Lecture Notes on Computer Algebra

Ziming Li

Abstract

These notes record seven lectures given in the computer algebra course in the fall of 2004. The theory of subresultants is *not* required for the final exam due to its complicated constructions.

1 Chinese remainder algorithm

1.1 Modular arithmetic

Let m be a positive integer greater than one. We discuss basic operations in $\mathbb{Z}/(m)$. The elements in $\mathbb{Z}/(m)$ can be represented in two ways:

- $\{0, 1, \ldots, m-1\}$ (nonnegative representation);
- $\{a \in \mathbb{Z} \mid -\frac{m}{2} < a \leq \frac{m}{2}\}$ (symmetric representation).

Let us fix a representation $\mathbb{Z}_m$ e.g. the nonnegative representation. Via division we have a canonical simplifier H_m from $\mathbb{Z}$ onto $\mathbb{Z}_m$.

Proposition 1 *For $n \in \mathbb{Z}$ with $L(n) \geq L(m)$, the time for computing $H_m(n)$ is dominated by $\mathcal{O}(L(m)(L(n) - L(m) + 1))$, where $L(n)$ denotes the length of the integer n.*

Proof. It follows from the cost estimation of integral division (see Theorem 2.1.7 in [8]). ∎

For two elements a, b in $\mathbb{Z}_m$, we compute $a + b$, $a - b$, and ab as if they were integers, and then apply H_m to the results. Thus, the costs for computing $(a+b)$ and $(a-b)$ are dominated by $\mathcal{O}(L(m))$, and the cost for ab dominated by $\mathcal{O}(L(m)^2)$. The latter estimation kills any hope to use fast multiplications for modular multiplication.

A particular operation in $\mathbb{Z}_m$ is the inversion. Given $a \in \mathbb{Z}_m$ with $\gcd(a, m) = 1$, compute an element $b \in \mathbb{Z}_m$ such that $ab \equiv 1 \bmod m$.

This can be done by half-extended Euclidean algorithm for m and a. Thus, the time complexity for computing the inverse is $\mathcal{O}(L(m)^2)$ (see Theorem 2.1.9 in [8]).

Example 1 *Compute the inverse of* 4 mod 7*.*

$$\{7 = 4 + 3, \quad v = 0 - 1 \cdot 1 = -1,\} \quad \{4 = 3 + 1 \quad 1 - 1 \cdot (-1) = 2\}.$$

It follows that

$$2 \cdot 4 \equiv 1 \mod 7.$$

At last, we consider modular exponentiation. Given $a \in \mathbb{Z}_m$ and $n \in \mathbb{N}$, compute $a^n \bmod m$. The time to compute $a^n \bmod m$ by repeated multiplication will be proportional to $nL(m)^2$.

Proposition 2 *The time to compute a^n mod m is $L(n)L(m)^2$.*

Proof. By binary exponentiation we get the recursion

$$E_n \sim E_{\lfloor n/2 \rfloor} + 2L(m)^2 \quad \text{with} \quad E_2 = L(m)^2.$$

It follows that $E_n \sim L(n)L(m)^2$. ■

Exercise. What is the time to compute a^n, where $a \in \mathbb{N}$, by binary exponentiation ?

1.2 Modular homomorphisms

The canonical simplifier H_m is a ring homomorphism from $\mathbb{Z}$ to $\mathbb{Z}/(m)$. Let R be a ring. For $r \in R$, $\phi_r : R[x] \to R$ is an evaluation that sends $f(x)$ to $f(r)$. Write

$$f(x) = (\ldots(f_n x + f_{n-1})x + \ldots)x + f_0,$$

which is called Horner's form of f. It takes n multiplications and n additions to compute $f(r)$.

Example 2 *Two ways to map $\mathbb{Z}[x]$ to $\mathbb{Z}/(m)$:*

$$\mathbb{Z}[x] \to \mathbb{Z}/(m)[x] \to \mathbb{Z}/(m)$$

and

$$\mathbb{Z}[x] \to \mathbb{Z} \to \mathbb{Z}/(m).$$

Horner's rule is good for evaluating a polynomial at a single point. When evaluating a polynomial at many points, e.g., the set of consecutive integers, there are better ways (see [7]).

Exercise. What is the time to evaluate a polynomial $f \in \mathbb{Z}[x]$ modulo m ?

1.3 The integer Chinese remainder algorithm

The Chinese Remainder Problem: Given $m_1, \ldots, m_k \in \mathbb{Z}^+$ and $r_1, \ldots r_k \in \mathbb{N}$, find integers that solve the congruence system:

$$\begin{cases} x \equiv r_1 \mod m_1 \\ \qquad \cdots \\ x \equiv r_k \mod m_k \end{cases} \tag{1}$$

Congruence systems often appears in the combining phase in the applications of the divide-conquer-combine principle.

Theorem 1 *If the nonunit moduli $m_1, \ldots, m_k$ are pairwise co-prime, then (1) has a unique integral solution r in $[0, M)$, where $M = \prod_{i=1}^{k} m_i$. Moreover, every integral solution of (1) is in form $(nM + r)$, for all $n \in \mathbb{Z}$.*

Proof. Let $M_i = M/m_i$ for $i = 1, \ldots, k$. Since $m_1, \ldots, m_k$ are pairwise co-prime, $\gcd(M_1, \ldots, M_k) = 1$. It follows that there are $a_1, \ldots, a_k \in \mathbb{Z}$ such that

$$a_1 M_1 + \cdots a_k M_k = 1.$$

Consequently,

$$a_i M_i \equiv 1 \mod m_i, \quad \text{for } i = 1, \ldots, k. \tag{2}$$

Set $R = r_1 a_1 M_1 + \cdots + r_k a_k M_k$. By (2) we deduce

$$R \equiv a_i r_i M_i \equiv r_i \mod m_i \quad \text{for } i = 1, \ldots, k.$$

Let r be the remainder of R and M. Then r is a solution of (1) in $[0, M)$. If r' is another solution of (1) in $[0, M)$, and suppose that $r - r' \geq 0$, then $(r - r') \equiv 0 \bmod m_i$ for $i = 1, \ldots, k$. Since the m_i's are pairwise co-prime, $(r - r')$ is divisible by M, so r is equal to r'. This shows the uniqueness. The last conclusion is obvious. ■

Although the above proof gives an algorithm for solving (1) in case of pairwise co-prime moduli, it is inefficient to work with $M_1, \ldots, M_k$, which are significantly large than the given ones. The key idea to solve (1) efficiently is to use small moduli whenever possible.

CRA2 Given nonunit $m_1, m_2 \in \mathbb{Z}^+$ with $\gcd(m_1, m_2) = 1$, and $r_1, r_2 \in \mathbb{N}$, compute the integer $r \in \mathbb{N}$ in $[0, m_1 m_2)$ solving the congruence system

$$x \equiv r_1 \bmod m_1 \quad \text{and} \quad x \equiv r_2 \bmod m_2.$$

1. [*compute inverse*] $c := m_1^{-1} \bmod m_2$;

2. [*modulo* m_2] $r_1' := \mathrm{rem}(r_1,\, m_2)$; $\quad s := \mathrm{rem}\left((r_2 - r_1')c,\, m_2\right)$;

3. [*build up the solution*] $r := r_1 + sm_1$;

CRA Given nonunit $m_1, m_2, \ldots, m_k \in \mathbb{Z}^+$ with $\gcd(m_i, m_j) = 1$ $(i \neq j)$, and $r_1, r_2, \ldots, r_k \in \mathbb{N}$, compute the integer $r \in \mathbb{N}$ in $[0,\ m_1 m_2 \cdots m_k)$ solving the congruence system (1).

1. [*initialize*] $M := m_1$; $r := r_1$;

2. [*loop*] for $i = 2, \ldots, k$, call **CRA2** with moduli M, m_i and residues r, r_i; set the solution to be r, and update M to be Mm_i.

3. [*return*] return r;

If all the moduli and residues are of length one, which is the usual case in practice, then the cost of **CRA2** in the loop of **CRA** is dominated by computing $H_{m_i}(M)$ proportional to $L(M) = i$. Thus, the overall cost of **CRA** is $\mathcal{O}(k^2)$.

Example 3 *Solve*

$$\{x \equiv 3 \bmod 4, \quad x \equiv 5 \bmod 7, \quad x \equiv 2 \bmod 3\}.$$

First, solve

$$\{x \equiv 3 \bmod 4, \quad x \equiv 5 \bmod 7\}.$$

$4^{-1} \bmod 7 = 2$. $x = 3 + 2 \cdot 4 \cdot (5 - 3) \bmod 7 = 19$.

Second, solve

$$\{x \equiv 19 \bmod 28, \quad x \equiv 2 \bmod 3\}.$$

$28^{-1} \bmod 3 = 1$. $x = 19 + 28 \cdot 1 \cdot (2 - 19) = -457 \equiv 47 \bmod 84$.

Exercise. Let m_1 and m_2 be two nonunit moduli with $g = \gcd(m_1, m_2)$. Let r_1 and r_2 be two residues such that $r_1 \equiv r_2 \bmod g$. Find a method to solve the congruence system

$$\{x \equiv r_1 \bmod m_1 \quad \text{and} \quad x \equiv r_2 \bmod m_2\}.$$

1.4 Interpolation

Let F be a field and $E = (e_1, v_1), \ldots, (e_k, v_k) \subset F \times F$ with $e_i \neq e_j$ $(i \neq j)$. We want to find a polynomial $f \in F[x]$ with degree less than k such that $f(e_i) = v_i$ for $i = 1, \ldots, k$. This problem can be stated in terms of a congruence system:

$$\begin{cases} f \equiv v_1 \mod (x - e_1) \\ \qquad \cdots \\ f \equiv v_k \mod (x - e_k) \end{cases} \tag{3}$$

Since the ideals $(x - e_1), \ldots, (x - e_k)$ are pairwise co-maximal and $F[x]$ is an Euclidean domain, the idea of CRA applies to solving (3). The proof of Theorem 1 leads to the Lagrangian formula for polynomial interpolation:

$$f = v_1 L_1 + \cdots v_k L_k \tag{4}$$

where $L_i = P_i(x)/P_i(e_i)$ with

$$P_i = \frac{\prod_{j=1}^{k}(x - e_j)}{x - e_i}.$$

Using the CRA we get

$$f = u_0 + u_1(x - e_1) + u_2(x_1 - e_1)(x_2 - e_2) + \cdots + u_{n-1} \prod_{i=1}^{n}(x - e_i),$$

where the u_i's are elements of F to be determined in the polynomial version of CRA.

The Hermite interpolation can also be stated as follows: Find a polynomial f with degree less than $(n_1 + \cdots + n_k)$ such that

$$\begin{cases} f \equiv v_1 \mod (x - e_1)^{n_1} \\ \qquad \cdots \\ f \equiv v_k \mod (x - e_k)^{n_k} \end{cases} \tag{5}$$

Thus, the polynomial version of CRA solves the Hermite interpolation problem, because the ideals

$$((x - e_1)^{n_1}), \ldots, ((x - e_k)^{n_k})$$

are co-maximal.

1.5 Additional notes

Modular arithmetic and the integer Chinese remainder algorithm are basic tools for efficient implementations of computer algebra algorithms. There is a very nice introduction to arithmetic in basic algebraic domains [3]. This paper and [7] are best materials for understanding basic arithmetic in computer algebra. The Chinese remainder theorem (problem) has an ideal-theoretic version. The Chinese remainder algorithm can be performed over an abstract Euclidean domain. Nonetheless, it is most important to understand this topic over integers. A good implementation of CRA requires a lot of careful work. For more details, the reader is referred to [8, page 57] and [6, page 176].

2 Reconstruction for rational functions and numbers

2.1 Rational function reconstruction

RFR Problem. *Let F be a field and M a polynomial with degree $n > 0$ over F. Let r be a residue in $F[x]/(M)$. For a given $k \in \{0, \ldots, n\}$, compute $a, b \in F[x]$ such that*

$$a \equiv b\,r \bmod M, \ \gcd(b, M) = 1, \ \deg\, a < k, \ \deg\, b \leq n - k. \tag{6}$$

The condition $\gcd(b, M) = 1$ implies that the polynomial b is invertible in $F[x]/(M)$, so that $ab^{-1} \equiv r \bmod M$.

Recall some properties of the extended Euclidean algorithm. For two polynomials $A_1, A_2 \in F[x]$ with $\deg A_1 \geq \deg A_2 > 0$. Applying the extended Euclidean algorithm to A_1, A_2, we get the following four sequences

1. Remainder sequence: A_1, A_2, A_3, $\cdots$, A_s with $A_i = \text{rem}(A_{i-2}, A_{i-1})$, for $i = 3, 4, \ldots, s$, and $\text{rem}(A_{s-1}, A_s) = 0$.

2. Quotient sequence: Q_3, $\ldots$, Q_s with $A_{i-2} = Q_i A_{i-1} + A_i$.

3. The first co-sequence: $U_1 = 1$, $U_2 = 0$, U_3, $\ldots$, U_s, and the second co-sequence: $V_1 = 0$, $V_2 = 0$, V_1, $\ldots$, V_s, with

$$U_i A_1 + V_i A_2 = A_i \quad \text{for } i = 1, \ldots, s.$$

Let $\deg A_i = d_i$ for $i = 1, \ldots, s$. We have the following degree sequences: $\deg Q_i = d_{i-2} - d_{i-1}$, $\deg U_i = d_2 - \deg A_{i-1}$ and $V_i = d_1 - \deg A_{i-1}$, where $i = 3, \ldots, n$. In addition $\gcd(U_i, V_i) = 1$ for $i = 1, \ldots, n$.

The next lemma is a consequence of the properties of these sequences (see [5])

Lemma 1 *Let A_1 and A_2 be two polynomials in $F[x]$ with $\deg A_1 = d_1 > \deg A_2 \geq d_2 > 0$. Suppose that $W = UA_1 + VA_2$ where $W, U, V \in F[x]$ and, moreover, $\deg W + \deg V < d_1$. Let $\{A_i\}$, $\{U_i\}$ and $\{V_i\}$ be the remainder sequence, the first and second co-sequences, respectively. If $\deg A_j \leq \deg W < \deg A_{j-1}$, then there exists $h \in F[x]$ such that*

$$W = hA_j, \quad U = hU_j \quad \text{and} \quad V = hV_j.$$

Proof. First, we claim that $UV_j = VU_j$. For otherwise, one could solve the system $\{UA_1 + VA_2 = W, U_jA_1 + V_jA_2 = A_j\}$ for A_1 and A_2 to get

$$(UV_j - VU_j)A_1 = WV_j - VA_j. \tag{7}$$

Let d be the degree of the right-hand side of (7). Then

$$\begin{aligned} d &\leq \max(\deg W + \deg V_j, \deg V + \deg A_j) \\ &\leq \max(\deg W + d_1 - \deg A_{j-1}, \deg V + \deg A_j) \\ &\leq \max(d_1, \deg V + \deg A_j) \\ &< \max(d_1, d_1 + \deg A_j - \deg W) \\ &= d_1. \end{aligned}$$

But the left hand-side of (7) is of degree no less than n, a contradiction. By the claim and the fact $\gcd(U_j, V_j) = 1$ we deduce that U is divisible by U_j. Let $U = hU_j$. Then $V = hV_j$ by the claim, and, hence, $W = hA_j$. ■

RFR: Given a modulus $M \in F[x]$ with degree $n > 0$, a residue $r \in F[x]/(M)$ as a polynomial with degree less than n, and $k \in \{0, 1, \ldots, n\}$, compute a pair (a, b) in $F[x]$ satisfying the conditions (6). If no such pair exists, FAIL, meaning there does not exist such a solution

```
1. [Initialize.] $A_1 := M$; $A_2 := r$; $V_1 := 0$; $V_2 := 1$; $i := 2$;
```

2. [*loop.*]

```
while true do
    if $\deg V_i > n - k$ then return(FAIL);
    if $\deg A_i < k$ and $\gcd(A_i, V_i) = 1$ then return $((A_i,\ V_i))$;
    $Q :=$ polynomial quotient of $A_{i-1}$ and $A_i$;
    $A_{i+1} := A_{i-1} - QA_i$; $V_{i+1} := V_{i-1} - QV_i$; $i := i + 1$;
```

We present two applications of RFR.

Cauchy Interpolation. Given $\{(e_1, v_1), \ldots, (e_n, v_n)\} \subset F \times F$ with $e_i \neq e_j$, and $k \in \{0, \ldots, n\}$, find $f, g \in F[x]$ such that

$$g(e_i) \neq 0,\ \frac{f(e_i)}{g(e_i)} = v_i,\ \deg f < k,\ \deg g < n - k. \tag{8}$$

Let h be the polynomial interpolating the set $\{(e_1, v_1), \ldots, (e_n, v_n)\}$ and $M = \prod_{i=1}^{n}(x - e_i)$. Such f and g exist if and only if $f \equiv gh \bmod (x - e_i)$ for $i = 1,\ \ldots,\ n$. Equivalently, there is a pair $(f,\ g)$ such that

$$\gcd(g,\ M) = 1,\ f \equiv gh \bmod M,\ \deg f < k,\ \deg g \leq n - k. \tag{9}$$

Padé Approximation. Given $s \in F[[x]]$, a positive integer n, and $k \in \{0, 1, \ldots, n\}$, find $f, g \in F[x]$ such that

$$\gcd(g, x) = 1,\ f \equiv gs \bmod x^n,\ \deg f < k,\ \deg g < n - k. \tag{10}$$

Example 4 *Find rational solutions of*

$$F(y, y') = 4y^4 + 4y^3 + 4yy' - y^2 - y' + 4y^2y' + 2(y')^2 = 0. \tag{11}$$

Let $S = \frac{\partial F}{\partial y'}$. *Then*

$$S(y, y')y^{(i)} = G_i(y, y', \ldots, y^{(i-1)}) \quad i = 2, 3, 4, \ldots,$$

Find $y_0, y_1 \in C$ *such that* $F(y_0, y_1) = 0$ *and* $S(y_0, y_1) \neq 0$. *We can find the values of* $y^{(i)}$ *at the point around which the formal power series solution is computed by the above equation. By Theorem 3 in [4], the maximum of the degrees of the numerator and denominator of a rational solution of (11) is equal to two. So we need to find a formal power series solution up to order four to recover a rational solution.*

Setting $y_0 = 0$ and $y_1 = \frac{1}{2}$, we get

$$s = \frac{1}{2}x - \frac{1}{2}x^2 + \frac{1}{4}x^3 + o(x^4).$$

Solving the congruence $r \equiv s \bmod x^5$ with the constraints $(3, 2)$, we get a solution $r_1 = \frac{x}{x^2+2x+2}$.

Setting $y_0 = -1$ and $y_1 = 1$, we get

$$s = -1 + x + x^2 - x^3 - x^4 + o(x^4).$$

From this we recover a rational solution $r_2 = \frac{x-1}{x^2+1}$. One can verify that $r_1(x-1) = r_2$.

2.2 Rational number reconstruction

RNR Problem. *Given $M \in \mathbb{Z}^+$ with $M > 1$ and $r \in \mathbb{N}$, find $a, b \in \mathbb{Z}$, with $b > 0$, such that*

$$a \equiv b\,r \bmod M,\ \gcd(b, M) = 1,\ |a|, b < \sqrt{\frac{M}{2}}. \tag{12}$$

The condition $\gcd(b, M) = 1$ implies that the number b is invertible in $\mathbb{Z}/(M)$, so that $ab^{-1} \equiv r \bmod M$. The constraints $|a|, b < \sqrt{\frac{M}{2}}$ implies the uniqueness of the solutions of the RNR problems.

RNR: Given a modulus $M \in \mathbb{Z}$ with $M > 1$, a residue $r \in \mathbb{Z}/(M)$, compute a pair (a, b) in $\mathbb{Z}$, with $b > 0$, satisfying the conditions (6). If no such pair exists, FAIL, meaning there does not exist such a solution

```
1. [Initialize.] $A_1 := M$; $A_2 := r$; $V_1 := 0$; $V_2 := 1$; $i := 2$;

2. [loop.]

   while true do
       if $\deg V_i \geq \sqrt{M/2}$ then return(FAIL);
       if $\deg A_i < \sqrt{M/2}$ and $\gcd(A_i, V_i) = 1$ then return $((A_i,\ V_i))$;
       $Q :=$ quotient of $A_{i-1}$ and $A_i$;
       $A_{i+1} := A_{i-1} - QA_i$; $V_{i+1} := V_{i-1} - QV_i$; $i := i + 1$;
```

The correctness of the algorithm is proved in [10]. This algorithm is only for classroom use. We refer to [2] and [9] for more efficient algorithms for solving the RNR problem.

2.3 Partial fraction decomposition

Let $r = \frac{a}{b} \in F(x)$ with $\deg a < \deg b$. Let $b = \prod_{i=1}^{k} b_i$ and $\gcd(b_i, b_j) = 1$ for all i, j with $1 \le i < j \le n$. Then there exist $a_i \in F[x]$, with $\deg a_i < \deg b_i$, for all i with $1 \le i \le k$, such that

$$r = \frac{a_1}{b_1} + \frac{a_2}{b_2} + \cdots + \frac{a_k}{b_k}. \tag{13}$$

Let $B_i = b/b_i$ for all i with $1 \le i \le n$. It follows from (13) that

$$a = a_1 B_1 + a_2 B_2 + \cdots + a_k B_k.$$

Thus, $a \equiv a_i B_i \bmod b_i$. Since $\gcd(B_i, b_i) = 1$, we have $a_i \equiv a B_i^{-1} \bmod b_i$. The degree constraint $\deg a_i < \deg b_i$ implies the uniqueness of $a_1, \ldots, a_k$. The existence of the a_i's follows from the fact $\gcd(B_1, \ldots, B_k) = 1$.

Let $r = \frac{a}{p^k} \in F(x)$ with $\deg a < k \deg p$. Then there exist $a_i \in F[x]$, with $\deg a_i < \deg p$, for all i with $1 \le i \le k$, such that

$$r = \frac{a_1}{p} + \frac{a_2}{p^2} + \cdots + \frac{a_k}{p^k}. \tag{14}$$

We compute the p-adic expansion of r to get

$$a = a_k + a_{k-1} p + \cdots + a_1 p^{k-1}$$

by polynomial division.

This two processes enable us to compute partial fraction decompositions of rational functions and prove the uniqueness.

3 A modular algorithm for computing gcd's

I can't find my lecture note on this topic, even worse, don't remember whether I ever wrote it. Nonetheless, here is a piece of Maple code I wrote when preparing the lecture.

```
ModularGCD1 := proc(A, B, x, L)
local a, b, l, g, t, i, p, Ap, Bp, Gp, gp, H, dp, d,
    M, G, G0, r;
1. initialize
  (a, b) := (lcoeff(A, x), lcoeff(B, x));
  (l, g) := (ilcm(a, b), igcd(a, b));
2. prepare for loop
```

```
t := true;
i := 2;
while t do
  p := ithprime(i);
  i := i + 1;
  if irem(l, p) = 0 then
    t := true;
  else
    t := false;
end if;
end do;
(Ap, Bp) := A mod p, B mod p;
Gp := Gcd(Ap, Bp) mod p;
dp := degree(Gp, x); gp := g mod p;
if dp = 0 then return 1 end if;
(d, M, G) := dp, p, gp*Gp;
if nargs = 4 then L1 := [[[p, Gp], [M, G]]]; end if;
```

3. loop

```
while true do
  t := true;
  while t do
    p := ithprime(i);
    i := i + 1;
    if irem(l, p) = 0 then
      t := true;
    else
      t := false;
    end if;
  end do;
  (Ap, Bp) := A mod p, B mod p;
  Gp := Gcd(Ap, Bp) mod p;
  dp := degree(Gp, x); gp := g mod p;
  if dp = 0 then return 1 end if;
  if dp < d then
    (d, M, G) := dp, p, gp*Gp;
    if nargs = 4 then
      L1 := [[[p, Gp], [M, G]]];
    end if;
  else
```

```
      if dp = d then
        G0 := G;
        G := chrem([gp*Gp, G], [p, M]);
        M := M*p;
        if nargs = 4 then
          L1 := [op(L1), [[p, Gp], [M, G]]];
        end if;
        t := expand(mods(G0, M)-mods(G0,M));
        H := mods(G, M);
        if t = 0 then
          r := rem(A, H, x);
          if r = 0 then
            r := rem(B, H, x);
            if r = 0 then
              if nargs = 4 then
                L := L1;
              end if;
              return primpart(H, x);
            end if;
          end if;
        end if;
      end if;
    end if;
  end do;
end proc;
```

4 Subresultants

4.1 Definitions

Let R be a domain, and

$$\mathcal{A} : A_1, A_2, \ldots, A_m \tag{15}$$

be a sequence in $R[x]$. We denote by $\deg \mathcal{A}$ the maximum of the degrees of the members in $\mathcal{A}$. Let $\deg \mathcal{A} = n \geq 0$ and write A_i as

$$A_i = \sum_{j=0}^{n} a_{ij} X^j, \quad (1 \leq i \leq m) \tag{16}$$

where each of the a_{ij}'s belongs to R. The *matrix associated with* $\mathcal{A}$ is defined to be the $m \times (n+1)$ matrix whose entry in the ith row and jth column is the coefficient of x^{n+1-j} in A_i, for $i = 1, \ldots, m$, and $j = 1, \ldots, n+1$. In other words, the matrix associated with $\mathcal{A}$ is

$$\begin{pmatrix} a_{1n} & a_{1,n-1} & \cdots & a_{10} \\ a_{2n} & a_{2,n-1} & \cdots & a_{20} \\ \cdots & \cdots & \cdots & \cdots \\ a_{mn} & a_{m,n-1} & \cdots & a_{m0} \end{pmatrix}. \tag{17}$$

This matrix is denoted by $\text{mat}(A_1, A_2, \ldots A_m)$ or $\text{mat}(\mathcal{A})$.

Let M be a matrix over R with r rows and c columns. Assume that $r \leq c$. The determinant polynomial of M is defined to be

$$\text{detpol}(M) = \det(M_{r-c})x^{r-c} + \cdots + \det(M_1)x + \det(M_0),$$

where M_i consists of the first $(r-1)$ columns and the $(c-i)$th column of M, $i = c-r, c-r-1, \ldots, 0$.

Lemma 2 *Let a sequence $\mathcal{A}$ be given in (15) with $n \geq m-1$. Then*

$$\text{detpol}(\text{mat}(\mathcal{A})) = \det \begin{pmatrix} a_{1n} & a_{1,n-1} & \cdots & a_{1,n-m+1} & A_1 \\ a_{2n} & a_{2,n-1} & \cdots & a_{2,n-m+1} & A_2 \\ \cdots & \cdots & \cdots & \cdots & \\ a_{mn} & a_{m,n-1} & \cdots & a_{m,n-m+1} & A_m \end{pmatrix}.$$

In addition,

$$\text{detpol}(\text{mat}(\mathcal{A})) \in (A_1, \ldots, A_m).$$

Proof. The determinant representation follows from the expressions $A_i = \sum_{i=0}^{n} a_{ij}x^j$ for $i = 1, \ldots, m$. The second statement is proved by expanding the determinant representation according to its last column. ■

Let $f, g \in R[x]$ with $\deg f = m$ and $\deg g = n$. Assume that $m > 0$ and $n > 0$. For j with $0 \leq j < \min(m, n)$, we define the jth subresultant of f and g is defined to be

$$\text{sres}_j(f, g) = \text{detpol}(\text{mat}(x^{n-j-1}f, \ldots, xf, f,\ x^{m-j-1}g, \ldots, xg, g)). \tag{18}$$

Note that the matrix in (18) has $(m+n-2j)$ rows and $(m+n-j)$ columns. Therefore, $\text{sres}_j(f, g)$ is well-defined. The resultant of f and g is defined to be $\text{sres}_0(f, g)$. Some basic properties of subresultants are given below.

Proposition 3 *Let $f, g \in R[x]$ with $\deg f = m$ and $\deg g = n$. Assume that $m > 0$ and $n > 0$. Then*

1. $\deg sres_j(f, g) \leq j$*;*

2. *there exist $u_j, v_j \in R[x]$ with $\deg u_j < (n - j)$ and $\deg v_j < (m - j)$ such that*

$$u_j f + v_j g = \mathrm{sres}_j(f, g) \quad \text{for } i = 0, 1, \ldots, \min(m, n) - 1;$$

3. $\mathrm{res}(f, g) = 0$ *if and only if f and g have a nontrivial gcd in $F[x]$, where F is the quotient field of R.*

4. *If $m \geq n$, then*

$$\mathrm{prem}(f, g) = (-1)^{m-n+1} \mathrm{sres}_{n-1}(f, g).$$

Proof. The first property follows from the definition of subresultants.

By Lemma 2 $\mathrm{sres}_j(f, g)$ is equal to a determinant of order $(m + n - 2j)$ whose first $(m + n - 2j - 1)$ columns consist of elements of R, and the last column is

$$(x^{n-j-1} f, \ldots, f, \, x^{m-j-1} g, \ldots, g)^\tau.$$

The second property follows from the expansion of the determinant according to its last column.

The second property implies that $u_0 f + v_0 g = \mathrm{res}(f, g)$ where $\deg u_0 < m$ and $\deg v_0 < n$. If $\gcd(f, g)$ is of positive degree, then it divides $\mathrm{res}(f, g)$, and, so $\mathrm{res}(f, g) = 0$. Conversely, we have g divides $u_0 f$ in $F[x]$. Since $\deg u_0 < n$, $\gcd(f, g)$ is of positive degree.

Since $\mathrm{sres}_{n-1}(f, g) = \mathrm{detpol}(\mathrm{mat}(f, x^{m-n} g, \ldots, xg, g))$, the pseudo-remainder formula implies that

$$\mathrm{lc}(g)^{m-n+1} \mathrm{sres}_{n-1}(f, g) = \mathrm{detpol}(\mathrm{mat}(\mathrm{prem}(f, g), \, x^{m-n} g, \ldots, g)).$$

Moving $\mathrm{prem}(f, g)$ to the last row of the above determinant, we get

$$\mathrm{lc}(g)^{m-n+1} \mathrm{sres}_{n-1}(f, g) = (-1)^{m-n+1} \mathrm{detpol}(\mathrm{mat}(x^{m-n} g, \ldots, g, \, \mathrm{prem}(f, g))).$$

Therefore, $\mathrm{sres}_{n-1}(f, g) = (-1)^{m-n+1} \mathrm{prem}(f, g)$. ■

5 Row reduction formula

For notational convenience, for a polynomial sequence $\mathcal{A}$ we use $|A|$ to denote the determinant polynomial $\mathrm{detpol}(\mathrm{mat}(\mathcal{A}))$.

Lemma 3 *Let $A, B \in R[x]$ with respective degrees m and n. Assume that $m \geq n > 0$. If there exist u, v and w in R and F, G in $R[x]$ such that $uB = vF$ and $\mathrm{sres}_{n-1}(A, B) = wG$, then*

$$u^{m-i}\mathrm{lc}(B)^{(m-n+1)(n-i)}\mathrm{sres}_i(A, B) = v^{m-i}w^{n-i}|x^{m-i-1}F, \ldots F, x^{n-i-1}G, \ldots G|. \tag{19}$$

Proof. Let $C = \mathrm{prem}(A, B)$. Since

$$\mathrm{sres}_i(A, B) = |x^{n-i-1}A, \ldots, A, x^{m-i-1}B, \ldots, B|,$$

$$\mathrm{lc}(B)^{(m-n+1)(n-i)}\mathrm{sres}_i(A, B) = |x^{n-i-1}C, \ldots, C, x^{m-i-1}B, \ldots, B|.$$

It follows that

$$\mathrm{lc}(B)^{(m-n+1)(n-i)}\mathrm{sres}_i(A, B) = (-1)^{(n-i)(m-i)}|x^{m-i-1}B, \ldots, B, x^{n-i-1}C, \ldots, C|.$$

By Proposition 3 $C = (-1)^{m-n+1}\mathrm{sres}_i(A, B)$, we get

$$\mathrm{lc}(B)^{(m-n+1)(n-i)}\mathrm{sres}_i(A, uF/v) = |x^{m-i-1}B, \ldots, B, x^{n-i-1}wG, \ldots, wG|.$$

Hence

$$u^{m-i}\mathrm{lc}(B)^{(m-n+1)(n-i)}\mathrm{sres}_i(A, F) = v^{m-i}w^{n-i}|x^{m-i-1}B, \ldots, B, x^{n-i-1}G, \ldots, G|.$$

The lemma is proved. ∎

6 Subresultant theorem

Let A and B be given above. Define $\mathrm{sres}_n(A, B) = S_n$ and set $S_i = \mathrm{sres}_i(A, B)$ for $i = n-1, \ldots, 0$. The subresultant S_i is said to be *regular* if $\deg S_i = i$. Otherwise, S_i is said to be *defective*.

Lemma 4 *Let $p_i = \mathrm{lc}(S_i)$ for i with $0 \leq i \leq n$. Let $q_n = \mathrm{lc}(S_n)^{m-n}$ and $q_i = p_i$ for i with $0 \leq i \leq n$. If S_{j+1} is regular and $\deg S_j = r$, then*

1. *if $S_j = 0$ then $S_i = 0$ for all i with $0 \leq i \leq j$;*

2. if $S_j \neq 0$ then

$$S_i = 0 \quad \textit{for all } i \textit{ with } r+1 \leq i < j, \tag{20}$$

$$q_{j+1}^{j-r} S_r = q_j^{j-r} S_j \tag{21}$$

and

$$p_{j+1}^{r-i} q_{j+1}^{j-i} S_i = \operatorname{sres}_i(S_{j+1}, S_j) \quad \textit{for all } i \textit{ with } 0 \leq i < r. \tag{22}$$

Proof. We proceed by induction on the sequence of regular subresultants. Begin with S_n, so $j = n-1$. Since $S_i = |x^{n-j-1}A, \ldots, A, x^{m-j-1}B, \ldots, B|$, (19) implies

$$p_n^{(m-n+1)(n-i)} S_i = |x^{m-1-i} S_n, \ldots, S_n, x^{n-1-i} S_{n-1}, \ldots, S_{n-1}|. \tag{23}$$

If $S_{n-1} = 0$, then $S_i = 0$ for all i with $0 \leq i \leq n-1$ by (23). Assume that $\deg S_{n-1} = r$. Then (23) implies

$$p_n^{(m-n+1)(n-i)} S_i = p_n^{m-r} q_{n-1}^{n-1-r} S_{n-1}.$$

A simplification shows

$$q_n^{n-1-r} S_r = q_{n-1}^{n-1-r} S_{n-1}.$$

For i with $0 \leq i \leq r-1$, (23) implies

$$p_n^{(m-n+1)(n-i)} S_i = p_n^{m-r} \operatorname{sres}_i(S_n, S_{n-1}).$$

A simplification then shows

$$p_n^{r-i} q_n^{n-1-i} S_n = \operatorname{sres}_i(S_n, S_{n-1}).$$

The base case is proved.

Assume that the lemma holds for j. Then the next regular subresultant is S_r. In addition Equation (21) implies

$$q_{j+1}^{j-r} p_r = q_j^{j-r} p_j. \tag{24}$$

Now assume that S_{r-1} is of degree t. We claim that

$$p_r^{j-i+1} q_r^{r-i-1} S_i = |x^{j-i} S_r, \ldots, S_r, x^{r-i-1} S_{r-1}, \ldots, S_{r-1}|. \tag{25}$$

Setting $i = r-1$, we find that (22) implies

$$\mathrm{sres}_{r-1}(S_{j+1}, S_j) = p_{j+1} q_{j+1}^{j-r+1} S_{r-1}. \tag{26}$$

Note that (22) is

$$p_{j+1}^{r-i} q_{j+1}^{j-i} S_i = |x^{r-i-1} S_{j+1}, \ldots, S_{j+1}, x^{j-i} S_j, \ldots, S_j|.$$

Using (21) and (26) and applying (19) to the above equation, one proves (25) by (24).

If $S_{r-1} = 0$, then $S_i = 0$ for all i with $0 \le i \le r-1$ by (25). Assume that $S_{r-1} \ne 0$. (26) implies that $S_i = 0$ for all i with $t+1 \le i \le r-2$. For $i = t$, (25) becomes

$$p_r^{j-t+1} q_r^{r-t-1} S_t = p_r^{j-t} p_{r-1}^{r-t} S_{r-1}.$$

It follows that $q_r^{r-t} S_t = q_r^{r-t} S_{r-1}$. For i with $0 \le i \le t-1$, (25) becomes

$$p_r^{j-i+1} q_r^{r-i-1} S_i = p_r^{j-t+1} \mathrm{sres}_i(S_r, S_{r-1}).$$

This shows $p_r^{t-i} q_r^{r-i-1} S_i = \mathrm{sres}_i(S_r, S_{r-1})$. ∎

Theorem 2 *Let $p_i = \mathrm{lc}(S_i)$ for $i = 0, \ldots, n$, $q_n = \mathrm{lc}(S_n)^{m-n}$ and $q_i = \mathrm{lc}(S_i)$ for $i = 0, \ldots, n-1$. If S_{j+1} is regular and $\deg S_j = r$, then*

1. *if $S_j = 0$ then $S_i = 0$ for all i with $0 \le i \le j$;*

2. *if $S_j \ne 0$ then*

$$S_i = 0 \quad \textit{for all } i \textit{ with } r+1 \le i < j, \tag{27}$$

$$q_{j+1}^{j-r} S_r = q_j^{j-r} S_j \tag{28}$$

and

$$p_{j+1} q_{j+1}^{j-r+1} S_{r-1} = (-1)^{j-r} \mathrm{prem}(S_{j+1}, S_j). \tag{29}$$

Proof. All the conclusions are the same as in Lemma (4) except the last one, which follows from (22) (setting $i = r-1$) and the last statement of Proposition 3. ∎

The subresultant sequence of the first kind consisting of the following elements: A, B and $S_j \ne 0$ if S_{j+1} is regular, for all regular S_{j+1}.

The subresultant sequence of the second kind consisting of the following elements A and all regular subresultants. If the second kind subresultant sequence of A and B consists of

$$A, B, S_{d_3}, S_{d_4}, \ldots, S_{d_k},$$

then the subresultant sequence of A and B consists of

$$A, B, S_{n-1}, S_{d_3-1}, S_{d_4-1}, \ldots, S_{d_{k-1}-1}.$$

By Theorem 2 S_{d_i} and $S_{d_{i-1}-1}$ are linearly dependent over R, for $i = 2, \ldots, k$.

Lemma 5 *Let A_1 and A_2 be in $R[x]$ with $\deg A_1 = n_1 \geq \deg A_2 = n_2 > 0$. Assume that the first kind subresultant sequence of A_1 and A_2 consists of*

$$A_1, A_2, A_3, \ldots, A_k,$$

and that the second kind subresultant sequence consists of

$$B_1, B_2, B_3, \ldots, B_k.$$

Let $b_2 = \mathrm{lc}(A_2)^{n_1-n_2}$, $a_i = \mathrm{lc}(A_i)$ and $b_i = \mathrm{lc}(B_i)$ for $i = 3, \ldots, k$. Set $m_i = \deg A_{i-1} - \deg A_i + 1$. Then

$$a_i^{m_i-2} A_i = b_{i-1}^{m_2-2} B_i.$$

In particular, $a_i^{m_i-1} = b_{i-1}^{m_i-2} b_i$.

Proof. Let $B_{i-1} = S_{j+1}$. Then $A_i = S_j$ by the definition of the first and second subresultant sequences. Let $\deg A_i = r$. Then $B_i = S_r$. Since A_{i-1} and B_{i-1} are R-linear dependent, $\deg A_{i-1} = \deg B_{i-1} = j + 1$. Note that $a_i = q_j$ and $b_{i-1} = q_{j+1}$. By (28) in Theorem 2, we get

$$a_i^{m_i-2} A_i = b_{i-1}^{m_i-2} B_i.$$

The leading coefficients of the left and right hand-sides of the above equation gives $a_i^{m_i-1} = b_{i-1}^{m_i-2} b_i$. ■

Subresultant algorithm. Given A_1 and A_2 be in $R[x]$ with $\deg A_1 = n_1 \geq \deg A_2 = n_2 > 0$, compute the first kind subresultant sequence of A_1 and A_2.

1. [*initialize.*]

$$(a_1, b_1) := (1, 1); (a_2, b_2) := \left(\mathrm{lc}(A_1), \mathrm{lc}(A_2)^{n_1-n_2}\right);$$

$$m_2 = \deg A_1 - \deg A_2 + 1;\ i = 3;$$

2. [*compute.*]

```
while true do
        $e_i := (-1)^{m_{i-1}} b_{i-2}^{m_{i-1}-1} a_{i-2}$;
        $A_i := \mathrm{prem}(A_{i-2}, A_{i-1})/e_i$;
        if $A_i = 0$ then return $A_1, A_2, \ldots A_{i-1}$;
        $m_i := \deg A_{i-1} - \deg A_i + 1$;
        $a_i := \mathrm{lc}(A_i)$; $b_i := a_i^{m_i-1}/b_{i-1}^{m_i-2}$;
        $i := i + 1$;
end do;
```

All the divisions in the subresultant algorithm are exact. The algorithm shows that the first kind subresultant sequence of A_1 and A_2 is a polynomial remainder sequence.

We verify the correctness of the subresultant algorithm. Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be the first and second kind subresultant sequences of A_1 and A_2, respectively. For $i = 3$, we have

$$A_3 = (-1)^{n_1-n_2+1}\mathrm{prem}(A_1, A_2).$$

It follows from the last property of Proposition 3 that A_3, if it is not zero, is the third member of $\mathcal{S}_1$. Let $\deg A_3 = d_3 \geq 0$. Then the third member of $\mathcal{S}_2$ is S_{d_3}. For $i = 4$, the fourth member of $\mathcal{S}_1$ is equal to S_{d_3-1}. By Theorem 2

$$p_n q_n^{n-d_3} S_{d_3-1} = (-1)^{n-d_3+1}\mathrm{prem}(S_n,\ S_{n-1}).$$

Note that $S_n = A_2$, $S_{n-1} = A_3$, $p_n = a_2$, $q_n = b_2$. So

$$e_4 = (-1)^{n-d_3+1} a_2 b_2^{n-d_3} = (-1)^{n-d_3+1} p_n q_n^{n-d_3}.$$

By the above two equations we see that A_4 is the fourth element of $\mathcal{S}_1$.

Assume that $A_1, A_2, \ldots, A_{k-1}$ computed in the subresultant algorithm are the first $(k-1)$th members of $\mathcal{S}_1$. Assume that $\deg A_i = d_i$ for $i = 1, \ldots, k-1$; Then $A_i = S_{d_{i-1}-1}$ and the first $(k-1)$th member of $\mathcal{S}_2$ are $A_1, A_2, S_{d_3}, \ldots, S_{d_{k-1}}$. Assume that the kth element of $\mathcal{S}_1$ is of degree d_k.

Then, by Theorem 2, the kth element of $\mathcal{S}_1$ is S_{d_k-1}. Set $j = d_{k-2} - 1$ and $r = d_{k-1}$. By (29) in Theorem 2

$$p_{d_{k-2}} q_{d_{k-2}}^{d_{k-2}-d_{k-1}} S_{d_{k-1}-1} = (-1)^{d_{k-2}-1-d_{k-1}} \mathrm{prem}(S_{d_{k-2}}, S_{d_{k-2}-1}).$$

Set $j = d_{k-3} - 1$ and $r = d_{k-2}$. By (28) in Theorem 2,

$$q_{d_{k-3}}^{d_{k-3}-1-d_{k-2}} S_{d_{k-2}} = q_{d_{k-3}-1}^{d_{k-3}-1-d_{k-2}} S_{d_{k-3}-1}.$$

Using the notation used in the subresultant algorithm, we find that the above two equations become

$$b_{k-2}^{m_{k-1}} S_{d_{k-1}-1} = (-1)^{m_{k-1}} \mathrm{prem}(S_{d_{k-2}}, A_{k-1})$$

and

$$b_{k-3}^{m_{k-2}-2} S_{d_{k-2}} = a_{k-2}^{m_{k-2}-2} A_{k-2}.$$

It follows that

$$b_{k-3}^{m_{k-2}-2} b_{k-2}^{m_{k-1}} S_{d_{k-1}-1} = (-1)^{m_{k-1}} a_{k-2}^{m_{k-2}-2} \mathrm{prem}(A_{k-2}, A_{k-1}). \qquad (30)$$

By the second conclusion of Lemma 5, the above equation becomes

$$\underbrace{\left((-1)^{m_{k-1}} a_{k-2} b_{k-2}^{m_{k-1}-1}\right)}_{e_k} S_{d_{k-1}-1} = \mathrm{prem}(A_{k-2}, A_{k-1}).$$

Thus, $A_k = S_{d_{k-1}-1}$. The correctness of the algorithm is verified.

7 Gröbner bases

7.1 Admissible orderings

Let K be a field and $R = K[x_1, \ldots, x_n]$. Let $\mathbf{X}$ be the monoid consisting of all monomials. For $S \subset \mathbf{X}$, a subset $B \subset S$ is called a basis of S if, for every $u \in S$, there exists $v \in B$ such that $v|u$.

Lemma 6 *(***Dickson***) Every subset of* $\mathbf{X}$ *has a finite basis.*

Proof. Let S be a nonempty subset of $\mathbf{X}$. We proceed by induction on n. The lemmas clearly holds for $n = 1$. Assume that the lemma holds for $(n-1)$. Let w be a monomial in S with $\deg_{x_i} w = d_i$, for $i = 1, \ldots, n$. Let

$$S_{ij} = \{u | u \in S, \deg_{x_i} u = j\} \quad \text{for all } i \text{ with } 1 \le i \le n \text{ and } 0 \le j \le d_i.$$

Then $T_{ij} = \{u/x_i^j | u \in S_{ij}\}$ is a subset of $\mathbf{X}$ involving $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$. By the induction hypothesis each T_{ij} has a finite basis, say C_{ij}. It follows that

$$B_{ij} = \{vx_i^j | v \in C_{ij}\}$$

is a finite basis of S_{ij}. Hence, $\{w\} \cup \left(\cup_{i=1}^{n} \cup_{j=1}^{d_i} B_{ij}\right)$ is a finite basis of S. ■

Definition 1 *A total ordering $\prec$ on $\mathbf{X}$ is called admissible if $1 \preceq u$ for all $u \in \mathbf{X}$, and $u \preceq v \Longrightarrow wu \preceq wv$ for all $u, v, w \in \mathbf{X}$.*

Theorem 3 *Every admissible ordering on $\mathbf{X}$ is Noetherian, that is, a strictly decreasing sequence is finite.*

Proof. Suppose the contrary, we have an infinite sequence S:

$$u_1 \succ u_2 \succ \cdots$$

in $\mathbf{X}$. Viewing S as a set, we have a finite basis

$$\{u_{i_1}, u_{i_2}, \ldots u_{i_k}\}$$

by Lemma 6. Let m be an integer greater than $\max(i_1, \ldots, i_k)$. Then u_m is a multiple of some u_{i_j}, contradicting the assumption that $u_{i_j} \succ u_m$. ■

From now on we fix an admissible ordering $\prec$ on $\mathbf{X}$. For every nonzero polynomial $f \in R$, we can write

$$f = c_1 M_1 + \cdots c_r M_r$$

where $c_i \in K$ is nonzero and $M_i \in \mathbf{X}$ with $M_1 \succ M_2 \succ \cdots \succ M_r$. We call M_1 the leading monomial of f and c_1 the leading coefficient of f. They are denoted by $\mathrm{lm}(f)$ and $\mathrm{lc}(f)$, respectively.

The admissible ordering $\prec$ induces a partial ordering on R as follows:

1. zero is less than every nonzero polynomial;

2. for two nonzero $f, g \in R$, $f \prec g$ if either

$$\mathrm{lm}(f) \prec \mathrm{lm}(g)$$

or

$$\mathrm{lm}(f) = \mathrm{lm}(g) \quad \text{and} \quad (f - \mathrm{lc}(f)\mathrm{lm}(f)) \prec (g - \mathrm{lc}(g)\mathrm{lm}(g)).$$

Theorem 4 *The induced partial ordering on R is Noetherian.*

Proof. Suppose that

$$f_1 \succ f_2 \succ f_3 \succ \cdots$$

is an infinite sequence in R. Then

$$\mathrm{lm}(f_1) \succeq \mathrm{lm}(f_2) \succeq \mathrm{lm}(f_3) \succ \cdots .$$

By Theorem 3 there exists an integer k_1 such that

$$\mathrm{lm}(f_{k_1}) = \mathrm{lm}(f_{k_1+1}) = \cdots .$$

Let $u_1 = \mathrm{lm}(f_{k_1})$. Then we have a new infinite sequence

$$(f_{k_1} - \mathrm{lc}(f_{k_1})u_1) \succ (f_{k_1+1} - \mathrm{lc}(f_{k_1+1})u_1) \succ \cdots .$$

Applying the same argument to the new sequence, we get a new integer $k_2 > k_1$ such that all leading monomials of the members in the new sequence with subscript $\geq k_2$ are equal to u_2. Moreover $u_1 \succ u_2$. So we can construct an infinite sequence

$$u_1 \succ u_2 \succ \cdots ,$$

which is a contradiction to Theorem 3. ■

7.2 Reduction

Let f and g be two nonzero polynomials in R. Let u be a monomial appearing effectively in f, and u be a multiple of $\mathrm{lm}(g)$. Assume that c is the coefficient of u in g, and $u = v\mathrm{lm}(g)$. Then

$$h = f - \frac{c}{\mathrm{lc}(g)} \cdot v \cdot g$$

is called a reduction of f with respect to g. Note that $h \prec f$. Let S be a subset of R and f be a nonzero polynomial of R. We write

$$f \rightarrow_S h$$

if h is obtained by a sequence of reductions with respect to members of G. Such a reduction sequence must be finite by Theorem 4. We call that h is obtained from f by a sequence of reductions (or by reduction) with respect to S. If h contains no monomial which is a multiple of the leading monomial of any members of S, then we say that h is irreducible with respect to S.

Theorem 5 (**Hilbert**) *Every ideal of R has a finite basis.*

Proof. Let I be an ideal R and S_I be the set of monomials which are leading monomials of polynomials in I. By Lemma 6, we have a finite basis $B_I = \{u_1, \dots, u_m\}$ of S_I. Let g_i be a polynomial in I with $\mathrm{lm}(g_i) = u_i$ for $i = 1, \dots, m$, and $G = \{g_1, \dots, g_m\}$. For every $f \in I$, and $f \rightarrow_G h$ which is irreducible with respect to G. Then h belongs to I, and so h has to be zero, for otherwise, $\mathrm{lm}(h)$ would be a multiple of some u_i, so h would be reducible with respect to G. Hence, G is a finite basis of I. ∎

Example 5 *Let $f_1 = xy + 1$ and $f_2 = y^2 - 1$. Write $F = \{f_1, f_2\}$. Let $f = xy^2 - x$. Compute*

$$f \rightarrow_{F,f_1} -(x + y)$$

and

$$f \rightarrow_{F,f_2} 0.$$

7.3 Gröbner bases

Definition 2 *Let I be an ideal of R and G be a subset of I. G is called a Gröbner basis of I if $\mathrm{lm}(G)$ is a basis of $\mathrm{lm}(I)$.*

From the proof of Theorem 5, one sees that finite Gröbner bases with respect to an admissible ordering always exist. Finite Gröbner bases solve two fundamental problems in the theory of polynomial ideals. Let I be an ideal of R and G a Gröbner bases of I.

1. (**ideal membership**) Given $f \in R$ decide whether $f \in I$.

$$f \in I \iff f \rightarrow_G 0.$$

2. (**normal forms**) If $f \rightarrow_G h_1$, $f \rightarrow_G h_2$, and both h_1 and h_2 are irreducible with respect to G, then $h_1 = h_2$.

An ideal I of R is said to be zero-dimensional if R/I is finite-dimensional linear space over K.

Proposition 4 *Let I be an ideal of R and G a Gröbner basis of I. Then I is zero-dimensional if and only if, for each i with $1 \leq i \leq n$, there exists an integer m_i such that $x_i^{m_i}$ is a leading monomial of some element of G.*

Proof. Let M be the set of monomials that are irreducible with respect to G. Then $\bar{M} = \{u+I|u \in M\}$ is a basis of the K-linear space R/I. If $x_i^{m_i}$ is the leading monomial of some element of G for $i = 1, \dots, n$, then $u \in M$ must be of degree in x_i less than m_i. There are only finitely many such monomials. Conversely, if any power of x_j is not a leading monomial of any element of G, then $1, x_j, x_j^2, \dots,$ are all in M. Hence, I is not zero-dimensional.

Hence, we can determine whether I is zero-dimensional by a Gröbner basis and compute a linear basis of R/I when I is zero-dimensional.

Gröbner's question: Given a zero-dimensional ideal I with a basis $\{e_1, \dots, e_m\}$, express e_ie_j as a linear combination of $e_1, \dots, e_n$.

Buchberger's answer: Compute a Gröbner basis of I. Get the set M of irreducible monomials with respect to G. Let $\bar{M}$ be the monomial basis of R/I. Let $\bar{u} = u + I$ and $\bar{v} = v + I$ be in $\bar{M}$, where u, v are irreducible monomials with respect to G. Compute

$$uv \to_G h = \sum_i c_i w_i, \quad \text{where } c_i \in K \text{ and } w_i \in M.$$

Then

$$(u + I)(v + I) = \sum_i c_i(w_i + I).$$

7.4 Buchberger's algorithm

Definition 3 *Let $f, g \in R$ with $fg \neq 0$. Let $w = \mathrm{lcm}(\mathrm{lm}(f), \mathrm{lm}(g))$, and $w = u\mathrm{lm}(f) = v\mathrm{lm}(g)$. The S-polynomial of f and g is defined to be*

$$S(f,\, g) = \frac{u}{\mathrm{lc}(f)} f - \frac{v}{\mathrm{lc}(g)} g.$$

Lemma 7 *Let $f_1, \dots f_m$ be nonzero polynomials in R with the same leading monomial u. If*

$$g = c_1 f_1 + \cdots + c_m f_m \tag{31}$$

where $\mathrm{lm}(g) \prec u$, and $c_1, \dots, c_m \in K$, which are nonzero. Then g can be written as a K-linear combination of S-polynomials $S(f_i, f_j)$, $1 \leq i < j \leq m$. Furthermore, each $S(f_i, f_j)$ has the leading monomial lower than u.

Proof. Induction on m. If $m = 2$, then $g = c_1 f_1 + c_2 f_2$. Since $\mathrm{lm}(g) \prec u$,

$$c_1 \mathrm{lc}(f_1) + c_2 \mathrm{lc}(f_2) = 0.$$

It follows that

$$g = c_1 \mathrm{lc}(f_1) S(f_1, f_2).$$

The lemma holds. Assume that the lemma holds for $m-1$. We rewrite (31) as:

$$g = c_1 \mathrm{lc}(f_1) S(f_1, f_2) + \left(\frac{c_1 \mathrm{lc}(f_1)}{\mathrm{lc}(f_2)} + c_2 \right) f_2 + \sum_{i=3}^{m} c_i f_i.$$

Let $h = (g - c_1 \mathrm{lc}(f_1) S(f_1, f_2))$, which has the leading monomial lower than u. Applying the induction hypothesis to

$$h = \left(\frac{c_1 \mathrm{lc}(f_1)}{\mathrm{lc}(f_2)} + c_2 \right) f_2 + \sum_{i=3}^{m} c_i f_i$$

yields the lemma. ■

Theorem 6 *Let $G = (g_1, \ldots, g_m)$ be an ideal I generated by nonzero g_1, …, g_m in R. Then G is a Gröbner basis of I if and only if $S(g_i, g_j) \rightarrow_G 0$ for all i, j with $1 \leq i < j \leq m$.*

Proof. "$\Longrightarrow$" is trivial. So we assume that $S(g_i, g_j) \rightarrow_G 0$ for all i, j with $1 \leq i < j \leq m$. For every nonzero $f \in I$, we have to show that $\mathrm{lm}(f)$ is a multiple of one of $\mathrm{lm}(g_1)$, …, $\mathrm{lm}(g_m)$. Since f is in L, there exists $h_1, \ldots, h_m \in R$ such that

$$f = h_1 g_1 + \cdots + h_m g_m. \tag{32}$$

Assume that

$$u = \max_{\prec}(\mathrm{lm}(h_1 g_1), \ldots, \mathrm{lm}(h_m g_m)).$$

Since admissible orderings are Noetherian, we may further assume that u is the lowest among all possible polynomials h_1, …, h_m such that (32) holds.

Suppose that $\mathrm{lm}(f) \prec u$. Rewrite (32) as

$$f = \sum_{j,\, w_j \in \mathbf{X},\, \mathrm{lm}(w_j g_j) = u} c_j w_j g_j + \sum_{k,\, \mathrm{lm}(a_k g_k) \prec u} a_k g_k. \tag{33}$$

Then the first sum is a polynomial whose leading monomial is lower than u. Thus it is a K-linear combination of $S(w_j g_j, w_k g_k)$, where $1 \leq i < k \leq m$, by Lemma 7. Since $\mathrm{lm}(w_j g_j) = \mathrm{lcm}(w_k g_k) = u$, $w \mathrm{lcm}(\mathrm{lm}(g_j), \mathrm{lm}(g_k)) = u$. Let $\mathrm{lcm}(\mathrm{lm}(g_j, \mathrm{lm}(g_k)) = v_j \mathrm{lm}(g_j) = v_k \mathrm{lm}(g_k)$. Then $u = w v_j \mathrm{lm}(g_j) =$

$wv_k\mathrm{lm}(g_k)$, and so $w_j = wv_j$ and $w_k = wv_k$. It follows that

$$\begin{aligned} S(w_jg_j,\, w_kg_k) &= \tfrac{1}{\mathrm{lc}(g_j)}w_jg_j - \tfrac{1}{\mathrm{lc}(g_k)}w_kg_k \\ &= w\left(\tfrac{v_j}{\mathrm{lc}(g_j)}g_j - \tfrac{v_k}{\mathrm{lc}(g_k)}g_k\right) \\ &= wS(g_j, g_k). \end{aligned}$$

Since $\mathrm{lm}(S(g_j, g_k)) \prec \mathrm{lcm}(\mathrm{lm}(g_j), \mathrm{lm}(g_k))$ and $S(g_j, g_k) \to_G 0$, $S(g_j, g_k)$ is an R-linear combination of the g_i's, in which each summand has the leading monomial lower that $\mathrm{lcm}(\mathrm{lm}(g_j), \mathrm{lcm}(g_k))$, so $S(w_jg_j,\, w_kg_k)$ is an R-linear combination of g_i's in which each summand has the leading monomial lower than u. Consequently, the first sum in (33) can be written as an R-linear combination of the g_i's, in which each summand has the leading monomials lower than u, a contradiction to the assumption that u is the lowest.

Hence there exist $h_1, \ldots, h_m \in R$ such that (32) holds and that $\mathrm{lm}(f)$ is equal to one of the $\mathrm{lm}(h_i)\mathrm{lm}(g_i)$. Hence $\mathrm{lm}(f)$ is a multiple of one of the g_i's. ■

Buchberger's algorithm: Given $f_1, \ldots, f_m \in R$ and an admissible ordering, compute a Gröbner basis of $(f_1, \ldots, f_m)$ with respect to $\prec$.

1. [*Initialize.*] $G := F$;
2. [*Compute.*]

```
while true do
    $G' := G$;
    for each pair $(p,\, q)$ in $G' \times G'$ do
        $S(p,\, q) \to'_G r$;   ($r$ is irreducible w.r.t. $G'$)
        if $r \neq 0$ then $G = G \subset \{r\}$; end if;
    end do;
    if $G = G'$ then return(G) end if;
end do;
```

Proof of termination. If the algorithm does not terminate, let G_i be the G in the ith iteration of the while-loop. Then

$$F = G_0 \subset G_1 \subset G_2 \subset G_3 \subset \cdots$$

and each G_i contains a polynomial whose leading monomial u_i is irreducible with respect to G_{i-1}. We have an infinite sequence u_1, u_2, …, in which

u_i and u_j are irreducible with respect to each other, a contradiction to Dickson's lemma.

Let G be a Gröbner basis of I. Let $U = \{u_1, \ldots, u_k\}$ be a monomial basis of $\mathrm{lm}(G)$. Furthermore we assume that u_i is not a divisor of u_j for all $m_i, m_j \in U$ with $u_i \neq u_j$. Let $g_i \in G$ with $\mathrm{lm}(g_i) = u_i$. Denote $\{g_1, \ldots, g_k\}$ by G_1. Let $g_i \rightarrow_{(G_1 \setminus \{g_i\})} \tilde{g}_i$ and $\tilde{g}_i$ be irreducible with respect to $(G \setminus \{g_i\})$. Then

$$\tilde{G} = \left\{ \frac{\tilde{g_1}}{\mathrm{lc}(\tilde{g_1})}, \ldots, \frac{\tilde{g_1}}{\mathrm{lc}(\tilde{g_1})} \right\}$$

is a Gröbner basis of I, which we call a reduced Gröbner basis of I. Such a reduced basis is unique.

Theorem 7 *The reduced Gröbner basis of an ideal with respect to an admissible ordering is unique.*

Proof. Let $G = \{g_1, \ldots, g_p\}$ and $H = \{h_1, \ldots, h_q\}$ be two reduced Gröbner bases with respect to a given admissible ordering $\prec$. Assume that $p \leq q$,

$$\mathrm{lm}(g_1) \prec \mathrm{lm}(g_2) \prec \cdots \prec \mathrm{lm}(g_p) \quad \text{and} \quad \mathrm{lm}(h_1) \prec \mathrm{lm}(h_2) \prec \cdots \prec \mathrm{lm}(h_q).$$

Since $g_1 \rightarrow_H 0$, there exists an integer s with $1 \leq s \leq q$ such that $\mathrm{lm}(h_s) | \mathrm{lm}(g_1)$. By the same reason there exists an integer t such that $\mathrm{lm}(g_t) | \mathrm{lm}(h_s)$. Hence, $\mathrm{lm}(g_t) | \mathrm{lm}(g_1)$. Since G_p is reduced, $t = s$, and so $\mathrm{lm}(g_1) = \mathrm{lm}(h_s)$. Similarly, we have $\mathrm{lm}(h_1) = \mathrm{lm}(g_r)$ for some r with $1 \leq r \leq p$. Hence, $t = s = r = 1$, that is $\mathrm{lm}(g_1) = \mathrm{lm}(h_1)$. Applying the same argument yields $\mathrm{lm}(g_i) = \mathrm{lm}(h_i)$ for $i = 1, \ldots, p$. If $p < q$, then $\mathrm{lm}(h_{p+1})$ would be a multiple of some $\mathrm{lm}(g_k)$ $(1 \leq k \leq p)$, which is equal to $\mathrm{lm}(h_k)$, a contradiction. Hence we have $\mathrm{lm}(G) = \mathrm{lm}(H)$. Note that $(h_i - g_i)$ is irreducible with respect to G_p, so $h_i = g_i$. We have proved that $G = H$. ■

7.5 First applications of Gröbner bases

7.5.1 Computation with ideals

Ideal membership. Given an ideal $I = (f_1, \ldots, f_m)$ of R, decide whether a polynomial $g \in I$.

Compute a Gröbner basis G of I. Then $g \in I \Longleftrightarrow g \rightarrow_G 0$.

Normal forms modulo an ideal. Given an ideal $I = (f_1, \ldots, f_m)$, define a normal form function NF from R to R such that $\mathrm{NF}(f) = \mathrm{NF}(g) \Longleftrightarrow f \equiv g \bmod I$.

Compute a Gröbner basis G of I. Define $\mathrm{NF}(f) = \tilde{f}$ such that $f \rightarrow_G \tilde{f}$ and $\tilde{f}$ is irreducible with respect to G.

Zero-dimensional ideals. Given an ideal $I = (f_1, \ldots, f_m)$, decide whether I is zero-dimensional and compute a basis of the K-vector space when I is zero-dimensional.

Lemma 8 *The ideal I is zero-dimensional if and only if, for all i with $1 \leq i \leq n$, there exists a univariate polynomial in x_i in I.*

Compute a Gröbner basis of I. The ideal I is zero-dimensional if and only if, for all i with $1 \leq i \leq n$, there exists a power of x_i in $\mathrm{lm}(G)$. Assume that I is zero-dimensional. Let

$$B = \{u \in X | u \text{ is not divisible by any monomial in } \mathrm{lm}(G)\}.$$

The set $\bar{B} = \{u + I | u \in B\}$ then forms a basis of R/I.

Contraction of ideals. Given an ideal I, compute a basis of $I \cap K[x_1, \ldots, x_m]$, where $0 < m \leq n$. Let $\prec$ be an admissible ordering such that any monomial free of $x_{m+1}, \ldots, x_n$ is lower than any monomial v in which a positive power of x_j, for some j with $m < j \leq n$, appears in v. Compute a Gröbner basis G of I with respect to $\prec$. Then $G \cap K[x_1, \ldots, x_m]$ is a basis of $I \cap K[x_1, \ldots, x_m]$.

Intersection of ideals. Given two ideals $I = (f_1, \ldots, f_p)$ and $J = (g_1, \ldots, g_q)$, compute a basis of $I \cap J$.

Lemma 9 *Let t be a new indeterminate. Let H be the ideals generated by tf_i and $(1-t)g_j$, $1 \leq i \leq p$ and $1 \leq j \leq q$, in $R[t]$. Then $I \cap J = H \cap R$.*

Proof. Let $H_0 = H \cap R$. If $h \in H_0$, then h is an $R[t]$-linear combination of the tf_i's and $(1-t)g_j$'s. Set $t = 0$. Then h becomes an R-linear combination of the f_i's, and hence $h \in I$. In the same vein, $h \in J$ (setting $t = 1$). Thus, $H_0 \subset I \cap J$. Conversely, if $h \in I \cap J$, then

$$h = \sum_{i=1}^{p} a_i f_i = \sum_{j=1}^{q} b_j g_j.$$

Thus

$$h = th + (1-t)h = \sum_{i=1}^{p} a_i t f_i + \sum_{j=1}^{q} b_j (1-t) g_j.$$

This shows that $h \in H_0$. ∎

To find a basis of $I \cap J$, we compute the contraction of the ideal $(tf_1, \ldots, tf_p, (1-t)g_1, \ldots, (1-t)g_q)$ in R.

7.5.2 Solving algebraic equations

Let I be an ideal of R. Define

$$\mathcal{V}(I) = \{(\xi_1, \ldots, \xi_n) \in \overline{K}^n | f(\xi_1, \ldots, \xi_n) = 0, \forall\ f \in I\}.$$

Theorem 8 *Let f be in R. Then $f \in \sqrt{I}$ if and only if f vanishes on $\mathcal{V}(I)$. Consequently, $\mathcal{V}(I) = \emptyset$ if and only if $1 \in I$.*

Solvability of algebraic equations. Let $f_1, \ldots, f_n$ be in R. Decide whether $\mathcal{V}(I) = \emptyset$.

Decide whether $1 \in I$.

Radical ideal membership. Decide whether $f \in \sqrt{I}$ where $I = (f_1, \ldots, f_m)$.

Lemma 10 *Let t be a new indeterminate. $J = (f_1, \ldots, f_m, tf - 1)$. Then*

$$f \in \sqrt{I} \Longleftrightarrow 1 \in J.$$

Proof. If $f \in \sqrt{I}$, then f vanishes on $\mathcal{V}(I)$ by Theorem 8. Thus $\mathcal{V}(J) = \emptyset$ so $1 \in J$ by, again, Theorem 8. Conversely, if $1 \in J$, then

$$a(tf-1) + \sum_{i=1}^{m} a_i f_i = 1, \quad \text{where } a, a_i \in R[t]$$

Substituting $1/f$ for t into the above equation, one gets

$$\sum_{i=1}^{m} b_i f_i = f^k \quad b_i \in R \text{ and } k \in \mathbb{N}.$$

Therefore, $f \in \sqrt{I}$. ∎

Saturation of ideals. Given $g \in R$ and an ideal $I \subset R$, compute a basis of

$$I : g^\infty = \{f \in R | \exists m \in \mathbb{N},\ g^m f \in I\}.$$

Lemma 11 *Let J be the ideal generated by I and $(tg-1)$ in $R[t]$. Then $I : g^\infty = J \cap R$.*

Proof. If $h \in (I : g^\infty)$, then there exists $m \in \mathbb{N}$ such that $g^k h \in I$. Then $(tg - 1 + 1)^k h \in J$. It follows that $h \in J$ so that $h \in J \cap R$. Conversely, if $h \in J \cap R$, then

$$h = a(tg-1) + \sum_i a_i f_i \quad \text{where } a, a_i \in R[t] \text{ and } f_i \in R.$$

Substituting $1/g$ for t into the above equation, we find $g^k h \in I$ for some $k \in \mathbb{N}$. ■

The next theorem illustrates a geometric interpretation of saturation of ideals.

Theorem 9 *Let I be an ideal of R and g in R. Let*

$$S = \{\xi \in \bar{K}^n | \xi \in \mathcal{V}(I) \text{ and } g(\xi) \neq 0\}.$$

Let $J = I : g^\infty$. Then $\mathcal{V}(J)$ is the smallest algebraic set containing S.

Proof. It is straightforward to show that $S \subset \mathcal{V}(J)$. Let H be an ideal of R such that $S \subset \mathcal{V}(H)$. We have to prove that $\mathcal{V}(J) \subset \mathcal{V}(H)$. By Theorem 8 we need only to show $\sqrt{H} \subset \sqrt{J}$. Assume that $h \in \sqrt{(}H)$, then h vanishes on S. Hence gh vanishes on $\mathcal{V}(I)$. Theorem 8 then implies that $(gh)^k \in I$ for some $k \in \mathbb{N}$. Hence, h^k is in J, that is, $h \in \sqrt{J}$. ■

References

[1] B. Buchberger, G. Collins, and R. Loos. (eds) *Computer Algebra, symbolic and algebraic computation.* Springer, 1982.

[2] G.E. Collins, M.J. Encarnacion. Efficient rational number reconstruction. *J. of Symbolic Computation* **20**, pp. 299–297, 1995.

[3] G. Collins, R. Loos, and F. Winkler. Arithmetic in basic algebraic domains. In [1], pages 189–220.

[4] R. Feng, X. Gao. Rational general solutions of algebraic ordinary differential equations. In the *Proceedings of ISSAC 2004*, pp. 155-161.

[5] J. von zur Gathen, J. Gerhard. *Modern Computer Algebra*, Cambridge Press, First Edition, 1999.

[6] K. Geddes, S. Czapor, and G. Labahn. *Algorithms for Computer Algebra.* Kluwer Academic Publisher, 1992.

[7] D. Kunth. *The Art of Computer Programming.* Vol. II, Addison-Wesley, 1981.

[8] F. Winkler. *Polynomial Algorithms in Computer Algebra.* Springer, 1996.

[9] M. Monagan. Maximal quotient rational reconstruction: an almost optimal algorithm for rational reconstruction. In the *Proceedings of ISSAC 2004*, pp. 243-249.

[10] P.S. Wang, J.T. Guy, J.H. Davenport. p-adic Reconstruction of rational numbers. *SIGSAM Bulletin*, **16**, No. 2, 1982.