MMA 数学特論 I

Algorithms for polynomial systems: elimination & Gröbner bases 多項式系のアルゴリズム: グレブナー基底 & 消去法

Lecture V: The Buchberger Algorithm

Xavier Dahan

Gröbner bases $\text{exist} \rightarrow \text{Dickson Lemma} + \text{Hilbert finite basis (Lect. IV)}$

Gröbner bases are **useful** \rightarrow Ideal membership (Theo. 4), and several other applications (next lectures).

Gröbner bases $\textbf{exist} \rightarrow \text{Dickson Lemma} + \text{Hilbert finite basis (Lect. IV)}$ Gröbner bases are **useful** \rightarrow Ideal membership (Theo. 4), and several other applications (next lectures).

Let $F = \{f_1, f_2, \ldots, f_s\}$ polynomial system in $\mathbb{k}[X_1, \ldots, X_n]$, and let $I = \langle f_1, \ldots, f_s \rangle$ the ideal it generates. Problem 1: Is *F* a Gröbner basis for *I* (w.r.t. to a monomial order \prec)?

Problem 2: If not, how to compute a Gröbner basis for *I*, starting from *F* ?

Problem 3: Is it easy to compute a Gröbner basis ? (**efficiency**)

Gröbner bases $\textbf{exist} \rightarrow \text{Dickson Lemma} + \text{Hilbert finite basis (Lect. IV)}$ Gröbner bases are **useful** \rightarrow Ideal membership (Theo. 4), and several other applications (next lectures).

Let $F = \{f_1, f_2, \ldots, f_s\}$ polynomial system in $\mathbb{k}[X_1, \ldots, X_n]$, and let $I = \langle f_1, \ldots, f_s \rangle$ the ideal it generates.

Problem 1: Is *F* a Gröbner basis for *I* (w.r.t. to a monomial order \prec)?

Problem 2: If not, how to compute a Gröbner basis for I , starting from F ? *→* Answer: use "*S*-polynomials" and Buchberger's criterion.

Problem 3: Is it easy to compute a Gröbner basis ? (**efficiency**)

Gröbner bases $\textbf{exist} \rightarrow \text{Dickson Lemma} + \text{Hilbert finite basis (Lect. IV)}$ Gröbner bases are **useful** \rightarrow Ideal membership (Theo. 4), and several other applications (next lectures).

Let $F = \{f_1, f_2, \ldots, f_s\}$ polynomial system in $\mathbb{k}[X_1, \ldots, X_n]$, and let $I = \langle f_1, \ldots, f_s \rangle$ the ideal it generates.

Problem 1: Is *F* a Gröbner basis for *I* (w.r.t. to a monomial order \prec)?

Problem 2: If not, how to compute a Gröbner basis for I , starting from F ? *→* Answer: use "*S*-polynomials" and Buchberger's criterion.

Problem 3: Is it easy to compute a Gröbner bais ? (efficiency)

→ Answer: Very hard. Many improvements possible *→* still active research topic.

The problem

Let $F = \{f_1, \ldots, f_s\}$ be a finite set of polynomials, \prec a monomial order. If *F* is not a Gröbner basis for $I = \langle F \rangle$, then:

 $\exists f \in I$, but $LM(f) \notin \langle LM(F) \rangle$ ($\Leftrightarrow LM(f_i) \nmid LM(f), \forall i$).

 \rightarrow LM(*F*) is "too small" for being a Gröbner basis (\Leftrightarrow $\langle LM(F) \rangle \subsetneq \langle LM(I) \rangle$).

 \rightarrow *(graphic of the example on Slide 5, Lect. IV on the blackboard...)*

The problem

Let $F = \{f_1, \ldots, f_s\}$ be a finite set of polynomials, \prec a monomial order. If *F* is not a Gröbner basis for $I = \langle F \rangle$, then:

 $\exists f \in I$, but $LM(f) \notin \langle LM(F) \rangle$ ($\Leftrightarrow LM(f_i) \nmid LM(f), \forall i$).

 \rightarrow LM(*F*) is "too small" for being a Gröbner basis (\Leftrightarrow $\langle LM(F) \rangle \subsetneq \langle LM(I) \rangle$). \rightarrow *(graphic of the example on Slide 5, Lect. IV on the blackboard...)*

How to extend $LM(F)$? (Try to) find $f \in I$, such that $LM(f) \notin \langle LM(F) \rangle$.

$$
\implies f = \sum_{i=1}^{s} h_i f_i \text{ such that } \text{LM}(f) = \text{LM}\left(\sum_{i=1}^{s} f_i h_i\right) \prec \max_{1 \le i \le s} \text{LM}(f_i h_i) \ (\star)
$$

Remember that... $LM_{\prec}(a_1 + a_2) \preccurlyeq max\{LM_{\prec}(a_1), LM_{\prec}(a_2)\}\$, with equality if $\text{LM}(a_1) \neq \text{LM}(a_2) \ldots$ and that $\text{LM}(f) \prec \text{LM}(f_i) \Rightarrow \text{LM}(f_i) \nmid \text{LM}(f).$

Conclusion: There is a term cancellation idenitity in (\star) .

*S***-polynomials**

Definition 1 *Given two non-zero polynomials* $f, g \in \mathbb{k}[X_1, \ldots, X_n]$ *, and a monomial order* \prec *, let* $X^{\alpha} = LM_{\prec}(f)$ *, and* $X^{\beta} = LM_{\prec}(g)$ *.*

The least common multiple *of* X^{α} *and* X^{β} *is* X^{γ} *, where* $\gamma = (\max\{\alpha_1, \beta_1\}, \dots, \max\{\alpha_n, \beta_n\})$, denoted $\text{lcm}(\text{lmm}\prec(f), \text{lmm}\prec(g)) = X^{\gamma}$.

The polynomial $S_{\prec}(f,g) := \frac{X^{\gamma}}{H^{\gamma}}$ $\overline{\text{LT}_{\prec}(f)}$ $f - \frac{X^{\gamma}}{2}$ lt*≺*(*g*) *g , is called the* S-polynomial *of f* and *g* (if it is clear what is \prec , we use simply $S(f,g)$ instead of $S_\prec(f,g)$).

*S***-polynomials**

Definition 1 *Given two non-zero polynomials* $f, g \in \mathbb{k}[X_1, \ldots, X_n]$ *, and a monomial order* \prec *, let* $X^{\alpha} = LM_{\prec}(f)$ *, and* $X^{\beta} = LM_{\prec}(g)$ *.*

The least common multiple *of* X^{α} *and* X^{β} *is* X^{γ} *, where* $\gamma = (\max\{\alpha_1, \beta_1\}, \ldots, \max\{\alpha_n, \beta_n\}), \ denoted \ \text{lcm}(\text{lmm}\times(f), \text{lmm}\times(g)) = X^{\gamma}.$

The polynomial $S_{\prec}(f,g) := \frac{X^{\gamma}}{H^{\gamma}}$ LT _≺ (f) $f - \frac{X^{\gamma}}{2}$ lt*≺*(*g*) *g , is called the* S-polynomial *of f* and *g* (if it is clear what is \prec , we use simply $S(f,g)$ instead of $S_\prec(f,g)$).

Comment: The *S*-polynomials control the "term cancellation identities":

Proposition 1 Let $T = \sum_{i=1}^{s} c_i f_i$, with $c_i \in \mathbb{k}$, and $\delta = \text{mdeg}_{\prec}(f_i)$ for all i. *If* mdeg_{\prec}(*T*) \prec *δ*, *then there exists* $c_{j,k} \in \mathbb{R}$ *such that*

$$
T = \sum_{1 \leq j,k \leq s} c_{j,k} S_{\prec} (f_j,f_k) \cdot \text{Moreover } \mathrm{mdeg}_{\prec} (S_{\prec} (f_j,f_k)) \prec \delta.
$$

PROOF: *(On the balckboard...)* \Box

Main theorem: Buchberger's criterion

The previous Proposition 1 is important for the following criterion (Theo. 1). Before, a definition... Let $G = \{g_1, \ldots, g_s\}$ be a polynomial system and *≺* a monomial order:

Definition 2 *A polynomial f is said to* reduce to 0 modulo *G, denoted* $f \rightarrow_G 0$ *if there exists* $a_1, \ldots, a_s \in \mathbb{k}[X_1, \ldots, X_n]$ *such that:* $f = a_1 g_1 + \cdots + a_s g_s$, with $LM(a_i g_i) \preccurlyeq LM(f)$ if $a_i \neq 0$.

Remark: If there exists (at least) one permutation $\sigma \in \mathfrak{S}_s$, such that:

$$
NF_{\prec}(f, [g_{\sigma(1)}, g_{\sigma(2)}, \ldots, g_{\sigma(s)}]) = 0,
$$

then $f \rightarrow_G 0$, but the reciprocal is not true in general.

Theorem 1 *G is a Gröbner basis of* $\langle G \rangle$ *, iff for all pairs* $i \neq j$ *,* $S(g_i, g_j) \rightarrow_G 0$.

PROOF: (On the blackboard...)

Is a polynomial system a Gröbner basis?

This is the problem 1 of Introduction.

The Buchberger criterion (Theorem 1), implies this algorithm to decide if a polynomial system F is a Gröbner basis or not.

- # Inputs: A polynomial system $F = \{f_1, \ldots, f_s\}$. A monomial order \prec . # Output: true if *F* is a Gröbner basis for $\langle F \rangle$, false else.
	- 1: for $p, q \in F, p \neq q$ do
	- 2: if NF _{\prec} $(S$ _{\prec} (p, q) , $F) \neq 0$ then return false; end if
	- 3: end for
	- 4: return true

Remark: It is just to show the power of *S*-polynomials. Else, it is very inefficient in practice, and not very useful.

Part II: The Algorithm **Version 1**

Version 1: very naive and slow.

- # Inputs: Non-zero polynomial system $F = \{f_1, \ldots, f_s\}$. A monomial order *≺*.
- # Output: A Gröbner basis $G = \{g_1, \ldots, g_t\}$ for $\langle F \rangle$, w.r.t. \prec .

\n- 1:
$$
G \leftarrow F
$$
\n- 2: $d \circ \{ G' \leftarrow G$
\n- 3: $f \circ r \ p, q \in G', p \neq q$ do
\n- 4: $S \leftarrow \text{NF}(S(p, q), G') \text{ // computed for any sequence order of } G'$
\n- 5: $if S \neq 0$ then $G \leftarrow G \cup \{ S \}$; end if
\n- 6: end for
\n- 7: $\}$ until $(G = G') \text{ // repeat from Step 2}$
\n- 8: $return G$
\n

Correctness - Termination

Correctness: Claim 1: we always have $F \subset G \subset I$ *(proof on the blackboard. . .)*

So, if $\langle F \rangle = I$, then $\langle G \rangle = I$.

Claim 2: When $G = G'$ (\Leftrightarrow exit the do/until loop \Leftrightarrow end of algorithm), we have $S = \text{NF}(S(p, q), G') = 0$ for all $p \neq q$ in *G*. By Buchberger's criterion (Theo. 1), G is a Gröbner basis.

Termination: If $LM(G') = LM(G)$ then $G = G'$. We have $\langle LM(G')\rangle \subset \langle LM(G)\rangle$, so the sequence $\{LM(G')\}$ verifies the "ascending chain condition" (Definition 4, Lect. IV), in $\mathbb{k}[X_1,\ldots,X_n]$. Because it is Nœtherian (Lect. IV, Theo. 3), after a finite number of steps, we have $LM(G) = LM(G')$.

Efficiency: detect useless *S***-polynomial**

! Computing a division (or normal form) can be slow: the size of the numbers can grow a lot.

- **!!** If *S*(*p, q*) reduces to 0 modulo *G*, then nothing happens in the algorithm !
- \rightarrow computing the division of $S(p,q)$ that gives a 0 remainder is useless.
- *⇒* Need to decrease as much as possible the number of divisions of S-polynomials computed at Step 4 of the Algo. version 1 (Slide 7)

Efficiency: detect useless *S***-polynomial**

! Computing a division (or normal form) can be slow: the size of the numbers can grow a lot.

- **!!** If *S*(*p, q*) reduces to 0 modulo *G*, then nothing happens in the algorithm !
- \rightarrow computing the division of $S(p,q)$ that gives a 0 remainder is useless.
- *⇒* Need to decrease as much as possible the number of divisions of S-polynomials computed at Step 4 of the Algo. version 1 (Slide 7)

Unnecessary pairs (1): Since $S(p,q) = -S(q,p)$: pair (p,q) already tested ⇒ need not consider the pair (q, p) (see definition of set *B* at Step 1, next slide).

Unnecessary pairs (2): If $S(p,q) \rightarrow G$ ^{*,*} 0, then $S(p,q) \rightarrow G' \cup \{S(a,b)\}$ 0 for any *S*-polynomial of $a, b \in G'$.

 \rightarrow Hence, the pair (p, q) needs not to be kept in the set *B* of all indices of pairs to be tested (see Step 10, next slide).

Buchberger algorithm: Version 2

Inputs: A polynomial system $F = \{f_1, \ldots, f_s\}$ # Output: A Gröbner basis $G = \{g_1, \ldots, g_t\}$ for $I = \langle F \rangle$.

\n- 1:
$$
G \leftarrow F; t \leftarrow s
$$
\n- $B \leftarrow \{(i,j), 1 \leq i < j \leq s\}$ // indices of pairs f_i, f_j to be tested
\n- 2: $\text{while } B \neq \emptyset$ do
\n- 3: $\text{for } (i,j) \in B$ do
\n- 4: $S \leftarrow \text{NF}(S(f_j, f_i), G)$
\n- 6: $\text{if } S \neq 0 \text{ then}$ // the S-pol. has not a 0 remainder
\n- 7: $t \leftarrow t + 1$; $f_t \leftarrow S$
\n- 8: $G \leftarrow G \cup \{f_t\}$ // then we add this remainder to $G \ldots$
\n- 9: $B \leftarrow B \cup \{(i,t), 1 \leq i \leq t-1\}$ // and the new indices.
\n- 10: $\text{else } B \leftarrow B - \{(i,j)\}$; end if // else the pair of index $i, j \ldots$
\n- 11: end for ; end while \ldots will always reduced to 0
\n- 12: return G
\n

Another criterion to detect useless pairs

This Proposition 2 permits to detect some pairs of polynomials *p, q* such that $S(p,q)$ will reduce to 0 modulo *G*.

→ permits to avoid useless computations (see Slide 14).

Proposition 2 Let G be finite set of polynomials. For a pair $f, g \in G$ and a $monomial\ order \prec$, if $LCM(LM_{\prec}(f), LM_{\prec}(g)) = LM_{\prec}(f)LM_{\prec}(g)$, then S ^{\prec (*f, q*) \rightarrow *G* 0*.*}

Proof: *(On the blackboard. . .)*

Application: This criterion is easy to check. (comparing to do a division).

Buchberger: Version 2.1

Inputs: A polynomial system $F = \{f_1, \ldots, f_s\}$

Output: A Gröbner basis $G = \{g_1, \ldots, g_t\}$ for $I = \langle F \rangle$.

12: return *G*

Part III: Syzygies Module over a ring

Let *R* be a commutative ring with 1_R for unit element, with addition $+$ and multiplication *·*.

An abelian group $(M,+)$ is an *R***-module** if, there is a map: $R \times M \rightarrow M$, $(r, m) \mapsto rm$, such that:

•
$$
1_R m = m
$$

• $(r \cdot r')m = r(r'm) = r(r'm)$

$$
\bullet \ (r+r')m=rm+r'm \qquad \bullet \ r(m+m')=rm+rm'
$$

Facts: If *R* is a field then *R*-modules are the vector spaces over *R*. If *R* is not a field, then a module *M* **has no base** in general. An *R*-module *M* is finitely generated if there exists some elements m_1, \ldots, m_s in *M* such that $\forall m \in M$, $\exists r_1, \ldots, r_s$ elements in *R* with: $m = r_1 m_1 + \cdots + r_s m_s$.

Examples: Let $I \subset R$ be an ideal. The quotient ring R/I is an R -module...

Syzygy (1/3)

Given an *R*-module *M*, the *first syzygy module* or the *syzygies* of *M* on a set of generators (m_1, \ldots, m_s) is the kernel the following presentation of M :

$$
R^{s} \xrightarrow{\times (m_1, \ldots, m_s)} M \to 0,
$$

$$
(r_1, \ldots, r_s) \longmapsto r_1 m_1 + \cdots + r_s m_s.
$$

then $Syz(m_1, ..., m_s) := \{(r_1, ..., r_s) \in R^s | \sum_{s=1}^s r_s$ $a_i m_i = 0$, so that $M \simeq R^s/Syz(m_1, \ldots, m_s).$

Definition 3 Let $F = (f_1, \ldots, f_s)$ a family of *s* polynomials. We simply *denoted by* $Syz(F)$ *the* syzygies on the leading terms of F:

$$
Syz(\text{LT}(f_1), \ldots, \text{LT}(f_s)) := \{(h_1, \ldots, h_s) \in \mathbb{k}[X_1, \ldots, X_n]^s \mid \sum_i h_i \text{LT}(f_i) = 0\}.
$$

Syzygy (2/3)

Homogeneous syzygy in $Syz(F)$ of (multi)degree $\alpha \in \mathbb{N}^n$:

 $(c_1 X^{\alpha(1)}, \ldots, c_s X^{\alpha(s)})$, where $c_i \neq 0 \Rightarrow X^{\alpha(i)} \text{LM}(f_i) = X^{\alpha}$.

Lemma 1 *Every syzygy of Syz*(*F*) *can be written uniquely as a linear combination over* k *of homogeneous syzygies.*

PROOF: *(On the blackboard...)*

Proposition 3 Let $F = (f_1, \ldots, f_s)$ be a family of polynomials, and $Syz(F)$ *be the syzygy module on the leading terms of* F *. For* $1 \leq i \leq j \leq s$ *, consider the pair* f_i, f_j *of* F *, and let* $X^\gamma := \text{LCM}(\text{LM}(f_i), \text{LM}(f_j))$ *. Define* $\mathbf{e}_1 = (1, 0, \ldots, 0)$, $\mathbf{e}_2 = (0, 1, 0, \ldots), \ldots, \mathbf{e}_r = (\ldots, 0, 1)$ and *X^γ X^γ*

$$
S_{ij} := \frac{A}{LT(f_i)} \mathbf{e}_i - \frac{A}{LT(f_j)} \mathbf{e}_j \in (\mathbb{k}[X_1,\ldots,X_n])^r,
$$

The syzygies $\{S_{ij}\}_{1\leq i,j\leq s}$ generate $Syz(F)$ as a $\mathbb{K}[X_1,\ldots,X_n]$ *-module.*

Syzygy (3/3)

PROOF:First we must check that S_{ij} are effectively syzygies on the leading terms of F (easy).

Next, we must show that each syzygy $S \in Syz(F)$ can be written:

$$
S = \sum_{i < j} p_{ij} S_{ij}, \quad p_{ij} \in \mathbb{k}[X_1, \dots, X_n]
$$

By Lemma 1 of the previous slide, we can assume that *S* is homogeneous of (multi)degree α . A syzygy $S \in Syz(F)$ *must* have at least two non-zero components, say $c_i X^{\alpha(i)}$ and $c_j X^{\alpha(j)}$ with $i < j$. By definition, we have $X^{\alpha(i)}$ ^{LM} $(f_i) = X^{\alpha(j)}$ ^{LM} $(f_j) = X^{\alpha}$, so $X^{\gamma}|X^{\alpha}$.

Claim: $S - c_i \text{LC}(f_i) X^{\alpha - \gamma} S_{ij}$ has its *i*-th component equal to zero, so has more zero components than *S*. By repeating this, we obtain that *S* is a $\mathbb{K}[X_1, \ldots, X_n]$ -linear combination of the S_{ij} , as required. \Box

The syzygy criterion

It gives another refinement of the Buchberger criterion that precises Theorem 1.

Theorem 2 *Let* $G = \{g_1, \ldots, g_s\}$ *be a family of polynomials, and* $Syz(G)$ *the Syzygy module on the leading terms of G. Let S be a homogeneous generating set of Syz*(*G*)*. We have:*

G is a Gröbner basis iff for all $S \in S$, $S \cdot G = \sum_{i=1}^{s} h_i g_i \rightarrow_G 0$.

PROOF:*(On the blackboard...)*

Remark: If we choose $S = \{S_{ij}, i < j\}$, as indicated in Proposition 3, then $S_{ij} \cdot G = S(g_i, g_j)$. Hence, Theorem 1 is a special case of the above one.

Practically ? The advantage of using this criterion is the possiblity to take a *smaller* generating set for $Syz(G)$ than the ${S_{ij}}$.

→ then we can avoid more useless pairs than the criterion of Proposition 2.

Choosing a smaller generating set

1) Start form $\{S_{ij}, i < j\}$ for a generating set of $Syz(G)$.

2) Suppose we have constructed a (smaller generating set) $S \subset Syz(G)$.

Proposition 4 If $\text{LM}(g_{\ell})|\text{LCM}(\text{LM}(g_i), \text{LM}(g_j))|$ and $S_{i\ell}, S_{j\ell} \in \mathcal{S}$, then $\mathcal{S} - \{S_{ij}\}\$ is a (smaller) basis of $Syz(G)$.

PROOF:Suppose $i < j < \ell$, and let $X^{\gamma_{i\ell}} := \text{LCM}(\text{LM}(g_i), \text{LM}(g_{\ell}))$ (and also let $X^{\gamma_{j\ell}}$, $X^{\gamma_{ij}}$ for the corresponding LCM). By assumption, both $X^{\gamma_{j\ell}}$ and $X^{\gamma_{i\ell}}$ divides *X^γij* .

$$
S_{ij} = \frac{X^{\gamma_{ij}}}{X^{\gamma_{i\ell}}}S_{i\ell} - \frac{X^{\gamma_{ij}}}{X^{\gamma_{j\ell}}}S_{j\ell}
$$

so S_{ij} is generated by $S_{i\ell}$ and $S_{j\ell}$ and can be removed from S .

Aim: We want to reduce the number of *pairs* to test. Let $[i, j] = (i, j)$ if *i* < *j* and $[i, j] = (j, i)$ if $i > j$. Let *B* ⊂ { (i, j) , 1 ≤ $i < j$ ≤ *s*}, such that *{Sab,* (*a, b*) *∈ B}* generate *Syz*(*G*).

Buchberger algorithm: Version 3

Define the boolean $Criterion(f_i, f_j, B)$ as true if $[i, \ell]$ and $[j, \ell]$ are not in *B*, and if $LM(f_{\ell})|LCM(LM(f_i),LM(f_j))$ and false else.

1:
$$
G \leftarrow F
$$
; $B \leftarrow \{(i,j), 1 \leq i < j \leq s\}$; $t \leftarrow s$ \n2: while $B \neq \emptyset$ do\n3: for $(i,j) \in B$ do\n4: if $\text{LCM}(\text{LM}(f_i), \text{LM}(f_j)) \neq \text{LM}(f_i)\text{LM}(f_j)$ and $!Criterion(f_i, f_j, B)$ then\n5: $S \leftarrow \text{NF}(S(f_j, f_i), G)$ \n6: if $S \neq 0$ then\n7: $t \leftarrow t + 1$; $f_t \leftarrow S$ \n8: $G \leftarrow G \cup \{f_t\}$ \n9: $B \leftarrow B \cup \{(i, t), 1 \leq i \leq t - 1\}$ \n10: end if ; end if\n11: $B \leftarrow B - \{(i, j)\}$ \n12: end for ; end while ; return G

Conclusion: Remarks about efficiency

 \dots *still a lot of research to compute Gröbner bases quickly...*

 $(Buchberger, 1985)$, (Gebauer-Möller, 1988) \rightarrow "Normal strategy" for choosing pairs to reduce and good reductors (will give a zero quickly).

(Giovanni, Mora *et al.*, 1991) "Sugar" and "Double sugar" strategy, refinement and heuristics.

J.-C. Faugère. A new efficient algorithm for computing Gröbner bases (F_4). J. Pure Appl. Algebra, pp:75–83, (1999, updated 2002).

Gröbner bases for grevlex are usually faster to compute $(Bayer-Stillman, 1987)$ \rightarrow monomial order conversion algorithm (to compute a lex GB, first compute a grevlex one and *convert it* into a lex). (Faugère, Gianni *et al.*, 1993), FGLM, change of order by linear algebra, $(Collart, Kalkbrener et al., 1993 97)$, "Gröbner walk" on different orders.