Using Computer Algebra

Zafeirakis Zafeirakopoulos

1 Introduction

We will start with a question that seems silly, but lies behind some of the most beautiful algebra of last century, as well as the birth of computer algebra.

Question 1.1 *Can we write 1013 as a linear combination of*

- *2,4,18?*
- *3,7?*
- *2,33?*

The question is not very well posed yet of course, since we need to specify what a linear combination is.

Let $\mathcal{R}$ be a ring. All rings in these notes are rings with 1, except if explicitly stated otherwise.

The question above becomes: given $p, g_1, g_2 \in \mathcal{R}$, can we write p as a linear combination of g_1 and g_2? Equivalently, are there $a_1, a_2 \in \mathcal{R}$ such that $p = a_1 g_1 + a_2 g_2$?

It is clear that the answer is yes if and only if the greatest common divisor of g_1 and g_2 divides p. In that case, there exist (and can be computed using the extended Euclidean algorithm) co-factors a_1 and a_2 such that $a_1 g_1 + a_2 g_2 = \gcd(g_1, g_2)$.

This is a direct consequence of

Theorem 1.2 (Bézout's identity) *Let $g_1, g_2 \in \mathbb{Z}$ with $\gcd(g_1, g_2) = d$. Then there exist integers a_1 and a_2 (Bézout coefficients) such that $a_1 g_1 + a_2 g_2 = d$. Moreover, if $c = b_1 g_1 + b_2 g_2$ for some $b_1, b_2 \in \mathbb{Z}$, then $d|c$.*

That was an easy answer, but we have to be careful. What are the assumptions we made so that we have a gcd and algorithm to compute it?

1.1 Notation and Terminology

Before we make the answer more precise let's fix some terminology and notation.

It will be useful to refer to the set of all linear combinations, since it is our main object of study.

Definition 1.3 (Ideal) *Let $\mathcal{R}$ be a commutative ring with 1. A subset $I \subseteq \mathcal{R}$ is called an ideal of $\mathcal{R}$, denoted by $I \lhd \mathcal{R}$, if*

- *for all $a, b \in I$ we have $a + b \in I$ (closure under addition)*
- *for all $a \in I$ and for all $r \in \mathcal{R}$ we have $ra \in I$ (closure under multiplication by scalar)*

The most usual way of describing an ideal is through a set of generators. The ideal generated by $g_1, g_2, \ldots, g_k \in \mathcal{R}$ will be denoted by $\langle g_1, g_2, \ldots, g_k \rangle$. From now on, when we say "given an ideal" we mean "given a set of generators".

Note that the set of generators is in general not unique.

Definition 1.4 (Principal Ideal) *Given $I \lhd \mathcal{R}$, if there exist $g\mathcal{R}$ such that $I = \langle g \rangle$, then I is a principal ideal.*

In the ring of integers, all ideals are principal ideals. This is not unique for that ring, but it is a very important property.

Definition 1.5 (Principal Ideal Domain (PID)) *If $\mathcal{R}$ is a commutative ring with 1, such that every ideal in $\mathcal{R}$ is a principal ideal, then $\mathcal{R}$ is a principal ideal domain.*

Now, let's move one step up.

Let $\mathcal{R} = \mathbb{K}[x]$, the univariate polynomial ring in x over a field $\mathbb{K}$. Given $p, g_1, g_2 \in \mathcal{R}$, can we write p as a linear combination of g_1, g_2? Equivalently, are there $a_1, a_2 \mathcal{R}$ such that $a_1 g_1 + a_2 g_2 = p$?

Univariate polynomials (over a field) behave like the integers, thus the answer is the same as before: There exist such a_1 and a_2 if and only if $\gcd(g_1, g_2) | p$.

The reason is that $\mathcal{R} = \mathbb{K}[x]$ is a PID and Theorem 1.2 holds true for any PID (and not only $\mathbb{Z}$).

Let's rephrase our main question in a more general and precise way.

[Ideal Membership] Given an ideal $I \lhd \mathcal{R}$ and $p \in \mathcal{R}$, decide if $p \in I$.

Historically, this is an important problem in the development of modern algebra (by Hilbert, Nöther, Hermann, etc). But its solution also provided a tool for the development of effective algebra in the second half of last century.

Let's first see the solution to the problem in the cases we studied till now.

Definition 1.6 *$\mathcal{R}$ (a commutative ring with 1) is a Euclidean ring if there is a map $\delta : \mathcal{R} - \{0\} \to \mathbb{N}$ such that for all $a, b \in \mathcal{R} - \{0\}$ there exist $q, r \in \mathcal{R}$ with*

$$\begin{cases} a = qb + r \\ \delta r < \delta b \end{cases}$$

The main consequence is that in a Euclidean ring we have a way to compute the greatest common divisor of two elements. Moreover, every Euclidean ring is a PID and Theorem 1.2 holds, and we can compute the Bezout coefficients through the extended Euclidean algorithm.

The transition from an existential theorem to an effective method is the transition to computer algebra.

One of the most important consequences of working in a Euclidean ring is that we have normal forms of elements. This essentially means that we have a consistent way of choosing a canonical representative of a class $a + \langle g \rangle$ in $\mathcal{R}/\langle g \rangle$. Namely, if $a = qg + r$ such that $\delta r < \delta g$, then the normal form of a is r.

Note that this implies $a \in \langle g \rangle \Leftrightarrow r = 0$, and solves Ideal Membership.

To wrap up, we conclude that in order solve Ideal Membership, we need a normal form (canonical representative) for elements in $\mathcal{R}/I$. The only problem is that in the next step we want to consider $\mathcal{R} = \mathbb{K}[x_1, x_2, \ldots, x_n]$ and now $\mathcal{R}$ is not a Euclidean ring, not even a PID.

But we know that what we need is:

"find a canonical representative of an element by dividing to get the remainder."

In the following, we will need to make the three red points concise and effective.

The main concept we will use is that of reduction, which mimics Euclidean division but can have more than one divisors.

Definition 1.7 (Reduction) *Let $\mathcal{R}$ be a commutative ring with 1 and $S \subseteq \mathcal{R}$. A reduction $\to_S$ is a function from $\mathcal{R}$ to $\mathcal{R}$ such that there exist $\delta : \mathcal{R} \to \mathbb{N}$ and for all $a\mathcal{R}$ we have that $a \to_I b$ implies $\delta b < \delta b$.*

Given a polynomial $p \in \mathcal{R} = \mathbb{K}[x_1, x_2, \ldots, x_n]$, we will denote by

- $lm(p)$ the leading monomial of p.
- $lt(p)$ the leading term of p.

Note that a monomial is the product of a term with a constant.

If we apply the algorithm to reduce $x^3 + 3x - 1$ by $x^2 + 1, x^3 - 1$, depending on the choice of divisor at every step, we obtain

- $x^3 + 3x - 1 \to_{x^2+1, x^3-1} x$ or
- $x^3 + 3x - 1 \to_{x^2+1, x^3-1} x^2 - 1$

Data: $f, g_1, g_2, \ldots, g_k \in \mathcal{R}$
Result: $r \in \mathcal{R}$ such that there exist $a_i\mathcal{R}$ with $f = r + \sum_{i=1}^{k} a_i g_i$
$r \leftarrow 0$
while $f \neq 0$ **do**
  Choose i such that $lt(g_i)|lt(f)$
  if *no such i exists* **then**
    $r \leftarrow r + lm(f)$ $f \leftarrow f - lm(f)$
  else
    $a \leftarrow a - \frac{lm(f)}{lm(g_i)} g_i$
  end
end
return r

Algorithm 1: The reduction algorithm

Note though that $\gcd\left(x^2+1, x^3-1\right) = 1$ and thus we know that $x^3 + 3x - 1 \in \left\langle x^2+1, x^3-1 \right\rangle$!

But in $\mathbb{K}[x_1, x_2, \ldots, x_n]$ we have even more choices, which is even worse.

We first have to decide if $xy > x^2$ for example, since we need to know the leading term of a polynomial in order to do reduction.

Let $[x] = [x_1, x_2, \ldots, x_n]$ be the monoid of all terms in $x_1, x_2, \ldots, x_n$. Note that $[x] \cong \mathbb{N}^n$.

We can define an order on $[x]$ by means of divisibility, but $|$ is only a partial order if $n \neq 1$.

Definition 1.8 (Term Order) *Given a term monoid T, a total order on T is called a term order if*

- $1 \leq a$ *for all* $a \in T$
- $a \leq b$ *implies* $ac \leq bc$ *for all* $a, b, c \in T$.

The most usual term orders are lexicographic $\leq_{lex}$ and degree reverse lexicographic $\leq_{degrevlex}$. They are defined as

$$x^a \leq_{lex} x^b \Leftrightarrow \exists i \in [n] : \begin{cases} \mathrm{a}_j = b_j \forall j \in [i-1] \\ \mathrm{a}_i < b_i \end{cases}$$

and

$$x^a \leq_{degrevlex} x^b \Leftrightarrow \begin{cases} \deg(\mathrm{x}^a) < deg(x^b) \\ \begin{cases} \deg(\mathrm{x}^a) = deg(x^b) \\ \exists i \in [n] : \begin{cases} \mathrm{a}_j = b_j \forall j \in [i+1, n] \\ \mathrm{a}_i < b_i \end{cases} \end{cases} \end{cases}$$

Finally, let us note that every term order can be given by a unimodular matrix $M \in \mathbb{Z}n \times m$ by defining $x^a \leq_{dM} x^b \Leftrightarrow x^{Ma} \leq_{lex} x^{Mb}$.

Let us fix a term order.For this example it will be the lexicographic order. Let $g_1 = x^2 + y^2 - 1$ and $g_2 = x - y$. If we reduce $x^4y + 3xy^3 - 1$ by $\{g1, g2\}$ we obtain

- $x^4y + 3xy^3 - 1 \to_{\{g_1,g_2\}} y^5 + 3y^4 - 2y^3 + y - 1$ or
- $x^4y + 3xy^3 - 1 \to_{\{g_1,g_2\}} y^5 + 3y^4 - 1$

Note that again the reduction is not unique.

Now let's consider $f = (x-2)g_1 + (xy^2 + y)g_2$. We know that $f \in \langle g_1, g_2 \rangle$ by construction.

If we reduce f by $\{g_1, g_2\}$ we get $2y^4 - y^2$, but we have no leading term to reduce that.

The main idea of Buchberger was to create those leading terms needed to continue the reduction.

Definition 1.9 (S-polynomial) *Let $f_1, f_2 \in I$ and $\ell = \operatorname{lcm}(lm(f_1, f_2))$. Then we define the s-polynomial of f_1 and f_2 as*

$$spol(f_1, f_2) = \frac{\ell}{lm(f_1)} f_1 - \frac{\ell}{lm(f_2)} f_2$$

Note that in the S-polynomial we eliminated the leading terms of f_1 and f_2, thus we computed a polynomial in the ideal with a new leading term.

For example, the S-polynomial of $g_1 = x^2 + y^2 - 1$ and $g_2 = x - y$ is $x^2 + xy - 1$. If we reduce it by $\{g_1, g_2\}$ we obtain $2y^2 - 1$ and now we can continue reducing $2y^4 - y^2$ which we couldn't before.

This lead to the main definition of these notes.

Definition 1.10 (Gröbner bases) *Fix a term order. A set $G = \{g_1, g_2, \ldots, g_k\}$ is a Gröbner basis if reduction by G is unique.*

Of course that is impossible to check since we have infinitely many polynomials in the ideal.

Bu the following two theorems by Buchberger solve this problem.

Theorem 1.11 (Buchberger) *Every ideal in $\mathbb{K}[x_1, x_2, \ldots, x_n]$ has a finite Gröbner basis.*

Theorem 1.12 (Buchberger) *G is a Gröbner basis if and only if for all $f, g \in G$ we have that $spol(f, g) \to_G 0$.*

Which lead to Buchberger's algorithm for the computation of Gröbner bases.

Note that the Gröbner basis depends on the term order we fix. But for any fixed term order, there exists a unique reduced Gröbner basis (where we interreduce the elemnts and we make the polynomials monic).

Data: $F = \{f_1, f_2, \ldots, f_k\} \subseteq \mathcal{R}$ and a term order $<$
Result: G a Gröbner basis for $\langle F \rangle$
$C \leftarrow \{(g_i, g_j) : 1, j \leq k\}$
while $C \neq \emptyset$ **do**
  Choose $(f, g) \in C$
  Remove (f, g) from C
  $spol(f, g) \rightarrow_G r$
  if 0 **then**
    $C \leftarrow C \cup (G \times \{r\})$ $G \leftarrow G \cup \{r\}$
  end
end
return G

Algorithm 2: Buchberger's algorithm

An important alternative definition of Gröbner bases is given by the following theorem.

Theorem 1.13 *G is a Gröbner basis of $I \lhd \mathcal{R}$ if and only if $\langle lt(G) \rangle = \langle lt(I) \rangle$.*

2 Applications

Let $\mathcal{R} = \mathbb{K}[x_1, x_2, \ldots, x_n]$. $\mathcal{R}$ is a $\mathbb{K}$-vector space and an important question is to find a vector space basis for the quotient $\frac{\mathcal{R}}{I}$ for $I \lhd \mathcal{R}$.

Since $\frac{\mathcal{R}}{I}$ is $\{p + I : p \in R\}$, if we reduce $p \rightarrow_G \bar{p}$ by a Gröbner basis G of I, then $\bar{p}$ contains terms that do not belong in $\langle lt(I) \rangle = \langle lt(G) \rangle$. Then a basis is the set of monomial in $[x] - lt(G)$.

Historical note: That was the question Gröbner gave to his PhD student Buchberger.

Among term orders, some have important properties we can exploit in applications.

Definition 2.1 (Elimination order) *A term order on $[x_1, x_2, \ldots, x_n]$ is an elimination order for $x_1, x_2, \ldots, x_k$ if $a > b$ when a contains any of $x_1, x_2, \ldots, x_k$ while b doesn't.*

Main example: lexicographic order.

Theorem 2.2 (Elimination Property of Gröbner bases) *Let $I \lhd \mathcal{R} = \mathbb{K}[x_1, x_2, \ldots, x_n]$ and fix an elimination order for $x_1, x_2, \ldots, x_k$. If G is a Gröbner basis of I for the order we fixed, then $G \cap \mathbb{K}[x_1, x_2, \ldots, x_k]$ is a Gröbner basis for $I \cap \mathbb{K}[x_1, x_2, \ldots, x_k]$ (with respect to the restriction of the order to $x_1, x_2, \ldots, x_k$).*

This means that for an elimination order a reduced Gröbner basis has a "triangular" form. Just like Gaussian elimination does for linear systems.

Note that having a system in triangular form is excellent for solving it (finding the roots). If we have an ideal which only has isolated roots (called zero dimensional), then we will have a univariate polynomial in the basis. This is easy to solve and we can substitute in the bivariate polynomial and do the same. Then we solve the problem by backtracking (just as we do for linear systems and Gaussian elimination).

Now we will see 3 applications that appear often in various fields.

2.1 Ideal Intersection

Given $I_1 = \langle g_1, g_2, \ldots, g_k \rangle$ and $I_2 = \langle f_1, f_2, \ldots, f_\ell \rangle$ ideals in $\mathcal{R}$, how do we compute a set of generators for the intersection $I_1 \cap I_2$?

Let us introduce a new variable t and consider the ideal

$$I = \langle tg_1, tg_2, \ldots, tg_k, (1-t)f_1, (1-t)f_2, \ldots, (1-t)f_\ell \rangle \lhd \mathbb{K}[x_1, x_2, \ldots, x_n, t].$$

The only polynomials that do not contain t are the one that belong to the intersection $I_1 \cap I_2$.

Thus we want $I \cap \mathbb{K}[x_1, x_2, \ldots, x_n]$, which we can do by computing a Gröbner basis with respect to an elimination order for $x_1, x_2, \ldots, x_n$ and then intersect it with $\mathbb{K}[x_1, x_2, \ldots, x_n]$.

Why would we care about intersection of ideals? One reason is that we can compute the multivariate greatest common divisor. Given $f, g \in \mathbb{K}[x_1, x_2, \ldots, x_n]$, we have that $\langle f \rangle \cap \langle g \rangle = \langle lcm(f, g) \rangle$. and then $gcd(f, g) = \frac{fg}{lcm(f,g)}$.

2.2 Implicitization

Let $r_1, r_2\mathbb{K}(t)$ be two rational functions. Then $C = \left\{ (r_1(t), r_2(t)) : t \in \bar{\mathbb{K}} \right\} \subseteq \bar{\mathbb{K}}^2$ is an algebraic curve in parametric form.

How do we find the implicit equation for C?

Let $r_i = \frac{f_i}{g_i}$ and introduce new variable x, y.

Then define $I = \langle g_1 - xf_1, g_2 - yf_2 \rangle \lhd \mathbb{K}[x, y, t]$.

The implicit equation is given by the generator of $I \cap \mathbb{K}[x, y]$, which we know how to compute now (elimination property).

2.3 Algebraic relations

Assume that given a set of polynomials $f_1, f_2, \ldots, f_\ell \in \mathbb{K}[x_1, x_2, \ldots, x_n]$, we want to compute the relations between them. Then we define $J = \langle t1 - f_1, t_2 - f_2, \ldots, t_\ell - f_\ell \rangle$ and intersect it with $\mathbb{K}[t_1, t_2, \ldots, t_\ell]$.

The resulting ideal will contain the algebraic relations between f_is as we essentially renamed them to t_is.

3 Final remarks

This is a very short introduction to Gröber bases with two goals:

- give the main ideas on why and how they were developed
- and how they can be used.

The two main textbooks one can follow are:

- "Ideals, Varieties and Algorithms" by David A. Cox , John Little , Donal O'Shea.
- "Modern Computer Algebra" by Joachim von zur Gathen and Jürgen Gerhard.

Moreover, there is software where most of the algorithms are implemented. For example: SageMath, Maple, Mathematica, Singular, Magma, Macaulay2.