Almost polynomial Complexity for Zero-dimensional Gröbner Bases

Amir Hashemi

co-author: Daniel Lazard

INRIA SALSA-project/ LIP6 CALFOR-team

http://www-calfor.lip6.fr/~hashemi/

Special semester on Gröbner Bases and related methods

Linz, Austria, 2006

Complexity of Gröbner basis computation

- [Lazard, 83]: Complexity $d^{O(n)}$
  - homogeneous zero-dim ideal
  - homogeneous regular sequence in generic coordinates
  - (degree) reverse lexicographical ordering

Complexity of Gröbner basis computation

- [Lazard, 83]: Complexity $d^{O(n)}$
  - homogeneous zero-dim ideal
  - homogeneous regular sequence in generic coordinates
  - (degree) reverse lexicographical ordering
- [Dickenstein et al., 91]: Bit complexity $d^{O(n^2)}$
  - zero-dim ideal
  - any ordering

Complexity of Gröbner basis computation

- [Lazard, 83]: Complexity $d^{O(n)}$
  - homogeneous zero-dim ideal
  - homogeneous regular sequence in generic coordinates
  - (degree) reverse lexicographical ordering
- [Dickenstein et al., 91]: Bit complexity $d^{O(n^2)}$
  - zero-dim ideal
  - any ordering
- [Lakshman, 91]: Arithmetic complexity $(nd^n)^{O(1)}$
  - zero-dim ideal
  - any ordering (using FGLM)

Motivation

Our objective:

- To have an algorithm to compute the Gröbner basis of a zero-dim ideal within a bit complexity $d^{O(n)}$:
  - To be able to extend it to regular sequences
    in positive dimension and in generic coordinates

Motivation

Our objective:

- To have an algorithm to compute the Gröbner basis of a zero-dim ideal within a bit complexity $d^{O(n)}$:
  - To be able to extend it to regular sequences
    in positive dimension and in generic coordinates

- To extend [Lazard, 83] to the non-homogeneous case
  by using a deformation method

  (already used in [Grigoriev, Chistov, 83], [Canny, 89], [Lakshman-Lazard, 91],...)

Notation

Input data:

- K : field, $R = K[x_1, \ldots, x_n]$: ring of polynomials
- $f_1, \ldots, f_k$: polynomials in R
- $I = \langle f_1, \ldots, f_k \rangle$
- $d_i = \deg(f_i)$ ordered in order that $d_2 \geq \cdots \geq d_k \geq d_1$

Measures of complexity:

- S : sum of the size of f_i in the dense representation
- $\mathrm{D} = (d_1 + \cdots + d_n)/n$ (if $i > k$ then $d_i = 1$)
- $\mathcal{T} = \max\{\mathrm{S}, D^n\}$

Notation

Monomial orderings for Gröbner bases:

- $\prec$: degree reverse lexicographic ordering s.t.

$$x_0 \prec x_n \prec \cdots \prec x_1$$

- $<$: any other ordering
- $\deg(I, <) =$ maximal degree of the elements of the reduced Gröbner basis of I

Complexity model

- $\mathrm{S} \leq \sum_{i=1}^{k} n h_i \binom{n+d_i}{n}$ where $h_i = \max\{$coefficients of $f_i\}$

Complexity model

- $\mathrm{S} \leq \sum_{i=1}^{k} n h_i \binom{n+d_i}{n}$ where $h_i = \max\{\text{coefficients of } f_i\}$
- We replace the bounds d^n and nd^n by:

$$\mathcal{T} = \max\{\mathrm{S}, \mathrm{D}^n\} \ll nhk(e\mathrm{D})^n$$

Complexity model

- $\mathrm{S} \leq \sum_{i=1}^{k} n h_i \binom{n+d_i}{n}$ where $h_i = \max\{$coefficients of $f_i\}$
- We replace the bounds d^n and nd^n by:

$$\mathcal{T} = \max\{\mathrm{S}, \mathrm{D}^n\} \ll nhk(e\mathrm{D})^n$$

- Bézout theorem: $\Longrightarrow$
  "Complexity $\geq$":

$$\max\{\mathrm{S}, d_1 \cdots d_n\} = \max\{\mathrm{S}, ((d_1 \cdots d_n)^{1/n})^n\}$$

Complexity model

- $\mathrm{S} \leq \sum_{i=1}^{k} n h_i \binom{n+d_i}{n}$ where $h_i = \max\{$coefficients of $f_i\}$
- We replace the bounds d^n and nd^n by:

$$\mathcal{T} = \max\{\mathrm{S}, \mathrm{D}^n\} \ll nhk(e\mathrm{D})^n$$

- Bézout theorem: $\Longrightarrow$
  "Complexity $\geq$":

$$\max\{\mathrm{S}, d_1 \cdots d_n\} = \max\{\mathrm{S}, ((d_1 \cdots d_n)^{1/n})^n\}$$

- The gap: geometric mean $\leftrightarrow$ arithmetic mean

Complexity model

- $\mathrm{S} \leq \sum_{i=1}^{k} n h_i \binom{n+d_i}{n}$ where $h_i = \max\{$coefficients of $f_i\}$
- We replace the bounds d^n and nd^n by:

$$\mathcal{T} = \max\{\mathrm{S}, \mathrm{D}^n\} \ll nhk(e\mathrm{D})^n$$

- Bézout theorem: $\Longrightarrow$
  "Complexity $\geq$":

$$\max\{\mathrm{S}, d_1 \cdots d_n\} = \max\{\mathrm{S}, ((d_1 \cdots d_n)^{1/n})^n\}$$

- The gap: geometric mean $\leftrightarrow$ arithmetic mean
- [Hashemi-Lazard, 05]: Complexity $\mathcal{T}^{O(1)}$
  for [Laz, 83], [Dick et al., 91], [Lak, 91]

Main results

$\boxed{I \text{ zero–dimensional } (k \geq n)}$

Main results

I zero–dimensional ($k \geq n$)

- $\deg(I, \prec) \leq d_1 + \cdots + d_n - n + 1 = n\mathrm{D} - n + 1$
- $\deg(I, <) \leq d_1 \cdots d_n \leq \mathrm{D}^n$
- Complexity $\mathcal{T}^{O(1)}$ to compute any Gröbner basis of I

Main results

$\boxed{I \text{ zero–dimensional } (k \geq n)}$

- $\deg(I, \prec) \leq d_1 + \cdots + d_n - n + 1 = n\mathrm{D} - n + 1$
- $\deg(I, <) \leq d_1 \cdots d_n \leq \mathrm{D}^n$
- Complexity $\mathcal{T}^{O(1)}$ to compute any Gröbner basis of I

$$\boxed{\begin{array}{c} f_1, \ldots, f_k \text{ regular sequence } (k \leq n) \\ x_{k+1}, \ldots, x_n \text{ in “generic position” for } I \end{array}}$$

Main results

I zero–dimensional ($k \geq n$)

- $\deg(I, \prec) \leq d_1 + \cdots + d_n - n + 1 = n\mathrm{D} - n + 1$
- $\deg(I, <) \leq d_1 \cdots d_n \leq \mathrm{D}^n$
- Complexity $\mathcal{T}^{O(1)}$ to compute any Gröbner basis of I

$f_1, \ldots, f_k$ regular sequence ($k \leq n$)
$x_{k+1}, \ldots, x_n$ in "generic position" for I

- A precise defi nition of generic position for this problem
- $\deg(I, \prec) \leq d_1 + \cdots + d_k - k + 1$
- Conjecture: Complexity $\mathcal{T}^{O(1)}$ to compute the Gröbner basis of I

Proof's idea

$\boxed{\text{Transform the problem}}$

for using [Lazard, 81] and [Lazard, 83]

$\Downarrow$

$\boxed{\text{Reduce back}}$

Proof's idea

Transform the problem

First transformation:

- Elimination of linear polynomials:
  - new system with a degree mean ≥ 2
- Denote it by $f_1, \ldots, f_k$ (abuse of notation)

Proof's idea

$\boxed{\text{Transform the problem}}$

Second transformation:

- Change of polynomials:
  - $f_1, \ldots, f_n$: a regular sequence
- If $|K| < \infty$ we do this change in $K(\alpha)$

Proof's idea

Transform the problem

Third transformation:

- Homogenization:

Proof's idea

Transform the problem

Third transformation:

- Homogenization:
  - $F_i = x_0^{\deg(f_i)} f_i(\frac{x_1}{x_0}, \dots, \frac{x_n}{x_0})$
  - $f_i = F_i(1, x_1, \dots, x_n)$

Proof's idea

Transform the problem

Third transformation:

- Homogenization:
  - $F_i = x_0^{\deg(f_i)} f_i(\frac{x_1}{x_0}, \dots, \frac{x_n}{x_0})$
  - $f_i = F_i(1, x_1, \dots, x_n)$

Problem: Introduces components $\subset \{x_0 = 0\}$:

- These "alien" components may have any dimension
- Thus one may not apply directly [Lazard, 83]

Proof's idea

Transform the problem

Fourth transformation:

- Deformation:
  - $G_i = (1-s)F_i + s x_i^{d_i}$ (s new indeterminate)

Proof's idea

Transform the problem

Fourth transformation:

- Deformation:
  - $G_i = (1-s)F_i + s x_i^{d_i}$ (s new indeterminate)
- Apply [Lazard, 83] for the G_i in $K(s)[x_0, \ldots, x_n]$

Proof's idea

Transform the problem

Fourth transformation:

- Deformation:
  - $G_i = (1-s)F_i + s x_i^{d_i}$ (s new indeterminate)
- Apply [Lazard, 83] for the G_i in $K(s)[x_0, \ldots, x_n]$

Problem: How to descend back to $K[x_1, \ldots, x_n]$?

Proof's idea

Transform the problem

Fourth transformation:

- Deformation:
  - $G_i = (1-s)F_i + s x_i^{d_i}$ (s new indeterminate)
- Apply [Lazard, 83] for the G_i in $K(s)[x_0, \ldots, x_n]$

Problem: How to descend back to $K[x_1, \ldots, x_n]$?

- With Gröbner basis: difficult to manage
- Thus we use "matrix Macaulay" in degree "regularity"

Proof's idea

$\boxed{\text{Transform the problem}}$

for using [Lazard, 81] and [Lazard, 83]

$\Downarrow$

$\boxed{\text{Reduce back}}$

Proof's idea

Reduce back

Substitution:

- $s = 0, x_0 = 1$

Proof's idea

$\boxed{\text{Reduce back}}$

Substitution:

- $s = 0, x_0 = 1$

$\boxed{\text{Problem: Divisions by } s \text{ in } K(s) \Longrightarrow \text{division by } 0}$

Proof's idea

$\boxed{\text{Reduce back}}$

Substitution:

- $s = 0, x_0 = 1$

$\boxed{\text{Problem: Divisions by } s \text{ in } K(s) \Longrightarrow \text{division by } 0}$

- Using Smith normal form over $K[s]$
  instead of Gauss-Jordan diagonalization in $K(s)$
  allow to divide by s the polynomials which are multiple of s
- Replacing $s \longrightarrow 0$ and $x_0 \longrightarrow 1$
  To show the conservation of Macaulay matrices properties

Macaulay matrix

☞ $S = K[s][x_0, \ldots, x_n]$

$\boxed{\text{Macaulay matrix in degree } \delta}$

$$\mathcal{M}ac_\delta(\langle G_1, \ldots, G_n \rangle) = \left[\begin{array}{c} \phi : S_{\delta - d_1} \times \cdots \times S_{\delta - d_n} \longrightarrow S_\delta \\ \text{where} \\ \phi(H_1, \ldots, H_n) = \sum_{i=1}^n H_i G_i \end{array} \right]$$

Macaulay matrix

☞ $S = K[s][x_0, \ldots, x_n]$

$$\boxed{\text{Macaulay matrix in degree } \delta}$$

$$\mathcal{M}ac_\delta(\langle G_1, \ldots, G_n \rangle) = \left[\begin{array}{c} \phi : S_{\delta-d_1} \times \cdots \times S_{\delta-d_n} \longrightarrow S_\delta \\ \text{where} \\ \phi(H_1, \ldots, H_n) = \sum_{i=1}^{n} H_i G_i \end{array} \right]$$

Quillen theorem: Includes all information about the ideal:

- Verify if "$\delta \geq$ regularity"
- Gröbner basis of I_δ
- $\ldots$

Algorithm

$\delta = n\mathrm{D} - n + 1$, $J = \langle G_1, \ldots, G_n \rangle$, $G_i = (1-s)F_i + s x_i^{d_i}$

- Compute the Smith normal form over $K[s]$ of $\mathcal{M}ac_\delta(J)$
- Divide by s, as much as possible,

  the columns of $\mathcal{M}ac_\delta(J)$

Algorithm

$\delta = n\mathrm{D} - n + 1$, $J = \langle G_1, \ldots, G_n \rangle$, $G_i = (1-s)F_i + s x_i^{d_i}$

- Compute the Smith normal form over $K[s]$ of $\mathcal{M}ac_\delta(J)$
- Divide by s, as much as possible,
  the columns of $\mathcal{M}ac_\delta(J)$
- $s \to 0 \Longrightarrow$ Macaulay matrix of $\tilde{I}$ s.t.

$$\langle F_1, \ldots, F_n \rangle \subset \tilde{I} \subset \langle F_1, \ldots, F_n \rangle : x_0^\infty$$

Algorithm

$\delta = n\mathrm{D} - n + 1$, $J = \langle G_1, \ldots, G_n \rangle$, $G_i = (1-s)F_i + s x_i^{d_i}$

- Compute the Smith normal form over $K[s]$ of $\mathcal{M}ac_\delta(J)$
- Divide by s, as much as possible,
  the columns of $\mathcal{M}ac_\delta(J)$
- $s \to 0 \Longrightarrow$ Macaulay matrix of $\tilde{I}$ s.t.

$$\langle F_1, \ldots, F_n \rangle \subset \tilde{I} \subset \langle F_1, \ldots, F_n \rangle : x_0^\infty$$

- Macaulay matrix of $\langle F_1, \ldots, F_n \rangle : x_0^\infty = \tilde{I} : x_0^\infty$:
  Gaussian elimination on a matrix formed by $D^{O(n)}$ of “Macaulay”

Algorithm

$\delta = n\mathrm{D} - n + 1$, $J = \langle G_1, \ldots, G_n \rangle$, $G_i = (1-s)F_i + s x_i^{d_i}$

- Compute the Smith normal form over $K[s]$ of $\mathcal{M}ac_\delta(J)$
- Divide by s, as much as possible,
  the columns of $\mathcal{M}ac_\delta(J)$
- $s \to 0 \Longrightarrow$ Macaulay matrix of $\tilde{I}$ s.t.

$$\langle F_1, \ldots, F_n \rangle \subset \tilde{I} \subset \langle F_1, \ldots, F_n \rangle : x_0^\infty$$

- Macaulay matrix of $\langle F_1, \ldots, F_n \rangle : x_0^\infty = \tilde{I} : x_0^\infty$:
  Gaussian elimination on a matrix formed by $D^{O(n)}$ of “Macaulay”
- $x_0 \to 1 \Longrightarrow$ the Gröbner basis of $\langle f_1, \ldots, f_n \rangle$

Algorithm

- Computing the basis of $\langle f_1, \ldots, f_n \rangle$ for any ordering
  - by [FGLM]

Algorithm

- Computing the basis of $\langle f_1, \ldots, f_n \rangle$ for any ordering
  - by [FGLM]
- If $k > n$:
  - compute the basis of the regular sequence $f_1, \ldots, f_n$
  - $f_{n+1}, \ldots, f_k$ used for up-to-date the basis
    by linear algebra (as [FGLM])

Conclusion

- An algorithm to compute the zero-dim Gröbner basis:
  - quasi-optimal complexity
  - bit complexity $\ll$ [Lakshman, 91]
  - arithmetic complexity $=$ [Lakshman, 91]
- This algorithm is not designed to be implemented:
  - does not verify the dimension zero
  - it uses the Smith normal form

whereas $\cdots$

  - F_5 (by Faugère) uses the echelon form on almost the smaller matrices (no counter-example yet known)