Gröbner Basis of a Module over $K[x_1,\dots,x_n]$ and Polynomial Solutions of a System of Linear Equations

A. Furukawa*), T. Sasaki**), H. Kobayashi***)

*) Department of Mathematics, Tokyo Metropolitan University
Fukazawa, Setagaya-ku, Tokyo 158, Japan

**) The Institute of Physical and Chemical Research
Wako-shi, Saitama 351-01, Japan

***) Department of Mathematics, Nihon University
Kanda, Chiyoda-ku, Tokyo 101, Japan

-EXTENDED ABSTRACT-

Many computations relating polynomial ideals are reduced to calculating polynomial solutions of a system of linear equations with polynomial coefficients[1]. Zacharias[2] pointed out that Buchberger's algorithm[3] for Gröbner basis can be applied to solving such a linear equation. From the computational viewpoint, Zacharias' method seems to be much better than the previous methods. Hence, we have generalized his method to solve a system of equations directly. After completing the paper, we knew that similar works had been done by several authors[4,5]. This paper describes our method briefly.

§1. Definitions of monoideal and order

Let $\mathbb{Z}_0$ be the set of nonnegative integers, and $\mathbb{Z}_0^n$ the Cartesian product of $\mathbb{Z}_0$. For an element $A = (\alpha_1,\dots,\alpha_n)$ in $\mathbb{Z}_0$, we define $|A| = \alpha_1 + \cdots + \alpha_n$. We write $(0,\dots,0)$ as 0. Let $K[x_1,\dots,x_n]$ (abbreviated to $K[x]$) be the ring of polynomials in n variables with coefficients in a field K. We express f in $K[x]$ as $f = \sum_A a_A x^A$, where $A = (\alpha_1,\dots,\alpha_n)$, $a_A \in K$, and $x^A = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$. We call $|A|$ the degree of x^A, i.e., $\deg(x^A) = |A|$.

BY $\vec{A}$ we denote an r-tuple $(A_1,\dots,A_r)$. Let $\vec{S} = (S_1,\dots,S_r)$ and $\vec{T} = (T_1,\dots,T_r)$ be r-tuples of subsets of $\mathbb{Z}_0^n$. In particular, we write $(\phi,\dots,\phi)$ as $\vec{\phi}$. Then, union and intersection of $\vec{S}$ and $\vec{T}$ are defined as

$$\vec{S} \cup \vec{T} = (S_1 \cup T_1, \dots, S_r \cup T_r),$$
$$\vec{S} \cap \vec{T} = (S_1 \cap T_1, \dots, S_r \cap T_r).$$

Furthermore, if $\vec{A} = (A_1,\dots,A_r) \in (\mathbb{Z}_0^n)^r$ is such that $A_1 \in S_1$, ..., and $A_r \in S_r$, then we write $\vec{A} \in \vec{S}$.

Definition 1 [monoideal]. A subset I_M of $\mathbb{Z}_0^n$ is a monoideal if $I_M + \mathbb{Z}_0^n = I_M$. □

Definition 2 [total-degree lexicographic order $\rhd$ in $\mathbb{Z}_0^n$].

Let $A = (\alpha_1,\dots,\alpha_n)$, $B = (\beta_1,\dots,\beta_n) \in \mathbb{Z}_0^n$. We define $A \rhd B$ iff either $|A| > |B|$ or $|A| = |B|$ and there is an integer i, $1 \leqq i \leqq n$, such that $\alpha_j = \beta_j$ for all j, $1 \leqq j < i$, and $\alpha_i > \beta_i$. We define $A \trianglerighteq B$ iff $A \rhd B$ or $A = B$. □

Definition 3 [exponent set exs(f), leading exponent lex(f), and head term ht(f), of $f \in K[x]$].

For nonzero f in $K[x]$, we define

$\mathrm{exs}(f) = \langle A \in \mathbb{Z}_0^n \mid A \text{ in } f = \sum a_A x^A,\ a_A \neq 0 \rangle$,

$\mathrm{lex}(f) \in \mathrm{exs}(f)$, where

$\mathrm{lex}(f) \rhd$ any other element of exs(f),

$\mathrm{ht}(f)$ = a term $a_A x^A$ of f, where $A = \mathrm{lex}(f)$.

Similarly, $\mathrm{exs}(0) = \phi$, $\mathrm{lex}(0) = \phi$, and $\mathrm{ht}(0) = 0$, and we consider that $\phi \lhd (0,\dots,0)$. □

Definition 4 [highest-order smallest-suffix component order $\rhd$ in $(\mathbb{Z}_0^n)^r$].

Let $\vec{A} = (A_1,\dots,A_r)$ and $\vec{B} = (B_1,\dots,B_r)$ be any elements of $(\mathbb{Z}_0^n)^r$. We reorder the components of $\vec{A}$ and define $\vec{A}' = (A_{i_1},\dots,A_{i_r})$ as follows: $(i_1,\dots,i_r)$ =

$\{1,\dots,r\}$ and $A_{i_1} \trianglerighteq A_{i_2} \trianglerighteq \cdots \trianglerighteq A_{i_r}$, where $\ell < m$ for any (ℓ,m) such that $A_{i_\ell} = A_{i_m}$. Similarly, we define $\vec{B}' = (B_{j_1},\dots,B_{j_r})$ by reordering the components of $\vec{B}$. Then, we define $\vec{A} \triangleright \vec{B}$ iff there is an integer k, $1 \le k \le r$, such that $[A_{i_\ell} = B_{j_\ell}$ and $i_\ell = j_\ell$ for all ℓ, $1 \le \ell < k]$ and [either $A_{i_k} \triangleright B_{j_k}$ or $A_{i_k} = B_{j_k}$ with $i_k < j_k$]. We define $\vec{A} \trianglerighteq \vec{B}$ iff $\vec{A} \triangleright \vec{B}$ or $\vec{A} = \vec{B}$. □

<u>Note</u>. We can define an order in $(\mathbb{Z}_0^n)^r$ variously. When solving a system of linear equations, however, the efficiency of calculation depends crucially on choice of the order.

<u>Definition 5</u> [head term $ht(\vec{f})$, head position $hp(\vec{f})$, and rest $rest(\vec{f})$, of $\vec{f}$].

Let $\vec{f} = (f_1,\dots,f_r) \in (K[x])^r$, $A_i = lex(f_i)$, $i=1,\dots,r$, and A_k be the highest-order smallest-suffix component of $(A_1,\dots,A_r)$. We define

$$ht(\vec{f}) = ht(f_k),$$
$$hp(\vec{f}) = k,$$
$$rest(\vec{f}) = \vec{f} - (0,\dots,0,\underset{\substack{\uparrow \\ k\text{-th component}}}{ht(\vec{f})},0,\dots,0). \quad \square$$

If $\vec{f} \ne \vec{0}$, we have $lex(\vec{f}) \triangleright lex(rest(\vec{f}))$. In the following, we say $\vec{f}$ is higher order than $\vec{g}$ if $lex(\vec{f}) \triangleright lex(\vec{g})$.

<u>Definition 6</u> [exponent set, leading exponent, and lex-monoideal $lmo(\vec{f})$, of $\vec{f}$].

With the notations in Def. 5, we define

$$exs(\vec{f}) = (exs(f_1),\dots,exs(f_r)),$$
$$lex(\vec{f}) = (0,\dots,0,\ lex(f_k)\ ,0,\dots,0),$$
$$lmo(\vec{f}) = (\phi,\dots,\phi,\underset{\substack{\uparrow \\ k\text{-th component}}}{lex(f_k)+\mathbb{Z}_0^n},\phi,\dots,\phi). \quad \square$$

§2. Gröbner basis of a module over $K[x_1,\dots,x_n]$

By a module $\vec{\Gamma} = (\vec{f}_1,\dots,\vec{f}_s)$ with $\vec{f}_i \in (K[x])^r$, $i=1,\dots,s$, we mean the set $\langle h_1\vec{f}_1 + \cdots + h_s\vec{f}_s \mid h_i \in K[x],\ i=1,\dots,s\rangle$.

<u>Definition 7</u> [reducibility].

Let $F = \{\vec{f}_1,\dots,\vec{f}_s\}$ be a subset of $(K[x])^r$, and put $\vec{E} = \bigcup_{i=1}^{s} lmo(\vec{f}_i)$. An element $\vec{h}$ of $(K[x])^r$ is called <u>reducible</u> with respect to F if $exs(\vec{h}) \cap \vec{E} \ne \vec{\phi}$, and $\vec{h}$ is called <u>irreducible</u> w.r.t. F if $exs(\vec{h}) \cap \vec{E} = \vec{\phi}$. □

<u>Definition 8</u> [reduct].

With the notations in Def. 7, let $\vec{h}' \in (K[x])^r$, and $\vec{h}'$ is called a <u>reduct</u> of $\vec{h}$ w.r.t. F and written as $\vec{h} \xrightarrow[F]{} \vec{h}'$ if one of the followings holds:

(a) $\vec{h}' = \vec{h}$ when $\vec{h}$ is irreducible w.r.t. F,

(b) $\vec{h}' = \vec{h} - c\cdot x^A\vec{f}_k$ when $exs(\vec{h}) \cap lmo(\vec{f}_k) \ne \vec{\phi}$,

where c and A are determined as follows: let $ht(\vec{f}_k) = a_{A_k}x^{A_k}$, hence the $hp(\vec{f}_k)$-th component of $\vec{h}$ contains a term $b_{A+A_k}x^{A+A_k}$, then $c = b_{A+A_k}/a_{A_k}$. In the case of (b), this reduct is called a genuine (one-step) reduct w.r.t. $\vec{f}_k$. □

<u>Definition 9</u> [normal form].

Suppose $\vec{h}$ in $(K[x])^r$ is reduced successively as $\vec{h} \xrightarrow[F]{} \vec{h}' \xrightarrow[F]{} \cdots \xrightarrow[F]{} \underline{\vec{h}}$, and if $\underline{\vec{h}}$ is irreducible w.r.t. F then $\underline{\vec{h}}$ is called a <u>normal form</u> of $\vec{h}$ w.r.t. F. We denote the above reduction sequence by $\vec{h} \xrightarrow[F]{*} \underline{\vec{h}}$. □

<u>Definition 10</u> [S-polynomial].

Let $\vec{f}$ and $\vec{g}$ be elements of $(K[x])^r$, and let $ht(\vec{f}) = a_Ax^A$ and $ht(\vec{g}) = b_Bx^B$. The <u>S-polynomial</u> of $\vec{f}$ and $\vec{g}$, to be abbreviated to $Sp(\vec{f},\vec{g})$, is defined by

$$Sp(\vec{f},\vec{g}) = \begin{cases} u\cdot\vec{f} - (a_A/b_B)v\cdot\vec{g} & \text{if } hp(\vec{f}) = hp(\vec{g}), \\ \vec{0} & \text{otherwise,} \end{cases}$$

where u and v are monomials satisfying $LCM(x^A,x^B) = u\cdot x^A = v\cdot x^B$, with LCM the least common multiple. □

<u>Theorem 1</u>. Let $G = \{\vec{g}_1,\dots,\vec{g}_t\}$ be a Gröbner basis of a module $\vec{\Gamma}$ in $(K[x])^r$, and $\vec{h} \in (K[x])^r$. Let $\underline{\vec{h}}_1$ and $\underline{\vec{h}}_2$ be normal forms of $\vec{h}$ w.r.t. G, then $\underline{\vec{h}}_1 = \underline{\vec{h}}_2$.

<u>Theorem 2</u>. Let $\vec{\Gamma} = (\vec{g}_1,\dots,\vec{g}_t)$ be a module in $(K[x])^r$ and put $G = \{\vec{g}_1,\dots,\vec{g}_t\}$. If $Sp(\vec{g}_i,\vec{g}_j) \xrightarrow[G]{*} \vec{0}$ for any pair $(\vec{g}_i,\vec{g}_j)$, $i \ne j$, $1 \le i,j \le t$, then G is a Gröbner basis of $\vec{\Gamma}$.

<u>Procedure BUCHBERGER</u>

input: a module $\vec{\Gamma} = (\vec{f}_1,\dots,\vec{f}_s)$ in $(K[x])^r$.

output: a Gröbner basis $G = \{\vec{g}_1,\dots,\vec{g}_t\}$ of $\vec{\Gamma}$.

```
G := {g_1:=f_1, ..., g_s:=f_s};
P := {(g_i,g_j) | g_i,g_j ∈ G, i ≠ j, hp(g_i) = hp(g_j)};
while P ≠ φ do begin
  p_ij := a pair (g_i,g_j) in P;
  P := P - {p_ij};
    g := a normal form of Sp(g_i,g_j) w.r.t. G;
    if g ≠ 0 then begin

      P := P ∪ {(g_i,g) | hp(g_i) = hp(g), g_i ∈ G};
      G := G ∪ {g};
      end;
  end.
```

<u>Theorem 3</u>. The procedure BUCHBERGER terminates and it gives us a Gröbner basis of the module $\vec{\Gamma}$.

§3. Solutions of a system of linear equations

Let us consider to calculate the solutions $(y_1,\dots,y_r)$ of the following system of linear equations:

$$y_1\vec{f}_1 + \cdots + y_s\vec{f}_s = \vec{f}_0, \qquad (1)$$

where $\vec{f}_i = (f_{i1},\dots,f_{ir})$ in $(K[x])^r$, $i=0,\dots,s$.

First, we consider the homogeneous equations:

$$z_1\vec{g}_1 + \cdots + z_t\vec{g}_t = \vec{0}, \qquad (2)$$

where $G = \{\vec{g}_1,\dots,\vec{g}_t\}$ is a Gröbner basis of the module $\vec{\Gamma} = (\vec{f}_1,\dots,\vec{f}_s)$.

Let $\vec{g}_i$ and $\vec{g}_j$ satisfy $hp(\vec{g}_i) = hp(\vec{g}_j)$, and let $ht(\vec{g}_i) = a_{A_i}x^{A_i}$ and $ht(\vec{g}_j) = b_{B_j}x^{B_j}$ with $a_{A_i}, b_{B_j} \in K$. Then, since G is a Gröbner basis, we have $Sp(\vec{g}_i,\vec{g}_j) \xrightarrow[G]{} \vec{0}$. This reduction relation can be rewritten as

$$u_{ij}\vec{g}_i - (a_{A_i}/b_{B_j})\cdot v_{ij}\vec{g}_j = \sum_{k=1}^{t} w_{ij,k}\vec{g}_k,$$

where u_{ij} and v_{ij} are monomials satisfying $LCM(x^{A_i},x^{B_j}) = u_{ij}x^{A_i} = v_{ij}x^{B_j}$ and $w_{ij,k}$ satisfies $lex(w_{ij,k}\vec{g}_k) \lhd lex(u_{ij}\vec{g}_i) = lex(v_{ij}\vec{g}_j)$.

Proposition 1. With the above notations, let $\vec{z}^{(ij)} = (w_{ij,1},\dots, w_{ij,i}-u_{ij}, \dots, w_{ij,j}+(a_{A_i}/b_{B_j})\cdot v_{ij}, \dots, w_{ij,t})$, (3) where we assumed $i<j$. Then, $(z_1,\dots,z_t) = \vec{z}^{(ij)}$ is the lowest order solution of (2) satisfying

$$lex(z_i\vec{g}_i) = lex(z_j\vec{g}_j) \rhd lex(z_k\vec{g}_k)$$

for all $k \neq i,j$.

Theorem 4. With the above notations,

$$\langle (z_1,\dots,z_t) = \vec{z}^{(ij)} \mid hp(\vec{g}_i) = hp(\vec{g}_j),\ i<j \rangle$$

constitutes the set of generators of the polynomial solutions of (2).

Now, consider the system (1). We note that $G = \{\vec{g}_1,\dots,\vec{g}_t\}$ is a Gröbner basis of $\vec{\Gamma} = (\vec{f}_1,\dots,\vec{f}_s)$. Tracing the construction of G from $\{\vec{f}_1,\dots,\vec{f}_s\}$, we can calculate polynomials $q_{ji} \in K[x]$, $j=1,\dots,t$, $i=1,\dots,s$, such that

$$\vec{g}_j = \sum_{i=1}^{s} q_{ji}\vec{f}_i, \quad j=1,\dots,t. \qquad (4)$$

Conversely, the reduction $\vec{f}_i \xrightarrow[G]{} \vec{0}$ gives us polynomials $p_{ij} \in K[x]$, $i=1,\dots,s$, $j=1,\dots,t$, such that

$$\vec{f}_i = \sum_{j=1}^{t} p_{ij}\vec{g}_j, \quad i=1,\dots,s. \qquad (5)$$

Using (5), we can transform the system (1) with $\vec{f}_0 = \vec{0}$

Algorithm SOLVE

Input: $y_1\vec{f}_1 + \cdots + y_s\vec{f}_s = \vec{f}_0$, $\vec{f}_i \in (K[x])^n$.

Output: generators of the solutions y_i, $1 \leq s$, in $K[x]$.

Step 1. By using procedure BUCHBERGER, calculate a Gröbner basis $G = \{\vec{g}_1,\dots,\vec{g}_t\}$ of $(\vec{f}_1,\dots,\vec{f}_s)$;

Step 2. By reducing $\vec{f}_0$ w.r.t. G, calculate polynomials $z_j^{(0)}$, $j=1,\dots,t$, such that $\vec{f}_0 = \sum_{j=1}^{t} z_j^{(0)}\vec{g}_j + \underline{\vec{f}}_0$, $\underline{\vec{f}}_0$ is irreducible w.r.t. G; If $\underline{\vec{f}}_0 \neq \vec{0}$ then return ϕ (no solution), else let $\vec{z}^{(0)} = (z_1^{(0)},\dots,z_t^{(0)})$ (particular solution);

Step 3. Calculate polynomials q_{ji} satisfying (4);

Step 4. For every pair $(\vec{g}_i,\vec{g}_j)$ in G such that $hp(\vec{g}_i) = hp(\vec{g}_j)$, calculate the generator $\vec{z}^{(ij)}$ by the formula (3);

Then, transform these generators and particular solution $\vec{z}^{(0)}$ by formula (7) and return the results.

References

[1] A. Seidenberg, "Constructions in algebra", Trans. Amer. Math. Soc. 197, pp.273-313, 1974.

[2] G. Zacharias, "Generalized Gröbner basis in commutative polynomial rings", Bachelor Thesis, M.I.T., Dept. of Comp. Scie., 1978.

[3] B. Buchberger, "Basic features and development of the critical-pair/completion procedure", Proc. First Intern'l Conf. on "Rewriting Techniques and Applications", France, 1985, Springer Lecture Notes Comput. Sci..

[4] H. M. Möller and F. Mora, "New constructive methods in classical ideal theory" Journ. Algebra, 99, 1986.

[5] D. Bayer, F. Mora, H. M. Möller, private communication.