Numerical Irreducible Decomposition

Jan Verschelde
Department of Math, Stat & CS
University of Illinois at Chicago
Chicago, IL 60607-7045, USA
\textit{e-mail:} jan@math.uic.edu
\textit{web:} www.math.uic.edu/~jan

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

1. We factor in three stages:
   (a) monodromy grouping of witness points;
   (b) certification of grouping with linear traces;
   (c) interpolation to get polynomials for the factors.

2. Special case: one single multivariate polynomial.
   We remove multiplicies by differentiation and use a theorem of Marden and Walsh for bound on precision.

3. Applications:
   (a) irreducible components of Griffis-Duffy platforms;
   (b) study singularities of Stewart-Gough platforms.
Recommended Background Literature


Factoring Solution Components

Input: $f(x) = 0$ polynomial system with a positive dimensional solution component, represented by witness set.

Coefficients of $f$ known approximately, work with limited precision

Wanted: decompose the component into irreducible factors,
for each factor, give its degree and multiplicity.

Symbolic-Numeric issue: essential numerical information
(such as degree and multiplicity of each factor),
is obtained much faster than the full symbolic representation.


The Riemann Surface of $z^3 - w = 0$:

Monodromy to Decompose Solution Components

Given: a system \( f(x) = 0 \); and \( W = (Z, L) \):

for all \( w \in Z : f(w) = 0 \) and \( L(w) = 0 \).

Wanted: partition of \( Z \) so that all points in a subset of \( Z \)
lie on the same irreducible factor.

Example: does \( f(x, y) = xy - 1 = 0 \) factor?

Consider \( H(x, y, \theta) = \begin{cases} 
xy - 1 = 0 & \text{for } \theta \in [0, 2\pi]. \\
x + y = 4e^{i\theta} & \text{for } \theta \in [0, 2\pi].
\end{cases} \)

For \( \theta = 0 \), we start with two real solutions. When \( \theta > 0 \), the
solutions turn complex, real again at \( \theta = \pi \), then complex until at
\( \theta = 2\pi \). Back at \( \theta = 2\pi \), we have again two real solutions, but their
order is permuted \( \Rightarrow \) irreducible.
Connecting Witness Points

1. For two sets of hyperplanes \( K \) and \( L \), and a random \( \gamma \in \mathbb{C} \)

\[
H(x, t, K, L, \gamma) = \begin{cases} 
  f(x) = 0 \\
  \gamma K(x)(1 - t) + L(x)t = 0
\end{cases}
\]

We start paths at \( t = 0 \) and end at \( t = 1 \).

2. For \( \alpha \in \mathbb{C} \), trace the paths defined by \( H(x, t, K, L, \alpha) = 0 \).

For \( \beta \in \mathbb{C} \), trace the paths defined by \( H(x, t, L, K, \beta) = 0 \).

Compare start points of first path tracking with end points of second path tracking. Points which are permuted belong to the same irreducible factor.

3. Repeat the loop with other hyperplanes.
Consider \( f(x, y(x)) = (y - y_1(x))(y - y_2(x))(y - y_3(x)) \)
\[ = y^3 - t_1(x)y^2 + t_2(x)y - t_3(x) \]

We are interested in the linear trace: \( t_1(x) = c_1x + c_0 \).

Sample the cubic at \( x = x_0 \) and \( x = x_1 \). The samples are 
\( \{(x_0, y_{00}), (x_0, y_{01}), (x_0, y_{02})\} \) and \( \{(x_1, y_{10}), (x_1, y_{11}), (x_1, y_{12})\} \).

Solve \( \begin{cases} y_{00} + y_{01} + y_{02} = c_1x_0 + c_0 \\ y_{10} + y_{11} + y_{12} = c_1x_1 + c_0 \end{cases} \) to find \( c_0, c_1 \).

With \( t_1 \) we can predict the sum of the \( y \)'s for a fixed choice of \( x \).
For example, samples at \( x = x_2 \) are \( \{(x_2, y_{20}), (x_2, y_{21}), (x_2, y_{22})\} \).
Then, \( t_1(x_2) = c_1x_2 + c_0 = y_{20} + y_{21} + y_{22} \).
Use \( \{(x_0, y_{00}), (x_0, y_{01}), (x_0, y_{02})\} \) and \( \{(x_1, y_{10}), (x_1, y_{11}), (x_1, y_{12})\} \) to find the linear trace \( t_1(x) = c_0 + c_1 x \).

At \( \{(x_2, y_{20}), (x_2, y_{21}), (x_2, y_{22})\} \): \( c_0 + c_1 x_2 = y_{20} + y_{21} + y_{22} \)?
Validation of Breakup with Linear Trace

Do we have enough witness points on a factor?

- We may not have enough monodromy loops to connect all witness points on the same irreducible component.
- For a $k$-dimensional solution component, it suffices to consider a curve on the component cut out by $k - 1$ random hyperplanes. The factorization of the curve tells the decomposition of the solution component.
- We have enough witness points on the curve if the value at the linear trace can predict the sum of one coordinate of all points in the set.

Notice: Instead of monodromy, we may enumerate all possible factors and use linear traces to certify. While the complexity of this enumeration is exponential, it works well for low degrees.
Special case: one single polynomial

- Input: $f(x) \in \mathbb{C}[x]$, $x = (x_1, x_2, \ldots, x_n)$.
  
  *coefficients known approximately, work with limited precision*

- Wanted: write $f$ as product of irreducible factors, as

$$f(x) = \prod_{i=1}^{N} q_i(x)^{\mu_i}, \quad \sum_{i=1}^{N} \mu_i \deg(q_i) = \deg(f),$$

  every irreducible factor $q_i$ occurs with multiplicity $\mu_i$.

Related Work


Dealing with Multiplicities

On a factor of degree $d$ and multiplicity $\mu$, we find $d$ clusters, each of $\mu$ witness points.

Choose $v = (v_1, v_2, \ldots, v_n)$ and compute

$$g(x) := \left( v_1 \frac{\partial}{\partial x_1} + v_2 \frac{\partial}{\partial x_2} + \cdots + v_n \frac{\partial}{\partial x_n} \right)^{\mu-1} f(x).$$

Then apply the techniques to the multiplicity one roots of $g(x)$ corresponding to the clusters.
Using a theorem of Marden and Walsh

Assume $d$ is the degree of $f(z)$, $f \in \mathbb{C}[z]$;

- $\mu$ is the multiplicity of a root of $f$;
- $z_0$ is the center of the cluster around the multiple root;
- $\Delta_r(z_0) = \{ z \in \mathbb{C} \mid |z - z_0| \leq r \}$ contains the cluster;
- $r$ is the radius of the disk $\Delta_r(z_0)$;
- $R$ is largest such that $\{ z \in \mathbb{C} \mid |z - z_0| \geq R \}$ contains all other $d - \mu$ roots of $f$.

If $\frac{R}{r} \geq \frac{2(\frac{d}{\mu})}{d-\mu+1}$, then $f^{(k)}$ has exactly $\mu - k$ roots in $\Delta_r(z_0)$,

for $k = 1, 2, \ldots, \mu - 1$.  

15
Applying the bound for $R/r$

Given a cluster of $\mu$ roots (and $d - \mu$ other roots), compute

- $z_0$ as the average of the roots in the cluster;
- $r$ as the largest distance of the roots in the cluster to $z_0$;
- $R$ as the smallest distance of the other $d - \mu$ roots to $z_0$.

\[
\frac{R}{r} \geq \frac{2^{(d)}_{\mu}}{d - \mu + 1} \implies r \leq R \left( \frac{d - \mu + 1}{2^{(d)}_{\mu}} \right)
\]

We obtain a bound on $r$, the precision of the roots in the cluster, in order for the successive derivatives of $f$ to be safe.
Numerical Limitations

- Evaluation of high degree polynomials is numerically unstable:

\[ f(x) = (x_0 + tv)^d = \sum_{k=0}^{d} \binom{d}{k} x_0^{d-k} v^k t^k = 0, \]

for example, \( d = 30 \) and \( k = 15 \): nine decimal places in \( \binom{d}{k} \).

- Working precision determines accuracy of factorization:

\[ f(x, y) = xy + 10^{-16} \]

- will factor when working with double precision floats;

- will not factor as soon as precision is high enough.
Application I: Architecturally Singular Platforms

Special Griffis-Duffy type

- Base and endplate are equilateral triangles.
- Legs connect vertices to midpoints.
Results of Husty and Karger


The special Griffis-Duffy platforms move:

- Case 1: Plates not equal, legs not equal.
  - Curve is degree 20 in Euler parameters.
  - Curve is degree 40 in position.

- Case 2: Plates congruent, legs all equal.
  - Factors are degrees $(4 + 4) + 6 + 2 = 16$ in Euler parameters.
  - Factors are degrees $(8 + 8) + 12 + 4 = 32$ in position.

Question: Can we confirm these results numerically?
## Components of Griffis-Duffy Platforms

### Solution components by degree

<table>
<thead>
<tr>
<th></th>
<th>Husty &amp; Karger</th>
<th>SVW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Euler</td>
<td>Position</td>
<td>Study</td>
</tr>
</tbody>
</table>

#### General Case

<table>
<thead>
<tr>
<th></th>
<th>20</th>
<th>40</th>
<th>28</th>
<th>40</th>
</tr>
</thead>
</table>

#### Legs equal, Plates equal

<table>
<thead>
<tr>
<th></th>
<th>4</th>
<th>8</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4</td>
<td>8</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>12</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>32</td>
<td>28</td>
<td>40</td>
</tr>
</tbody>
</table>
Case A: One irreducible component of degree 28 (general case).

Case B: Five irreducible components of degrees 6, 6, 6, 6, and 4.

<table>
<thead>
<tr>
<th>user cpu on 800Mhz</th>
<th>Case A</th>
<th>Case B</th>
</tr>
</thead>
<tbody>
<tr>
<td>witness points</td>
<td>1m 12s 480ms</td>
<td></td>
</tr>
<tr>
<td>monodromy breakup</td>
<td>33s 430ms</td>
<td>27s 630ms</td>
</tr>
<tr>
<td>Newton interpolation</td>
<td>1h 19m 13s 110ms</td>
<td>2m 34s 50ms</td>
</tr>
</tbody>
</table>

32 decimal places used to interpolate polynomial of degree 28

Linear traces replace Newton interpolation:

⇒ time to factor independent of geometry!
Griffis-Duffy Platforms: an Animation
Application II: three Stewart-Gough platforms

General platform, fixed position

Planar base, planar platform

Parallel base and platform

At singularity, rigidity of device is lost, allowing finite motion which cannot be controlled by leg lengths (*disaster!*).

Denote

\[ p \in \mathbb{C}^3 \] position of platform;

\[ q \in \mathbb{P}^3 \] quaternion defines a rotation;

\[ a_i, b_i \in \mathbb{C}^3 \] ball joints at platform and base, \( i = 1, 2, \ldots, 6 \);

\[ J \in \mathbb{C}^{6 \times 6} \] Jacobian matrix of mapping from platform motion to leg lengths.

Then the condition on a singular configuration is \( \det J = 0 \).

\( \det J \) is a polynomial of degree 1728 in 43 variables: \( p, q, a_i, b_i \).


First general case of a Stewart-Gough platform

- Case of almost all manipulators
  - $p, a_i,$ and $b_i$ are randomly chosen
- $\deg(\det \mathbf{J}) = 12,$ homogeneous in $\mathbf{q}$
  - The expanded $\det \mathbf{J}$ has 910 terms
- $\det \mathbf{J} = F_1(\mathbf{q})(F_2(\mathbf{q}))^3$
  - $\mathbf{q} = (q_0, q_1, q_2, q_3)$ quaternion
  - $\deg(F_1) = 6$
  - $F_2(\mathbf{q}) = q_0^2 + q_1^2 + q_2^2 + q_3^2$
    - $F_2$ has no physical significance
Computational results for first platform

<table>
<thead>
<tr>
<th>cluster</th>
<th>$r$</th>
<th>$R$</th>
<th>$R/r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>one</td>
<td>1.7E-05</td>
<td>3.4E-01</td>
<td>2.0E+04</td>
</tr>
<tr>
<td>two</td>
<td>4.9E-06</td>
<td>1.7E-01</td>
<td>3.6E+04</td>
</tr>
</tbody>
</table>

Lower bound on $R/r$ evaluates to 44.

Elapsed user CPU times on 2.4Ghz WindowsXP

<table>
<thead>
<tr>
<th>stage</th>
<th>time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. monodromy grouping</td>
<td>0h 6m 40s 469ms</td>
</tr>
<tr>
<td>2. linear traces certification</td>
<td>0h 0m 30s 672ms</td>
</tr>
<tr>
<td>3. interpolation at factors</td>
<td>1h 41m 53s 78ms</td>
</tr>
<tr>
<td>4. multiplication validation</td>
<td>0h 0m 8s 156ms</td>
</tr>
<tr>
<td>total time for all 4 stages</td>
<td>1h 49m 12s 391ms</td>
</tr>
</tbody>
</table>
second case: planar base and platform

- ball joints $a_i$ lie in planar platform
- ball joints $b_i$ lie in planar base
- $\deg(\det{J}) = 12$, homogeneous in $q$
  - the expanded $\det{J}$ has 910 terms
- $\det{J} = F_1(q)(F_2(q))^3$
  - $q = (q_0, q_1, q_2, q_3)$ quaternion
  - $\deg(F_1) = 6$  $\deg(F_2) = 2$
Computational results for second platform

<table>
<thead>
<tr>
<th>cluster</th>
<th>$r$</th>
<th>$R$</th>
<th>$R/r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>one</td>
<td>6.2E−05</td>
<td>2.4E−01</td>
<td>3.8E+04</td>
</tr>
<tr>
<td>two</td>
<td>4.8E−05</td>
<td>6.0E−01</td>
<td>1.2E+04</td>
</tr>
</tbody>
</table>

Lower bound on $R/r$ evaluates to 44.

Elapsed user CPU times on 2.4Ghz WindowsXP

<table>
<thead>
<tr>
<th></th>
<th>time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. monodromy grouping</td>
<td>0h 17m 34s 735ms</td>
</tr>
<tr>
<td>2. linear traces certification</td>
<td>0h 0m 27s 359ms</td>
</tr>
<tr>
<td>3. interpolation at factors</td>
<td>1h 24m 45s 766ms</td>
</tr>
<tr>
<td>4. multiplication validation</td>
<td>0h 0m 8s 172ms</td>
</tr>
<tr>
<td>total time for all 4 stages</td>
<td>1h 42m 56s 32ms</td>
</tr>
</tbody>
</table>
third case: parallel base and platform

- ball joints $a_i, b_i$ in parallel planes, position $p$ is variable, $q_1 = q_2 = 0$
- $	ext{deg}(\text{det } J) = 15$, in $(p, q)$
  expanded $\text{det } J$ has 24 terms, much sparser, as $24 << 910$
- $\text{det } J = ap_3^3(q_0 + bq_3)(q_0 + cq_3)(q_0 + iq_3)^5(q_0 - iq_3)^5$
  where the constants $a, b, c$
  depend on the choice of $a_i, b_i$
Computational results for third platform

<table>
<thead>
<tr>
<th>cluster</th>
<th>$r$</th>
<th>$R$</th>
<th>$R/r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>one</td>
<td>5.1E-07</td>
<td>1.0E+00</td>
<td>2.0E+06</td>
</tr>
<tr>
<td>two</td>
<td>7.3E-04</td>
<td>3.4E-01</td>
<td>4.7E+02</td>
</tr>
<tr>
<td>three</td>
<td>4.0E-03</td>
<td>7.2E-01</td>
<td>1.8E+02</td>
</tr>
</tbody>
</table>

Lower bound on $R/r$ evaluates to 546.

Elapsed user CPU times on 2.4GHz Windows XP

<table>
<thead>
<tr>
<th>Stage</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. monodromy grouping</td>
<td>1m 13s  656ms</td>
</tr>
<tr>
<td>2. linear traces certification</td>
<td>0m 3s   891ms</td>
</tr>
<tr>
<td>3. interpolation at factors</td>
<td>0m 4s   734ms</td>
</tr>
<tr>
<td>4. multiplication validation</td>
<td>0m 1s   657ms</td>
</tr>
<tr>
<td>total time for all 4 stages</td>
<td>1m 23s  938ms</td>
</tr>
</tbody>
</table>
Monodromy Compared to the Enumeration Method

Enumeration of all possible factors certified by linear traces outperforms the monodromy algorithm for our application:

<table>
<thead>
<tr>
<th>User CPU times on 2.4Ghz Windows XP</th>
<th>monodromy</th>
<th>enumeration</th>
</tr>
</thead>
<tbody>
<tr>
<td>case</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>6m 40s 460ms</td>
<td>40s 750ms</td>
</tr>
<tr>
<td>2</td>
<td>17m 34s 735ms</td>
<td>31s 657ms</td>
</tr>
<tr>
<td>3</td>
<td>1m 13s 656ms</td>
<td>3s 0ms</td>
</tr>
</tbody>
</table>

Random irreducible polynomials of five monomials:

<table>
<thead>
<tr>
<th>User CPU times on 2.4Ghz Windows XP</th>
<th>monodromy</th>
<th>enumeration</th>
</tr>
</thead>
<tbody>
<tr>
<td>degree</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>5s 484ms</td>
<td>312ms</td>
</tr>
<tr>
<td>15</td>
<td>8s 187ms</td>
<td>1s 453ms</td>
</tr>
<tr>
<td>16</td>
<td>16s 63ms</td>
<td>2s 875ms</td>
</tr>
</tbody>
</table>
Exercises

- Apply phc -f to factor

\[ x^{**6} - x^{**5}y + 2x^{**5}z - x^{**4}y^{**2} - x^{**4}yz + x^{**3}yz + 3y^{**3}z \\
- 4x^{**3}y^{**2}z + 3x^{**3}yz^{**2} - 2x^{**3}z^{**3} + 3x^{**2}y^{**3}z^{**2} \\
- 6x^{**2}y^{**2}z^{**2} + 5x^{**2}yz^{**3} - x^{**2}z^{**4} + 3x^{**}y^{**3}z^{**2} \\
- 4x^{y^{**}2}z^{**3} + 2x^{y^z**4} + y^{**3}z^{**3} - y^{**2}z^{**4}; \]

- Consider the adjacent minors of a general 2 × 4-matrix:

\[
\begin{bmatrix}
  x_{11} & x_{12} & x_{13} & x_{14} \\
  x_{21} & x_{22} & x_{23} & x_{24}
\end{bmatrix}
\]

\[
f(x) = \begin{cases} 
  x_{11}x_{22} - x_{21}x_{12} = 0 \\
  x_{12}x_{23} - x_{22}x_{13} = 0 \\
  x_{13}x_{24} - x_{23}x_{14} = 0
\end{cases}
\]

Compute the irreducible decomposition of \( f^{-1}(0) \).