Math 533 Winter 2021, Lecture 17: Multivariate polynomials

website: `https://www.cip.ifi.lmu.de/~grinberg/t/21w/`

1. Polynomials II

1.1. Division with remainder and Gröbner bases (cont'd)

Last time (i.e., in Lecture 16), we started discussing quotient rings of multivariate polynomial rings.

Recall our standing conventions: R is a commutative ring; $n \in \mathbb{N}$; we set $P = R[x_1, x_2, \ldots, x_n]$.

We have stated the following theorem (and sketched its proof):

Theorem 1.1.1. Let $b \in P$ be a nonzero polynomial whose leading coefficient $\operatorname{LC} b$ is a unit of R. Let $a \in P$ be any polynomial.

Then, there is a **unique** pair (q, r) of polynomials in P such that

$$a = qb + r \qquad \text{and} \qquad r \text{ is } \operatorname{LM} b\text{-reduced}.$$

Here, a polynomial $r \in P$ is said to be $\mathfrak{m}$**-reduced** (where $\mathfrak{m}$ is a monomial) if no monomial divisible by $\mathfrak{m}$ appears in r.

We used this theorem to find a basis of the R-module $P/b = P/bP$.

1.1.1. The case of arbitrary ideals

Now what if we want to know how P/I looks like for a non-principal ideal I, say $I = b_1P + b_2P + \cdots + b_kP$ for some $b_1, b_2, \ldots, b_k \in P$? Can we divide a polynomial by I with remainder? Can we check whether a polynomial belongs to I ? (Remember: If $I = bP$ is a principal ideal, then this means checking whether the polynomial is divisible by b. We have seen how to do this using Theorem 1.1.1)

We can try to replicate the above "division with remainder" logic.

Example 1.1.2. Let $n = 2$, and let us write x, y for the indeterminates x_1, x_2. Let $R = \mathbb{Q}$ (just to be specific), and let $I = b_1P + b_2P$ with

$$b_1 = xy + 1,$$
$$b_2 = y + 1.$$

Let $a \in P$ be any polynomial. We try to divide a by I with remainder. This means writing a in the form $a = i + r$ where $i \in I$ and r is a "remainder".

Here, a "**remainder**" (modulo b_1 and b_2) means a polynomial that is both $\operatorname{LM} b_1$-reduced and $\operatorname{LM} b_2$-reduced, i.e., that contains neither multiples of $\operatorname{LM} b_1$ nor multiples of $\operatorname{LM} b_2$ among its monomials. We can achieve this by subtracting multiples of b_1 and multiples of b_2 from a until no such remain. In more detail: Whenever some monomial that is a multiple of $\operatorname{LM} b_1$ appears in our polynomial, we can subtract an appropriate multiple of b_1 from our polynomial to remove this monomial. (Namely, the multiple of b_1 that we choose is the one whose leading term would cancel the multiple of $\operatorname{LM} b_1$ we want to remove from our polynomial.) Similarly we get rid of multiples of $\operatorname{LM} b_2$. When no more monomials that are multiples of $\operatorname{LM} b_1$ or multiples of $\operatorname{LM} b_2$ remain in our polynomial, then we have found our "remainder".

We refer to this procedure as the **division-with-remainder algorithm**. Note that this is a nondeterministic algorithm, in the sense that you often have a choice of which step you make. For instance, if your polynomial contains a monomial that is a multiple of both $\operatorname{LM} b_1$ and $\operatorname{LM} b_2$ at the same time, do you remove it by subtracting a multiple of b_1 or by subtracting a multiple of b_2 ? Thus, the "remainder" at the end might fail to be unique.

Let us check this on a specific example. Let $a = xy - y \in P$. Here is one way to perform our division-with-remainder algorithm:

$$\begin{aligned} a = xy - y & \xrightarrow[\text{to get rid of the } xy \text{ monomial}]{\text{subtract } 1b_1} (xy - y) - (xy + 1) = -y - 1 \\ & \xrightarrow[\text{to get rid of the } y \text{ monomial}]{\text{subtract } -1b_2} (-y - 1) - (-1)(y + 1) = 0. \end{aligned}$$

Here is another way to do it:

$$\begin{aligned} a = xy - y & \xrightarrow[\text{to get rid of the } xy \text{ monomial}]{\text{subtract } xb_2} (xy - y) - x(y + 1) = -x - y \\ & \xrightarrow[\text{to get rid of the } y \text{ monomial}]{\text{subtract } -1b_2} (-x - y) - (-1)(y + 1) = -x - 1. \end{aligned}$$

Both results we have obtained are both $\operatorname{LM} b_1$-reduced and $\operatorname{LM} b_2$-reduced, so they qualify as "remainders" of a modulo b_1 and b_2. However, they are not equal! So the remainder is not unique this time (unlike in Theorem 1.1.1). In particular, the first remainder we obtained was 0, which showed that $a \in I$ (because this remainder was obtained from a by subtracting multiples of b_1 and b_2, and of course these multiples all belong to I); but the second remainder was not 0, thus allowing no such conclusion. So we don't have a sure way of telling whether a polynomial belongs to I or not; if we are unlucky, we get a nonzero remainder even for a polynomial that does belong to I.

This is bad!

Example 1.1.3. A simpler example: Let $n = 2$ and $I = b_1 P + b_2 P$ with

$$b_1 = xy + x,$$
$$b_2 = xy + y.$$

The polynomial $x - y$ lies in I (since $x - y = b_1 - b_2$), but it is both $\operatorname{LM} b_1$-reduced and $\operatorname{LM} b_2$-reduced, so we cannot see this from our division-with-remainder algorithm no matter what choices we make (because the algorithm does nothing: $x - y$ already is a "remainder"). We could, of course, subtract b_1 from $x - y$ (to obtain $(x - y) - (xy + x) = -y - xy$), but this would be a "step backwards", as it would increase the leading monomial (and even the degree) of our polynomial. The idea of the division-with-remainder algorithm is to reduce the polynomial step by step, always "walking downhill", rather than having to "cross a mountain" first (temporarily increasing the leading monomial).

Example 1.1.3 might give you an idea of what is standing in our way here: It is the fact that when we compute $b_1 - b_2$, the leading terms xy cancel. It means, in a sense, that our b_1 and b_2 are "unnecessarily convoluted"; we should perhaps fix this by replacing b_2 by the smaller polynomial $b_2 - b_1 = y - x$. This simplifies b_2 but does not change I (since $b_1 P + b_2 P = b_1 P + (b_2 - b_1) P$). This is similar to one of the row-reduction steps involved in bringing a matrix to row echelon form.

What does it mean in general that a list $(b_1, b_2, \ldots, b_k)$ of polynomials is "unnecessarily convoluted"? The xy cancellation in $b_1 - b_2$ above was easy to see; what other cancellations can hide in a list of polynomials?

Let me formalize this question. The following definition will be a bit long-winded but it is just giving names to the kind of observations you would have made when trying to discuss the above algorithm:

Definition 1.1.4. Let $\mathbf{b} = (b_1, b_2, \ldots, b_k)$ be a list of nonzero polynomials in P whose leading coefficients are units of R.

(a) Given two polynomials $c, d \in P$, we write $c \xrightarrow[\mathbf{b}]{} d$ (and say "c can be **reduced** to d in a **single step** using $\mathbf{b}$") if

- some monomial $\mathfrak{m}$ appearing in c is a multiple of $\operatorname{LM} b_i$ for some $i \in \{1, 2, \ldots, k\}$;
- we have
$$d = c - \frac{[\mathfrak{m}]\, c}{\operatorname{LC} b_i} \cdot \frac{\mathfrak{m}}{\operatorname{LM} b_i} \cdot b_i.$$
(This equation essentially says that we obtain d from c by subtracting the appropriate multiple of b_i to get rid of the monomial $\mathfrak{m}$. The multiple is $\dfrac{[\mathfrak{m}]\, c}{\operatorname{LC} b_i} \cdot \dfrac{\mathfrak{m}}{\operatorname{LM} b_i} \cdot b_i$, since the $\dfrac{\mathfrak{m}}{\operatorname{LM} b_i}$ factor is needed to turn the

leading monomial of b_i into $\mathfrak{m}$, whereas the $\frac{[\mathfrak{m}]\, c}{\operatorname{LC} b_i}$ factor serves to make the coefficient of this monomial the same as that in c. Note that the fraction $\frac{[\mathfrak{m}]\, c}{\operatorname{LC} b_i} \in R$ is well-defined since $\operatorname{LC} b_i$ is a unit, whereas the fraction $\frac{\mathfrak{m}}{\operatorname{LM} b_i} \in C^{(n)}$ is well-defined since $\mathfrak{m}$ is a multiple of $\operatorname{LM} b_i$.)

For instance, using the notations of Example 1.1.2 and setting $\mathbf{b} = (b_1, b_2)$, we have

$$xy - y \underset{\mathbf{b}}{\longrightarrow} -x - y,$$

because we obtain $-x - y$ from $xy - y$ by subtracting the multiple $1b_1$ of b_1 (which kills the xy monomial). Likewise, for the same $\mathbf{b}$, we have

$$5x^2y^3 \underset{\mathbf{b}}{\longrightarrow} -5xy^2,$$

because we obtain $-5xy^2$ from $5x^2y^3$ by subtracting the multiple $5xy^2b_1$ of b_1 (which kills the x^2y^3 monomial).

(b) Given two polynomials $c, d \in P$, we write $c \underset{\mathbf{b}}{\overset{*}{\longrightarrow}} d$ (and say "c can be **reduced** to d in **many steps** using $\mathbf{b}$") if there is a sequence $(c_0, c_1, \ldots, c_m)$ of polynomials in P such that $c_0 = c$ and $c_m = d$ and

$$c_i \underset{\mathbf{b}}{\longrightarrow} c_{i+1} \qquad \text{for each } i \in \{0, 1, \ldots, m-1\}.$$

Note that this sequence can be trivial (i.e., we can have $m = 0$), in which case of course we have $c = d$. Thus, $c \underset{\mathbf{b}}{\overset{*}{\longrightarrow}} c$ for any $c \in P$. (Like any true algebraists, we understand "many steps" to allow "zero steps".) We also can have $m = 1$; thus, $c \underset{\mathbf{b}}{\overset{*}{\longrightarrow}} d$ holds if $c \underset{\mathbf{b}}{\longrightarrow} d$. (That is, "many steps" allows "one step".)

As an example of a nontrivial many-steps reduction, we observe that using the notations of Example 1.1.2 and setting $\mathbf{b} = (b_1, b_2)$, we have

$$5x^2y^3 \underset{\mathbf{b}}{\longrightarrow} -5xy^2 \underset{\mathbf{b}}{\longrightarrow} 5y \underset{\mathbf{b}}{\longrightarrow} -5$$

and thus $5x^2y^3 \underset{\mathbf{b}}{\overset{*}{\longrightarrow}} -5$.

(c) We say that a polynomial $r \in P$ is **$\mathbf{b}$-reduced** if it is $\operatorname{LM} b_i$-reduced for all $i \in \{1, 2, \ldots, k\}$. This is equivalent to saying that there exists no polynomial $s \in P$ with $r \underset{\mathbf{b}}{\longrightarrow} s$ (that is, "r cannot be reduced any further using $\mathbf{b}$").

(d) A **remainder** of a polynomial $a \in P$ modulo $\mathbf{b}$ means a $\mathbf{b}$-reduced polynomial $r \in P$ such that $a \underset{\mathbf{b}}{\overset{*}{\longrightarrow}} r$. Such a remainder always exists (this is not hard to show), but is not always unique (as we have seen in Example 1.1.2).

(e) We say that the list $\mathbf{b}$ is a **Gröbner basis** if any $a \in P$ has a **unique** remainder modulo $\mathbf{b}$.

(Don't take the word "basis" in "Gröbner basis" to heart. It is closer to "generating set" or "spanning set" than to any sort of "basis" in linear algebra. In particular, a Gröbner basis can be R-linearly dependent or even contain the same polynomial twice.)

So we have seen that not every list of nonzero polynomials is a Gröbner basis. But here are the more interesting questions:

- Can we **tell** whether a list of nonzero polynomials is a Gröbner basis? (We cannot afford to check every $a \in P$ and every way of reducing it modulo $\mathbf{b}$.)
- If a list is not a Gröbner basis, can we at least **find** a Gröbner basis that generates the same ideal as the list?

If R is not a field, then the answers to these questions are "no" for reasons that should be familiar from the univariate case (non-unit leading coefficients, etc.).

When R is a field, Bruno Buchberger has answered both questions in the positive in the 1960s. The algorithms he found are one of the pillars of modern computer algebra. I will state the main results without proof, but you can find proofs in the literature (e.g., [DF, §9.6] or [deGraa20, Chapter 1]).

We will need the notion of an **S-polynomial**:

Definition 1.1.5. Let $f, g \in P$ be nonzero polynomials whose leading coefficients are units of R.

Let $\mathfrak{p} = x_1^{p_1} x_2^{p_2} \cdots x_n^{p_n} = \operatorname{LM} f$ and $\mathfrak{q} = x_1^{q_1} x_2^{q_2} \cdots x_n^{q_n} = \operatorname{LM} g$ be their leading monomials, and let $\lambda = \operatorname{LC} f$ and $\mu = \operatorname{LC} g$ be their leading coefficients. So

$$f = \lambda\mathfrak{p} + (\text{smaller terms});$$
$$g = \mu\mathfrak{q} + (\text{smaller terms}).$$

Let

$$\mathfrak{m} = x_1^{\max\{p_1,q_1\}} x_2^{\max\{p_2,q_2\}} \cdots x_n^{\max\{p_n,q_n\}}.$$

(This is the lcm of $\mathfrak{p}$ and $\mathfrak{q}$ among the monomials; it is the smallest-degree monomial that is divisible by both $\mathfrak{p}$ and $\mathfrak{q}$.) Note that $\dfrac{\mathfrak{m}}{\mathfrak{p}}$ and $\dfrac{\mathfrak{m}}{\mathfrak{q}}$ are well-defined monomials (since $\mathfrak{p} \mid \mathfrak{m}$ and $\mathfrak{q} \mid \mathfrak{m}$).

The **S-polynomial** (short for **syzygy polynomial**) of f and g is defined to be the polynomial

$$S(f,g) := \frac{1}{\lambda} \cdot \frac{\mathfrak{m}}{\mathfrak{p}} \cdot f - \frac{1}{\mu} \cdot \frac{\mathfrak{m}}{\mathfrak{q}} \cdot g \in P.$$

Here is the intuition behind this: $S(f,g)$ is the simplest way to form a P-linear combination of f and g in which the leading terms of f and g cancel. Namely, in order to obtain such a P-linear combination, we must first

rescale f and g so that their leading coefficients become equal (this can be achieved by scaling f by $\frac{1}{\lambda}$ and scaling g by $\frac{1}{\mu}$); then we must multiply them with appropriate monomials to make their leading monomials equal (this can be achieved by multiplying f by $\frac{\mathfrak{m}}{\mathfrak{p}}$ and multiplying g by $\frac{\mathfrak{m}}{\mathfrak{q}}$, so that both leading monomials become $\mathfrak{m}$). The resulting two polynomials have equal leading terms (namely, $\mathfrak{m}$), so their leading terms cancel out when we subtract them. The result of this subtraction is $S(f,g)$. To be more specific, when we multiplied f and g with appropriate monomials to make their leading monomial equal, we made sure to choose the latter monomials as low-degree as possible; this is why we took $\mathfrak{m}$ to be the lcm of $\mathfrak{p}$ and $\mathfrak{q}$ and not some other monomial divisible by $\mathfrak{p}$ and $\mathfrak{q}$ (such as the product $\mathfrak{pq}$).

Example 1.1.6. For $n = 2$ (and denoting x_1, x_2 by x, y as usual), we have

$$S\left(x^2y+1, xy^2+1\right) = y\left(x^2y+1\right) - x\left(xy^2+1\right) = y - x$$

and

$$S(xy+1, 2x) = 1(xy+1) - \frac{1}{2}\cdot y \cdot 2x = 1.$$

Note that the cancellation of the leading terms in the construction of $S(f,g)$ is precisely the sort of cancellation that prevented us from having a unique remainder in our above examples.

The following crucial theorem says that these cancellations are a canary in the mine: If they don't happen, then the list is a Gröbner basis.

Theorem 1.1.7 (Buchberger's criterion)**.** Let $\mathbf{b} = (b_1, b_2, \ldots, b_k)$ be a list of nonzero polynomials in P whose leading coefficients are units of R.

Then, $\mathbf{b}$ is a Gröbner basis if and only if every $i < j$ satisfy

$$S\left(b_i, b_j\right) \xrightarrow[\mathbf{b}]{*} 0.$$

The idea behind this theorem is that a list of polynomials (whose leading coefficients are units) is a Gröbner basis if and only if any S-polynomial of two polynomials in the list reduces to 0 modulo the list. Note that "reduces to 0 modulo the list" means that there is some way to get the remainder 0 when applying the division-with-remainder algorithm to this S-polynomial; we are not requiring that **every** way of applying the division-with-remainder algorithm to it will give 0. (But this will follow automatically if we have shown that $\mathbf{b}$ is a Gröbner basis.)

Example 1.1.8. Let $n = 2$, and write x, y for x_1, x_2. Let $I = b_1 P + b_2 P$, where

$$\begin{aligned} b_1 &= xy + 1, \\ b_2 &= y + 1. \end{aligned}$$

We already know from Example 1.1.2 that (b_1, b_2) is not a Gröbner basis, but let us now see this using Buchberger's criterion:

$$S(b_1, b_2) = 1(xy + 1) - x(y + 1) = 1 - x.$$

This polynomial $1 - x$ is already $\mathbf{b}$-reduced (where $\mathbf{b} = (b_1, b_2)$), and it is not 0, so we **don't** have $S(b_1, b_2) \xrightarrow[\mathbf{b}]{*} 0$. Thus, Theorem 1.1.7 confirms again that our $\mathbf{b}$ is not a Gröbner basis.

Example 1.1.9. Let $n = 3$, and write x, y, z for x_1, x_2, x_3. Let $I = b_1 P + b_2 P + b_3 P$, where

$$\begin{aligned} b_1 &= x^2 - yz, \\ b_2 &= y^2 - zx, \\ b_3 &= z^2 - xy. \end{aligned}$$

Is $\mathbf{b} := (b_1, b_2, b_3)$ a Gröbner basis? We check this using Buchberger's criterion. First, we rewrite b_1, b_2, b_3 in a way that their leading terms are up front:

$$\begin{aligned} b_1 &= x^2 - yz, \\ b_2 &= -zx + y^2, \\ b_3 &= -xy + z^2. \end{aligned}$$

(It is generally advised to always write the terms of a polynomial in the deglex order, from highest to lowest, when performing division-with-remainder or computing S-polynomials. Otherwise, it is too easy to get confused about which terms are leading!)

Now, we compute remainders of $S(b_i, b_j)$ modulo $\mathbf{b}$ for all $i < j$:

- We have

$$\begin{aligned} S(b_1, b_2) &= S\left(x^2 - yz, -zx + y^2\right) \\ &= z\left(x^2 - yz\right) - (-x)\left(-zx + y^2\right) = xy^2 - yz^2 \\ &\xrightarrow[\mathbf{b}]{} \left(xy^2 - yz^2\right) - (-y)\left(-xy + z^2\right) \\ &\qquad \left(\begin{array}{c} \text{here, we subtracted } -yb_3 \\ \text{in order to remove the } xy^2 \text{ monomial} \end{array}\right) \\ &= 0, \end{aligned}$$

so that $S(b_1, b_2) \xrightarrow[\mathbf{b}]{*} 0$.

- We have

$$\begin{aligned} S(b_1, b_3) &= S\left(x^2 - yz, -xy + z^2\right) \\ &= y\left(x^2 - yz\right) - (-x)\left(-xy + z^2\right) = xz^2 - y^2 z \\ &\xrightarrow[\mathbf{b}]{} \left(xz^2 - y^2 z\right) - (-z)\left(-zx + y^2\right) \\ &\qquad \left(\begin{array}{c} \text{here, we subtracted } -zb_2 \\ \text{in order to remove the } xz^2 \text{ monomial} \end{array}\right) \\ &= 0, \end{aligned}$$

so that $S(b_1, b_3) \xrightarrow[\mathbf{b}]{*} 0$.

- We have

$$\begin{aligned} S(b_2, b_3) &= S\left(-zx + y^2, -xy + z^2\right) = y\left(-zx + y^2\right) - z\left(-xy + z^2\right) \\ &= y^3 - z^3 \text{ is } \mathbf{b}\text{-reduced and not } 0. \end{aligned}$$

Thus, we **do not** have $S(b_2, b_3) \xrightarrow[\mathbf{b}]{*} 0$. This shows that (b_1, b_2, b_3) is **not** a Gröbner basis.

(This example was a bit unusual in that our many-step reductions were actually one-step reductions. But it is certainly not unusual in that we have wasted a lot of work before getting the answer "no".)

Buchberger's criterion is proved (e.g.) in [DF, p. 324] and in [deGraa20, proof of Theorem 1.1.33]. The "only if" part is obvious; the "if" part is interesting.

Gröbner bases help us better understand ideals of P:

Definition 1.1.10. Let I be an ideal of P. A **Gröbner basis of** I means a Gröbner basis $(b_1, b_2, \ldots, b_k)$ that generates I (that is, that satisfies $I = b_1 P + b_2 P + \cdots + b_k P$).

Corollary 1.1.11 (Macaulay's basis theorem)**.** Let $\mathbf{b} = (b_1, b_2, \ldots, b_k)$ be a list of nonzero polynomials in P whose leading coefficients are units of R. Assume that $\mathbf{b}$ is a Gröbner basis.

Let I be the ideal $b_1 P + b_2 P + \cdots + b_k P$ of P. Then, each element of P/I can be uniquely written in the form

$$\sum_{\substack{\mathfrak{m} \text{ is a } \mathbf{b}\text{-reduced} \\ \text{monomial}}} a_{\mathfrak{m}} \overline{\mathfrak{m}} \qquad \text{with } a_{\mathfrak{m}} \in R$$

(where all but finitely many $\mathfrak{m}$ satisfy $a_{\mathfrak{m}} = 0$). Equivalently, the family $\left(\overline{\mathfrak{m}}\right)_{\mathfrak{m} \text{ is a } \mathbf{b}\text{-reduced monomial}}$ is a basis of the R-module P/I. If none of the polynomials $b_1, b_2, \ldots, b_k$ is constant, then the ring P/b contains "a copy of R".

Proof. LTTR. □

To summarize: If we know a Gröbner basis of an ideal I of P, then we know a lot about I (in particular, we can tell when a polynomial belongs to I, and we can find a basis for P/I). But how do we find a Gröbner basis of an ideal? Is there always one?

Not for arbitrary R. But if R is a field, then yes:

Theorem 1.1.12 (Buchberger's theorem)**.** Let R be a field. Let I be an ideal of the polynomial ring $P = R[x_1, x_2, \ldots, x_n]$. Then, I has a Gröbner basis.

Moreover, if $b_1, b_2, \ldots, b_k$ are nonzero polynomials such that $I = b_1P + b_2P + \cdots + b_kP$, then we can construct a Gröbner basis of I by the following algorithm (**Buchberger's algorithm**):

- Initially, let $\mathbf{b}$ be the list $(b_1, b_2, \ldots, b_k)$.
- As long as there exist two entries of $\mathbf{b}$ whose S-polynomial has a nonzero remainder modulo $\mathbf{b}$, we append this remainder to the list. (It is enough to compute one remainder for each pair of entries of $\mathbf{b}$.)
- Once no such two entries exist any more, we are done: $\mathbf{b}$ is a Gröbner basis of I.

This algorithm always terminates after finitely many steps (i.e., we don't keep adding new entries to $\mathbf{b}$ forever).

We won't prove this, but we will give an example:

Example 1.1.13. Let $n = 3$, and write x, y, z for x_1, x_2, x_3. Let $I = b_1P + b_2P + b_3P$, where

$$b_1 = x^2 - yz,$$
$$b_2 = y^2 - zx,$$
$$b_3 = z^2 - xy.$$

We want to find a Gröbner basis of this ideal I.

As we have seen before, the list $\mathbf{b} := (b_1, b_2, b_3)$ is not a Gröbner basis, since $S(b_2, b_3) = y^3 - z^3$ does not have remainder 0 modulo $\mathbf{b}$. Its remainder is $y^3 - z^3$ itself. Thus, following Buchberger's algorithm, we append this remainder to the list. That is, we set $b_4 = y^3 - z^3$, and continue with the list (b_1, b_2, b_3, b_4). We call this list $\mathbf{b}$ again.

Since $\mathbf{b}$ has grown, we must now also check whether the new S-polynomials

$$S(b_1, b_4), \quad S(b_2, b_4), \quad S(b_3, b_4)$$

reduce to 0 modulo $\mathbf{b}$. Fortunately, they do. Thus, our new list $\mathbf{b} = (b_1, b_2, b_3, b_4)$ is a Gröbner basis of I.

Example 1.1.14. Let $n = 3$, and write x, y, z for x_1, x_2, x_3. Let $I = b_1P + b_2P + b_3P$, where

$$\begin{aligned} b_1 &= x^2 + xy, \\ b_2 &= y^2 + yz, \\ b_3 &= z^2 + zx. \end{aligned}$$

Then, again, it is not hard to see that (b_1, b_2, b_3) is not a Gröbner basis of I. Using Buchberger's algorithm, we can easily compute its Gröbner basis. For example, I has Gröbner basis

$$\left(x^2 + xy, \quad y^2 + yz, \quad xz + z^2, \quad yz^2 - z^3, \quad z^4\right).$$

(Note that the Gröbner basis of an ideal is not unique, so you might get a different one if you perform Buchberger's algorithm differently. When there are several pairs (b_i, b_j) whose S-polynomial does not reduce to 0, you have a choice of which of these pairs you handle first.)

This Gröbner basis reveals that $z^4 \in I$ but $z^3 \notin I$ (since z^3 is reduced modulo the above Gröbner basis). Just working from the original definition of I, this would be far from obvious!

You can do Gröbner basis computations with most computer algebra systems (e.g., SageMath, Mathematica, Singular, SymPy). For example, here is SageMath code for the Gröbner basis of the above ideal. Note that we took $R = \mathbb{Q}$ in this computation (the "`QQ`" means the ring of rational numbers), but the same computation works over any field R (and, because our ideal is rather nice, even over any commutative ring R; this is not automatic).

1.1.2. Term orders

We have so far been using the deg-lex order on the monomials. There are many other total orders that share most of its nice properties and are often more suited for specific problems.

Let me only mention the **lexicographic order**, which is defined just as the deg-lex order but without taking the degree into account. That is:

Definition 1.1.15. We define a total order $\prec$ (called the **lexicographic order**, or – for short – the **lex order**) on the set $C^{(n)}$ of all monomials as follows:

For two monomials $\mathfrak{m} = x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$ and $\mathfrak{n} = x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n}$, we declare that $\mathfrak{m} \prec \mathfrak{n}$ if and only if

- there is an $i \in \{1, 2, \ldots, n\}$ such that $a_i \neq b_i$, and the **smallest** such i satisfies $a_i < b_i$.

Recall the proposition from Lecture 16 where we collected properties of the deg-lex order:

Proposition 1.1.16. **(a)** The deg-lex order really is a total order on $C^{(n)}$.

(b) If $\mathfrak{m}, \mathfrak{n}, \mathfrak{p} \in C^{(n)}$ satisfy $\mathfrak{m} \prec \mathfrak{n}$, then $\mathfrak{mp} \prec \mathfrak{np}$.

(c) We have $1 \preccurlyeq \mathfrak{m}$ for any $\mathfrak{m} \in C^{(n)}$.

(d) Let $\mathfrak{m} \in C^{(n)}$ be any monomial. Then, there are only finitely many monomials $\mathfrak{p}$ such that $\mathfrak{p} \prec \mathfrak{m}$.

(e) There are no infinite decreasing chains $\mathfrak{m}_0 \succ \mathfrak{m}_1 \succ \mathfrak{m}_2 \succ \cdots$ of monomials.

(f) If T is a nonempty finite set of monomials, then T has a largest element with respect to $\prec$ (that is, an element $\mathfrak{t} \in T$ such that $\mathfrak{m} \preccurlyeq \mathfrak{t}$ for all $\mathfrak{m} \in T$).

(g) If T is a nonempty set of monomials, then T has a smallest element with respect to $\prec$ (that is, an element $\mathfrak{t} \in T$ such that $\mathfrak{m} \succcurlyeq \mathfrak{t}$ for all $\mathfrak{m} \in T$).

All of these properties except for part **(d)** hold for the lex order as well. Part **(d)** fails (for $n > 1$), because x_1 is larger (with respect to the lex order) than **any** power of x_2 (and, of course, there are infinitely many powers of x_2). Part **(e)** is still true, but its proof is harder. However, the theory of Gröbner bases does not use part **(d)**, so it still can be done with the lex order. This yields new (in general, different) Gröbner bases.

Example 1.1.17. Let $n = 3$ and let $I = \left(x^2 - y\right) P + \left(y^2 - z\right) P + \left(z^2 - x\right) P$ (where we write x, y, z for x_1, x_2, x_3). Then, a Gröbner basis of I with respect to the deg-lex order is

$$\left(x^2 - y, y^2 - z, z^2 - x\right)$$

(this is precisely the list of generators that we started with). But this is not a Gröbner basis with respect to the lex order. Instead, a Gröbner basis of I with respect to the lex order is

$$\left(x - z^2, y - z^4, z^8 - z\right).$$

Example 1.1.18. Let $n = 3$ and let $I = \left(x^2 - y^3\right) P + \left(y^4 - z^2\right) P + \left(z^2 - x^5\right) P$ (where we write x, y, z for x_1, x_2, x_3). Then, a Gröbner basis of I with respect to the deg-lex order is

$$\Big(z^6 - yz^2, x^3z^2 - yz^2, xy^2z^2 - z^2, xz^4 - y^2z^2, \\ yz^4 - xz^2, x^4 - y^2z^2, x^2y - z^2, y^3 - x^2\Big).$$

But a Gröbner basis of I with respect to the lex order is

$$\left(x^2 - y^3, xz^2 - z^8, y^4 - z^2, yz^2 - z^6, z^{16} - z^2\right).$$

In the SageMath computer algebra system, you can signal the use of the lex order (as opposed to the deg-lex order, which is used by default) by replacing "`PolynomialRing(QQ)`" by "`PolynomialRing(QQ, order="lex")`".

This last example illustrates one reason to vary the total order on monomials: Gröbner bases can often be rather long (even if the ideal is easy to write down). The size of a Gröbner basis can be doubly exponential in the number of generators of I (I believe). In real life, this worst case doesn't happen very often, but when it does, switching to a different monomial order will often make it easier. (It's essentially a way of rolling the dice again if you got an unlucky roll.)

1.2. Solving polynomial systems using Gröbner bases

Another occasion to use Gröbner bases (and the lex order in particular) is solving systems of polynomial equations. Polynomial equations are closely connected to ideals:

Definition 1.2.1. Let $b_1, b_2, \ldots, b_k$ be k polynomials in P, and let A be a commutative R-algebra. Then, a **root** (or, alternatively, a **common root**) of $(b_1, b_2, \ldots, b_k)$ in A means an n-tuple $(a_1, a_2, \ldots, a_n) \in A^n$ such that

$$b_i(a_1, a_2, \ldots, a_n) = 0 \qquad \text{for all } i \in \{1, 2, \ldots, k\}.$$

This definition generalizes the standard notion of a root of a polynomial to multiple variables and multiple polynomials.

Thus, solving systems of polynomial equations means finding roots of lists of polynomials. It turns out that the list of polynomials doesn't really matter; what does is the ideal it generates:

Proposition 1.2.2. Let $b_1, b_2, \ldots, b_k$ be k polynomials in P, and let A be a commutative R-algebra.

Then, the roots of $(b_1, b_2, \ldots, b_k)$ in A depend only on the ideal generated by $b_1, b_2, \ldots, b_k$, rather than on the polynomials $b_1, b_2, \ldots, b_k$ themselves.

More concretely: If $I = b_1P + b_2P + \cdots + b_kP$ is the ideal of P generated by $b_1, b_2, \ldots, b_k$, then the roots of $(b_1, b_2, \ldots, b_k)$ are precisely the n-tuples $(a_1, a_2, \ldots, a_n) \in A^n$ such that

$$f(a_1, a_2, \ldots, a_n) = 0 \qquad \text{for all } f \in I.$$

Proof. Easy, LTTR. (You have to prove that if $(a_1, a_2, \ldots, a_n) \in A^n$ is a root of $(b_1, b_2, \ldots, b_k)$, then $f(a_1, a_2, \ldots, a_n) = 0$ for all $f \in I$. But this is easy: Each $f \in I$ is a P-linear combination $c_1b_1 + c_2b_2 + \cdots + c_kb_k$ of $(b_1, b_2, \ldots, b_k)$, and therefore satisfies

$$\begin{aligned} & f(a_1, a_2, \ldots, a_n) \\ &= c_1(a_1, a_2, \ldots, a_n) \underbrace{b_1(a_1, a_2, \ldots, a_n)}_{=0} + c_2(a_1, a_2, \ldots, a_n) \underbrace{b_2(a_1, a_2, \ldots, a_n)}_{=0} \\ &\qquad + \cdots + c_k(a_1, a_2, \ldots, a_n) \underbrace{b_k(a_1, a_2, \ldots, a_n)}_{=0} \\ &= 0. \end{aligned}$$

The converse is even more obvious, since the polynomials $b_1, b_2, \ldots, b_k$ all belong to I.) □

Thus, if we want to find the roots of $(b_1, b_2, \ldots, b_k)$, we can replace $(b_1, b_2, \ldots, b_k)$ by any other tuple of polynomials that generates the same ideal of P. (This is just the polynomial analogue of the classical "addition" strategy for solving systems of linear equations.)

One of the most useful ways to do this is to replace $(b_1, b_2, \ldots, b_k)$ by a Gröbner basis of the ideal it generates – particularly, by a Gröbner basis with respect to the lex order. Let us see how this helps on an example:

Example 1.2.3. Recall Exercise 8 on homework set #0:

> Solve the following system of equations:
>
> $$\begin{aligned} a^2 + b + c &= 1, \\ b^2 + c + a &= 1, \\ c^2 + a + b &= 1 \end{aligned}$$
>
> for three complex numbers a, b, c.

Let us formalize this in terms of polynomials and roots. We set $R = \mathbb{Q}$ and $n = 3$, and we write x, y, z for x_1, x_2, x_3. Thus, the exercise is asking for the roots of

$$\left(x^2 + y + z - 1,\ \ y^2 + z + x - 1,\ \ z^2 + x + y - 1\right)$$

in the $\mathbb{Q}$-algebra $\mathbb{C}$.

Let I be the ideal of $P = \mathbb{Q}[x,y,z]$ generated by the three polynomials $x^2+y+z-1$, $y^2+z+x-1$, $z^2+x+y-1$. Using a computer (or a lot of patience), we can easily find a Gröbner basis of I with respect to the lex order. We get

$$\left(x+y+z^2-1,\ y^2-y-z^2+z,\ yz^2+\frac{1}{2}z^4-\frac{1}{2}z^2,\ z^6-4z^4+4z^3-z^2\right).$$

We observe that the last polynomial in this Gröbner basis only involves the variable z ! Thus, the c entry in each of the solutions (a,b,c) of our system must be a root of this polynomial $z^6-4z^4+4z^3-z^2$. We can therefore find all possibilities for c by finding the roots of this polynomial (I am here assuming that you can solve univariate polynomials; we will learn a bit more about this in Lecture 18). In our concrete case, we can easily do this:

$$z^6-4z^4+4z^3-z^2 = z^2(z-1)^2\left(z^2+2z-1\right).$$

Thus, the options for c are $0, 1, \sqrt{2}-1, \sqrt{2}+1$.

Now let us find b. Either we use the symmetry of the original system to argue that the options for b must be the same as for c; or we use the second-to-last polynomial in our Gröbner basis (or the second one) to compute b now that c is known. At last, we get to a in a similar way.

In the end, we get finitely many options for (a,b,c). We need to check which of these options actually are solutions of the original system. This is a lot of work, but a computer can do it.

Of course, there are more elegant ways to solve the above exercise (otherwise, I would not have posed it on homework set #0). However, the way we just showed is generalizable. In general, if a system of polynomial equations over $\mathbb{C}$ has only finitely many solutions, then we can find them all in this way (provided that we have an algorithm for finding all roots of a univariate polynomial).[1] Thus, using Gröbner bases with respect to the lex order, we can (often) reduce solving systems of polynomial equations in multiple variables to

[1] If a system of polynomial equations has infinitely many solutions, then this strategy usually will not work. For example, if we try to use it to solve the system

$$\begin{aligned} ab &= 0, \\ bc &= 0, \\ ca &= 0, \end{aligned}$$

then we find the Gröbner basis (xy, yz, xz), which doesn't get us any closer to the solutions. Blame this on the problem, not on the Gröbner basis: The system has a more complicated combinatorial structure (its solution set is the union of the three axes in 3D space; there are infinitely many options for each of a, b, c).

solving polynomial equations in a single variable.

Some things don't look like systems of polynomial equations, but yet boil down to such systems. Here is an example:

Example 1.2.4. Recall Exercise 4 on homework set #0:

Simplify $\sqrt[3]{2+\sqrt{5}}+\sqrt[3]{2-\sqrt{5}}$.

There are various ways of solving this using some creativity or lucky ideas. Let us try to be more methodical here. We set

$$a=\sqrt[3]{2+\sqrt{5}}, \qquad b=\sqrt[3]{2-\sqrt{5}}, \qquad c=\sqrt[3]{2+\sqrt{5}}+\sqrt[3]{2-\sqrt{5}}.$$

Thus, we want to find a simpler expression for c. A good first step would be to find a polynomial whose root c is (since we would then have a chance of finding c by root-finding techniques). We see that a, b, c satisfy the following system of equations:

$$\begin{aligned}\left(a^3-2\right)^2-5&=0,\\ \left(b^3-2\right)^2-5&=0,\\ a+b-c&=0.\end{aligned}$$

(Indeed, the first equation comes from "unraveling" $a=\sqrt[3]{2+\sqrt{5}}$, and likewise for the second; the third comes from the obvious fact that $c=a+b$.)

We try to solve this system using Gröbner bases. Thus, we consider the ideal

$$I:=\left(\left(x^3-2\right)^2-5\right)P+\left(\left(y^3-2\right)^2-5\right)P+(x+y-z)\,P$$

of the polynomial ring $P=\mathbb{Q}[x,y,z]$. Using SageMath, we can easily find a Gröbner basis of this ideal I with respect to the lex order. Its last entry is a polynomial that involves only the variable z, so we can narrow down the options for c to the roots of this polynomial.

This looks nice in theory, but in practice you will realize that this last entry is

$$z^{21}-40z^{18}+218z^{15}-72z^{12}-9931z^9-5216z^6+19136z^3-4096.$$

Eek. With a good computer algebra system, you can factor this polynomial, but there will be some degree-4 factors irreducible over $\mathbb{Q}$. The polynomial has 5 real roots, so c must be one of them, but we need some harder work to find out which one. This is all not very convenient.

But our approach can be salvaged. We have been "throwing away" information about our a, b, c; no wonder that we got so many options for c.

Indeed, the equation $\left(a^3-2\right)^2-5=0$ doesn't really mean $a=\sqrt[3]{2+\sqrt{5}}$; it only means that a is **some** cube root of (2 plus **some** square root of 5). Here, we are using the word "root" in the wider sense, so a nonzero complex number has two square roots and three cube roots; thus, there are 6 possibilities in total for a. Likewise for b. Our system of equations above allows c to be a sum of any of the 6 possible a's with any of the 6 possible b's. Unsurprisingly, this leaves lots of different options for c.

Thus, we need to integrate a bit more information about the actual values of a, b into our system. Of course, we know that a is the **real** cube root of the **positive** square root of 5. But this is not the kind of information we can easily integrate into a system of equations.

However, we can observe that

$$\begin{aligned} ab &= \sqrt[3]{2+\sqrt{5}} \cdot \sqrt[3]{2-\sqrt{5}} = \sqrt[3]{\left(2+\sqrt{5}\right)\cdot\left(2-\sqrt{5}\right)} \\ &\qquad \left(\text{since } \sqrt[3]{u}\cdot\sqrt[3]{v} = \sqrt[3]{uv} \text{ for any } u, v \in \mathbb{R}\right) \\ &= \sqrt[3]{-1} = -1. \end{aligned}$$

Thus, we can extend our system to

$$\begin{aligned} \left(a^3-2\right)^2-5 &= 0, \\ \left(b^3-2\right)^2-5 &= 0, \\ a+b-c &= 0, \\ ab+1 &= 0. \end{aligned}$$

This is a different system and has a smaller set of solutions than the previous one, but that's good news, since the solution we are looking for is one of its solutions.

Now, solving this new system using the Gröbner basis technique, we find that c is a root of the polynomial z^3+3z-4 (since this polynomial is the last entry of the Gröbner basis we find). But the roots of this polynomial are easy to find: The factorization

$$z^3+3z-4 = (z-1)\underbrace{\left(z^2+z+4\right)}_{\text{always positive on } \mathbb{R}}$$

shows that its only real root is 1, so that c must be 1 (since c is real by definition). Thus our exercise is solved.

See [CoLiOs15] for more about solving systems of polynomial equations, and for further applications of Gröbner bases.

References

[CoLiOs15] David A. Cox, John Little, Donal O'Shea, *Ideals, Varieties, and Algorithms*, Undergraduate Texts in Mathematics, 4th edition, Springer 2015.
https://dx.doi.org/10.1007/978-3-319-16721-3

[deGraa20] Willem de Graaf, *Computational Algebra*, 22 October 2020.
https://www.science.unitn.it/~degraaf/algnotes/compalg.pdf