Stable methods for solving polynomial equations

M. Alonso, B. Mourrain, Ph. Trébuchet
Univ. of Madrid,
GALAAD, INRIA Méditerrannée
LIP6, Paris

June 20, 2008

Equations with approximate coefficients

Input: $f_1, \ldots, f_m \in R := \mathbb{Q}[x_1, \ldots, x_n]$, $I := (f_1, \ldots, f_m) \subset R$.

❒ We consider a **neighbourhood** of the system $\mathbf{f}$,
or a family of systems depending on parameters, of the same "shape".

❒ Around a **regular** value of the parameters,

- **continuity** of the solution set.
- **continuity** of the algebraic structure.

❒ At a **singular** value of the parameters, all sort of bad things may happen.

☞ **Stability** searched at **regular** values.

The objective(s):

- Develop methods which are stable (work with approximate coefficients) and efficient ?

Problems to be solved:

- If we start with a perturbation of the input, do we get nearby output ?
- If the approximation is not sufficient, how can we improve it ?

How to proceed ?

Exploit the algebraic structure **and** numerical information.

Related works

- Janet basis: 20'
- Grobner basis: 60'
- H-basis: Macaulay'16, Möller-Sauer'00
- Border basis:

  Cartan'45

  Kuranishi'57

  ...

  Mourrain, Trébuchet: '99 (characterisation), '00, '02 (algorithm), '02 (Ph.D.), '05 (issac), '06 (syzygies), '08 (syzygies).

  Stetter'04.

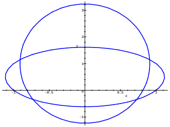
  Kehrein, Kreuzer, Robbiano: '05, '05 (characterisation), '06 (algorithm), '08.

  Huibregste '06 (syzygies).

GB are going boink, aren't they ?

A system:

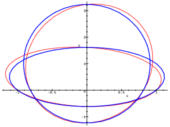
$$\begin{cases} p_1 := a\,x_1^2 + b\,x_2^2 + l_1(x_1, x_2) \\ p_2 := c\,x_1^2 + d\,x_2^2 + l_2(x_1, x_2) \end{cases}$$



Basis of
$\mathcal{A} = \mathbb{K}[\mathbf{x}_1, x_2]/(p_1, p_2)$:
$(1, x_1, x_2, x_1\,x_2)$.

A small perturbation:

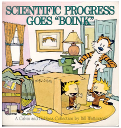
$$\begin{cases} \tilde{p}_1 := p_1 + \epsilon_1\, x_1 x_2 \\ \tilde{p}_2 := p_2 + \epsilon_2\, x_1 x_2 \end{cases}$$



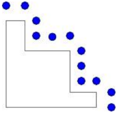
Basis of
$\tilde{\mathcal{A}} = \mathbb{K}[x_1, x_2]/(\tilde{p}_1, \tilde{p}_2)$:
$(1, x_1, x_2, x_2^2)$.

Catastroph !

- A new set of monomials for the basis.
- Big coefficients (in $\frac{1}{\epsilon_i}$) appear.

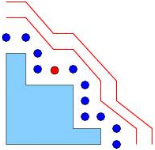


Border basis



Notations:

- $I=(f_1,\ldots,f_s)$, $\mathcal{A}=\mathbb{K}[\mathbf{x}]/I$,
- B a set of monomials **connected** to 1 $(m\in B\text{-}\{1\}\Rightarrow \exists m'\in B,\ i\in[1,n]$ st. $m=m'x_i)$.
- $B^+ = B\cup x_1B\cup\cdots\cup x_nB$, $\delta B = B^+ - B$.



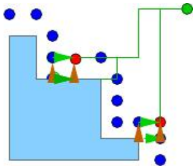
Suppose B is a basis of $\mathcal{A}$, then

- Each $\mathbf{x}^\alpha\in\delta B$ yields a rewritting rule
$$f_\alpha=\mathbf{x}^\alpha-\sum_{\beta\in B}\lambda_{\alpha,\beta}\mathbf{x}^\beta.$$
- The rewritting rules of δB allow to reduce any $p\in\mathbb{K}[\mathbf{x}]$ to $\langle B\rangle$.

Definition

A **border basis** of B for I is a set of relations of the form f_α for $\alpha\in\delta B$, such $I=(f_\alpha)$, $\langle B\rangle\cap I=\{0\}$.

Normal form criterion



- Many possible reductions of m on $\langle B \rangle$.
- Not necessarly, $\mathbb{K}[\mathbf{x}] = \langle B \rangle \oplus I$ (or $\langle B \rangle \cap I = \{0\}$).
- **How to check normal form ?**

❒ The rewritting family defines a projection $N : \langle B^+ \rangle \to \langle B \rangle$.

Theorem (-'99)

Let B be connected to 1 and $M_i : \langle B \rangle \to \langle B \rangle$ such that $M_i(b) = N(x_i b)$.

$$N \text{ normal form modulo } I = (Ker(N))$$
$$\Leftrightarrow B \text{ basis of } \mathcal{A} = R/I$$
$$\Leftrightarrow M_i \circ M_j = M_j \circ M_i, i, j = 1, \ldots, n.$$

Normal form criterion

❒ **A choice function** $\gamma : \mathbb{K}[\mathbf{x}] \to (\mathbf{x}^*)$ refining a grading Λ, such that $\forall p \in \mathbb{K}[\mathbf{x}]$, $\gamma(p)$ is a monomial of the support of p.

❒ **A rewritting family** $(f_i)_{i \in I}$ for B is s.t.:

- $Supp(f_i) \subset B^+$,
- f_i **has exactly one monomial** $\gamma(f_i)$ **in** δB,
- **if** $\gamma(f_i) = \gamma(f_j)$ **then** $i = j$,
- $\forall m \in \delta B$, $\exists i \in I \mid \gamma(f_i) = m$.

❒ For any $p_1, p_2 \in \mathbb{K}[\mathbf{x}]$, $C(p_1, p_2) = \frac{lcm(\gamma(p_1), \gamma(p_2))}{\gamma(p_1)} p_1 - \frac{lcm(\gamma(p_1), \gamma(p_2))}{\gamma(p_2)} p_2$.

Theorem

Let F be a normalising family for a set B of monomials, **connected** *to 1, and let N be the projection of $\langle B^+ \rangle \to \langle B \rangle$ along $\langle F \rangle$. Then,*

$$N \textit{ extends uniquely to } \tilde{N} : \mathbb{K}[\mathbf{x}] \to \langle B \rangle \textit{ s.t. } \ker(\tilde{N}) = (F).$$

iff

$$\forall f, f' \in F \textit{ s.t. } C(f, f') \in \langle B^+ \rangle, N(C(f, f')) = 0$$

Algorithm (Normal form computations)

INPUT: *$f_1, \ldots, f_m$ defining and ideal of dimension 0, and γ a choice function refining the degree.*

INITILIZATION:

Choose the f_{i_0} of minimal degree k,

$B_k = (\gamma(f_{i_0}))^c$, $k = deg(f_{i_0})$, $P_k = \{f_{i_0}\}$, $M_k = \{\gamma(f_{i_0})\}$.

CORE LOOP: *While $\cup_k M_k \neq B_k^+ - B_k$ do*

- *Compute $P_{k+1} = (P_k^+ \cup F_k) \cap B_k^+$,*
- $M_{k+1} = \{M_k^+ \cap B_k^+\}$,
- $F_{k+1} = RewrittingFamily(P_{k+1}, M_{k+1})$,
- *Reduce $C_{k+1} =\{\ C(f, f')$ of degree $k+1$, $f, f' \in P_k\}$ by $F = \cup_{1 \leq j \leq k+1} F_j$.*
- *According to $r = \#M_{k+1} - \#F_{k+1}$ and $c = \#\{C - polynomials\ of\ C_{k+1}\ non\ reduced\ to\ 0\}$, update B and P_{k+1}.*

OUTPUT: *$F = \cup_j F_j$ a normalising family for (f_i).*

Generalized normal form

- It **generalises Gröbner basis** computation.
  If γ is a monomial ordering,
  - the output is a grobner basis,
  - the C-polynomials "are" the S-polynomials.
- **Linear algebra** on vector spaces of polynomials.
- It allows pivoting on the **rows and columns**, according to numerical criterions.
- Use of **generic sparse lu** decomposition (superlu).
- Extension to **Laurent polynomials**.

Experimentations (MATHEMAGIX, Ph. Trébuchet)

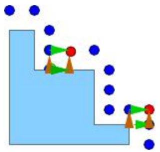
Katsura(6):

choice function	number of bits	time	$max(\|\|f_i\|\|_\infty)$
grevlex	128	1.98s	10^{-28}
dlex	128	2.62s	10^{-24}
mac	128	1.64s	10^{-30}
grevlex	80	1.35s	10^{-20}
dlex	80	3.98s	10^{-15}
mac	80	0.95s	10^{-19}
grevlex	64		–
dlex	64		–
mac	64	0.9s	10^{-11}

Paralell robot:

choice function	number of bits	time	$max(\|\|f_i\|\|_\infty)$
dlex	250	11.16s	$0.42 * 10^{-63}$
mac	250	11.62s	$0.46 * 10^{-63}$
dinvlex	250	13.8s	$0.135 * 10^{-60}$
dlex	128	9.13s	$0.3 * 10^{-24}$
dinvlex	128	11.1s	$0.3 * 10^{-23}$
mac	128	9.80s	$0.1 * 10^{-24}$
dlex	80	-	-
dinvlex	80	-	-
mac	80	6.80s	10^{-12}

Syzygies:



Proposition (-'06;H'06;-,T'08)

The syzygies of $(f_\alpha)_{\alpha\in\delta B}$ *are generated by the relations:*

$$x_\alpha f_\alpha - x_{\alpha'} f_{\alpha'} - \sum_{\gamma\in\delta B} \lambda_\gamma f_\gamma = 0$$

for $x_\alpha \mathbf{x}^\alpha = x_{\alpha'}\mathbf{x}^{\alpha'} = x_\alpha x_{\alpha'} m$, $m \in B$, *deduced from the reduction of the* C*-polynomials.*

Stability:

- A grading Λ of $\mathbb{K}[\mathbf{x}]$.
- $N_\epsilon(\mathbf{f}) = \{(h_1,\ldots,h_s) \in \mathbb{K}[\mathbf{x}], \Lambda(h_i) \leq \Lambda(f_i), ||h_i - f_i||_\infty < \epsilon\}$
- γ a choice function refining the grading Λ.
- $\gamma_\epsilon(p) = \gamma(p_\epsilon)$ where $||p - p_\epsilon||_\infty < \epsilon$, p_ϵ of smallest support.

Theorem (-,T'08)

Let $\mathbf{f} = (f_1,\ldots,f_s)$ *be a zero dimensional s.t. that* $\forall \mathbf{f}' \in U(\mathbf{f})$, $\mathbf{f}'$ *has* D *complex roots, counted with multiplicities. Then* $\forall\epsilon > 0$ *small enough, there exists* $\nu_0 > 0$ *s.t.* $\forall \mathbf{f}' \in N_{\nu_0}(\mathbf{f}) \subset U(\mathbf{f})$, *the basis* B *computed with* γ *for the system* $\mathbf{f}$ *is also the basis computed with* γ_ϵ *for* $\mathbf{f}'$.

Perturbations

The algebraic set $\mathcal{H}_B$ of quotient algebras with basis B.

- B a given set of monomials connected to 1.
- A quotient algebra $\mathcal{A}$ with basis B described by

$$\mathbf{z} = (z_{\alpha,\beta}) \in \bar{\mathbb{K}}^{\delta B \times B},$$

the coefficient vectors of the border relations:

$$h^{\mathbf{z}}_{\alpha}(\mathbf{x}) := \mathbf{x}^{\alpha} - \sum_{\beta \in B} z_{\alpha,\beta} \mathbf{x}^{\beta}.$$

- $M^{\mathbf{z}}_{x_i}$:= multiplication tables modulo $h^{\mathbf{z}}_{\alpha}(\mathbf{x})$.

Theorem

The polynomials $h^{\mathbf{z}}_{\alpha}(\mathbf{x})$ are the border relations of a quotient algebra $\mathcal{A}^{\mathbf{z}}$ iff $M^{\mathbf{z}}_{x_i} \circ M^{\mathbf{z}}_{x_j} - M^{\mathbf{z}}_{x_j} \circ M^{\mathbf{z}}_{x_i} = 0 \ {}_{(1 \leqslant i < j \leqslant n)}$.

$\mathcal{H}_B := \{\mathbf{z} = (z_{\alpha,\beta}) \in \mathbb{K}^{\delta B \times B}; M^{\mathbf{z}}_{x_i} \circ M^{\mathbf{z}}_{x_j} - M^{\mathbf{z}}_{x_j} \circ M^{\mathbf{z}}_{x_i} = 0 \ {}_{1 \leqslant i < j \leqslant n}\}$

☞ Affine chart of **the Hilbert Scheme** $\mathcal{H}^{n}_{|B|}$.

The tangent space to $\mathcal{H}_B$

A point of $\mathcal{H}_B$:

- $(f_k^0) = I_0$,
- $(h_\alpha^0)_{\alpha \in \partial B}$ a border basis of I_0 for B connected to 1.
- N_0 the normal form for I_0 on $\langle B \rangle$.

A controlled perturbation: $f_k^\varepsilon = f_k^0 + \varepsilon\, \mathbf{f_k^1}\ _{(k=1\ldots s)}$, $(f_k^\varepsilon) = I_\varepsilon$.

Proposition

Suppose that $\mathbb{K}[\mathbf{x}]/I^\epsilon$ contains a subalgebra $\mathcal{A}^\varepsilon$ with basis B, then the border basis of $\mathcal{A}^\varepsilon$ is of the form $h_\alpha^\varepsilon = h_\alpha^0 + \varepsilon\, \mathbf{h_\alpha^1} + \mathcal{O}(\varepsilon^2)$ with

- $\mathbf{M_{x_i}^1} \circ M_{x_j}^0 + M_{x_i}^0 \circ \mathbf{M_{x_j}^1} - \mathbf{M_{x_j}^1} \circ M_{x_i}^0 - M_{x_j}^0 \circ \mathbf{M_{x_i}^1} = 0_{(1 \leqslant i < j \leqslant n)}, \quad (1)$
- $N_0(f_k^1) - \sum_{\alpha \in \partial B} N_0(q_{\alpha,k}^0 \mathbf{h_\alpha^1}) = 0, \quad (2)$

where

$$\mathbf{N^1}(\mathbf{x}^\alpha) = \mathbf{h_\alpha^1},\ \mathbf{N^1}(\mathbf{x}^\beta) = 0,\ \mathbf{M_{x_i}^1}(\mathbf{x}^\beta) = \mathbf{N^1}(x_i\, \mathbf{x}^\beta)\ _{(\alpha \in \delta, \beta \in B)},$$
$$f_k^0 = \sum_{\alpha \in \partial B} q_{\alpha,k}^0 h_\alpha^0.$$

Initially: $f_1^0 = y^2, f_1^0 = x^3 - x^2 y$.

- $B = \{1, x, x^2, y, xy, x^2 y\}$.
- Border basis: $\{y^2, xy^2, x^2y^2, x^3 - x^2 y, x^3 y\}$.

Perturbation: $f_1^\varepsilon = f_1^0 - \varepsilon(x^2 y + 1), f_2^\varepsilon := f_2^0 - \varepsilon$.

- $h^\varepsilon_{y^2} = y^2 - \varepsilon(x^2 y + 1)$.
- $h^\varepsilon_{x^3} = x^3 - x^2 y - \varepsilon$.
- $h^\varepsilon_{x\,y^2} = xy^2 - \varepsilon x + \mathcal{O}(\varepsilon^2)$,
- $h^\varepsilon_{x^2 y^2} = x^2 y^2 - \varepsilon x^2 + \mathcal{O}(\varepsilon^2)$.
- $h^\varepsilon_{x^3 y} = x^3 y - \varepsilon(z_1 + z_2 x + z_3 x^2 + z_4 y + z_5 xy + z_6 x^2 y) + \mathcal{O}(\varepsilon^2)$.

Equation (1) yields

$$M_x^\varepsilon M_y^\varepsilon - M_y^\varepsilon M_x^\varepsilon = \varepsilon \begin{pmatrix} 0 & z_1 & 0 & 0 & 0 & 0 \\ 0 & z_2 & 0 & 0 & 0 & 0 \\ 0 & z_3 - 1 & 0 & 0 & 0 & 0 \\ 0 & z_4 - 1 & 0 & 0 & 0 & -z_1 \\ 0 & z_5 & 0 & 0 & 0 & -z_5 \\ 0 & z_6 & 0 & 0 & 0 & 0 \end{pmatrix} + \mathcal{O}(\varepsilon^2).$$

$h^\varepsilon_{x^3 y} = x^3 y - \varepsilon(x^2 + y) + \mathcal{O}(\varepsilon^2)$.

- $T_{\mathbf{h}_0} = \{(h^1_\alpha)_{\alpha\in\delta B} \in \langle B\rangle^{\delta B}$, which satisfies (1)$\}$.

$$\begin{array}{rcl} \phi : T_{\mathbf{f}_0} & \rightarrow & Hom_R(I_0, R/I_0) \\ (h^1_\alpha) & \mapsto & \phi(h^1_\alpha) : h^0_\alpha \mapsto h^1_\alpha \end{array}$$

is an isomorphism of $\mathbb{K}$-vector spaces.

- H_B contains a component of dimension $n \times |B|$ parametrised by

$$\begin{array}{rcl} \mathfrak{H}_B : \mathbb{C}^{n\times|B|} & \rightarrow & \mathbb{C}^{\partial B\times B} \\ \mathfrak{Z} = \{\zeta_1, \ldots, \zeta_{|B|}\} & \mapsto & (\rho_{\alpha,\beta}(\mathfrak{Z}))_{\alpha\in\partial B, \beta\in B}. \end{array}$$

where $h_\alpha(\mathfrak{Z}, \mathbf{x}) = \frac{R_\alpha(\mathfrak{Z},\mathbf{x})}{V_B(\mathfrak{Z})} = \mathbf{x}^\alpha - \sum_{\beta\in B} \rho_{\alpha,\beta}(\mathfrak{Z})\mathbf{x}^\beta$ and

$$V_B(\mathfrak{Z}) = \begin{vmatrix} \zeta_1^{\beta_1} & \cdots & \zeta_1^{\beta_{|B|}} \\ \vdots & \vdots & \vdots \\ \zeta_{|B|}^{\beta_1} & \cdots & \zeta_{|B|}^{\beta_{|B|}} \end{vmatrix}, \ R_\alpha(\mathfrak{Z}, \mathbf{x}) = \begin{vmatrix} \mathbf{x}^\alpha(\mathbf{x}) & \mathbf{x}^{\beta_1} & \cdots & \mathbf{x}^{\beta_{|B|}} \\ \mathbf{x}^\alpha(\zeta_1) & \zeta_1^{\beta_1} & \cdots & \zeta_1^{\beta_{|B|}} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{x}^\alpha(\zeta_{|B|}) & \zeta_{|B|}^{\beta_1} & \cdots & \zeta_{|B|}^{\beta_{|B|}} \end{vmatrix}.$$

- Weierstrass iteration, explicit inversion of $d\mathfrak{H}_B$ [-,R'03].
- H_B irreducible for $n \leq 2$, not always irreducible for $n > 3$ [I'72].
- The irreducible components of H_B are not known for $|B| > 8$.