x_4x_{12} + c_8^{(k)}x_3x_6 + c_9^{(k)}x_4^2 + c_{10}^{(k)}x_9 &= 0, \quad k = 1, 2, \ldots, 9 \\
\end{align*}
\]

Random coefficients chosen on the complex unit circle.

Despite the high degrees, only 100 well conditioned solutions.
Solve a “binomial” system $x^A = b$ via Hermite

**Hermite normal form** of $A$: $MA = U$, $\det(M) = \pm 1$,

$U$ is upper triangular, $|\det(U)| = |\det(A)| = \#solutions$.

Let $x = z^M$, then $x^A = z^{MA} = z^U$, so solve $z^U = b$.

$n = 2$:

$$
\begin{bmatrix}
  u_{11} & u_{12} \\
  0 & u_{22}
\end{bmatrix}
\begin{bmatrix}
  z_1 \\
  z_2
\end{bmatrix}
= \begin{bmatrix}
  b_1 \\
  b_2
\end{bmatrix}.
$$

\[
\begin{cases}
  z_1^{u_{11}} = b_1 & |b_k| = 1 \Rightarrow |z_i| = 1 \\
  z_1^{u_{12}} z_2^{u_{22}} = b_2 & \text{numerically well conditioned}
\end{cases}
\]
Reduce a “simplicial” system $Cx^A = b$ via LU

\[ C = LU \quad \Rightarrow \quad (1) \quad LUy = b \quad \text{linear system} \]

\[ \text{assume } \det(C) \neq 0 \quad (2) \quad x^A = y \quad \text{binomial system} \]

This is a numerically unstable algorithm!

Randomly chosen coefficients for $C$ and $b$ on complex unit circle,
but still, varying magnitudes in $y$ do occur.

High powers, e.g.: 50, magnify the imbalance

$\rightarrow$ numerical underflow or overflow in binomial solver.
Separate Magnitudes from Angles

Solve \( x^A = y \) via Hermite: \( MA = U \Rightarrow x = z^M : z^U = y \).

\[ z = |z|e_z, \quad e_z = \exp(i\theta_z), \quad y = |y|e_y, \quad e_y = \exp(i\theta_y), \quad i = \sqrt{-1}. \]

Solve \( z^U = y \): \( |z|^U e_z^U = |y|e_y \leftrightarrow \begin{cases} e_z^U = e_y & \text{well conditioned} \\ |z|^U = |y| & \text{find magnitudes} \end{cases} \)

To solve \( |z|^U = |y| \), consider: \( U \log(|z|) = \log(|y|) \).

Even as the magnitude of the values \( y \) may be extreme, \( \log(|y|) \) will be modest in size.
We solve \( Cx^A = b \) by

1. LU factorization of \( C \rightarrow x^A = y \), where \( Cy = b \).

2. Use Hermite normal form of \( A \): \( MA = U \), \( \det(M) = \pm 1 \),
   to solve binomial system \( e^U_z = e_y \), \( z = |z|e_z \), \( y = |y|e_y \).

3. Solve upper triangular linear system \( U \log(|z|) = \log(|y|) \).

4. Compute magnitude of \( x = z^M \) via \( \log(|x|) = M \log(|z|) \).

5. As \( |e_z| = 1 \), let \( e_x = e^M_z \).

Even as \( z \) may be extreme, we deal with \( |z| \) at a logarithmic scale
and never raise small or large number to high powers.

Only at the very end do we calculate \( |x| = 10^{\log(|x|)} \) and \( x = |x|e_x \).
E. applications

**Design of Serial Chains I**

![Diagram of a PRS serial chain]

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

Design of Serial Chains II

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

Design of Serial Chains III

For more about these problems:


### Results on Mechanical Design Problems

**joint with Yan Zhuang**

<table>
<thead>
<tr>
<th>Surface</th>
<th>Bounds on #Solutions</th>
<th>Wall Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bézout</td>
<td>linear-product</td>
</tr>
<tr>
<td>elliptic cylinder</td>
<td>2,097,152</td>
<td>247,968</td>
</tr>
<tr>
<td>circular torus</td>
<td>2,097,152</td>
<td>868,352</td>
</tr>
<tr>
<td>general torus</td>
<td>4,194,304</td>
<td>448,702</td>
</tr>
</tbody>
</table>

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.