Polyhedral Homotopies

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

e-mail: jan@math.uic.edu
web: www.math.uic.edu/~jan

CIMPA Summer School, Buenos Aires, Argentina
22 July 2003
Plan of the Lecture

1. Geometric Root Counting  
   *why consider mixed volumes?*

2. The Theorems of Bernshtein  
    *sharp root counts + deficiency criterion*

3. Mixed Volumes  
    *mixed subdivisions visualize Minkowski’s theorem*

4. Polyhedral End Games  
    *finding certificates for divergence*

5. Polyhedral Continuation  
    *solving sparse system in two stages*

6. Software and Applications  
    *outline of blackbox solver*
Recommended Background Literature


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 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

\[ f_i(x) = \sum_{a \in A_i} c_{ia} x^a \]

\[ c_{ia} \in \mathbb{C}^* = \mathbb{C} \setminus \{0\} \]

\[ f = (f_1, f_2, \ldots, f_n) \]

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

Newton polytope

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

<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_1 x^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) \geq V(\mathcal{P}) )</td>
</tr>
<tr>
<td>( L(f) = L(f_1(x^U a), \ldots, f_n(x^U a)) )</td>
<td>( V(U P_1, \ldots, U P_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.
Systems, Supports, and Newton Polytopes

\[ f = (f_1, f_2) \]
\[ = \begin{cases} 
  x_1^3 x_2 + x_1 x_2^2 + 1 = 0 \\
  x_1^4 + x_1 x_2 + 1 = 0 
\end{cases} \]
\[ A = (A_1, A_2) \]
\[ A_1 = \{(3, 1), (1, 2), (0, 0)\} \]
\[ A_2 = \{(4, 0), (1, 1), (0, 0)\} \]

The sparse structure of \( f \) is modeled by the tuple \( A = (A_1, A_2) \).

\( A_1 \) and \( A_2 \) are the supports of \( f_1 \) and \( f_2 \) respectively.

The Newton polytopes are the convex hulls of the supports.
The Cayley polytope

Place one polytope at level 0, the other at level 1.
A triangulation of the Cayley polytope
A mixed subdivision induced by a triangulation of the Cayley polytope
Mixed Volumes

Mixed subdivisions visualize Minkowski’s theorem:

\[
\text{area}(\lambda_1 P_1 + \lambda_2 P_2) = V(P_1, P_1)\lambda_1^2 + 2V(P_1, P_2)\lambda_1 \lambda_2 + V(P_2, P_2)\lambda_2^2 \\
= 5\lambda_1^2 + 2 \times 8\lambda_1 \lambda_2 + 5\lambda_2^2
\]
Newton Polytopes and Real Solutions


**Bernshtein’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 $\mathcal{P} = (P_1, P_2, \ldots, P_n)$ and mixed volume $V(\mathcal{P})$.

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

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

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

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

The Newton polytopes:

\[ V(P_1, P_2) = 4 \text{ always exact} \]

for all nonzero coefficients
Power Series

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

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

\[
\begin{align*}
  x_i(s) &= b_is^{\omega_i}(1 + O(s)), \quad i = 1, 2, \ldots, n \\
  t(s) &= 1 - s^m \\
\end{align*}
\]

for \( t \approx 1, s \approx 0 \)

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

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

\( z(2\pi m) = z(0), \quad h(z(\theta), t(\theta)) = 0, \quad t = 1 + (t_0 - 1)e^{i\theta}, \quad t_0 \approx 1. \)
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
  
  $h(x(s), t(s)) = f(x(s)) + s^m(g(x(s)) - f(x(s))) = 0$

  dominant as $s \to 0$

- $f_i(x) = \sum_{a \in A_i} c_{ia}x^a \to f_i(x(s)) = \sum_{a \in A_i} c_{ia} \prod_{i=1}^{n} b_i^{a_i} s^{\langle a, \omega \rangle} (1 + O(s))$

  $\partial_\omega f_i(x(s))$ dominant

  face $\partial_\omega A_i := \{ a \in A_i \mid \langle a, \omega \rangle = \min_{a' \in A_i} \langle a', \omega \rangle \}$

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

  key idea in proof of Berenshtein’s second theorem
Richardson Extrapolation for $\omega$ and $m$

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

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

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

Geometric sampling \(0 < h < 1\)
\[
\begin{align*}
1 - t_k &= h(1 - t_k) = \cdots = h^k (1 - t_0) \\
1 - t_{k+1} &= h^{1 / m} s_{k-1} = \cdots = h^{k / m} s_0 
\end{align*}
\]

Extrapolation on samples
\[
\begin{align*}
v_{kk+1} &= v_k - 1 + \frac{v_{k+1} - v_k}{1 - h} \\
\omega_i &= m \frac{v_0}{\log(h)} + O(s_0^r) 
\end{align*}
\]

Extrapolation on errors
\[
\begin{align*}
e_i^{(k..l)} &= e_i^{(k+1..l)} + \frac{e_i^{(k..l-1)} - e_i^{(k+1..l)}}{1-h_{k..l}} \\
h_{kk..l} &= h^{(l-k-1) / m_{k..l}} \\
\log(h) &= \frac{\log(h)}{e_i^{(k..l)}} + O(h (l-k) k / m) 
\end{align*}
\]
the system of Cassou-Noguès

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

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

\[ \partial_{(0,0,0,-1)} f(b, c, d, e) = \]
\[ m = 2 \]
\[
\begin{align*}
-8b^2c^2e - 28b^2cde + 36b^2d^2e &= 0 \\
-32de^2c + 16d^2e^2 + 16e^2c^2 &= 0 \\
-80de^2c + 40d^2e^2 + 40e^2c^2 &= 0 \\
22ec - 22de &= 0
\end{align*}
\]
Some further recommended reading

B. Huber and J. Verschelde: **Polyhedral end games for polynomial continuation.**

J. Verschelde: **Toric Newton Method for Polynomial Homotopies.**
Sparsity and Unimodular Transformations

\[ f(\mathbf{x}) = \begin{cases} \mathbf{x}_1^3 \mathbf{x}_2^{-1} + c_1 = 0 \\ \mathbf{x}_1 \mathbf{x}_2^2 + c_2 = 0 \end{cases} \] \[ f(\mathbf{x} = \mathbf{y}^U) = \begin{cases} y_1 + c_1 = 0 \\ y_1^{-2} y_2^7 + c_2 = 0 \end{cases} \]

The substitution \( \mathbf{x}^V = (\mathbf{y}^U)^V = \mathbf{y}^{VU} = \mathbf{y}^L \) is elaborated as

\[
\begin{pmatrix} x_1^3 \cdot x_2^{-1} \\ x_1 \cdot x_2^2 \end{pmatrix} = \begin{pmatrix} (y_1^0 y_2^1)^3 \cdot (y_1^{-1} y_2^3)^{-1} \\ (y_1^0 y_2^1)^1 \cdot (y_1^{-1} y_2^3)^2 \\ y_1^{3 \cdot 0-1 \cdot (-1)} \cdot y_2^{3 \cdot 1-1 \cdot 3} \\ y_1^{1 \cdot 0+2 \cdot (-1)} \cdot y_2^{1 \cdot 1+2 \cdot 3} \end{pmatrix} = \begin{pmatrix} y_1^1 \cdot y_2^0 \\ y_1^{-2} \cdot y_2^7 \end{pmatrix}.
\]

Factorization \( VU = L \):

\[
\begin{bmatrix} 3 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -2 & 7 \end{bmatrix}
\]

\( (U \text{ unimodular, } \det(U)=1) \)
Pivoting to update a triangulation

\( c_0 = (0, 0) \quad x = (2, 3) \quad x = +\frac{1}{8}c_0 + \frac{1}{4}c_1 + \frac{5}{8}c_2 \)

\( c_1 = (3, 2) \quad y = (5, 1) \quad y = -\frac{1}{3}c_0 + \frac{9}{4}c_1 - \frac{7}{8}c_2 \)

\( c_2 = (2, 4) \quad z = (1, 5) \quad z = +\frac{1}{8}c_0 - \frac{3}{4}c_1 + \frac{13}{8}c_2 \)

barycentric coordinates
Incremental Polyhedral Continuation

\[ g(x) = \left\{ \begin{array}{l}
    c_{111}x_1x_2 + c_{110}x_1 + c_{101}x_2 + c_{100} = 0 \\
    c_{211}x_1x_2 + c_{210}x_1 + c_{201}x_2 + c_{200} = 0
\end{array} \right. \]

\[ g_1(x, t) = \left\{ \begin{array}{l}
    c_{111}x_1x_2 + c_{110}x_1 + c_{101}x_2t + c_{100} = 0 \\
    c_{211}x_1x_2 + c_{210}x_1 + c_{201}x_2t + c_{200} = 0
\end{array} \right. \]

\[ v_1 = (0, 0, 1) \]

\[ x_1 = \tilde{x}_1 t \quad x_2 = \tilde{x}_2 t^{-1} \]

\[ v_2 = (1, -1, 1) \]

\[ g_2(\tilde{x}, t) = \left\{ \begin{array}{l}
    c_{111}\tilde{x}_1\tilde{x}_2 + c_{110}\tilde{x}_1 + c_{101}\tilde{x}_2 + c_{100} = 0 \\
    c_{211}\tilde{x}_1\tilde{x}_2 + c_{210}\tilde{x}_1 + c_{201}\tilde{x}_2 + c_{200} = 0
\end{array} \right. \]
Mixed Cell Configurations and Normal Fans

Normals to mixed cells are in the intersections of normal cones to the edges.

Find all $v$ satisfying

\[
\begin{align*}
\langle \hat{a}, v \rangle &= \langle \hat{b}, v \rangle \\
\langle \hat{a}, v \rangle &\geq \langle \hat{b}, v \rangle \\
\forall \hat{a}, \hat{b} &\in \partial_v \hat{A}_i \\
\forall \hat{a} &\in \hat{A}_i \setminus \partial_v \hat{A}_i, \forall \hat{b} \in \partial_v \hat{A}_i \\
v_{n+1} &= 1
\end{align*}
\]

$i = 1, 2, \ldots, n$
Polyhedral Homotopies

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
   I.2 mixed cells
   I.3 volume of mixed cell

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

III. Solve the specific system $f(x) = 0$:

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

coefficient-parameter continuation
Some references on polyhedral methods


Recent computational advances

more efficient use of linear programming:


and parallel mixed-volume computations:


Blackbox Solving and Benchmarking

Building a simple blackbox solver:

\texttt{phc -b} first computes various root counts based on versions of Bézout’s theorem and mixed volumes. The start system is based on the smallest root count, and in case of equal counts, using the least complicated method.

The collection of test systems:

available at \url{http://www.math.uic.edu/~jan/demo.html} blackbox strategy opened up a wide range of applications
Exercises

- Take any polynomial system, solve it with the blackbox solver of PHCpack (as \texttt{phc -b input output}), and see what root count was used to build the start system.

- Explore the options of \texttt{phc -m}. In particular, the Cayley trick is efficient when there are only few different Newton polytopes. Find such an example where dynamic lifting with the Cayley trick outperforms the static lifting techniques.