PMATH446 FINAL PROJECT GRÖBNER BASIS

JIAHUI HUANG

Contents

1. Introduction

Through out this note we will denote $R = k[x_1, ..., x_n]$ where k is a field. By a basis of an ideal, we mean a generating set of the ideal (not necessarily independent).

Given an ideal $I = \langle f_1, ..., f_m \rangle$ of R, consider the following problems:

Question 1.1 (Division and ideal membership problem)**.** *Given $f \in R$, determine if $f \in I$. If not, express the image of f in R/I in terms of a basis for the vector space R/I over k.*

Question 1.2 (The problem of solving polynomial equations)**.** *Find solutions in k^n of the equations*

$$f_1(x_1, ..., x_n) = ... = f_m(x_1, ..., x_n) = 0$$

namely compute the variety $V(I)$.

When there is only one variable, problem (1) can be solved using Euclid's algorithm, which computes the unique monic generator g of the ideal $I = \langle f_1(x), ..., f_m(x) \rangle$. We can write $f = gq + r$ using division algorithm, where r is the remainder, namely the image of f in R/I, expressed in the basis $\{1, x, ..., x^{\deg(g)-1}\}$. When the polynomials $f_1, ..., f_m$ are linear, problem (2) is solved using Gaussian elimination that reduces the matrix for the system of equations into row echelon form.

We shall try to generalize the division algorithm with the idea of "initial ideals". During which we will be able to define the Gröbner basis. We will then discuss its properties and relations to the above problems. We will see how to find such a basis and its applications in solving question 1.2.

2. Monomial Orderings

Both Euclid's algorithm and Gaussian elimination take in $I = \langle f_1, ..., f_m \rangle$ and outputs a generating set for the ideal that is simpler to deal with. Observe that in each step, we are reducing the polynomias with respect to some ordering of the terms.

Euclid's algorithm uses division to find the greatest common denominator, where the degree of the polynomial is reduced. This uses the ordering

$$1 < x < x^2 < ...$$

Gaussian elimination eliminates the number of variables, namely a system of linear equations in echelon form satisfies that the i-th equation does not contain the terms $x_1, ..., x_{i-1}$. This corresponds to the ordering

$$x_n < x_{n-1} < ... < x_1$$

In general we can define similar orderings on all monomials of R.

Definition 2.1. A **monomial order** is a total order $>$ on the set of monomials in $\{x^\alpha = x_1^{a_1}...x_n^{a_n} : \alpha = (a_1, ..., a_n) \in \mathbb{N}^n\} \subseteq R$ such that for multi-indices α, β, γ,

(1) $x^\alpha < x^\beta \Rightarrow x^{\alpha+\gamma} < x^{\beta+\gamma}$
(2) $1 < x^\alpha$ for all $\alpha \neq (0, ..., 0)$.

From the conditions we see if x^α is divisible by x^β, then $x^\alpha \geq x^\beta$ as a result of $x^{\alpha-\beta} \geq 1$. This shows that the a monomial order $>$ is a well-ordering:

Lemma 2.2 (Eis, Lemma 15.2)**.** *Given any monomial order, every set of monomials have a least element.*

Proof. Let X be a set of monomials. Since R is Noetherian, the ideal generated by X is generated by a finite subset $Y \subseteq X$. Then every element in X is a multiple of some element in Y. So each element in X is greater than or equal to some element in Y. Hence the least element of X would be the least element of the finite set Y. □

Example 2.3 (Lexicographic ordering). $x^\alpha > x^\beta$ if and only if the first non-zero entry of $\alpha - \beta$ is positive. This is the order by index when the polynomials are linear. ◇

Example 2.4 (Degree lexicographic ordering). $x^\alpha > x^\beta$ if and only if $\deg x^\alpha > \deg x^\beta$ or $\deg x^\alpha = \deg x^\beta$ and the first non-zero entry of $\alpha - \beta$ is positive. This is the order by degree when there is only one variable. ◇

Example 2.5 (Reverse lexicographic ordering). $x^\alpha > x^\beta$ if and only if $\deg x^\alpha > \deg x^\beta$ or $\deg x^\alpha = \deg x^\beta$ and the *last* non-zero entry of $\alpha - \beta$ is *positive*. This ordering favours larger total degree, but when total degree is equal, it favours smaller power on the rightmost variable. ◇

Ideals generated by monomials are called **monomial ideals**. Similar to homogeneous ideals in a graded ring, working with monomial ideals simplifies many problems. For example, x^α is a member of the ideal $I = \langle x^{\alpha_1}, ..., x^{\alpha_m} \rangle$ if and only if $x^\alpha = \sum_i f_i x^{\alpha_i}$ for some $f_i \in R$, and by comparing coefficients we see this holds if and only if x^α is divisible by some x^{α_i}; similarly, an arbitrary polynomial f is a member of I if and only if each term of f is divisible by some x^{α_i} [Cox, Lemma 4.2.3]. Finding greatest common divisors and least common multiples for monomials is also trivialized to

$$\text{GCD}(x^\alpha, x^\beta) = x_1^{\min(\alpha_1,\beta_1)}...x_n^{\min(\alpha_n,\beta_n)}$$
$$\text{LCM}(x^\alpha, x^\beta) = x_1^{\max(\alpha_1,\beta_1)}...x_n^{\max(\alpha_n,\beta_n)}$$

To describe a division algorithm on R, we use the following terminology:

Definition 2.6. Let $>$ be a monomial order and $f \in R$. The **initial term** of f, written as $\text{in}_>(f)$ or $\text{in}(f)$ (when the order is clear), is the greatest term of f with respect to $>$. The **initial ideal** of an ideal I is the ideal generated by all initial terms of elements in I:

$$\text{in}_>(I) = \langle \text{in}_>(f) : f \in I \rangle$$

When working with monomials, since k is a field, scalar multiples do not affect the ideals they generate. Sometimes we will apply definitions or results to monomials that are not monic, this would be interpreted as first multiplying a scalar to cancel the coefficient, and then apply the results.

Theorem 2.7 (Division Algorithm)**.** *Let $>$ be a monomial order and let $f_1, ..., f_s \in R$. Then every $f \in R$ can be written in form*

$$f = q_1 f_1 + ... + q_s f_s + r$$

where $q_i, r \in R$ and either $r = 0$ or r is a linear combination of monomials that are not in $\langle \text{in}(f_1), ..., \text{in}(f_s) \rangle$. Furthermore, for each i,

$$\text{in}(f) \geq \text{in}(q_i f_i)$$

Proof. We begin by setting $p = f$ and $q_1 = ... = q_s = r = 0$, and repeat the following steps until $p = 0$:

(1) If some $\text{in}(f_i)$ divides $\text{in}(p)$, then subtract $\frac{\text{in}(p)}{\text{in}(f_i)} f_i$ from p and add $\frac{\text{in}(p)}{\text{in}(f_i)}$ to q_i.
(2) If no $\text{in}(f_i)$ divides $\text{in}(p)$, then subtract $\text{in}(p)$ from p and add it to r.

First we show that the algorithm terminates. Suppose we performed step (1), setting p to be

$$p' = p - \frac{\text{in}(p)}{\text{in}(f_i)} f_i$$

Observe that the initial term of $\frac{\text{in}(p)}{\text{in}(f_i)} f_i$ is the same as the initial term of p, and it will be cancelled out in the difference. Thus the initial term of the new value of p is strictly smaller. Suppose we performed step (2), then we also see that the initial term of p become strictly smaller.

By Lemma 2.2, a decreasing chain of initial terms would eventually stabilize after finitely many steps, which means that we would not be able to perform either step (1) or (2). Thus we reach the case $p = 0$ and the algorithm terminates.

It remains to show the algorithm outputs the required polynomials. During each step, the following holds

$$f = q_1 f_1 + ... + q_s f_s + p + r$$

because the changes on the right hand side cancel out. Note that in each step, the initial term of p is at most the initial term of f. As a result, terms added to q_i satisfy $\text{in}(f) \geq \text{in}\left(\frac{\text{in}(p)}{\text{in}(f_i)} f_i\right)$, so $\text{in}(f) \geq \text{in}(q_i f_i)$. At the end of the algorithm, we have $p = 0$ and terms are added to r only if they are not divisible by any $\text{in}(f_i)$, so the final expression

$$f = q_1 f_1 + ... + q_s f_s + r$$

satisfies the requirements. □

Note that for an ideal $I = \langle f_1, ..., f_s \rangle$, the ideals $\langle \text{in}(f_1), ..., \text{in}(f_s) \rangle$ and $\text{in}(I)$ may be different, as seen in the following example.

Example 2.8. Let $I = \langle xy, y^2 - x \rangle$, then

$$y \cdot (xy) - x \cdot (y^2 - x) = x^2 \in I$$

Using degree lexicographic ordering, we have $x^2 \in \text{in}(I)$. But

$$\langle \text{in}(xy), \text{in}(y^2 - x) \rangle = \langle xy, y^2 \rangle$$

does not contain x^2. ◇

When the ideals $\langle \text{in}(f_1), ..., \text{in}(f_s) \rangle$ and $\text{in}(I)$ are the same, the remainder in the division algorithm is expressed in a basis for R/I by the following theorem, thus giving us an answer to Question 1.1. In particular, f is inside I if and only if the remainder is 0. We shall see later that such a basis $f_1, ..., f_s$ for I always exists, and is called a **Gröbner basis**.

Theorem 2.9 (Macaulay's Theorem [Eis, Thm 15.3])**.** *Let I be any ideal of R, for any monomial order $>$, the set B of all monomials not in* $\text{in}(I)$ *form a basis for R/I.*

Proof. Suppose B is not linearly independent. Let $\sum_i u_i x^{\beta_i} \in I$ be a non-trivial linear relation where $0 \neq u_i \in k$, $x^{\beta_i} \in B$. Say $u_i x^{\beta_i}$ is the initial term of this expression, then $x^{\beta_i} \in \text{in}(I)$, contraidtcting the definition of B.

Suppose B does not span R/I. By Lemma 2.2 there is an element f with minimal initial term such that $f + I$ is not in the span of B. If $\text{in}(f) \in B$, then subtracting $\text{in}(f)$ gives an element with smaller initial term, whose image is also not in the span of B, contradicting the minimality of f. If $\text{in}(f) \notin B$, so $\text{in}(f) \in \text{in}(I)$. Say $\text{in}(f) = \text{in}(g)$ for some $g \in I$. Then $f - g$ has the same image in R/I as f, but with a smaller initial term, again contradicting minimality of f. □

3. Gröbner Basis

Theorem 3.1 (Dickson's Lemma, [Cox, Theorem 2.4.5])**.** *Let $I = \langle x^\alpha : \alpha \in A \rangle$ be a monomial ideal. Then $I = \langle x^{\alpha_1}, ..., x^{\alpha_s} \rangle$ for some $\alpha_1, ..., \alpha_s \in A$.*

Proof. When $n = 1$, $A \subseteq \mathbb{N}$, so we can take its least element α and then $I = \langle x^\alpha \rangle$.

Suppose for an induction that the theorem is true for $n - 1$. Let

$$J = \langle x^\beta : \beta = (\alpha(1), ..., \alpha(n-1)), \alpha \in A \rangle \subseteq k[x_1, ..., x_{n-1}]$$

By induction hypothesis J is generated by finitely many β's, say $J = \langle x^{\beta_1}, ..., x^{\beta_t} \rangle$. For each i, we can find some $\alpha_i \in A$ such that $\alpha_i = (\beta_i, \alpha_i(n))$. Take $m = \max_i\{\alpha_i(n)\}$, and for each $0 \leq j \leq m - 1$, define ideals

$$J_j = \langle x^\beta : \beta = (\alpha(1), ..., \alpha(n-1)), \alpha \in A, \alpha(n) = j \rangle$$

Now by induction hypothesis again, $J_j = \langle x^{\beta_1^{(j)}}, ..., x^{b_{t_j}^{(j)}} \rangle$. Now for each $\alpha \in A$, let $\beta = (\alpha(1), ..., \alpha(n-1))$. Then $x^\beta \in J$ and is a multiple of x^{β_i} for some i. If $\alpha(n) \geq m > \alpha_i(n)$, then x^α would be a multiple of x^{α_i}. If $\alpha(n) \leq m - 1$, then $x^\beta \in J_{\alpha(n)}$ and is a multiple of $x^{\beta_i^{(\alpha(n))}} x_n^{\alpha(n)}$. Therefore I is generated by a subset of the finite set

$$\{x^{\alpha_i} : i = 1, ..., t\} \cup \bigcup_{j=0}^{m-1} \{x^{\beta_i^{(j)}} x_n^j : i = 1, ..., t_j\} \subseteq I$$

and the powers on x are indeed a subset of A, finishing the induction. □

Note that $\text{in}(I)$ is a monomial ideal, so applying Dickson's lemma yields the following corollary

Corollary 3.2. *Let $I \subseteq R$ be a non-zero ideal.* $\text{in}(I)$ *has a finite basis of form* $\{\text{in}(g_1), ..., \text{in}(g_t)\}$ *for some $g_1, ..., g_t \in I$.*

Something interesting to note is that given a finite basis for a monomial ideal I, if any element is divisible by another, we can remove it and obtain a smaller basis. Repeating this process we get a basis whose elements do not divide each other. Such a "minimal" basis of (monic) monomials is unique because given two such bases A and B, each element in A would be divisible by an element in B and vise versa, giving us $A \subseteq B$ and $B \subseteq A$.

Theorem 3.3 (Gröbner basis theorem [Cox, Thm 2.5.4])**.** *Let $I \subseteq R$ be an ideal. Then $I = \langle g_1, ..., g_t \rangle$ for $g_1, ..., g_t \in I$ from the previous Corollary.*

Proof. Say I is non-zero. Pick any monomial order. By the previous corollary, we can write $\text{in}(I) = \langle \text{in}(g_1), ..., \text{in}(g_t) \rangle$ for some $g_1, ..., g_t \in I$. We shall proceed to show that $I = \langle g_1, ..., g_t \rangle$.

Let $f \in I$, apply division algorithm and get

$$f = q_1 g_1 + ... + q_t g_t + r$$

where either the monomials in r is not divisible by any $\text{in}(g_i)$ or $r = 0$. However $r = f - q_1 g_1 - ... - q_t g_t \in I$ implies $\text{in}(r) \in \text{in}(I) = \langle \text{in}(g_1), ..., \text{in}(g_t) \rangle$, which means $\text{in}(r)$ is divisible by some $\text{in}(g_i)$, so we must be in the case $r = 0$ and $f \in \langle g_1, ..., g_t \rangle$. □

Definition 3.4. Let $>$ be a monomial order and $I \subseteq R$. A **Gröbner basis** of I with respect to $>$ is a finite set of elements $g_1, ..., g_t \in I$ such that

$$\text{in}(I) = \langle \text{in}(g_1), ..., \text{in}(g_t) \rangle$$

We know from the previous theorem that a Gröbner basis for I is a basis for I. Recall from last section that such a basis gives a solution to Question 1.1 when used with division algorithm. By Dickson's lemma, such a basis always exists. Although the proof of Dickson's lemma is technically constructive, it is too complicated practically. To fully answer Question 1.1, we shall develop an algorithm to compute for a Gröbner basis in the next section.

4. Buchberger's Criterion and Algorithm

Macaulay's Theorem shows that when dividing using a Gröbner basis, the remainder is always unique, but in general, the remainder from division algorithm depends on the choice of i in step (1) (as in the proof of Theorem 2.8). We shall make this process determinate by taking i to be maximal. This requires us to treat the set $\{f_1, ..., f_s\}$ as an ordered s-tuple.

Definition 4.1. Write $\overline{f}^F$ for the unique remainder from the determinate division algorithm, when dividing f by ordered s-tuple $F = (f_1, ..., f_s)$.

Definition 4.2. Let $0 \neq f, g \in R$ be polynomials. The **S-polynomial** of f and g is

$$S(f, g) = \frac{\text{LCM}(\text{in}(f), \text{in}(g))}{\text{in}(f)} f - \frac{\text{LCM}(\text{in}(f), \text{in}(g))}{\text{in}(g)} g \in R$$

Similar to two integers x, y satisfy the relation $\frac{\text{lcm}(x,y)}{x} x - \frac{\text{lcm}(x,y)}{y} y = 0$, the S-polynomials are constructed to match the initial terms of f and g and subtract them, so that the initial terms cancel out. In particular,

$$(\star) \qquad \text{in}(S(f, g)) < \text{in}\left(\frac{\text{LCM}(\text{in}(f), \text{in}(g))}{\text{in}(f)} f \right)$$

With these notations, we can state Buchberger's Criterion for when a basis is a Gröbner basis.

Theorem 4.3 (Buchberger's Criterion)**.** *Let $I \subseteq R$ be an ideal. A basis $G = \{g_1, ..., g_t\} \subseteq I$ is a Gröbner basis if and only if for all pairs $i \neq j$, we can view G as an ordered t-tuple in some order, and have*

$$\overline{S(g_i, g_j)}^G = 0$$

Before giving the proof, we shall first show a useful property of S-polynomials which will be an essential step of the proof.

Lemma 4.4. *Suppose we have a sum $\sum_{i=1}^t p_i$ where the initial terms of p_i differ only by scalars. If $\text{in}\left(\sum_{i=1}^t p_i\right) < \text{in}(p_i)$, then $\sum_{i=1}^t p_i$ is a linear combination of $S(p_1, p_j)$ with coefficients in k.*

Proof. Since all p_i have the same initial term and $\text{in}\left(\sum_{i=1}^t p_i\right) < \text{in}(p_i)$ suggests that the initial term gets canceled, let d_i be the leading coefficient of p_i, we must have $\sum_{i=1}^t d_i = 0$. Observe that

$$\frac{\text{LCM}(\text{in}(p_i), \text{in}(p_j))}{\text{in}(p_i)} = \frac{1}{d_i}$$

So

$$S(p_i, p_j) = \frac{1}{d_i} p_i - \frac{1}{d_j} p_j$$

Now

$$\begin{aligned}
\sum_{j=2}^{t} -d_j S(p_1, p_j) &= \sum_{j=2}^{t} p_j - \frac{d_j}{d_1} p_1 \\
&= \sum_{j=2}^{t} p_j + \frac{-\sum_{j=2}^{t} d_j}{d_1} p_1 \\
&= \sum_{j=2}^{t} p_j + \frac{d_1}{d_1} p_1, \text{ since } \sum_{j=1}^{t} d_j = 0 \\
&= \sum_{i=1}^{t} p_i
\end{aligned}$$

□

proof of Buchberger's Criterion. Let $G = \{g_1, ..., g_t\}$ be a basis of I. Suppose G is a Gröbner basis. Since $S(g_i, g_j) \in I$, by the division algorithm we automatically have $\overline{S(g_i, g_j)}^{G} = 0$.

Conversely, suppose all $S(g_i, g_j)$ have remainder 0. Suppose for a contradiction that G is not a Gröbner basis. Then there exists some expression

$$f = \sum_{u=1}^{t} f_u g_u \in I$$

such that $\text{in}(f) \notin \langle \text{in}(g_1), ..., \text{in}(g_t) \rangle$. Let $m = \max(\text{in}(f_u g_u))$. By Lemma 2.2 we may assume that this expression for f is chosen so that m is as small as possible. If $m = \text{in}(f)$ (up to a scalar mutiple), then $\text{in}(f) = \text{in}(f_v g_v)$ for some v, and it follows $\text{in}(g_v)$ divides $\text{in}(f)$, which contradicts out assumption. Thus $m > \text{in}(f)$.

Rearrange the summands so that the first t' of them have initial terms are scalar multiples of m. We have

$$\begin{aligned}
f = \sum f_u g_u &= \sum_{v \le t'} f_v g_v + \sum_{u > t'} f_u g_u \\
&= \underbrace{\sum_{v} \text{in}(f_v) g_v}_{\text{initial term of each summand is scalar multiple of } m} + \underbrace{\sum_{v} (f_v - \text{in}(f_v)) g_v + \sum_{u > t'} f_u g_u}_{\text{their initial terms are strictly less than } m}
\end{aligned}$$

Our goal is to replace these summands by something smaller, thus finding another expression of f with a smaller m and contradicting its minimality. Since $m > \text{in}(f)$, the initial terms (scalar multiples of m) in the above sum must cancel out, meaning

$$\text{in}\left(\sum_{v} \text{in}(f_v) g_v \right) < m$$

Let $p_v = \text{in}(f_v) g_v$, then $\text{in}(p_v)$ are scalar multiples of m for all $v \in V$, and we can apply Lemma 4.4 to get

$$\sum_{i} \text{in}(f_v) g_v = \sum a_v S(p_1, p_v)$$

for some scalars a_v. Let d_v be the leading coefficient of p_v, then

$$\begin{aligned} S(p_1, p_v) =& \frac{1}{d_1} p_1 - \frac{1}{d_v} p_v \\ =& \frac{1}{d_1} \text{in}(f_1) g_1 - \frac{1}{d_v} \text{in}(f_v) g_v \end{aligned}$$

Note that this expression is very similar to the S-polynomial, namely it matches the initial terms of g_1, g_v by multiplying monomials (in this case $\frac{1}{d_1}\text{in}(f_1)$ and $\frac{1}{v}\text{in}(f_v)$) and subtracts them. The only difference is that we are matching the initial polynomials to m in this case instead of $\text{LCM}(\text{in}(g_1), \text{in}(g_v))$, therefore we observe

$$S(p_1, p_v) = \frac{m}{\text{LCM}(\text{in}(g_1), \text{in}(g_v))} S(g_1, g_v) := m_v S(g_1, g_v)$$

Recall our assumption that $S(g_i, g_j)$ has remainder 0 when divide by G for all i, j, so division algorithm gives

$$S(g_1, g_v) = \sum_{u=1}^{t} f_u^{(v)} g_u \text{ where } \text{in}(f_u^{(v)} g_u) \leq \text{in}(S(g_1, g_v))$$

Now equation $(\star)$ from Definition 4.2 says

$$\text{in}(S(g_1, g_v)) < \text{in}\left(\frac{\text{LCM}(\text{in}(g_1), \text{in}(g_v))}{\text{in}(g_1)} g_1 \right)$$

Therefore for each u,

$$\text{in}(m_v f_u^{(v)} g_u) \leq \text{in}(m_v S(g_1, g_v)) < \text{in}\left(\frac{m}{\text{in}(g_1)} g_1 \right) = m$$

Finally, we have the expression

$$\begin{aligned} f =& \sum_v \text{in}(f_v) g_v + \sum_v (f_v - \text{in}(f_v)) g_v + \sum_{u > t'} f_u g_u \\ =& \sum_{v=1}^{t'} \sum_{u=1}^{t} a_v m_v f_u^{(v)} g_u + \sum_{v=1}^{t'} (f_v - \text{in}(f_v)) g_v + \sum_{u > t'} f_u g_u \end{aligned}$$

where each term of the summand has initial term strictly less than m, giving us the desired contradiction. □

Theorem 4.5 (Buchberger's Algorithm, [Eis 15.9])**.** *Let $F = \{f_1, ..., f_t\} \subseteq I$ be a basis for I. A Gróbner basis can be constructed using the following algorithm:*

If all the $\overline{S(f_i, f_j)}^F = 0$, then they form a Gröbner basis. If it is non-zero for some i, j, add it to the set F and repeat.

Proof. The correctness follows from Buchberger's criterion. To show that it terminates, recall from division algorithm that if $\text{in}(\overline{S(f_i, f_j)}^F) \neq 0$, then $\langle \text{in}(f_1), ..., \text{in}(f_t), \text{in}(\overline{S(f_i, f_j)}^F) \rangle$ is a strictly larger ideal than $\langle \text{in}(f_1), ..., \text{in}(f_t) \rangle$, and will eventually be equal to $\text{in}(I)$ since it is Noetherian.

□

5. Applications of Gröbner Basis

Given a system of polynomials in $k[x, y]$, we could try to take linear combinations of the polynomials to get a polynomial in $k[x]$, then solve the one variable polynomial and substitute the solution for x into the other polynomials to solve for y. More generally, given ideal $I \subseteq S = k[x_1, ..., x_n, y_1, ..., y_m]$, we would like to find $J = I \cap R = I \cap k[x_1, ..., x_n]$. This can be

computed easily using Gröbner basis, with respect to an **elimination order (with respect to the variables $y_1, ..., y_m$)**, that is an order on S satisfying

$$\text{if } f \in S \text{ and } \text{in}(f) \in R, \text{ then } f \in R$$

The lexicographic ordering (considering $y > x$) is one such order.

Proposition 5.1 (Eis, Prop 15.29)**.** *Let $>$ be an elimination order on S. Let I be an ideal of S. If $g_1, ..., g_t$ is a Gröbner basis for I and $g_1, ..., g_u$ are those that do not involve the variables y_i, then $g_1, ..., g_u$ is a Gröbner basis in R for $J = I \cap R$.*

Proof. Let $x^\alpha \in \text{in}(J) \subseteq \text{in}(I)$ be a monomial. Then x^α is a multiple of $\text{in}_>(g_i)$ for some i since $g_1, ..., g_t$ is a Gröbner basis. Since $x^\alpha \in R$ and $>$ is an elimination order, we must have $i \leq u$. Thus

$$\text{in}(J) \subseteq \langle \text{in}(g_1), ..., \text{in}(g_u) \rangle$$

The other inclusion follows from that $g_1, ..., g_u \in I \cap R = J$. Therefore $\text{in}(J) = \langle \text{in}(g_1), ..., \text{in}(g_u) \rangle$. Thus by definition 3.4 $g_1, ..., g_u$ is a Gröbner basis for J. □

Now given an ideal $I \subseteq k[x_1, ..., x_n]$, we can find $V(I)$ by first finding $I_1 = I \cap k[x_1]$, solve and substitute into $I_2 = I \cap k[x_1, x_2]$, and so on. Note that this way each point in $V(I_i)$ can be extended to a point in $V(I_j)$ for $j > i$. If $V(I)$ is finite, then I_i must contain some equation that involves x_i. By the above proposition, if we compute a Gröbner basis and order them increasingly with respect to the lexicographic ordering, then the first few equations will only contain x_1, followed with equations involving only x_1, x_2, and so on. Computing $V(I)$ is then simplified to computing roots for a series of one-variable polynomials.

Example 5.2. Find complex roots to the equations

$$12x^2 - 3y = 0$$

$$6xy - 10y^2 - 3y = 0$$

Solution: Let $I = \langle 12x^2 - 3y, 6xy - 10y^2 - 3y \rangle$. Using lexicographic ordering with $x > y$, the S-polynomial of the basis is

$$S(12x^2 - 3y, 6xy - 10y^2 - 3y) = \frac{1}{12}y(12x^2 - 3y) - \frac{1}{6}x(6xy - 10y^2 - 3y) = \frac{5xy^2}{3} + \frac{xy}{2} - \frac{y^2}{4}$$

Its remainder after division algorithm is

$$\frac{25y^3}{9} + \frac{17y^2}{12} + \frac{y}{4}$$

Thus we write

$$I = \langle 12x^2 - 3y, 6xy - 10y^2 - 3y, \frac{25y^3}{9} + \frac{17y^2}{12} + \frac{y}{4} \rangle$$

Solving the cubic equation $\frac{25y^3}{9} + \frac{17y^2}{12} + \frac{y}{4} = 0$ gives $y = 0$ or $y = -\frac{51}{200} \pm \frac{3i\sqrt{111}}{200}$. Substitute back to the original equations yields

$$(x, y) = (0, 0) \text{ or } (\frac{3 \pm i\sqrt{111}}{40}, -\frac{51}{200} \pm \frac{3i\sqrt{111}}{200})$$

◇

Example 5.3. Find the defining equation over $\mathbb{Q}$ for the algebraic number z satisfying

$$z^5 + \sqrt{2}z - a^2z + a = 0$$

where $a^3 + a - 1 = 0$.

Solution: Extend the basis $F = \{x^2 - 2, a^3 + a - 1, z^5 + xz - a^2z + a\}$ to a Gröbner basis using lexicographic ordering with $z < a < x$. Then we shall obtain equations that only involve z. From those equations we can then find the minimal polynomial of z over $\mathbb{Q}$.

$\diamond$

In general, Proposition 5.1 can be used to compute the closure of the image of an algebraic variety under a projection:

Lemma 5.4. *Let $I \subseteq k[x_1, ..., x_n]$ be an ideal. Let $\pi : k^n \to k^m$ be the projection onto the first m coordinates. If k is algebraically closed, then $V(I \cap k[x_1, ..., x_m])$ is the smallest variety containing $\pi(V(I))$.*

Proof. Let $X = V(I), J = I \cap k[x_1, ..., x_m]$. Observe that $\pi(X) \subseteq V(J)$ since polynomials in J are those in I that do not involve $x_{m+1}, ..., x_n$. So $V(I(\pi(X))) \subseteq V(J)$.

Conversely, suppose $f \in I(\pi(X))$, then $f(a_1, ..., a_m) = 0$ for all $(a_1, ..., a_m) \in \pi(X)$. View f as an element of $k[x_1, ..., x_n]$ shows that $f \in I(X) = \sqrt{I}$ by Nullstellensatz. So $f^N \in I$ for some N, and since $f \in J$, $f^N \in J$ for some N and $f \in \sqrt{J}$. Hence $I(\pi(X)) \subseteq \sqrt{J}$ and

$$V(I(\pi(X))) \supseteq V(\sqrt{J}) = V(J)$$

$\square$

As a result of Macaulay's Theorem, one can also compute the number of roots when the variety has dimension 0 (finite) using a Gröbner basis. We state the following Theorem as a fact:

Theorem 5.5 (Affine Bézout's inequality)**.** *Let I be an ideal of R. If $V(I)$ is finite, then $|V(I)| \leq \dim_k(R/I)$. When k is algebraically closed, $|V(I)| = \dim_k(R/I)$ if and only if I is radical.*

References

[Cox] David Cox, John Little, Donal O'Shea, *Ideals, Varieties, and Algorithms An Introduction to Computational Algebraic Geometry and Commutative Algebra*, Undergraduate Texts in Mathematics. Springer, Cham, fourth edition, 2015.

[Eis] David Eisenbud, *Commutative Algebra with a View Toward Algebraic Geometry*, Springer-Verlag, New York-Heidelberg, 1989.