MAG1 Algebraic Geometry Notes

Brian Chan

May 22, 2024

Contents

0.1 Purpose

The purpose of this document is to record my own detailed notes for the introductory course MAG1. MAG1 consists of eight 1.5 hour lectures, presented by Dan Murfet and Ken Chan on Metauni. The course serves as an introduction to algebraic geometry. The main reference used is [CLO15].

Chapter 1

Introduction

1.1 Polynomials

This section is based on [CLO15, Chapter 1 §1]. Throughout this document, we will use k to denote a field — one of $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$. We will mostly work with $\mathbb{C}$.

Definition 1.1.1. Let $x_1, x_2, \ldots, x_n$ be variables. A **monomial** in the variables $x_1, x_2, \ldots, x_n$ is a product of the form

$$x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}$$

where $\alpha_i \in \mathbb{Z}_{\geq 0}$ for $i \in \{1, 2, \ldots, n\}$.

If $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in (\mathbb{Z}_{\geq 0})^n$ then the monomial x^α is defined by

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}.$$

The **total degree** of the monomial x^α is the sum $|\alpha| = \alpha_1 + \cdots + \alpha_n$.

Monomials are the basic building blocks of polynomials, which motivates the definition of the ring of polynomials.

Definition 1.1.2. Let k be a field and $x_1, x_2, \ldots, x_n$ be variables. The **ring of polynomials** in $x_1, x_2, \ldots, x_n$ with coefficients in k is defined by the set

$$k[x_1, x_2, \ldots, x_n] = \{ \sum_{\alpha \in (\mathbb{Z}_{\geq 0})^n} a_\alpha x^\alpha \mid a_\alpha \in k \}$$

where the sum above is a *finite* sum over $(\mathbb{Z}_{\geq 0})^n$. The two operations which make the set $k[x_1, x_2, \ldots, x_n]$ a commutative ring is the operations of addition and multiplication of polynomials.

From the above definition, we infer that polynomials are finite linear combinations of monomials in $x_1, x_2, \ldots, x_n$. Indeed, monomials are the building blocks of polynomials that most people are familiar with. It is worth mentioning that $\mathbb{C}[x_1, \ldots, x_n]$ in particular has another basis — the basis of *Macdonald polynomials*, which are again indexed by $(\mathbb{Z}_{\geq 0})^n$. The theory of Macdonald polynomials is in a sense, parallel to the theory of monomials and the representation theory of S_n and $GL_n(\mathbb{C})$. However, this digression takes us too far afield for these notes. See [Ram22] for the definition of Macdonald polynomials.

Another remark we will make here is that $k[x_1, x_2, \ldots, x_n]$ is actually a Noetherian ring. This means that any ascending chain of ideals in $k[x_1, \ldots, x_n]$ must terminate/stabilise or equivalently, every ideal in $k[x_1, \ldots, x_n]$ is finitely generated. This is a consequence of the *Hilbert basis theorem*, which we will prove later in these notes.

Next, we will define some more terminology associated with polynomials.

Definition 1.1.3. Let $f = \sum_\alpha a_\alpha x^\alpha$ be a polynomial in $k[x_1, \ldots, x_n]$. We call a_α the **coefficient of the monomial** x^α.

If $a_\alpha \neq 0$ then we call $a_\alpha x^\alpha$ a **term** of f.

The **total degree** of $f \neq 0$, denoted by $\deg(f)$, is the quantity

$$\deg(f) = \max\{|\alpha| \mid \alpha \in (\mathbb{Z}_{\geq 0})^n, a_\alpha \neq 0\}.$$

In other words, the total degree of $f \neq 0$ is the largest degree of any monomial which appears in the monomial expansion of f.

One of the fundamental settings for algebraic geometry is the notion of an affine space.

Definition 1.1.4. Let k be a field and $n \in \mathbb{Z}_{>0}$. The n-dimensional **affine space** over k is the set

$$k^n = \{(a_1, \ldots, a_n) \mid a_1, \ldots, a_n \in k\}.$$

The sets $k^1 = k$ and k^2 are usually referred to as the **affine line** and the **affine plane** respectively.

If $f = \sum_\alpha a_\alpha x^\alpha \in k[x_1, \ldots, x_n]$ and $\beta_1, \ldots, \beta_n \in k$ then the following map defines a ring homomorphism:

$$\begin{array}{rccc} ev_{\beta_1,\ldots,\beta_n}: & k[x_1,\ldots,x_n] & \to & k \\ & f=\sum_\alpha a_\alpha x^\alpha & \mapsto & \sum_\alpha a_\alpha \beta^\alpha \end{array}$$

where $\beta^\alpha = \beta_1^{\alpha_1}\ldots\beta_n^{\alpha_n}$. Thus, every polynomial $f \in k[x_1,\ldots,x_n]$ can be viewed a function from k^n to k by defining the map

$$\begin{array}{rccc} \varphi_f: & k^n & \to & k \\ & \beta=(\beta_1,\ldots,\beta_n) & \mapsto & \sum_\alpha a_\alpha \beta^\alpha. \end{array}$$

This is a consequence of the *universal property* satisfied by $k[x_1,\ldots,x_n]$, which we state below as a theorem.

Theorem 1.1.1. *Let k be a field and A be a commutative k-algebra. Let $\beta_1,\ldots,\beta_n \in A$. Then, there exists a unique algebra morphism*

$$ev_{\beta_1,\ldots,\beta_n}: k[x_1,\ldots,x_n] \to A$$

such that

$$ev_{\beta_1,\ldots,\beta_n}(\sum_\alpha a_\alpha x^\alpha) = \sum_\alpha a_\alpha \beta^\alpha$$

where $\beta^\alpha = \beta_1^{\alpha_1}\ldots\beta_n^{\alpha_n}$.

Proof. Assume that k is a field and A is a commutative k-algebra. Assume that $\beta_1,\ldots,\beta_n \in A$. It is straightforward to verify that $ev_{\beta_1,\ldots,\beta_n}$ is an algebra morphism.

To see that $ev_{\beta_1,\ldots,\beta_n}$ is unique, assume that we have another algebra morphism $\Phi: k[x_1,\ldots,x_n] \to A$ such that if $f=\sum_\alpha a_\alpha x^\alpha \in k[x_1,\ldots,x_n]$ then

$$\Phi(f) = \Phi(\sum_\alpha a_\alpha x^\alpha) = \sum_\alpha a_\alpha \beta^\alpha$$

and consequently, $\Phi = ev_{\beta_1,\ldots,\beta_n}$. So, the morphism $ev_{\beta_1,\ldots,\beta_n}$ is unique. □

As stated concisely in [CLO15, Page 3], the ability to view a polynomial $f: k[x_1,\ldots,x_n] \to k$ as a *polynomial function* $\varphi_f: k^n \to k$ is the reason why we can link algebra and geometry together in the study of algebraic geometry.

First, we would like to know when a polynomial is the zero polynomial. It turns out that the answer is obvious when k is an infinite field.

Theorem 1.1.2. *Let k be an infinite field and $f \in k[x_1, \ldots, x_n]$. Then, $f = 0$ as polynomials in $k[x_1, \ldots, x_n]$ if and only if the polynomial function $f : k^n \to k$ is the zero function (take note of the abuse of notation here!).*

Proof. Assume that k is an infinite field and $f \in k[x_1, \ldots, x_n]$. First, assume that f is the zero polynomial. Then, $f(\alpha_1, \ldots, \alpha_n) = 0$ for $(\alpha_1, \ldots, \alpha_n) \in k^n$ and consequently, the function f is the zero function.

We will prove the converse statement by induction on n. For the base case, assume that $n = 1$ and that $f(a) = 0$ for $a \in k$. Then, f has infinitely many roots because k is an infinite field. By the fundamental theorem of algebra, f must be the zero polynomial. This proves the base case.

Now assume that if $g(a_1, \ldots, a_l) = 0$ for $(a_1, \ldots, a_l) \in k^l$ then g is the zero polynomial. Assume that $f \in k[x_1, \ldots, x_{l+1}]$ and that $f(a_1, \ldots, a_{l+1}) = 0$ as a function from k^{l+1} to k. The key idea is that we can rewrite the polynomial f as a finite sum of powers of x_{l+1} so that

$$f = \sum_{i=0}^{N} g_i(x_1, \ldots, x_l) x_{l+1}^i$$

for some $N \in \mathbb{Z}_{>0}$ and $g_i \in k[x_1, \ldots, x_l]$.

Now let $(\beta_1, \ldots, \beta_l) \in k^l$. We obtain the polynomial $f(\beta_1, \ldots, \beta_l, x_{l+1}) \in k[x_{l+1}]$. From the base case and the assumption on f, $f(\beta_1, \ldots, \beta_l, x_{l+1})$ vanishes as a function from k to k and thus, $f(\beta_1, \ldots, \beta_l, x_{l+1})$ must be the zero polynomial in $k[x_n]$.

Hence, the coefficients $g_i \in k[x_1, \ldots, x_l]$ must satisfy $g_i(\beta_1, \beta_2, \ldots, \beta_n) = 0$ for $i \in \{1, 2, \ldots, N\}$. By the inductive hypothesis, g_i must be the zero polynomial in $k[x_1, \ldots, x_l]$. So, f must be the zero polynomial in $k[x_1, \ldots, x_{l+1}]$, which completes the proof. □

Note that in the above proof, we were able to apply the fundamental theorem of algebra because k was an infinite field. The following example demonstrates that Theorem 1.1.2 does not hold when k is a finite field.

Example 1.1.5. Let $p \in \mathbb{Z}_{>0}$ be a prime number and $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ be a finite field with p elements. Consider the polynomial $f(x) = x^p - x \in \mathbb{F}_p[x]$. This polynomial is not the zero polynomial. However, $f(0) = 0$ and if $a \in \mathbb{F}_p - \{0\}$ then $a^p \equiv a \bmod p$ by Fermat's little theorem and consequently, $f(a) = 0$. So, f is the zero function on $\mathbb{F}_p$, but is not the zero polynomial in $\mathbb{F}_p[x]$.

A consequence of Theorem 1.1.2 is that in an infinite field k, two polynomials f and g are equal in $k[x_1, \ldots, x_n]$ if and only if they define the same function from k^n to k.

Theorem 1.1.3. *Let k be an infinite field and let $f, g \in k[x_1, \ldots, x_n]$. Then, $f = g$ in $k[x_1, \ldots, x_n]$ if and only if f and g are the same function from k^n to k.*

Proof. Assume that k is an infinite field and $f, g \in k[x_1, \ldots, x_n]$. Then, $f = g$ in $k[x_1, \ldots, x_n]$ if and only if $f - g$ is the zero polynomial. By Theorem 1.1.2, this holds if and only if $f - g$ is the zero function from k^n to k. Consequently, $f = g$ in $k[x_1, \ldots, x_n]$ if and only if $f = g$ as functions from k^n to k. □

Although we have already used it, the fundamental theorem of algebra warrants a mention.

Theorem 1.1.4 (Fundamental theorem of algebra)**.** *Let $f \in \mathbb{C}[x]$ be a non-constant polynomial. Then, f must have a root in $\mathbb{C}$.*

The Hilbert Nullstellensatz is a powerful generalisation of the fundamental theorem of algebra, which we will prove later.

1.2 Affine varieties and the Zariski topology

We will jump into the definition of an affine variety, which is a central object of study in algebraic topology.

Definition 1.2.1. Let k be a field and $S \subseteq k[x_1, \ldots, x_n]$. The **affine variety** defined by the set S is the set

$$V(S) = \{(a_1, \ldots, a_n) \in k^n \mid f_i(a_1, \ldots, a_n) = 0 \text{ for } f \in S\}.$$

As an example of an affine variety, consider the polynomial ring $\mathbb{R}[x, y]$. The affine variety $V(x^2 + y^2 - 1) \subseteq \mathbb{R}^2$ is depicted graphically as the unit circle in $\mathbb{R}^2$.

The first glimpse of the topological structure defined by the affine varieties is seen by the fact that affine varieties are closed under intersections and unions.

Theorem 1.2.1. *Let k be a field and $V, W \subseteq k^n$ be affine varieties. Then, $V \cup W$ and $V \cap W$ are also affine varieties.*

Proof. Assume that k is a field and $V, W \subseteq k^n$. Then, there exists subsets $S, T \subseteq k[x_1, \ldots, x_n]$ such that $V = V(S)$ and $W = V(T)$.

To show: (a) $V \cap W = V(S \cup T)$.

(b) $V \cup W = V(\{fg \mid f \in S, g \in T\})$.

(a) Assume that $(a_1, \ldots, a_n) \in V \cap W \subseteq k^n$. If $f \in S$ and $g \in T$ then $f(a_1, \ldots, a_n) = g(a_1, \ldots, a_n) = 0$ and consequently, $(a_1, \ldots, a_n) \in V(S \cup T)$.

Now assume that $(a_1, \ldots, a_n) \in V(S \cup T)$. If $f \in S \cup T$ then $f(a_1, \ldots, a_n) = 0$. So, $f \in V(S) = V$ and $f \in V(T) = W$. Subsequently, $f \in V \cap W$ and $V \cap W = V(S \cup T)$.

(b) Assume that $(b_1, \ldots, b_n) \in V \cup W$. Then, either $f(b_1, \ldots, b_n) = 0$ or $g(b_1, \ldots, b_n) = 0$ for some $f \in S$ or $g \in T$. If $f(b_1, \ldots, b_n) = 0$ then $fg' = 0$ for any $g' \in T$. Similarly, if $g(b_1, \ldots, b_n) = 0$ then $f'g = 0$ for any $f' \in S$. In either case, we find that $(b_1, \ldots, b_n) \in V(\{fg \mid f \in S, g \in T\})$.

Now assume that $(b_1, \ldots, b_n) \in V(\{fg \mid f \in S, g \in T\})$. There are two cases to consider here.

Case 1: $f(b_1, \ldots, b_n) = 0$ for some $f \in S$.

If $f(b_1, \ldots, b_n) = 0$ for some $f \in S$ then $(b_1, \ldots, b_n) \in V(S)$ by definition of an affine variety.

Case 2: If $f \in S$ then $f(b_1, \ldots, b_n) \neq 0$.

Suppose that $f(b_1, \ldots, b_n) \neq 0$ for any $f \in S$. Since $(b_1, \ldots, b_n) \in V(\{fg \mid f \in S, g \in T\})$,

$$f(b_1, \ldots, b_n)g(b_1, \ldots, b_n) = 0$$

for any $g \in T$ and since k is an integral domain, $g(b_1, \ldots, b_n) = 0$ for any $g \in T$ and $(b_1, \ldots, b_n) \in W = V(T)$.

By combining these two cases, we find that $(b_1, \ldots, b_n) \in V \cup W$ and consequently, $V \cup W = V(\{fg \mid f \in S, g \in T\})$. □

The affine varieties of k^n are the closed sets of the *Zariski topology* on k^n. We will formally define the Zariski topology below and prove that it is a

topology on k^n.

Definition 1.2.2. Let k be a field. The **Zariski topology** on k^n, denoted by τ_{Zar}, is defined by setting the closed subsets of k^n to be affine varieties of the form $V(S)$, where S is a subset of $k[x_1, \ldots, x_n]$.

Theorem 1.2.2. *Let $n \in \mathbb{Z}_{>0}$ and k be a field. Let τ_{Zar} denote the Zariski topology on k^n. Then, the pair (k^n, τ_{Zar}) is a topological space.*

Proof. Assume that the τ_{Zar} is the topology on k^n defined as above.

To show: (a) $\emptyset \in \tau_{Zar}$ and $k^n \in \tau_{Zar}$.

(b) $\bigcap_{i \in I} V(S_i) \in \tau_{Zar}$.

(c) $\bigcup_{i=1}^{n} V(S_i) \in \tau_{Zar}$.

(a) In order to see that $\emptyset \in \tau_{Zar}$, let $S = \{1\}$ where 1 is the constant polynomial. Then, from the definition of $V(S)$, $V(S) = \emptyset$. So, $\emptyset \in \tau_{Zar}$. Now, let $S = \{0\}$ where 0 is the zero polynomial. Consequently,

$$V(S) = \{(\lambda_1, \ldots, \lambda_n) \in k^n \mid 0(\lambda_1, \ldots, \lambda_n) = 0\} = k^n.$$

So, $k^n \in \tau_{Zar}$.

(b) Assume that $S_i \subseteq k[x_1, \ldots, x_n]$ for $i \in I$.

To show: (ba) $\bigcap_{i \in I} V(S_i) = V(\bigcup_{i \in I} S_i)$.

(ba) Suppose that $(\lambda_1, \ldots, \lambda_n) \in \bigcap_{i \in I} V(S_i)$. If $f \in S_i$ then $f(\lambda_1, \ldots, \lambda_n) = 0$. This is true if and only if $f(\lambda_1, \ldots, \lambda_n) = 0$ whenever $f \in \bigcup_{i \in I} S_i$. So, $(\lambda_1, \ldots, \lambda_n) \in V(\bigcup_{i \in I} S_i)$ and $\bigcap_{i \in I} V(S_i) = V(\bigcup_{i \in I} S_i)$.

(b) From this, it follows that $\bigcap_{i \in I} V(S_i) \in \tau_{Zar}$.

(c) Assume that $S_i \subseteq k[x_1, \ldots, x_n]$ for $i \in \{1, \ldots, n\}$. Define the set S by

$$S = \{\prod_{i=1}^{n} p_i \mid p_i \in S_i\}.$$

To show: (ca) $\bigcup_{i=1}^{n} V(S_i) = V(S)$.

(ca) To show: (caa) $\bigcup_{i=1}^{n} V(S_i) \subseteq V(S)$.

(cab) $V(S) \subseteq \bigcup_{i=1}^{n} V(S_i)$.

(caa) Assume that $(\lambda_1, \ldots, \lambda_n) \in \bigcup_{i=1}^{n} V(S_i)$. Then, there exists a $j \in \{1, \ldots, n\}$ such that $(\lambda_1, \ldots, \lambda_n) \in V(S_j)$. If $f \in S_j$ then $f(\lambda_1, \ldots, \lambda_n) = 0$. Now assume that $g \in S$. Then, $g = q_1 \ldots q_n$ where $q_i \in S_i$. In particular, since $q_j \in S_j$, $q_j(\lambda_1, \ldots, \lambda_n) = 0$. Therefore, $g(\lambda_1, \ldots, \lambda_n) = 0$ as well. Hence, $(\lambda_1, \ldots, \lambda_n) \in V(S)$. So, $\bigcup_{i=1}^{n} V(S_i) \subseteq V(S)$.

(cab) Assume that $(\lambda_1, \ldots, \lambda_n) \in V(S)$. Suppose for the sake of contradiction that $(\lambda_1, \ldots, \lambda_n) \notin \bigcup_{i=1}^{n} V(S_i)$. If $i \in \{1, 2, \ldots, n\}$ and $f \in S_i$ then $f(\lambda_1, \ldots, \lambda_n) \neq 0$. Now suppose that $g \in S$. Once again, write $g = p_1 \ldots p_n$ where $p_i \in S_i$. Consequently, $g(\lambda_1, \ldots, \lambda_n) \neq 0$ because $p_i(\lambda_1, \ldots, \lambda_n) \neq 0$ for all $p_i \in S_i$. This contradicts the assumption that $(\lambda_1, \ldots, \lambda_n) \in V(S)$. Therefore, $(\lambda_1, \ldots, \lambda_n) \in \bigcup_{i=1}^{n} V(S_i)$ and as a result, $V(S) \subseteq \bigcup_{i=1}^{n} V(S_i)$.

(c) This means that $V(S) = \bigcup_{i=1}^{n} V(S_i)$. Subsequently, we deduce that $\bigcup_{i=1}^{n} V(S_i) \in \tau_{Zar}$. Hence, (k^n, τ_{Zar}) is a topological space. □

There is one key difference separating the Zariski topology τ_{Zar} on k^n from the more standard Euclidean topology on k^n.

Theorem 1.2.3. *Let $n \in \mathbb{Z}_{>0}$ and k be an infinite field. Then, the Zariski topology τ_{Zar} on k^n is not Hausdorff.*

Proof. Assume that $n \in \mathbb{Z}_{>0}$. Assume that k is an infinite field. Let $S \subseteq k[x_1, \ldots, x_n]$. The open sets in the Zariski topology τ_{Zar} are of the form

$$k^n \backslash V(S) = \{(a_1, \ldots, a_n) \in k^n \mid \text{There exists } f \in S \text{ such that } f(a_1, \ldots, a_n) \neq 0\}.$$

Notably, the set $V(S)$ is finite so that the complement $k^n \backslash V(S)$ is a cofinite set.

We claim that if $S, T \subseteq k[x_1, \ldots, x_n]$ and $V(S)^c = k^n \backslash V(S)$ then $V(S)^c \cap V(T)^c \neq \emptyset$. Assume that S and T are subsets of $k[x_1, \ldots, x_n]$. Since $V(S)$ and $V(T)$ are finite sets, $V(S) \cup V(T)$ must also be a finite set. Consequently, $V(S) \cup V(T) \neq k^n$ and by taking complements, $V(S)^c \cap V(T)^c \neq \emptyset$. Therefore, the Zariski topology on k^n is not Hausdorff. □

It is useful to observe that τ_{Zar} has a basis of open sets given by

$$V(f) = \{(a_1, \ldots, a_n) \in k^n \mid f(a_1, \ldots, a_n) \neq 0\}$$

where $f \in k[x_1, \ldots, x_n]$.

In line with the emphasis of [CLO15] on examples, we will end the introduction to affine varieties by giving an involved example of an affine variety.

Example 1.2.3. Consider the graph in $\mathbb{R}^2$ defined by the equation $r = \sin 2\theta$. This is the four leaved rose.

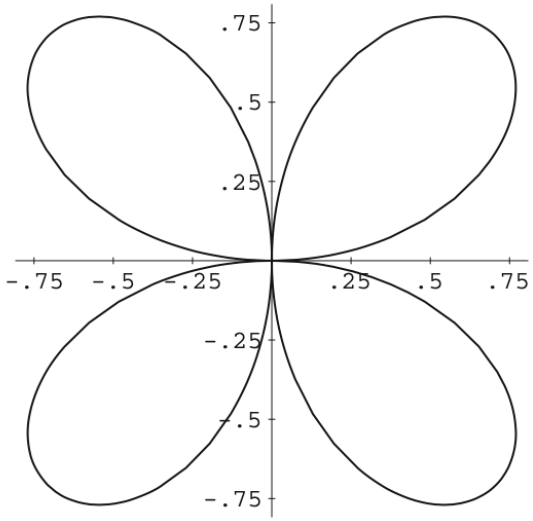


Figure 1.1: The four leaved rose. Figure is from [CLO15, Page 12]

Recall that polar coordinates in $\mathbb{R}^2$ are defined by the equations $x = r\cos\theta$ and $y = r\sin\theta$. We claim that the four leaved rose is the affine variety $V((x^2+y^2)^3 - 4x^2y^2) \subseteq \mathbb{R}^2$.

First, assume that the point $(x, y) \in \mathbb{R}^2$ lies on the four leaved rose. The idea is to rewrite the equation $r = \sin 2\theta$ as $r = 2\sin\theta\cos\theta$. Squaring both sides, we find that $r^2 = 4\sin^2\theta\cos^2\theta$. Now multiply both sides by r^4 to find that

$$(x^2+y^2)^3 = r^6 = 4r^4\sin^2\theta\cos^2\theta = 4x^2y^2.$$

So, $(x^2+y^2)^3 - 4x^2y^2 = 0$ and consequently, $(x,y) \in V((x^2+y^2)^3 - 4x^2y^2)$.

Now assume that $(p,q) \in V((x^2+y^2)^3 - 4x^2y^2)$ so that $(p^2+q^2)^3 - 4p^2q^2 = 0$ and $(p^2+q^2)^3 = 4p^2q^2$. Squaring both sides, we obtain in polar coordinates

$$r^6 = 4r^4 \sin^2\theta \cos^2\theta$$

where $r^2 = p^2+q^2$, $p = r\cos\theta$ and $q = r\sin\theta$. Hence, we have the equation $r^2 = \sin^2 2\theta$ and

$$r = |\sin 2\theta|.$$

Now, we have two cases to consider.

Case 1: $r = \sin 2\theta$.

In this case, we are done since this is the equation of the four leaved rose. We conclude that (p,q) is a point on the four leaved rose.

Case 2: $r = -\sin 2\theta$.

If we have the equation $r = -\sin 2\theta$ then $r = -\sin 2\theta = \sin(-2\theta)$. Hence, (p,q) is a point on the four leaved rose.

By combining both cases, we conclude that the four leaved rose is the affine variety $V((x^2+y^2)^3 - 4x^2y^2)$ in $\mathbb{R}^2$ as required.

1.3 Parametrisation and stereographic projection

Suppose that $f_1, \ldots, f_s \in k[x_1, \ldots, x_n]$ are polynomials. Consider the affine variety

$$V(f_1, \ldots, f_s) = \{(a_1, \ldots, a_n) \in k^n \mid f_i(a_1, \ldots, a_n) = 0 \text{ for } i \in \{1, 2, \ldots, s\}\}.$$

How do we describe the points of $V(f_1, \ldots, f_s)$? This amounts to writing down the solutions of the system of polynomial equations $f_1 = \cdots = f_s = 0$. If there are finitely many solutions then we can simply list them all. However, when there are infinitely many solutions, the main technique we use is **parametrisation**.

Example 1.3.1. Here is a basic example from linear algebra. We will work in the field $\mathbb{R}$ and describe the points in the affine variety $V(x+y+z-1, x+2y-z-3)$.

This amounts to solving the equations $x+y+z=1$ and $x+2y-z=3$ simultaneously. By row reduction, we obtain the equivalent relations $x+3z=-1$ and $y-2z=2$. Now let $z=t$, where $t \in \mathbb{R}$ is some parameter. Then, the points in the affine variety $V(x+y+z-1, x+2y-z-3)$ are given by

$$x=-1-3t, \qquad y=2+2t, \qquad \text{and} \qquad z=t$$

for $t \in \mathbb{R}$.

In line with [CLO15, §3], we will define two special types of parametrisations below.

Definition 1.3.2. Let k be a field. The field of **rational functions** in $x_1, \ldots, x_n$, denoted by $k(x_1 \ldots, x_n)$, is the field of fractions of $k[x_1, \ldots, x_n]$. That is,

$$k(x_1 \ldots, x_n) = \{\frac{f}{g} \mid f, g \in k[x_1, \ldots, x_n], g \neq 0\}.$$

Definition 1.3.3. Let $f_1, \ldots, f_s \in k[x_1, \ldots, x_n]$ and $V = V(f_1, \ldots, f_s) \subseteq k^n$ be an affine variety. A **rational parametric representation** of V is a set of rational functions $r_1, \ldots, r_n \in k(x_1, \ldots, x_n)$ such that the points $(y_1, \ldots, y_n) \in k^n$ given by the equations

$$y_i = r_i(a_1, \ldots, a_n)$$

for $i \in \{1, 2, \ldots, n\}$ and $(a_1, \ldots, a_n) \in k^n$ all lie in V.

A **polynomial parametric representation** of V is a rational parametric representation such that $r_1, \ldots, r_n \in k[x_1, \ldots, x_n]$ are polynomials.

The original defining equations $f_1 = \cdots = f_s = 0$ of V is called an **implicit representation** of V.

As mentioned in [CLO15, §3], we often work with both parametrisations and implicit representations, using the one which is more useful for a given task. For instance, if we wanted to determine if a point is in an affine variety, it is generally much easier to use the implicit representation rather than a parametrisation. On the other hand, parametrisations are easier to

deal with when plotting an affine variety.

Since it is useful to have both a parametrisation and an implicit representation of an affine variety, there are two main questions one can ask about them

1. (Parametrisation) Does every affine variety have a rational parametric representation?

2. (Implicitisation) Given a parametrisation of an affine variety, can we find an implicit representation?

The answer to the first question is no in general. In fact, it is difficult to even tell when an affine variety has a rational parametric representation. On the other hand, the answer to the second question is yes. We will see this later when we study elimination theory.

Example 1.3.4. We will work in the field $\mathbb{R}$. Define $V = V(x^2 + y^2 - 1)$. Then, the equation $x^2 + y^2 = 1$ gives an implicit representation of the affine variety V.

The purpose of this example is to derive the well-known *stereographic projection* of the circle S^1, which is a rational parametric representation of V.

The idea is to draw a line from the point $(0, 1)$ to $(t, 0)$, where $t \in \mathbb{R}$. If the line is not horizontal then the line intersects the circle S^1 at a unique point, which we will denote by $(x, y) \in S^1$. Note that $(x, y) \neq (0, 1)$.

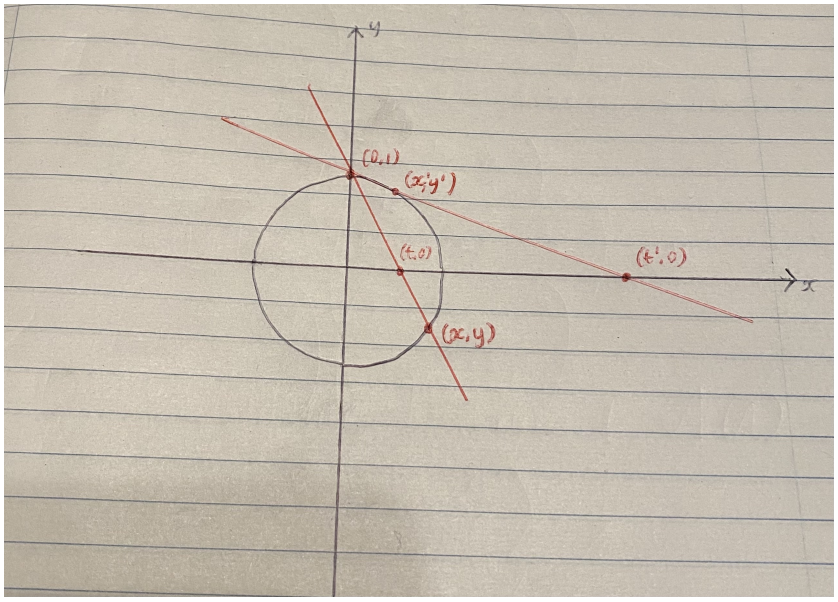


Figure 1.2: Stereographic projection as applied to the unit circle S^1

Our goal is to write x and y in terms of t. The line connecting the points $(0,1)$ and $(t,0)$ is for $t \neq 0$

$$y = -\frac{1}{t}x + 1.$$

Solving for x, we obtain $x = t(1-y)$. We can substitute this into the implicit equation $x^2 + y^2 = 1$ in order to obtain the following quadratic equation:

$$(1+t^2)y^2 - 2t^2 y + (t^2 - 1) = 0$$

Now we can solve for y via the quadratic formula

$$y = \frac{2t^2 \pm \sqrt{4t^4 - 4(t^4-1)}}{2(1+t^2)} = \frac{t^2 \pm 1}{1+t^2}.$$

Notice that one of the roots is $y = 0$. This gives the point $(x,y) = (0,1)$ again, which is not the point (x,y) we are looking for. So, we set $y = \frac{-1+t^2}{1+t^2}$ and upon substitution into $x = t(1-y)$, we find that

$$(x,y) = (\frac{2t}{1+t^2}, \frac{-1+t^2}{1+t^2}) \tag{1.1}$$

for $t \neq 0$. Luckily, setting $t = 0$ in the above parametric representation yields the point $(x,y) = (0,-1)$. This corresponds to the intersection of the

circle with a vertical line passing through $(0,1)$ and the origin.

Equation (1.1) yields a parametrisation of the affine variety V, with the exception of the north pole $(0,1)$. Clearly, we can repeat the above argument, but instead of drawing a line from $(0,1)$, we can start from the south pole $(0,-1)$ instead. This time, we obtain the parametrisation

$$(x,y) = (\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2}) \tag{1.2}$$

for $t \in \mathbb{R}$.

The reason why stereographic projection is important is because it generalises to the n-sphere $S^n \subseteq \mathbb{R}^{n+1}$. The n-sphere S^n is a smooth manifold of dimension n. It has two charts, which are given by stereographic projection from the north and south poles. Let $N, S \in S^n$ denote the north and south poles of S^n. Then, the charts of S^n are given explicitly by

$$\begin{array}{rccc} \varphi_N : & S^n \backslash \{N\} & \to & \mathbb{R}^n \\ & (x_1, \ldots, x_{n+1}) & \mapsto & \frac{1}{1-x_{n+1}}(x_1, \ldots, x_n) \end{array}$$

which has inverse

$$\begin{array}{rccc} \varphi_N^{-1} : & \mathbb{R}^n & \to & S^n \backslash \{N\} \\ & (x_1, \ldots, x_n) & \mapsto & (\frac{2x_1}{1+x_1^2+\cdots+x_n^2}, \ldots, \frac{2x_n}{1+x_1^2+\cdots+x_n^2}, \frac{-1+x_1^2+\cdots+x_n^2}{1+x_1^2+\cdots+x_n^2}) \end{array}$$

and

$$\begin{array}{rccc} \varphi_S : & S^n \backslash \{S\} & \to & \mathbb{R}^n \\ & (x_1, \ldots, x_{n+1}) & \mapsto & \frac{1}{1+x_{n+1}}(x_1, \ldots, x_n) \end{array}$$

which has inverse

$$\begin{array}{rccc} \varphi_S^{-1} : & \mathbb{R}^n & \to & S^n \backslash \{S\} \\ & (x_1, \ldots, x_n) & \mapsto & (\frac{2x_1}{1+x_1^2+\cdots+x_n^2}, \ldots, \frac{2x_n}{1+x_1^2+\cdots+x_n^2}, \frac{1-x_1^2-\cdots-x_n^2}{1+x_1^2+\cdots+x_n^2}) \end{array}$$

It is straightforward, but tedious to check that φ_N and φ_S are diffeomorphisms.

The diffeomorphisms φ_N and φ_S tell us that if we remove a point from S^n, the resulting space is homeomorphic to $\mathbb{R}^n$. The point we remove from S^n can be arbitrary because the n-sphere possesses rotational symmetry. Conversely, if we have $\mathbb{R}^n$ and add a point at infinity then the space is homeomorphic to S^n. As a result, we call S^n the **one point compactification** of $\mathbb{R}^n$.

1.4 Ideals of an affine variety

We begin with a familiar definition.

Definition 1.4.1. Let R be a commutative ring. An **ideal** I is a subset of R such that

1. $0 \in I$
2. If $f, g \in I$ then $f + g \in I$
3. If $f \in I$ and $h \in R$ then $hf \in I$.

The **ideal generated by** $r_1, \ldots, r_n \in R$ is defined by

$$(r_1, \ldots, r_n) = \{\sum_{i=1}^{n} s_i r_i \mid s_i \in R\}.$$

We will sometimes use $\langle r_1, \ldots, r_n \rangle$ to denote the ideal generated by $r_1, \ldots, r_n$. The set $\{r_1, \ldots, r_n\}$ is called a **generating set** for the ideal $(r_1, \ldots, r_n)$.

From every affine variety $V \subseteq k^n$, one can construct an ideal in the following manner.

Definition 1.4.2. Let k be a field, $n \in \mathbb{Z}_{>0}$ and $V = V(S)$ be the affine variety associated with a subset $S \subseteq k[x_1, \ldots, x_n]$. The **ideal of** V, denoted by $I(V)$, is defined by

$$I(V) = \{f \in k[x_1, \ldots, x_n] \mid f(a_1, \ldots, a_n) = 0 \text{ for } (a_1, \ldots, a_n) \in V\}.$$

Notably, $I(V)$ is an ideal of $k[x_1, \ldots, x_n]$.

It follows straight from the definition of an ideal that $I(V)$ is an ideal of $k[x_1, \ldots, x_n]$.

Example 1.4.3. We will work in the field $\mathbb{R}$ and with the polynomial ring $\mathbb{R}[x, y]$. Let $V = \{(0, 0)\}$ be the variety consisting of the origin in $\mathbb{R}^2$. We claim that $I(V) = (x, y)$.

Suppose that $h(x, y) \in (x, y)$. Then, there exists $f, g \in \mathbb{R}[x, y]$ such that $h(x, y) = f(x, y)x + g(x, y)y$. Since $h(0, 0) = 0$, $h \in I(V)$ and $(x, y) \subseteq I(V)$.

Conversely, assume that $p(x,y) \in I(V)$. Write $p(x,y) = \sum_{i,j} a_{ij}x^i y^j$. Since $p(0,0) = 0$, the constant term $a_{00} = 0$. So,

$$p(x,y) = \sum_{i,j \in \mathbb{Z}_{\geq 0}, (i,j) \neq (0,0)} a_{ij}x^i y^j \in (x,y).$$

Therefore, $I(V) = (x,y)$ as required. Note that this example generalises easily to the fields $\mathbb{C}, \mathbb{Q}$ and polynomials in n variables $x_1, x_2, \ldots, x_n$.

The rest of this section is dedicated to proving important properties about ideals of affine varieties.

Theorem 1.4.1. *Let $f_1, \ldots, f_s \in k[x_1, \ldots, x_n]$. Then, $(f_1, \ldots, f_s) \subseteq I(V(f_1, \ldots, f_s))$.*

Proof. Assume that $f_1, \ldots, f_s \in k[x_1, \ldots, x_n]$. Assume that $h \in (f_1, \ldots, f_s)$. If $i \in \{1, 2, \ldots, s\}$ then there exists $h_i \in k[x_1, \ldots, x_n]$ such that

$$h = \sum_{i=1}^{s} h_i f_i.$$

The ideal $I(V(f_1, \ldots, f_s))$ is given by

$$I(V(f_1, \ldots, f_s)) = \{f \in k[x_1, \ldots, x_n] \mid f(a_1, \ldots, a_n) = 0 \text{ for } (a_1, \ldots, a_n) \in V(f_1, \ldots, f_s)\}.$$

If $(a_1, \ldots, a_n) \in V(f_1, \ldots, f_s)$ then $f_i(a_1, \ldots, a_n) = 0$ for $I \in \{1, 2, \ldots, s\}$ and consequently, $h(a_1, \ldots, a_n) = 0$. Therefore, $h \in I(V(f_1, \ldots, f_s))$ and $(f_1, \ldots, f_s) \subseteq I(V(f_1, \ldots, f_s))$. □

In general, $I(V(f_1, \ldots, f_s)) \neq (f_1, \ldots, f_s)$ as ideals of $k[x_1, \ldots, x_n]$. This is demonstrated by the next example.

Example 1.4.4. Consider the ideal $(x^2, y^2) \subseteq k[x,y]$. From the above theorem, we have $(x^2, y^2) \subseteq I(V(x^2, y^2))$. However, the affine variety $V(x^2, y^2) = \{(0,0)\}$ and $I(V(x^2, y^2)) = (x,y)$. Thus,

$$I(V(x^2, y^2)) = (x,y) \neq (x^2, y^2).$$

The next property tells us what happens to the corresponding ideals when we have two affine varieties V, W such that $V \subseteq W$.

Theorem 1.4.2. *Let k be a field, $n \in \mathbb{Z}_{>0}$ and $V, W \subseteq k^n$ be affine varieties. Then, $V \subseteq W$ if and only if $I(W) \subseteq I(V)$.*

Proof. Assume that k is a field, $n \in \mathbb{Z}_{>0}$ and $V, W \subseteq k^n$ be affine varieties.

Assume that $V \subseteq W$. Assume that $f \in I(W) \subseteq k[x_1, \ldots, x_n]$ so that f vanishes on W. Since $V \subseteq W$, f must vanish on V and consequently, $f \in I(V)$. So, $I(W) \subseteq I(V)$.

Conversely, assume that $I(W) \subseteq I(V)$ and $(a_1, \ldots, a_n) \in V \subseteq k^n$. If $f \in I(V)$ then $f(a_1, \ldots, a_n) = 0$. In particular, since $I(W) \subseteq I(V)$, $g(a_1, \ldots, a_n) = 0$ for any $g \in I(W)$. So, $(a_1, \ldots, a_n) \in W$ and $V \subseteq W$ as required. □

A consequence of Theorem 1.4.2 is that $V = W$ as affine varieties in k^n if and only if $I(V) = I(W)$ as ideals in $k[x_1, \ldots, x_n]$. Here are the three main questions about ideals that the course is concerned with:

1. (Ideal description) Is every ideal $I \subseteq k[x_1, \ldots, x_n]$ finitely generated (has a finite generating set)? That is, is $k[x_1, \ldots, x_n]$ a Noetherian ring? The answer is yes, as a consequence of the Hilbert basis theorem we will prove later.

2. (Ideal membership) If $f_1, \ldots, f_s \in k[x_1, \ldots, x_n]$ then is there a systematic algorithm to decide whether a given polynomial $f \in k[x_1, \ldots, x_n]$ lies in the ideal $(f_1, \ldots, f_s)$? We will answer this question in the affirmative for polynomials in a single variable in the next few sections.

3. (Nullstellensatz) What is the exact relation between $(f_1, \ldots, f_s)$ and $I(V(f_1, \ldots, f_s))$?

By definition of an ideal, a polynomial $f \in k[x_1, \ldots, x_n]$ is an element of the ideal $(f_1, \ldots, f_s)$ if it can be written as a $k[x_1, \ldots, x_n]$-linear combination of the polynomials $f_1, \ldots, f_s$. The point here is that in very particular circumstances, we can use linear algebra to solve the ideal membership problem and tell whether two ideals are equal to each other.

Theorem 1.4.3. *Let k be a field, $n \in \mathbb{Z}_{>0}$ and $(f_1, \ldots, f_s), (g_1, \ldots, g_s) \subseteq k[x_1, \ldots, x_n]$ be two ideals. Then, $(f_1, \ldots, f_s) = (g_1, \ldots, g_s)$ if and only if there exists an invertible matrix $A \in GL_s(k[x_1, \ldots, x_n])$ such that*

$$A \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_s \end{pmatrix} = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_s \end{pmatrix}.$$

Proof. Assume that k is a field, $n \in \mathbb{Z}_{>0}$ and $(f_1, \ldots, f_s), (g_1, \ldots, g_s)$ be two ideals of $k[x_1, \ldots, x_n]$.

To show: (a) If there exists $A \in GL_s(k[x_1, \ldots, x_n])$ such that $A[f_1, \ldots, f_s]^T = [g_1, \ldots, g_s]^T$ then $(f_1, \ldots, f_s) = (g_1, \ldots, g_s)$.

(b) If $(f_1, \ldots, f_s) = (g_1, \ldots, g_s)$ then there exists $A \in GL_s(k[x_1, \ldots, x_n])$ such that $A[f_1, \ldots, f_s]^T = [g_1, \ldots, g_s]^T$.

(a) Assume that there exists $A = (a_{ij}) \in GL_s(k[x_1, \ldots, x_n])$ such that $A[f_1, \ldots, f_s]^T = [g_1, \ldots, g_s]^T$. If $i \in \{1, 2, \ldots, s\}$ then

$$\sum_{k=1}^{s} a_{ik} f_k = g_i$$

and $g_i \in (f_1, \ldots, f_s)$. So, $(g_1, \ldots, g_s) \subseteq (f_1, \ldots, f_s)$.

Now let $B = (b_{ij}) = A^{-1}$. If $i \in \{1, 2, \ldots, s\}$ then

$$\sum_{k=1}^{s} b_{ik} g_k = f_i$$

and $f_i \in (g_1, \ldots, g_s)$. Therefore, $(f_1, \ldots, f_s) \subseteq (g_1, \ldots, g_s)$ and $(f_1, \ldots, f_s) = (g_1, \ldots, g_s)$.

(b) Assume that $(f_1, \ldots, f_s) = (g_1, \ldots, g_s)$. For each $i \in \{1, 2, \ldots, s\}$, there exists polynomials $p_{i1}, \ldots, p_{is} \in k[x_1, \ldots, x_n]$ such that

$$\sum_{\ell=1}^{s} p_{i\ell} f_\ell = g_i.$$

Consequently, if $P = (p_{ij}) \in M_{s\times s}(k[x_1, \ldots, x_n])$ then

$$P \begin{pmatrix} f_1 \\ \vdots \\ f_s \end{pmatrix} = \begin{pmatrix} g_1 \\ \vdots \\ g_s \end{pmatrix}.$$

Similarly, there exists $D = (d_{ij}) \in M_{s\times s}(k[x_1, \ldots, x_n])$ such that

$$D \begin{pmatrix} g_1 \\ \vdots \\ g_s \end{pmatrix} = \begin{pmatrix} f_1 \\ \vdots \\ f_s \end{pmatrix}.$$

From these two equations, it is easy to check that $PD = DP = I_s$, where I_s is the $s \times s$ identity matrix with elements in $k[x_1, \ldots, x_n]$. So, P is invertible as required. □

We finish with an example of Theorem 1.4.3 in action.

Example 1.4.5. This example is taken from [CLO15, §4 Exercise 3c]. We want to prove the equality

$$(2x^2 + 3y^2 - 11, x^2 - y^2 - 3) = (x^2 - 4, y^2 - 1)$$

of ideals in $\mathbb{Q}[x, y]$.

After some inspection, we find that

$$2x^2 + 3y^2 - 11 = 2(x^2 - 4) + 3(y^2 - 1)$$

and

$$x^2 - y^2 - 3 = (x^2 - 4) - (y^2 - 1).$$

Consequently, we obtain the matrix equation

$$\begin{pmatrix} 2 & 3 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x^2 - 4 \\ y^2 - 3 \end{pmatrix} = \begin{pmatrix} 2x^2 + 3y^2 - 11 \\ x^2 - y^2 - 3 \end{pmatrix}.$$

Now observe that the determinant

$$\begin{vmatrix} 2 & 3 \\ 1 & -1 \end{vmatrix} = -5$$

and -5 is an invertible element of $\mathbb{Q}[x, y]$. By Theorem 1.4.3, we find that

$$(2x^2 + 3y^2 - 11, x^2 - y^2 - 3) = (x^2 - 4, y^2 - 1).$$

1.5 Polynomials in one variable

This section is dedicated to answering the ideal membership problem for the polynomial ring $k[x]$ of one variable. The idea here is that if k is a field then $k[x]$ is a Euclidean domain. By contrast, the polynomial ring of n variables $k[x_1, \ldots, x_n]$ is not even a PID. What does it mean for $k[x]$ to be a Euclidean domain? It means that we can divide polynomials of one variable in the same way we divide positive integers. Also, we can perform

the **Euclidean algorithm**.

Let k be a field and $g \in k[x]$ with $g \neq 0$. If $f \in k[x]$ then there exists unique $q, r \in k[x]$ such that either $r = 0$ or $\deg(r) < \deg(q)$ and

$$f = qg + r.$$

The purpose of the polynomial division algorithm is to find q and r from f and g. We require a quick definition for the purpose of describing the algorithm.

Definition 1.5.1. Let

$$f(x) = c_0x^m + c_1x^{m-1} + \cdots + c_m \in k[x]$$

where $c_0 \neq 0$ and $m \in \mathbb{Z}_{>0}$. We define c_0x^m to be the **leading term** of f. The leading term is written as $LT(f) = c_0x^m$.

We will now describe the polynomial division algorithm for $k[x]$ below:

1. Suppose that we are given two polynomials $f, g \in k[x]$ such that $g \neq 0$.

2. First, define $q_0 = 0$ and $r_0 = f$. If either $r_0 = 0$ or $LT(g)$ does not divide $LT(r_0)$, set $q = q_0$ and $r = r_0$ and terminate the algorithm. Otherwise, proceed to Step 3.

3. Define

$$q_1 = q_0 + \frac{LT(r_0)}{LT(g)} \qquad \text{and} \qquad r_1 = r_0 - \frac{LT(r_0)}{LT(g)}g.$$

4. If If either $r_1 = 0$ or $LT(g)$ does not divide $LT(r_1)$, set $q = q_1$ and $r = r_1$ and terminate the algorithm. Otherwise, go to Step 5.

5. For $i \in \mathbb{Z}_{\geq 0}$, define

$$q_{i+1} = q_i + \frac{LT(r_i)}{LT(g)} \qquad \text{and} \qquad r_{i+1} = r_i - \frac{LT(r_i)}{LT(g)}g.$$

6. If either $r_{i+1} = 0$ or $LT(g)$ does not divide $LT(r_{i+1})$, set $q = q_{i+1}$ and $r = r_{i+1}$. Terminate the algorithm. Otherwise, repeat Step 5.

The main sticking point we will have to address is: why does the above algorithm work? We will break our reasoning up into different parts:

1. Do the polynomials $q, r \in k[x]$ satisfy $f = qg + r$?

The answer is yes and we will prove this by induction on $i \in \mathbb{Z}_{\geq 0}$. For the base case, assume that $i = 0$. Then, $q_0 = 0$, $r_0 = f$ and $f = q_0 g + r_0$.

For the inductive hypothesis, assume that $f = q_m g + r_m$ for some $m \in \mathbb{Z}_{>0}$. Then,

$$\begin{aligned} q_{m+1}g + r_{m+1} &= (q_m + \frac{LT(r_m)}{LT(g)})g + (r_m - \frac{LT(r_m)}{LT(g)}g) \\ &= q_m g + r_m = f. \end{aligned}$$

This completes the induction. Hence, at every step of the algorithm, $f = q_i g + r_i$ for $i \in \mathbb{Z}_{\geq 0}$. So, $f = qg + r$ because $q = q_j$ and $r = r_j$ for some $j \in \mathbb{Z}_{\geq 0}$.

2. Does $r \in k[x]$ satisfy either $r = 0$ or $\deg(r) < \deg(g)$?

The algorithm always terminates when either $r_j = 0$ or $LT(g)$ does not divide $LT(r_j)$ for some $j \in \mathbb{Z}_{\geq 0}$. So, either $r = 0$ or $LT(g)$ does not divide $LT(r)$. Notice that $LT(g)$ does not divide $LT(r)$ if and only if $\deg(r) < \deg(g)$.

3. Does the algorithm always terminate?

The key observation here is that if $i \in \mathbb{Z}_{>0}$ then either $r_i = 0$ or $\deg(r_i) < \deg(r_{i-1})$. To see why this is the case, fix $m \in \mathbb{Z}_{>0}$. Let $r_{m-1} = c_0 x^p + \cdots + c_p$ and $g = d_0 x^q + \cdots + d_q$ so that $p \geq \ell$. The leading terms are $LT(r_{m-1}) = c_0 x^p$ and $LT(g) = d_0 x^q$. Computing r_m directly, we obtain

$$r_m = r_{m-1} - \frac{LT(r_{m-1})}{LT(g)}g = (c_0 x^p + \dots) - (\frac{c_0}{d_0}x^{p-q} \cdot d_0 x^q + \dots)$$

Therefore, $\deg(r_m) < \deg(r_{m-1})$ or $r_m = 0$. Since the degree of a polynomial is a non-negative integer, the degree can only drop at most finitely many times. So, the algorithm must terminate.

4. Are the polynomials $q, r \in k[x]$ unique?

Assume that $f = qg + r = q'g + r'$, where $\deg(r) < \deg(g)$ and $\deg(r') < \deg(g)$. We want to show that $q = q'$ and $r = r'$. Suppose for the sake of contradiction that $r \neq r'$. We know that $\deg(r' - r) < \deg(g)$. Now observe that

$$(q - q')g = r' - r$$

so that $q - q' \neq 0$ and consequently,

$$\deg(r' - r) = \deg((q - q')g) = \deg(q - q') + \deg(g) \geq \deg(g).$$

This contradicts the previous finding that $\deg(r' - r) < \deg(g)$. So, $r' = r$ and since $(q - q')g = r' - r$, $q = q'$. This demonstrates that q and r are unique.

We have finally demonstrates that the polynomial division algorithm works in the polynomial ring $k[x]$. We summarise this finding as the following theorem.

Theorem 1.5.1. *Let k be a field. Then, $k[x]$ is a Euclidean domain.*

By Theorem 1.5.1, $k[x]$ is also a PID, UFD and a Noetherian ring. In particular, the observation that $k[x]$ is a PID is crucial for answering the ideal membership problem for $k[x]$. Recall that this means that every ideal of $k[x]$ is generated by exactly one polynomial in $k[x]$. Below, we give a direct proof of the fact that if k is a field then $k[x]$ is a PID.

Theorem 1.5.2. *Let k be a field. Then, $k[x]$ is a PID.*

Proof. Assume that k is a field. Assume that I is an ideal in $k[x]$. If $I = 0$ then I is generated by 0 and is hence, principal. So, suppose that $I \neq 0$. Let $g(x) \in I$ be a polynomial of minimal degree amongst all non-zero elements in I. Since k is a field, we can assume that $g(x)$ is a monic polynomial.

To show: (a) $I = (g(x))$.

(a) To show: (aa) $I \subset (g(x))$.

(ab) $(g(x)) \subset I$.

(aa) Assume that $f(x) \in I$. From the polynomial division algorithm with $f(x)$ and $g(x)$, we can deduce the existence of unique $q(x), r(x) \in k[x]$ such that

$$f(x) = q(x)g(x) + r(x)$$

where $\deg(r(x)) < \deg(g(x))$ or $r(x) = 0$. Suppose for the sake of contradiction that $r(x) \neq 0$. Then, $\deg(r(x)) < \deg(g(x))$. Observe that $r = f(x) - q(x)g(x) \in I$. However, this contradicts the fact that $g(x)$ is the polynomial of minimal degree amongst all non-zero elements in I. Therefore, $r(x) = 0$ and consequently, $f(x) = q(x)g(x)$. Hence, $I \subset (g(x))$.

(ab) Since $g(x) \in I$, $(g(x)) \subset I$.

Thus, $I = (g(x))$. So, $k[x]$ is a PID. □

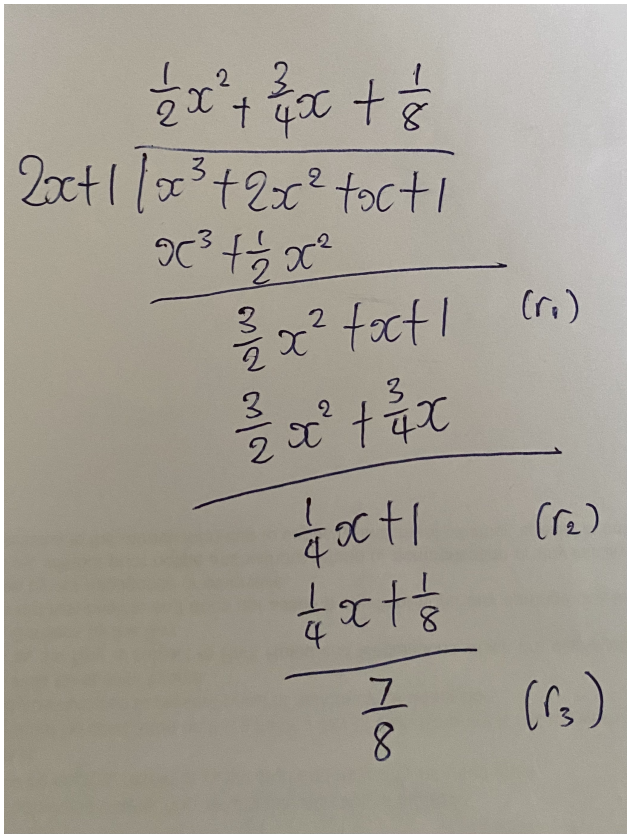
Example 1.5.2. We will apply the polynomial division algorithm to the polynomials $f(x) = x^3 + 2x^2 + x + 1$ and $g(x) = 2x + 1$ in $\mathbb{Q}[x]$. We list the relevant steps of the polynomial division algorithm below:

1. $q_0 = 0, r_0 = f$
2. $q_1 = \frac{1}{2}x^2$, $r_1 = \frac{3}{2}x^2 + x + 1$
3. $q_2 = \frac{1}{2}x^2 + \frac{3}{4}x$, $r_2 = \frac{1}{4}x + 1$
4. $q_3 = \frac{1}{2}x^2 + \frac{3}{4}x + \frac{1}{8}$, $r_3 = \frac{7}{8}$

Since $\deg(r_3) < \deg(g)$, the algorithm terminates with $q = q_3$ and $r = r_3$. One can check directly that

$$x^3 + 2x^2 + x + 1 = (\frac{1}{2}x^2 + \frac{3}{4}x + \frac{1}{8})(2x + 1) + \frac{7}{8}.$$

Note that each step in the algorithm corresponds to a step in the classic polynomial division below:



Let us return to the observation that $k[x]$ is a PID. This means that there is a well-defined notion of a greatest common divisor.

Definition 1.5.3. Let k be a field and $f_1, \ldots, f_s \in k[x]$. The **greatest common divisor** of $f_1, \ldots, f_s$, denoted by $\gcd(f_1, \ldots, f_s)$, is a polynomial in $k[x]$ such that

1. If $i \in \{1, 2, \ldots, s\}$ then $\gcd(f_1, \ldots, f_s) | f_i$
2. If $h \in k[x]$ divides f_i for every $i \in \{1, 2, \ldots, s\}$ then $h | \gcd(f_1, \ldots, f_s)$.

We will now show that the greatest common divisor of two polynomials exists and is unique up to multiplication by a non-zero constant in k.

Theorem 1.5.3. *Let k be a field and $f, g \in k[x]$. Then, $\gcd(f, g)$ exists and is unique up to multiplication by a non-zero constant in k. Moreover, $\gcd(f, g)$ is a generator of the ideal (f, g).*

Proof. Assume that k is a field and $f, g \in k[x]$. Since $k[x]$ is a PID, there exists a polynomial $h \in k[x]$ such that $(h) = (f, g)$.

To show: (a) $h = \gcd(f, g)$.

(a) Since $f, g \in (h)$, there exists polynomials $p_1, p_2 \in k[x]$ such that $f = p_1 h$ and $g = p_2 h$. Therefore, h divides both f and g.

Now assume that there exists a polynomial $q \in k[x]$ such that $q|f$ and $q|g$. Then, there exists $a, b \in k[x]$ such that $f = aq$ and $g = bq$. Since $h \in (f, g)$, there exists $c, d \in k[x]$ such that $h = cf + dg = caq + dbq = (ca + db)q$. So, $q|h$.

So, h satisfies the properties of the greatest common divisor of f and g. Hence, $h = \gcd(f, g)$, which means that the greatest common divisor of f and g exists and is the generator of the ideal (f, g).

It remains to show that $\gcd(f, g)$ exists and is unique up to multiplication by a non-zero constant in k. Suppose that h and h' are both the greatest common divisors of f and g. Then, $h|h'$ and $h'|h$. Hence, h and h' are associates and consequently, $h = h'j$ for some non-zero constant $j \in k$. So, $\gcd(f, g)$ is unique up to multiplication by a non-zero constant in k. □

We emphasise that the proof of Theorem 1.5.3 extends readily to an arbitrary PID. We can also generalise Theorem 1.5.3 with an almost identical proof.

Theorem 1.5.4. *Let k be a field and $f_1, \ldots, f_s \in k[x]$. Then, $\gcd(f_1, f_2, \ldots, f_s)$ exists and is unique up to multiplication by a non-zero constant in k. Moreover, $\gcd(f_1, f_2, \ldots, f_s)$ is a generator of the ideal $(f_1, f_2, \ldots, f_s)$.*

There is a simple method of finding the greatest common divisor of more than two polynomials.

Theorem 1.5.5. *Let k be a field and $f_1, \ldots, f_s \in k[x]$, where $s \in \mathbb{Z}_{>2}$. Then,*

$$\gcd(f_1, \ldots, f_s) = \gcd(f_1, \gcd(f_2, \ldots, f_s)).$$

Proof. Assume that $s \in \mathbb{Z}_{>2}$ and $f_1, \ldots, f_s \in k[x]$. Let $h = \gcd(f_2, f_3, \ldots, f_s)$.

To show: (a) $(f_1, h) = (f_1, f_2, \ldots, f_s)$.

(a) First assume that $g \in (f_1, h)$ so that there exists $p_1, p_2 \in k[x]$ such that $g = p_1 f_1 + p_2 h$. Since $(h) = (f_2, \ldots, f_s)$, we can write h as a linear combination of the polynomials $f_2, \ldots, f_s$. So, g is a linear combination of polynomials $f_1, \ldots, f_s$ and consequently, $g \in (f_1, f_2, \ldots, f_s)$.

Now assume that $j \in (f_1, f_2, \ldots, f_s)$. Then, j is a linear combination of the polynomials $f_1, \ldots, f_s$. Note that h divides f_i for $i \in \{2, 3, \ldots, s\}$. So, j is a linear combination of f_1 and h. Hence, $j \in (f_1, h)$. So, $(f_1, h) = (f_1, f_2, \ldots, f_s)$.

The equality of ideals $(f_1, h) = (f_1, f_2, \ldots, f_s)$ reveals that $\gcd(f_1, h) = \gcd(f_1, f_2, \ldots, f_s)$ as required. □

The Euclidean algorithm is a systematic method for finding the greatest common divisor of two polynomials $f, g \in k[x]$, where $g \neq 0$.

1. Suppose that we are given two polynomials $f, g \in k[x]$ where $g \neq 0$.
2. Use the polynomial division algorithm to write $f = q_1 g + r_1$ for unique $q_1, r_1 \in k[x]$. If $r_1 = 0$ then $\gcd(f, g) = g$ and the algorithm terminates. Otherwise, proceed to Step 3.
3. Use the polynomial division algorithm to write $g = q_2 r_1 + r_2$ for unique $q_2, r_2 \in k[x]$. If $r_2 = 0$ then $\gcd(f, g) = r_1$ and the algorithm terminates. Otherwise proceed to Step 4.
4. For $i \in \mathbb{Z}_{>0}$, use the polynomial division algorithm to write $r_i = q_{i+2} r_{i+1} + r_{i+2}$ (beginning with $i = 1$).
5. If $r_{i+2} = 0$ then $\gcd(f, g) = r_{i+1}$ and the algorithm terminates. Otherwise, increase i by 1 and repeat step 4.

Let us explain why the Euclidean algorithm gives us the greatest common divisor. In step 2, we used polynomial division to write $f = q_1 g + r_1$. If $r_1 \neq 0$ then in step 3, we used polynomial division to write $g = q_2 r_1 + r_2$. The point here is that in transitioning from step 2 to step 3,

$$\gcd(f, g) = \gcd(f - q_1 g, g) = \gcd(r_1, g).$$

This is true because as ideals, $(f, g) = (f - q_1 g, g)$. In step 4 with $i = 1$, we write $r_1 = q_3 r_2 + r_3$. When we transition from step 3 to step 4, we find that

$$\gcd(r_1, g) = \gcd(r_1, g - q_2 r_1) = \gcd(r_1, r_2).$$

Therefore,

$$\gcd(f,g) = \gcd(g,r_1) = \gcd(r_1,r_2) = \cdots = \gcd(r_i,r_{i+1}) = \ldots$$

and by the polynomial division algorithm,

$$\deg(g) > \deg(r_1) > \deg(r_2) > \deg(r_3) > \ldots$$

Since the degree is a positive integer, the above chain of inequalities tells us that there exists $j \in \mathbb{Z}_{\geq 0}$ such that $r_j = 0$ (we set $r_0 = g$). Then, the algorithm terminates and

$$r_{j-1} = \gcd(r_{j-1}, 0) = \gcd(r_{j-1}, r_j) = \gcd(f,g).$$

Therefore, the Euclidean algorithm must terminate and yields the greatest common divisor for f and g.

To reiterate, if we want to compute the gcd of more than two polynomials, we can use Theorem 1.5.5 multiple times, in tandem with the Euclidean algorithm.

Finally, we will describe an algorithm for answering the ideal membership problem for $k[x]$. Let $f_1, \ldots, f_s, g \in k[x]$. Can we determine if $g \in (f_1, \ldots, f_s)$?

1. Let $f_1, \ldots, f_s, g \in k[x]$.
2. Use Theorem 1.5.5 and the Euclidean algorithm to compute $\gcd(f_1, \ldots, f_s)$. By Theorem 1.5.4, $(\gcd(f_1, \ldots, f_s)) = (f_1, \ldots, f_s)$ as ideals of $k[x]$.
3. Now use the polynomial division algorithm to write $g = q \gcd(f_1, \ldots, f_s) + r$ for unique $q, r \in k[x]$. If $r = 0$ then $\gcd(f_1, \ldots, f_s)$ divides g and $g \in (f_1, \ldots, f_s)$. If $r \neq 0$ then $g \notin (f_1, \ldots, f_s)$.

We will now apply the ideal membership algorithm to a concrete example.

Example 1.5.4. We will work in the polynomial ring $\mathbb{Q}[x]$. We want to determine whether $x^2 - 4$ is an element of the ideal

$$I = (x^3 + x^2 - 4x - 4, x^3 - x^2 - 4x + 4, x^3 - 2x^2 - x + 2).$$

The first step is to compute the gcd of the polynomials which generate I. We find that

$$
\begin{aligned}
&\gcd(x^3 + x^2 - 4x - 4, x^3 - x^2 - 4x + 4, x^3 - 2x^2 - x + 2) \\
&= \gcd(\gcd(x^3 + x^2 - 4x - 4, x^3 - x^2 - 4x + 4), x^3 - 2x^2 - x + 2) \\
&= \gcd(2x^2 - 8, x^3 - 2x^2 - x + 2) \\
&= 3x - 6.
\end{aligned}
$$

So, $I = (3x - 6) = (x - 2)$. The equality $(3x - 6) = (x - 2)$ follows from the fact that the gcd is unique up to multiplication by a non-zero constant in k. Hence, $x^2 - 4 = (x + 2)(x - 2) \in I$.

In the next chapter, we will tackle the ideal membership problem for the polynomial ring $k[x_1, \ldots, x_n]$.

Chapter 2

Gröbner bases

2.1 Orderings on the monomials in $k[x_1, \ldots, x_n]$

Before we dive into the ideal membership problem for $k[x_1, \ldots, x_n]$, we want to highlight a particular feature of polynomial division in $k[x]$. The observation that interests us the most is the fact that polynomial division in $k[x]$ relies on an ordering of the terms in the polynomial. In particular, the polynomials we begin with are generally written as a sum of terms with decreasing degree. We begin with the leading term and end with the constant term.

For division in $k[x]$, we are dealing with the degree ordering on the monomials

$$\cdots > x^{m+1} > x^m > \cdots > x^2 > x > 1.$$

In fact, the success of the polynomial division algorithm in $k[x]$ hinges on working with the leading terms of both polynomials, rather than using arbitrary terms from both polynomials.

Another example where an ordering of the variables plays a subtle, but important role is the row reduction algorithm, used to convert a matrix into its row echelon form. Generally, the algorithm begins with the leftmost column with non-zero entries.

These examples, as outlined in [CLO15], tell us that if we want to extend polynomial division to $k[x_1, \ldots, x_n]$ for the purposes of handling the ideal membership problem then one must place a particular order on the terms

comprising a polynomial in $k[x_1, \ldots, x_n]$. In this section, we will discuss some of these orderings.

First, we want an idea of the desirable properties an ordering on monomials in $k[x_1, \ldots, x_n]$ should satisfy.

Definition 2.1.1. Let k be a field. A **monomial ordering** $>$ on $k[x_1, \ldots, x_n]$ is a relation $>$ on $(\mathbb{Z}_{\geq 0})^n$ or equivalently, a relation on the set of monomials $\{x^\alpha \mid \alpha \in (\mathbb{Z}_{\geq 0})^n\}$ such that

1. $>$ is a total/linear ordering on $(\mathbb{Z}_{\geq 0})^n$ (and not a partial ordering; we want to be able to compare every pair of n-tuples)

2. If $\alpha > \beta$ and $\gamma \in (\mathbb{Z}_{\geq 0})^n$ then $\alpha + \gamma > \beta + \gamma$, where addition on $(\mathbb{Z}_{\geq 0})^n$ is defined componentwise.

3. $>$ is a well-ordering on $(\mathbb{Z}_{\geq 0})^n$. This means that if a subset $A \subseteq (\mathbb{Z}_{\geq 0})^n$ is non-empty then there exists $\alpha \in A$ such that if $\beta \in A - \{\alpha\}$ then $\beta > \alpha$.

The well-ordering property of a monomial ordering will play a crucial role in showing that various algorithms terminate. To see why this is the case, we will provide an equivalent formulation of the well-ordering property.

Theorem 2.1.1. *Let $n \in \mathbb{Z}_{>0}$. An order $>$ on $(\mathbb{Z}_{\geq 0})^n$ is a well-ordering if and only if every strictly decreasing sequence in $(\mathbb{Z}_{\geq 0})^n$*

$$\alpha(1) > \alpha(2) > \alpha(3) > \ldots$$

terminates.

Proof. Assume that $n \in \mathbb{Z}_{>0}$ and $>$ is an order on $(\mathbb{Z}_{\geq 0})^n$.

To show: (a) If $>$ is not a well-ordering then there exists an infinite strictly decreasing sequence in $(\mathbb{Z}_{\geq 0})^n$.

(b) If there exists an infinite strictly decreasing sequence in $(\mathbb{Z}_{\geq 0})^n$ then $>$ is not a well-ordering.

(a) Assume that $>$ is not a well-ordering. Then, there exists a non-empty subset $B \subseteq (\mathbb{Z}_{\geq 0})^n$ such that B does not contain a minimal element with respect to the ordering $>$. Now pick an element $\beta(1) \in B$. It is not minimal with respect to $>$. Next, pick an element $\beta(2) \in B$ such that $\beta(1) > \beta(2)$.

The element $\beta(2)$ is not minimal with respect to $>$.

Continuing in this fashion, we obtain an strictly decreasing infinite sequence

$$\beta(1) > \beta(2) > \dots$$

in B and hence, in $(\mathbb{Z}_{\geq 0})^n$.

(b) Suppose that there exists an infinite strictly decreasing sequence in $(\mathbb{Z}_{\geq 0})^n$ with respect to $>$. From the definition of well-ordering, we deduce that $>$ is not a well-ordering. □

If we have an algorithm which uses a well-ordering and we construct a strictly decreasing sequence $\beta(1) > \beta(2) > \dots$ at each step of the algorithm then Theorem 2.1.1 tells us that the algorithm must terminate.

Example 2.1.2. The usual order on $\mathbb{Z}_{\geq 0}$

$$\dots > m+1 > m > \dots > 2 > 1 > 0$$

is a monomial ordering. Equivalently, the degree ordering on monomials of $k[x]$

$$\dots > x^{m+1} > x^m > \dots > x^2 > x^1 > x^0 = 1$$

is a monomial ordering.

We will now provide a first example of a monomial ordering on $(\mathbb{Z}_{\geq 0})^n$, where $n \in \mathbb{Z}_{>0}$.

Definition 2.1.3. Let $\alpha = (\alpha_1, \dots, \alpha_n)$ and $\beta = (\beta_1, \dots, \beta_n)$ be elements of $(\mathbb{Z}_{\geq 0})^n$. We say that $\alpha >_{lex} \beta$ if the leftmost non-zero entry of the difference $\alpha - \beta \in \mathbb{Z}^n$ is positive.

We write $x^\alpha >_{lex} x^\beta$ in $k[x_1, \dots, x_n]$ if $\alpha >_{lex} \beta$. The order $>_{lex}$ is called the **lexicographical order**.

Lexicographical order is sometimes called a dictionary order because it is the same order used to organise a dictionary. For instance, the word "bread" appears before "brioche" in the dictionary.

Notice that in lexicographical ordering,

$$(1, 0, \dots, 0) >_{lex} (0, 1, \dots, 0) >_{lex} \dots >_{lex} (0, 0, \dots, 1)$$

so that $x_1 >_{lex} x_2 >_{lex} \cdots >_{lex} x^n$. Now we will show that the lexicographical ordering on $(\mathbb{Z}_{\geq 0})^n$ is a monomial ordering.

Theorem 2.1.2. *The lexicographical ordering $>_{lex}$ is a monomial ordering.*

Proof. The fact that $(\mathbb{Z}_{\geq 0})^n$ is a total ordering follows from the fact that the usual order $>$ on $\mathbb{Z}_{\geq 0}$ is a total ordering and from the definition of the lexicographical ordering (which uses the order $>$).

Assume that $\alpha >_{lex} \beta$ so that there exists $i \in \mathbb{Z}_{>0}$ such that i is minimal and $\alpha_i - \beta_i > 0$. Assume that $\gamma \in (\mathbb{Z}_{\geq 0})^n$. Since $\alpha_i + \gamma_i > \beta_i + \gamma_i$, $\alpha + \gamma >_{lex} \beta + \gamma$.

Finally, suppose for the sake of contradiction that $>_{lex}$ is not a well-ordering. By Theorem 2.1.1, there exists an infinite strictly decreasing series

$$\alpha(1) >_{lex} \alpha(2) >_{lex} \alpha(3) >_{lex} \dots$$

Consider the first entry of each $\alpha(i)$ in the above sequence. Due to the definition of $>_{lex}$, the first entries of each $\alpha(i)$ form a non-increasing sequence in $\mathbb{Z}_{\geq 0}$. So, there exists $\ell \in \mathbb{Z}_{\geq 0}$ such that if $i, j \geq \ell$ then the first entry of $\alpha(i)$ is equal to the first entry of $\alpha(j)$.

Now consider the infinite strictly descending sequence $\alpha(\ell) >_{lex} \alpha(\ell+1) >_{lex} \dots$ The lexicographical order is determined by the second entry of each member of the sequence. The second entries of $\alpha(\ell), \alpha(\ell+1), \dots$ form a non-increasing sequence. By repeating the same reasoning, we deduce that the second entries of the sequence eventually stabilises.

Continuing in the same way, we deduce eventually that there exists $m \in \mathbb{Z}_{>0}$ such that if $i, j \geq m$ then $\alpha(i) = \alpha(j)$. This contradicts the fact that $\alpha(m) >_{lex} \alpha(m+1)$. Therefore, $>_{lex}$ is a well-ordering and hence, a monomial ordering. □

Example 2.1.4. Here is a practical use of lexicographical order. Let R be a commutative ring and $A \in M_{n\times n}(R)$. Let $k \in \{1, 2, \dots, n-1\}$. Then, $\Lambda^k(A)$ is the $\binom{n}{k} \times \binom{n}{k}$ matrix whose elements are the determinants of $k \times k$ minors of A.

The point is that the rows and columns of $\Lambda^k(A)$ are indexed by the $\binom{n}{k}$ subsets of $\{1, 2, \dots, n\}$ with cardinality k. The order in which the rows and

columns are indexed does not matter, as long as the rows and columns are indexed in the same manner.

The lexicographical order on the $\binom{n}{k}$ subsets of $\{1, 2, \ldots, n\}$ with cardinality k provides us with a systematic way of indexing the rows and columns of $\Lambda^k(A)$. Let $R = \mathbb{Z}$, $n = 4$, $k = 2$ and

$$A = \begin{pmatrix} 1 & 2 & 5 & -2 \\ 0 & 4 & 2 & 6 \\ 5 & -3 & 9 & 7 \\ -8 & -2 & -1 & 2 \end{pmatrix}$$

With lexicographical ordering, we have

$$\{1,2\} <_{lex} \{1,3\} <_{lex} \{1,4\} <_{lex} \{2,3\} <_{lex} \{2,4\} <_{lex} \{3,4\}$$

and the matrix $\Lambda^2(A)$ is

$$\Lambda^2(A) = \begin{array}{c} \begin{matrix} \{1,2\} & \{1,3\} & \{1,4\} & \{2,3\} & \{2,4\} & \{3,4\} \end{matrix} \\ \begin{pmatrix} 4 & 2 & 6 & -16 & 20 & 34 \\ -13 & -16 & 17 & 33 & 8 & 53 \\ 14 & 39 & -14 & 8 & 0 & 8 \\ -20 & -10 & -30 & 42 & 46 & -40 \\ 32 & 16 & 48 & 0 & 20 & 10 \\ -34 & 67 & 66 & 21 & 8 & 25 \end{pmatrix} \end{array} \begin{matrix} \{1,2\} \\ \{1,3\} \\ \{1,4\} \\ \{2,3\} \\ \{2,4\} \\ \{3,4\} \end{matrix}$$

with rows and columns indexed in lexicographical order. For instance, $(\Lambda^2(A))_{\{1,2\},\{1,3\}}$ is the determinant of the 2×2 minor of A formed from the first and second rows of A and the first and third columns of A. So,

$$(\Lambda^2(A))_{\{1,2\},\{1,3\}} = \begin{vmatrix} 1 & 5 \\ 0 & 2 \end{vmatrix} = 2.$$

This example is from [Cha22, Example 2.3.2].

So far, we observe that the lexicographic order is defined so that $x_1 >_{lex} x_2 >_{lex} \cdots >_{lex} x_n$. The point is that given any ordering of the variables $x_1, \ldots, x_n$, there is a corresponding lexicographic order. There are $n!$ such lexicographic orders.

One of the features of lexicographic order is that a variable dominates any monomial containing smaller variables. For instance $(1, 0, 0) >_{lex} (0, 10, 2)$, which means that as monomials, $x >_{lex} y^{10}z^2$. As evidenced by polynomial

division in $k[x]$, we sometimes want to order based on total degree. This is where our next ordering comes into play.

Definition 2.1.5. Let $\alpha, \beta \in (\mathbb{Z}_{\geq 0})^n$. We say that $\alpha >_{grlex} \beta$ if

$$|\alpha| = \sum_{i=1}^{n} \alpha_i > |\beta| = \sum_{i=1}^{n} \beta_i \qquad \text{or} \qquad |\alpha| = |\beta| \text{ and } \alpha >_{lex} \beta.$$

The order $>_{grlex}$ is called the **graded lexicographic order**.

As examples, we have $(1, 0, 0) <_{grlex} (0, 10, 2)$ and $(3, 2, 1) >_{grlex} (2, 3, 1)$. The graded lexicographic order orders by total degree first and then "breaks ties" with the lexicographic order. As with lexicographic order, there are $n!$ different graded lexicographic orders stemming from the order imposed on the variables $x_1, \ldots, x_n$.

The graded lexicographic order is a monomial ordering. The key reason is because the usual order $>$ on $\mathbb{Z}_{\geq 0}$ and $>_{lex}$ on $(\mathbb{Z}_{\geq 0})^n$ are all monomial orderings.

The final monomial ordering we will state is similar to the lexicographic order, but it breaks ties in a different way.

Definition 2.1.6. Let $\alpha, \beta \in (\mathbb{Z}_{\geq 0})^n$. We say that $\alpha >_{grevlex} \beta$ if

$$|\alpha| = \sum_{i=1}^{n} \alpha_i > |\beta| = \sum_{i=1}^{n} \beta_i$$

or

$|\alpha| = |\beta|$ and the rightmost non-zero entry of $\alpha - \beta \in \mathbb{Z}^n$ is negative.

The order $>_{grevlex}$ is called the **graded reverse lexicographic order**.

The grevlex ordering is indeed a monomial ordering (see [CLO15, Chapter 2 §2, Exercise 12]). Here are a few examples for the purpose of understanding grevlex ordering.

Example 2.1.7. As our first example, $(4, 7, 1) >_{grevlex} (4, 2, 3)$ because the total degree of $(4, 7, 1)$ is 12, while the total degree of $(4, 2, 3)$ is 9.

As a second example, we claim that $(1, 5, 2) >_{grevlex} (4, 1, 3)$. The 3-tuples $(1, 5, 2)$ and $(4, 1, 3)$ have the same total degree. However, $(1, 5, 2) - (4, 1, 3) = (-3, 4, -1)$. By the grevlex ordering, we have $(1, 5, 2) >_{grevlex} (4, 1, 3)$ as required.

Just like the lex and grlex orderings, we have

$$x_1 >_{grevlex} x_2 >_{grevlex} \cdots >_{grevlex} x_n.$$

Also, there are $n!$ such grevlex orderings, corresponding to the orderings of $x_1, \ldots, x_n$.

Finally, we will show how monomial ordering can be applied to polynomials. If $f = \sum_\alpha a_\alpha x^\alpha \in k[x_1, \ldots, x_n]$ and we have a monomial ordering on $k[x_1, \ldots, x_n]$ then (ignoring constant coefficients) we can rearrange the sum so that the terms appear in decreasing order, with respect to the monomial order.

Example 2.1.8. Let $f = 4xy^2z + 4z^2 - 5x^3 + 7x^2x^2 \in k[x, y, z]$. Then, f with respect to the lexicographical order is

$$f = -5x^3 + 7x^2z^2 + 4xy^2z + 4z^2.$$

The polynomial f with respect to grlex order is

$$f = 7x^2z^2 + 4xy^2z - 5x^3 + 4z^2.$$

Finally, the polynomial f with respect to grevlex order is

$$f = 4xy^2z + 7x^2z^2 - 5x^3 + 4z^2$$

Now we will establish specific terminology for monomial ordering.

Definition 2.1.9. Let $f = \sum_\alpha a_\alpha x^\alpha$ be a non-zero polynomial in $k[x_1, \ldots, x_n]$ and let $>$ be a monomial order.

The **multidegree** of f is defined by

$$multideg(f) = \max\{\alpha \in (\mathbb{Z}_{\geq 0})^n \mid a_\alpha \neq 0\}.$$

The **leading coefficient** of f is

$$LC(f) = a_{multideg(f)} \in k.$$

The **leading monomial** of f is

$$LM(f) = x^{multideg(f)}.$$

The **leading term** of f is $LT(f) = LC(f) \cdot LM(f)$.

We will now state a useful property of the multidegree.

Theorem 2.1.3. *Let $f, g \in k[x_1, \ldots, x_n]$ be non-zero polynomials. Then,*

$$multideg(fg) = multideg(f) + multideg(g).$$

Moreover, if $f + g \neq 0$ then

$$multideg(f + g) \leq \max(multideg(f), multideg(g)).$$

If, in addition, $multideg(f) \neq multideg(g)$ then equality occurs.

It is straightforward to prove the above theorem from the definition of the monomial ordering. For instance, the first statement follows from the second property of monomial orderings — if $\alpha, \beta, \gamma \in (\mathbb{Z}_{\geq 0})^n$ and $\alpha > \beta$ then $\alpha + \gamma > \beta + \gamma$. In the second statement, if $LT(f) = -LT(g)$ then inequality occurs.

2.2 The division algorithm in $k[x_1, \ldots, x_n]$

In this section, we would like to generalise the polynomial division algorithm for $k[x]$ to the polynomial ring of several variables $k[x_1, \ldots, x_n]$. Our goal is to divide some $f \in k[x_1, \ldots, x_n]$ by $f_1, \ldots, f_s \in k[x_1, \ldots, x_n]$, which amounts to expressing f in the form

$$f = q_1 f_1 + \cdots + q_s f_s + r$$

where $q_1, \ldots, q_s, r \in k[x_1, \ldots, x_n]$. Expecting a polynomial division algorithm on $k[x_1, \ldots, x_n]$ to have the same nice properties as the polynomial division algorithm on $k[x]$ is rather naive, as the following example demonstrates.

Example 2.2.1. Consider the polynomial ring $\mathbb{R}[x, y]$. We will prove that $\mathbb{R}[x, y]$ is not a PID by showing that the ideal (x, y) is not principal.

Suppose for the sake of contradiction that the ideal $(x, y) \in \mathbb{R}[x, y]$ is principal. Then, there exists a polynomial $p(x, y) \in \mathbb{R}[x, y]$ such that $(p(x, y)) = (x, y)$. So, it must be the case that

$$p(x, y) = q(x, y)x + r(x, y)y$$

where $q(x, y), r(x, y) \in \mathbb{R}[x, y]$. It is important to note from this that $p(x, y)$ cannot be a constant polynomial.

To show: (a) $\gcd(x, y) = p(x, y)$.

(a) To show: (aa) $(x) \subset (p(x, y))$.

(ab) $(y) \subset (p(x, y))$.

(aa) Since $(p(x, y)) = (x, y)$, $(x) \subset (x, y) = (p(x, y))$.

(ab) Since $(p(x, y)) = (x, y)$, $(y) \subset (x, y) = (p(x, y))$.

(a) So, $p(x, y)$ must divide both x and y.

To show: (ac) If $s(x, y)$ divides both x and y then $(p(x, y)) \subset (s(x, y))$.

(ac) Assume $s(x, y) \in \mathbb{R}[x, y]$ divides both x and y. Then, $x = s_1(x, y)s(x, y)$ and $y = s_2(x, y)s(x, y)$ for some $s_1(x, y), s_2(x, y) \in \mathbb{R}[x, y]$. Assume $a(x, y) \in (p(x, y))$. Since $(p(x, y)) = (x, y)$, $a(x, y) = r_1(x, y)x + r_2(x, y)y$ for some $r_1(x, y), r_2(x, y) \in \mathbb{R}[x, y]$. But,

$$\begin{aligned} a(x, y) &= r_1(x, y)x + r_2(x, y)y \\ &= r_1(x, y)s_1(x, y)s(x, y) + r_2(x, y)s_2(x, y)s(x, y) \\ &= (r_1(x, y)s_1(x, y) + r_2(x, y)s_2(x, y))s(x, y). \end{aligned}$$

Hence, $a(x, y) \in (s(x, y))$ and so, $(p(x, y)) \subset (s(x, y))$.

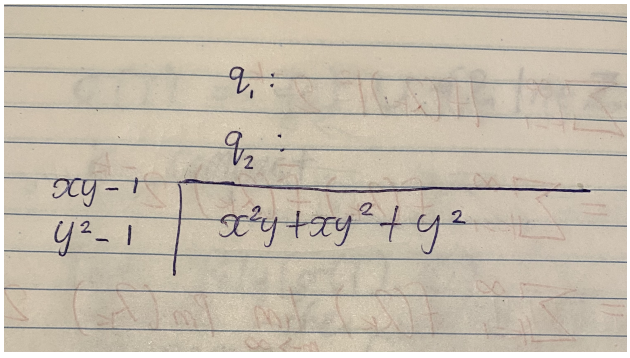
(a) Consequently, $\gcd(x, y) = p(x, y)$, where $p(x, y)$ is not a constant polynomial. However, this contradicts the fact that $\gcd(x, y) = 1$ (x and y are relatively prime). Hence, the ideal (x, y) is not principal.

The above example can be easily adapted to show that $k[x_1, \ldots, x_n]$ is not a PID and thus, not a Euclidean domain. Hence, any remainders we obtain from a division algorithm on $k[x_1, \ldots, x_n]$ cannot be expected to be unique.

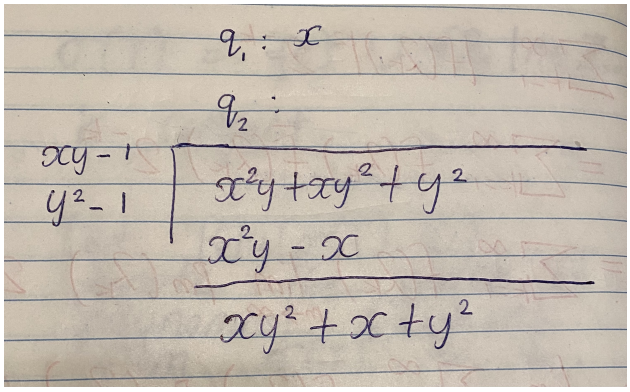
As stated in [CLO15], the basic idea of division in $k[x_1, \ldots, x_n]$ is to cancel the leading term of f by multiplying f_i by a monomial and subtracting. Then, the multiple of f_i becomes a term in the corresponding q_i. Before we state the division algorithm on $k[x_1, \ldots, x_n]$, we will work through the example [CLO15, Chapter 2 §3 Example 2] in detail so that we have a better idea of what to expect.

Example 2.2.2. We will work in the polynomial ring $k[x, y]$. Our goal is to divide $f = x^2y + xy^2 + y^2$ by the two polynomials $f_1 = xy - 1$ and $f_2 = y^2 - 1$. We impose the lexicographic order on $k[x, y]$ with $x >_{lex} y$. The terms in the polynomials f, f_1 and f_2 are already arranged in lexicographic order. Hence, we do not have to worry about first arranging the terms in the polynomials involved by lexicographic order.

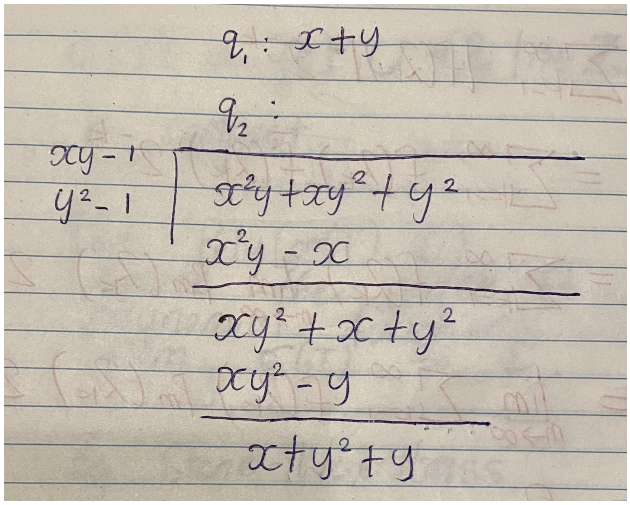
This time, we have two divisors and two quotients q_1 and q_2, which we depict graphically by



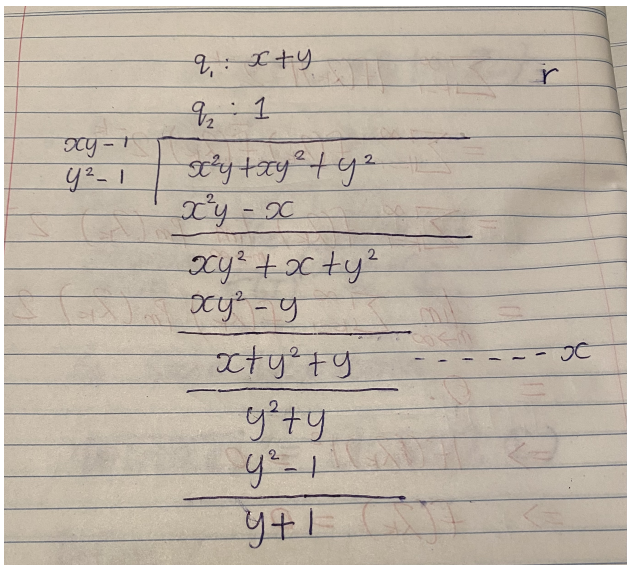
The leading term $LT(xy - 1) = xy$ divides $LT(x^2y + xy^2 + y^2) = x^2y$. We multiply $xy - 1$ by x and then subtract from $x^2y + xy^2 + y^2$.



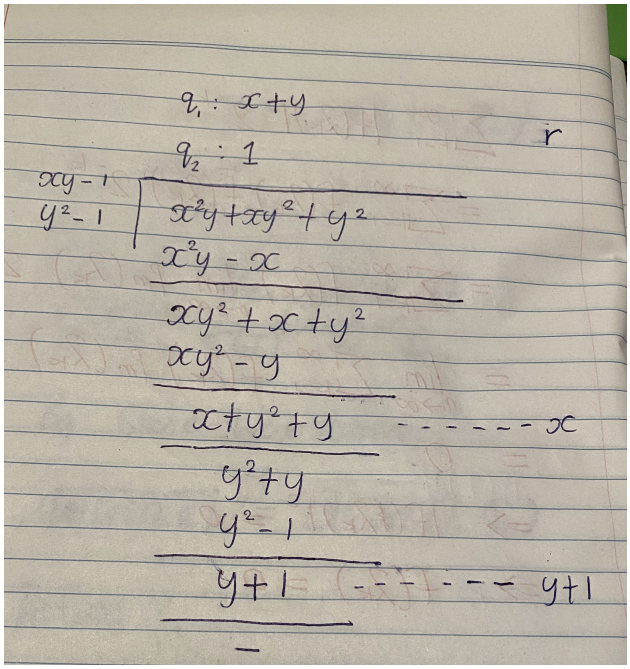
Now $LT(xy - 1)$ and $LT(y^2 - 1)$ both divide xy^2. We will choose the first polynomial $xy - 1$, multiply it by y and then subtract from $xy^2 + x + y^2$. The result is



This time, neither $LT(xy - 1)$ nor $LT(y^2 - 1)$ divide $LT(x + y^2 + y) = x$. Here, the remainder makes it entrance into the algorithm. We could declare $x + y^2 + y$ as the remainder and terminate the algorithm here. However, $LT(y^2 - 1)$ does divide $LT(y^2 + y) = y^2$. So, we can shove the x term into the remainder r and then proceed with the division by dividing $y^2 + y$ by $y^2 - 1$. Our working out now becomes



Again $LT(xy - 1)$ and $LT(y^2 - 1)$ do not divide $LT(y + 1) = y$. But this time, $LT(xy - 1)$ and $LT(y^2 - 1)$ do not divide any of y or 1. Hence, we can bring the entire expression $y + 1$ into the remainder column and add it to the remainder. We are left with no polynomial. So, the algorithm terminates and our final working is



One can check that

$$f = q_1 f_1 + q_2 f_2 + r = (x+y)(xy-1) + 1(y^2-1) + (x+y+1).$$

Keeping this example in mind, we will now describe the polynomial division algorithm for $k[x_1, x_2, \ldots, x_n]$.

Step 1: Suppose that we want to divide a polynomial $f \in k[x_1, \ldots, x_n]$ by some polynomials $f_1, \ldots, f_s \in k[x_1, \ldots, x_n]$. Suppose that we have a monomial ordering $<$ on $k[x_1, \ldots, x_n]$. We first apply the monomial ordering to arrange the terms of the polynomials $f, f_1, \ldots, f_s$ in decreasing order. Set the polynomials $q_1 = \cdots = q_s = 0$, $p = f$ and $r = 0$.

Step 2a: If there exists a $j \in \{1, 2, \ldots, s\}$ such that $LT(f_j)$ divides $LT(p)$ then compute

$$k = \min\{j \in \{1, 2, \ldots, s\} \mid LT(f_j) | LT(p)\}.$$

Then, subtract the quantity $\frac{LT(p)}{LT(f_k)} f_k$ from p and add $\frac{LT(p)}{LT(f_k)}$ to q_k.

Step 2b: If there does not exist $j \in \{1, 2, \ldots, s\}$ such that $LT(f_j)$ divides $LT(p)$ then subtract $LT(p)$ from p. Add $LT(p)$ to r.

Step 3: If $p \neq 0$ then repeat steps 2a and 2b. Otherwise, the algorithm terminates and the resulting polynomials $q_1, \ldots, q_s, r$ satisfy

$$f = q_1 f_1 + \cdots + q_s f_s + r,$$

where the polynomials $f, f_1, \ldots, f_s$ are the ones we started with. Once again, we have to justify why the above algorithm works. Similarly to the polynomial division algorithm in $k[x]$, we will break down our reasoning into different parts.

1. Does $f = q_1 f_1 + \cdots + q_s f_s + p + r$ at every step of the algorithm?

   The answer is yes. Suppose that we have just commenced the algorithm and we are on step 1. Since $q_1 = \cdots = q_s = 0$, $p = f$ and $r = 0$, we can check directly that $f = q_1 f_1 + \cdots + q_s f_s + p + r$.

   To complete this argument, we need to show that steps 2a and 2b preserve the equality $f = q_1 f_1 + \cdots + q_s f_s + p + r$. To see that step 2a preserves this equality, we compute directly that if $j \in \{1, 2, \ldots, s\}$ then

   $$f = q_1 f_1 + \cdots + (q_j + \frac{LT(p)}{LT(f_j)}) f_j + \cdots + q_s f_s + (p - \frac{LT(p)}{LT(f_j)} f_j) + r.$$

   To see that step 2b preserves the equality, note that

   $$f = q_1 f_1 + q_s f_s + (p - LT(p)) + (r + LT(p)).$$

   Therefore, $f = q_1 f_1 + \cdots + q_s f_s + p + r$ at every step of the algorithm given above.

2. Do we actually obtain $f = q_1 f_1 + \cdots + q_s f_s + r$ when the algorithm terminates?

   If $p = 0$ then the algorithm terminates. Since $f = q_1 f_1 + \cdots + q_s f_s + p + r$ at every step of the algorithm, we find that $f = q_1 f_1 + \cdots + q_s f_s + r$ once the algorithm is terminated (because $p = 0$).

3. Do the terms of the remainder r divide any of the leading terms $LT(f_1), \ldots, LT(f_s)$?

   The answer is no. Notice that in step 2b, we add terms to r which are divisible by none of the leading terms $LT(f_1), LT(f_2), \ldots, LT(f_s)$. Therefore, the terms of r cannot divide any leading terms.

4. Does the algorithm always terminate?

   This question is where the properties of a monomial ordering on $k[x_1, \ldots, x_n]$ come into play. We claim that in steps 2a and 2b, the multidegree of p decreases. In step 2b, this is obvious because we always subtract the leading term $LT(p)$ from p.

   In step 2a, we subtract the quantity

   $$\frac{LT(p)}{LT(f_i)} f_i$$

   from p. We know that

   $$LT(\frac{LT(p)}{LT(f_i)} f_i) = \frac{LT(p)}{LT(f_i)} LT(f_i) = LT(p)$$

   Since $\frac{LT(p)}{LT(f_i)} f_i$ has the same leading term as p, the mutlidegree of p must be strictly smaller after step 2a is executed.

   Suppose for the sake of contradiction that the division algorithm never terminates. By repeated applications of steps 2a and 2b, we obtain an infinite decreasing sequence of multidegrees; decreasing with respect to the monomial ordering $<$ on $k[x_1, \ldots, x_n]$. However, this contradicts Theorem 2.1.1. Thus, the algorithm must terminate.

5. How does $multideg(q_i f_i)$ compare to $multideg(f)$?

   By step 2a, every term of q_i is of the form $LT(p)/LT(f_i)$ for some $p \in k[x_1, \ldots, x_n]$. We know that in steps 2a and 2b, the multidegree of p always decreases. So, $multideg(p) \leq multideg(f)$ and consequently, if $q_i f_i \neq 0$ then by Theorem 2.1.3,

$$multideg(q_i f_i) \leq multideg(\frac{LT(p)}{LT(f_i)} f_i) = multideg(p) \leq multideg(f).$$

The polynomial division algorithm on $k[x_1, \ldots, x_n]$ constitutes a proof of the following generalisation of Theorem 1.5.1.

Theorem 2.2.1. *Let $>$ be a monomial order on $(\mathbb{Z}_{\geq 0})^n$ and $f_1, \ldots, f_s \in k[x_1, \ldots, x_n]$. If $f \in k[x_1, \ldots, x_n]$ then there exists $q_1, \ldots, q_s, r \in k[x_1, \ldots, x_n]$ such that*

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

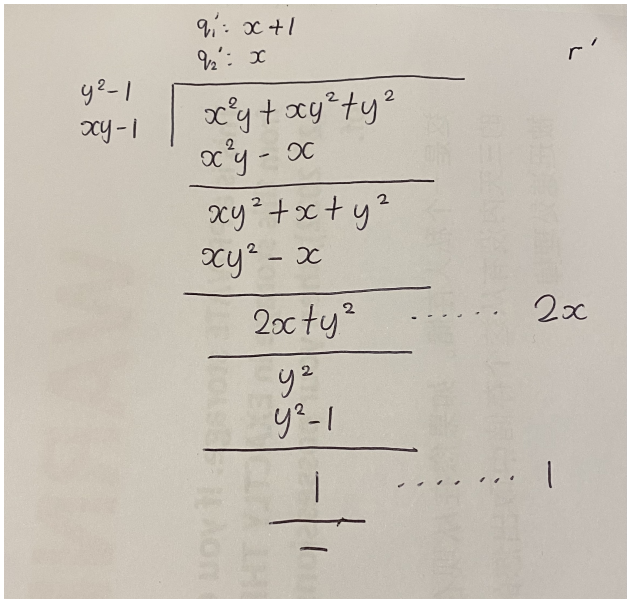
Either $r = 0$ or r is a k-linear combination of monomials, none of which are divisible by any of the leading terms $LT(f_1), LT(f_2), \ldots, LT(f_s)$. Moreover, if $q_i f_i \neq 0$ then $multideg(q_i f_i) \leq multideg(f)$.

Observe that in Theorem 2.2.1, the polynomials $q_1, \ldots, q_s, r \in k[x_1, \ldots, x_n]$ are not unique.

Example 2.2.3. Recall that in the previous example, we imposed the lexicographic order on $k[x, y]$ with $x >_{lex} y$. We used the polynomial division algorithm on $k[x, y]$ to divide $f = x^2y + xy^2 + y^2$ by $f_1 = xy - 1$ and $f_2 = y^2 - 1$. We found that

$$f = q_1 f_1 + q_2 f_2 + r = (x + y)(xy - 1) + 1(y^2 - 1) + (x + y + 1).$$

For clarity, $q_1 = x + y$, $q_2 = 1$ and $r = x + y + 1$. We will repeat the polynomial division algorithm for this particular scenario, but instead we will reverse the order of the two polynomials f_1 and f_2. Let $f_1' = f_2$ and $f_2' = f_1$. The resulting division is displayed below:



So,

$$f = q_1'f_1' + q_2'f_2' + r' = (x+1)(y^2-1) + x(xy-1) + (2x+1).$$

It is clear that $q_1 \neq q_1'$, $q_2 \neq q_2'$ and $r \neq r'$.

To top this section off, let us briefly discuss how the polynomial division algorithm addresses the ideal membership problem for $k[x_1, \ldots, x_n]$. Suppose that we have an ideal $(f_1, \ldots, f_s)$ in $k[x_1, \ldots, x_n]$ and are given a polynomial $f \in k[x_1, \ldots, x_n]$. We want to determine whether $f \in (f_1, \ldots, f_s)$.

Just like in the single variable case, we use Theorem 2.2.1 to find $q_1, \ldots, q_s, r \in k[x_1, \ldots, x_n]$ such that

$$f = q_1f_1 + \cdots + q_sf_s + r.$$

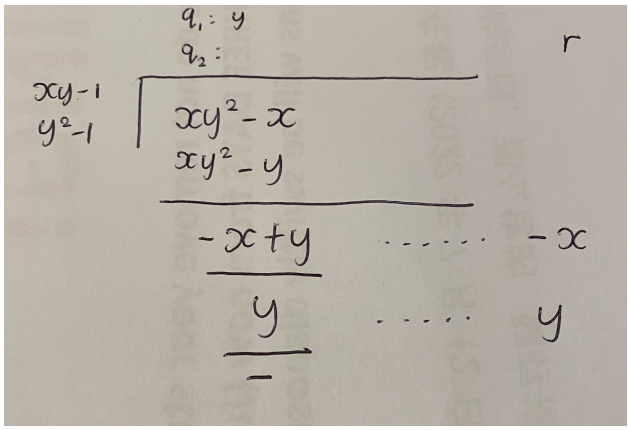
If $r = 0$ then $f \in (f_1, \ldots, f_s)$. Thus, we have shown that if we apply the polynomial division algorithm to divide f by $f_1, \ldots, f_s$ and we obtain a remainder of $r = 0$ then $f \in (f_1, \ldots, f_s)$.

One might ask if the converse statement holds. That is, if $f \in (f_1, \ldots, f_s)$ then is it the case that the remainder $r = 0$ if we use the division algorithm to divide f by $f_1, \ldots, f_s$? In the one variable case, the answer is yes, but for the general case, the answer is **no**, as demonstrated by the following example.

Example 2.2.4. We work in the polynomial ring $k[x, y]$ and impose the lexicographic order on $k[x, y]$. Let $f = xy^2 - x$, $f_1 = xy - 1$ and $f_2 = y^2 - 1$. Notice that $f \in (f_1, f_2)$ because

$$f = x(y^2 - 1) + 0(xy - 1).$$

But, if we apply the polynomial division algorithm to divide f by f_1 and f_2, we obtain



The remainder we obtain is $-x + y \neq 0$.

So, unlike in the single variable case, the polynomial division algorithm in $k[x_1, \ldots, x_n]$ provides a partial and imperfect answer to the ideal membership problem. How do we improve on this? The idea we will pursue lies with Theorem 1.4.3. If we have an ideal $(f_1, \ldots, f_s)$ in $k[x_1, \ldots, x_n]$ then we could potentially find a nicer generating set for $(f_1, \ldots, f_s)$, nicer in the sense that if we divide a polynomial f by the new generators then the remainder r is unique and the condition $r = 0$ is equivalent to being an element of $(f_1, \ldots, f_s)$. Later on, we will see that this idea can be realised by the use of a Grobner basis for the ideal.

2.3 Dickson's lemma

In this section, we will put aside the ideal membership problem for now and turn our attention to the ideal description problem — is every ideal $I \subseteq k[x_1, \ldots, x_n]$ finitely generated? We will take a big leap towards proving this statement in the affirmative by proving a result known as Dickson's lemma.

The approach towards the ideal description problem is to solve the problem first for monomial ideals, which we will define below.

Definition 2.3.1. Let k be a field. An ideal $I \subseteq k[x_1, \ldots, x_n]$ is a **monomial ideal** if there exists a subset $A \subseteq (\mathbb{Z}_{\geq 0})^n$ such that

$$I = (x^\alpha \mid \alpha \in A).$$

That is, there exists a subset $A \subseteq (\mathbb{Z}_{\geq 0})^n$ such that I is generated by monomials x^α for $\alpha \in A$.

Let us prove some properties satisfied by monomial ideals. Firstly, we observe that it is easy to tell whether a given monomial is an element of a monomial ideal.

Lemma 2.3.1. *Let $I = (x^\alpha \mid \alpha \in A)$ be a monomial ideal in $k[x_1, \ldots, x_n]$. Let $\beta \in (\mathbb{Z}_{\geq 0})^n$. Then, $x^\beta \in I$ if and only if there exists $\alpha \in A$ such that x^β is divisible by x^α.*

Proof. Assume that $I = (x^\alpha \mid \alpha \in A)$ is a monomial ideal in $k[x_1, \ldots, x_n]$. Assume that $\beta \in (\mathbb{Z}_{\geq 0})^n$.

To show: (a) If $x^\beta \in I$ then there exists $\alpha \in A$ such that x^β is divisible by x^α.

(b) If there exists $\alpha \in A$ such that x^β is divisible by x^α then $x^\beta \in I$.

(a) Assume that $x^\beta \in I$. Then,

$$x^\beta = \sum_{i=1}^{s} h_i x^{\alpha(i)}$$

where $h_i \in k[x_1, \ldots, x_n]$ and $\alpha(i) \in A$ for $i \in \{1, 2, \ldots, s\}$. The idea is to expand each h_i so that $h_i = \sum_j c_{i,j} x^{\beta(i,j)}$ and

$$x^\beta = \sum_{i=1}^{s} h_i x^{\alpha(i)} = \sum_{i,j} c_{i,j} x^{\beta(i,j)} x^{\alpha(i)}.$$

If we collect the terms of $\sum_{i,j} c_{i,j} x^{\beta(i,j)} x^{\alpha(i)}$ with the same multidegree then every term on the RHS of the above equation is divisible by some $x^{\alpha(i)}$. So, x^β must be divisible by $x^{\alpha(i)}$.

(b) Assume that there exists $\alpha \in A$ such that x^β is divisible by x^α. Then, there exists $f \in k[x_1, \ldots, x_n]$ such that $x^\beta = f \cdot x^\alpha$. So, $x^\beta \in I$. □

The next lemma tells us that whether a given polynomial f lies is a monomial ideal can be determined by looking at the monomials of f.

Lemma 2.3.2. *Let $I \subseteq k[x_1, \ldots, x_n]$ be a monomial ideal and $f \in k[x_1, \ldots, x_n]$. Then, the following are equivalent:*

1. *$f \in I$.*
2. *Every term of f is an element of I.*
3. *f is a k-linear combination of the monomials in I.*

Proof. Assume that I is a monomial ideal and $f \in k[x_1, \ldots, x_n]$. Directly from the definitions of a monomial ideal, the implications $3 \implies 2 \implies 1$ and $2 \implies 3$ are trivial.

To see that $1 \implies 2$, assume that $f \in I$. Since I is a monomial ideal, there exists a subset $A \subseteq (\mathbb{Z}_{\geq 0})^n$ such that $I = (x^\alpha \mid \alpha \in A)$.

So, we can write f as

$$f = \sum_{i=1}^{s} f_i x^{\alpha(i)}$$

where $f_i \in k[x_1, \ldots, x_n]$ and $\alpha(i) \in A$ for $i \in \{1, 2, \ldots, s\}$. Since $x^{\alpha(i)} \in I$, $f_i \in x^{\alpha(i)}$. So, every term of f is contained in I. □

One consequence of the third statement in Lemma 2.3.2 is that a monomial ideal is uniquely determined by its monomials. That is, two monomial ideals are the same if and only if they contain the same monomials.

We have now arrived at Dickson's lemma, which states that every monomial ideal of $k[x_1, \ldots, x_n]$ is finitely generated. Due to the importance of Dickson's lemma, we will state it as a theorem.

Theorem 2.3.3 (Dickson's lemma)**.** *Let $I = (x^\alpha \mid \alpha \in A)$ be a monomial ideal in $k[x_1, \ldots, x_n]$. Then, $I = (x^{\alpha(1)}, \ldots, x^{\alpha(s)})$, where $\alpha(1), \ldots, \alpha(s) \in A$ and $s \in \mathbb{Z}_{>0}$ is finite.*

Proof. We will Dickson's lemma by induction on n, the number of variables in our polynomial ring $k[x_1, \ldots, x_n]$.

For the base case, assume that $n = 1$ so that $I = (x_1^\alpha \mid \alpha \in A)$, where $A \subseteq \mathbb{Z}_{\geq 0}$. Let $\beta = \min\{\alpha \mid \alpha \in A\}$. If $\alpha \in A$ then x_1^β divides x_1^α and

consequently, $I = (x_1^\beta)$. This proves the base case.

For the inductive hypothesis, assume that $n > 1$ and that Dickson's lemma holds for n. Let $I \subseteq k[x_1, \ldots, x_n, y]$ be a monomial ideal. Then,

$$I = (u^\alpha \mid \alpha \in A \subseteq (\mathbb{Z}_{\geq 0})^{n+1})$$

where $u^\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n} y^{\alpha_{n+1}}$. If $A \subseteq (\mathbb{Z}_{\geq 0})^{n+1}$ contains a least element μ — that is, if $\alpha \in A$ then $u^\mu | u^\alpha$ — then $I = (u^\mu)$ and we are done. However, such a least element does not always exist. See [Mur22, Figure 3.1] for an example.

We will call a monomial u^α **minimal** if there does not exist $\beta \in A$ distinct from α such that $u^\beta | u^\alpha$. Define

$$\min(A) = \{\alpha \in A \mid u^\alpha \text{ is minimal}\}.$$

We claim that $I = (u^\alpha \mid \alpha \in A) = (u^\beta \mid \beta \in \min(A))$. Since these are both monomial ideals, it suffices to show that they contain the same monomials.

Assume that $u^\gamma \in (u^\beta \mid \beta \in \min(A))$. Since $\min(A) \subseteq A$, if $\beta \in \min(A)$ then $u^\beta \in I$ and consequently, $u^\gamma \in I$.

Now assume that $u^\gamma \in I$. If $\gamma \in \min(A)$ then $u^\gamma \in (u^\beta \mid \beta \in \min(A))$ by definition. If on the other hand $\gamma \in A - \min(A)$ then there exists $\delta \in A$ such that $\delta \neq \gamma$ and $u^\delta | u^\gamma$. If $\delta \in A - \min(A)$ then there exists $\delta_2 \in A$ such that $\delta \neq \delta_2$ and $u^{\delta_2} | u^\delta$.

Continuing in this fashion, we obtain a decreasing sequence $\gamma > \delta > \delta_2 > \delta_3 > \ldots$ This sequence must eventually terminate. So, there exists $\epsilon \in \min(A)$ such that $u^\epsilon | u^\gamma$ and $u^\gamma \in (u^\beta \mid \beta \in \min(A))$.

We conclude that $I = (u^\beta \mid \beta \in \min(A))$. We now have to show that the set $\min(A) \subseteq A$ is finite.

To show: (a) $\min(A)$ is finite.

(a) Define the ideal

$$J = (x^\beta \mid x^\beta y^m \in I \text{ for some } m \in \mathbb{Z}_{\geq 0}).$$

and the set

$$B = \{\beta \in (\mathbb{Z}_{\geq 0})^n \mid (\beta, m) \in A \text{ for some } m \in \mathbb{Z}_{\geq 0}\}$$

For clarity, J is an ideal of $k[x_1, \ldots, x_n]$. Consider the subset $\min(B)$ of B, defined similarly to $\min(A)$.

To show: (aa) If $\beta \in \min(B)$ then there exists a unique $m \in \mathbb{Z}_{\geq 0}$ such that $\alpha = (\beta, m) \in \min(A)$.

(aa) Assume that $\beta \in \min(B) \subseteq B$. Then, there exists $m \in \mathbb{Z}_{\geq 0}$ such that $x^\beta y^m \in I$ and $(\beta, m) \in A$. Suppose for the sake of contradiction that $(\beta, m) \notin \min(A)$. Then, there exists $(\beta', m') \in A$ such that $x^{\beta'} y^{m'} | x^\beta y^m$. So, $x^{\beta'} | x^\beta$, which contradicts the assumption that $\beta \in \min(B)$. So, $(\beta, m) \in \min(A)$.

To see that m is unique, suppose that $(\beta, m), (\beta, m') \in \min(A)$. Suppose for the sake of contradiction that $m \neq m'$. Let $\ell = \max\{m, m'\}$ and $\ell' = \min\{m, m'\}$. Then, $x^\beta y^{\ell'} | x^\beta y^\ell$, which contradicts the fact that $(\beta, m), (\beta, m') \in \min(A)$. So, $m = m'$ and m is unique.

By part (aa), we obtain an injective map $\Phi_{B,A}$ from $\min(B)$ to $\min(A)$, which sends $\beta \in \min(B)$ to $\alpha = (\beta, m) \in \min(A)$. By the inductive hypothesis, J is generated by a finite subset $B_0 \subseteq B$. Since $\min(B) \subseteq B_0$, $\min(B)$ must also be finite.

Now let $\min(B) = \{\beta_1, \ldots, \beta_r\}$ and $(\beta_1, m_1), \ldots, (\beta_r, m_r) \in \min(A)$. Define $M = \max\{m_1, \ldots, m_r\}$. We claim that if $(\gamma, k) \in \min(A)$ then $k < M$.

To see why this is the case, assume that $(\gamma, k) \in \min(A)$. Then, there exists $\beta \in \min(B)$ such that $\gamma \geq \beta$. Let $\alpha = \Phi_{B,A}(\beta) = (\beta, m) \in \min(A)$. Suppose for the sake of contradiction that $k > m$. Then, $x^\beta y^m | x^\gamma y^k$. This contradicts the assumption that $(\gamma, k) \in \min(A)$. Therefore, $k \leq m$.

The above argument inspires us to define for $k \in \{0, 1, \ldots, M-1\}$ the set

$$B_k = \{\beta \in (\mathbb{Z}_{\geq 0})^n \mid (\beta, k) \in A\}$$

and the ideal of $k[x_1, \ldots, x_n]$

$$J_k = (x^\beta \mid \beta \in B_k).$$

By the same argument as before, $\min(B_k)$ is a finite set. Furthermore, if $\beta \in \min(B_k)$ then there exists a unique $n \in \mathbb{Z}_{\geq 0}$ such that

$\alpha = (\beta, n) \in \min(A)$ and consequently, an injective map $\Phi_{B_k,A}$ from B_k to A, which sends $\beta \in \min(B_k)$ to $(\beta, n) \in \min(A)$. Note that it is possible that $n < k$.

Now assume that $(\gamma, k) \in \min(A)$. Let $\beta \in \min(B_k)$ such that $\beta \leq \gamma$. Suppose for the sake of contradiction that $\beta \neq \gamma$. Then $x^\beta y^k \in I$ and $x^\beta y^k | x^\gamma y^k$. This contradicts the assumption that $(\gamma, k) \in \min(A)$. Hence, $\beta = \gamma \in \min(B_k)$ and $(\gamma, k) \in \text{im } \Phi_{B_k,A}$.

Since every element in $\min(A)$ belongs to the image of one of $\Phi_{B,A}, \Phi_{B_0,A}, \ldots, \Phi_{B_{M-1},A}$ and each of the sets $\min(B), \min(B_0), \ldots, \min(B_{M-1})$ are finite, we deduce that $\min(A)$ is also finite, which completes the proof. □

Before we proceed, we remark that Dickson's lemma has a few consequences in the context of monomial orderings on $(\mathbb{Z}_{\geq 0})^n$.

Theorem 2.3.4. *Let $>$ be a relation on $(\mathbb{Z}_{\geq 0})^n$ satisfying*

1. *$>$ is a total/linear ordering on $(\mathbb{Z}_{\geq 0})^n$ (and not a partial ordering; we want to be able to compare every pair of n-tuples)*
2. *If $\alpha > \beta$ and $\gamma \in (\mathbb{Z}_{\geq 0})^n$ then $\alpha + \gamma > \beta + \gamma$, where addition on $(\mathbb{Z}_{\geq 0})^n$ is defined componentwise.*

Then, $>$ is a well-ordering if and only if for $\alpha \in (\mathbb{Z}_{\geq 0})^n$, $\alpha \geq 0$.

Proof. Assume that $>$ is a relation on $(\mathbb{Z}_{\geq 0})^n$ which satisfies the two properties outlined above.

To show: (a) If $>$ is a well-ordering then for $\alpha \in (\mathbb{Z}_{\geq 0})^n$, $\alpha \geq 0$.

(b) If for $\alpha \in (\mathbb{Z}_{\geq 0})^n$, $\alpha \geq 0$ then $>$ is a well-ordering.

(a) Assume that $>$ is a well-ordering. Then, there exists $\alpha_0 \in (\mathbb{Z}_{\geq 0})^n$ such that if $\beta \in (\mathbb{Z}_{\geq 0})^n$ then $\beta \geq \alpha_0$. Since $>$ satisfies the second property stated above, it suffices to show that $\alpha_0 \geq 0$.

Suppose for the sake of contradiction that $\alpha_0 < 0$. By adding α_0 to both sides (componentwise), we find that $2\alpha_0 < \alpha_0$. However, this contradicts the assumption that α_0 is the minimal element in $(\mathbb{Z}_{\geq 0})^n$.

So, $\alpha_0 \geq 0$ and if $\beta \in (\mathbb{Z}_{\geq 0})^n$ then $\beta \geq 0$.

(b) Assume that if $\alpha \in (\mathbb{Z}_{\geq 0})^n$ then $\alpha \geq 0$. Assume that $A \subseteq (\mathbb{Z}_{\geq 0})^n$ is non-empty.

To show: (ba) The subset A has a minimal element with respect to the relation $>$.

(ba) The ideal $(x^\alpha \mid \alpha \in A)$ is a monomial ideal in $k[x_1, \ldots, x_n]$. By Dickson's lemma (see Theorem 2.3.3),

$$(x^\alpha \mid \alpha \in A) = (x^{\alpha(1)}, x^{\alpha(2)}, \ldots, x^{\alpha(m)})$$

for some $m \in \mathbb{Z}_{>0}$. By relabelling if necessary, we can assume that $\alpha(1) < \alpha(2) < \cdots < \alpha(m)$. We claim that $\alpha(1)$ is the minimal element of A.

Assume that $\alpha \in A$. Then, x^α is an element of the ideal $(x^\alpha \mid \alpha \in A)$ and by Lemma 2.3.1, there exists $i \in \{1, 2, \ldots, m\}$ such that $x^{\alpha(i)}$ divides x^α. So, $\alpha = \alpha(i) + \gamma$ for some $\gamma \in (\mathbb{Z}_{\geq 0})^n$ and because $\gamma \geq 0$,

$$\alpha = \alpha(i) + \gamma \geq \alpha(i) + 0 = \alpha(i) \geq \alpha(1).$$

(b) From part (ba), we found that $\alpha(1)$ is the minimal element of A. Therefore, $>$ is a well-ordering as required. □

The utility of Theorem 2.3.4 is that it makes it much easier to verify whether an ordering on $(\mathbb{Z}_{\geq 0})^n$ is a monomial ordering.

In light of Theorem 2.3.4, one might wonder whether Dickson's lemma itself can be reformulated as a statement about monomial orderings. In the next theorem, we will reformulate Dickson's lemma so that it applies to monomial orderings.

Theorem 2.3.5. *Dickson's lemma is equivalent to the following statement:*

If $n \in \mathbb{Z}_{>0}$ and $A \subseteq (\mathbb{Z}_{\geq 0})^n$ is a non-empty subset then there exists finitely many elements $\alpha(1), \ldots, \alpha(s) \in A$ such that if $\alpha \in A$ then there exists $i \in \{1, 2, \ldots, s\}$ and $\gamma \in (\mathbb{Z}_{\geq 0})^n$ such that $\alpha = \alpha(i) + \gamma$.

Proof. Assume that $n \in \mathbb{Z}_{\geq 0}$ and $A \subseteq (\mathbb{Z}_{\geq 0})^n$ is a non-empty set.

To show: (a) If Theorem 2.3.3 is satisfied then there exists finitely many elements $\alpha(1), \ldots, \alpha(s) \in A$ such that if $\alpha \in A$ then there exists

$i \in \{1, 2, \ldots, s\}$ and $\gamma \in (\mathbb{Z}_{\geq 0})^n$ such that $\alpha = \alpha(i) + \gamma$.

(b) If there exists finitely many elements $\alpha(1), \ldots, \alpha(s) \in A$ such that if $\alpha \in A$ then there exists $i \in \{1, 2, \ldots, s\}$ and $\gamma \in (\mathbb{Z}_{\geq 0})^n$ such that $\alpha = \alpha(i) + \gamma$ then Theorem 2.3.3 holds.

(a) Assume that Dickson's lemma holds (see Theorem 2.3.3). Then, the monomial ideal $I = (x^\alpha \mid \alpha \in A)$ must be finitely generated. So, $I = (x^{\alpha(1)}, \ldots, x^{\alpha(s)})$ for some finite $s \in \mathbb{Z}_{>0}$. Now assume that $\alpha \in A$. Then, $x^\alpha \in I$. By Lemma 2.3.1, there exists $i \in \{1, 2, \ldots, s\}$ such that $x^{\alpha(i)} | x^\alpha$. Consequently, there exists $\gamma \in (\mathbb{Z}_{\geq 0})^n$ such that $\alpha = \alpha(i) + \gamma$.

(b) Assume that there exists $\alpha(1), \ldots, \alpha(s) \in A$ such that if $\alpha \in A$ then there exists $i \in \{1, 2, \ldots, s\}$ and $\gamma \in (\mathbb{Z}_{\geq 0})^n$ such that $\alpha = \alpha(i) + \gamma$. We will show that $I = (x^\alpha \mid \alpha \in A) = (x^{\alpha(1)}, \ldots, x^{\alpha(s)})$. Since $x^{\alpha(i)} \in I$ for $i \in \{1, 2, \ldots, s\}$, $(x^{\alpha(1)}, \ldots, x^{\alpha(s)}) \subseteq I$.

Now assume that $f = \sum_{j=1}^{p} a_j x^{\beta(j)} \in I$, where $\beta(p) \in (\mathbb{Z}_{\geq 0})^n$. By assumption, for $j \in \{1, 2, \ldots, p\}$, there exists $i_j \in \{1, 2, \ldots, s\}$ and $\gamma_j \in (\mathbb{Z}_{\geq 0})^n$ such that $\beta(j) = \alpha(i_j) + \gamma_j$. So,

$$f = \sum_{j=1}^{p} a_j x^{\beta(j)} = \sum_{j=1}^{p} a_j x^{\gamma_j} x^{\alpha(i_j)}.$$

Since $x^{\alpha(i_j)} \in (x^\alpha \mid \alpha \in A)$, the terms of f

$$a_j x^{\gamma_j} x^{\alpha(i_j)} \in (x^\alpha \mid \alpha \in A).$$

By Lemma 2.3.2, $f \in (x^\alpha \mid \alpha \in A)$. Therefore, $I \subseteq (x^{\alpha(1)}, \ldots, x^{\alpha(s)})$ and the monomial ideal I must be finitely generated, which in turn gives Dickson's lemma in Theorem 2.3.3. □

Finally, we will single out a particular basis for a monomial ideal.

Theorem 2.3.6. *Let k be a field and $I \subseteq k[x_1, \ldots, x_n]$ be a monomial ideal. Then, there exists a unique finite basis $x^{\alpha(1)}, \ldots, x^{\alpha(s)}$ of I such that $I = (x^{\alpha(1)}, \ldots, x^{\alpha(s)})$ and if $i, j \in \{1, 2, \ldots, s\}$ are distinct then $x^{\alpha(i)}$ does not divide $x^{\alpha(j)}$.*

Proof. Assume that k is a field and $I \subseteq k[x_1, \ldots, x_n]$ is a monomial ideal. By Dickson's lemma (see Theorem 2.3.3), I is finitely generated so that $I = (x^{\alpha(1)}, \ldots, x^{\alpha(t)})$ for some $t \in \mathbb{Z}_{>0}$.

If $x^{\alpha(p)}$ divides $x^{\alpha(q)}$ for some $p, q \in \{1, 2, \ldots, t\}$ then we can discard $x^{\alpha(q)}$ from the basis and the resulting set still generates I. By doing this, we obtain a basis $x^{\alpha(1)}, \ldots, x^{\alpha(s)}$ such that if $i, j \in \{1, 2, \ldots, s\}$ are distinct then $x^{\alpha(i)}$ does not divide $x^{\alpha(j)}$.

To see that the basis $x^{\alpha(1)}, \ldots, x^{\alpha(s)}$ is unique, assume that $x^{\beta(1)}, \ldots x^{\beta(u)}$ is another basis for I such that if $i, j \in \{1, 2, \ldots, u\}$ are distinct then $x^{\beta(i)}$ does not divide $x^{\beta(j)}$. By Lemma 2.3.1, there exists $j_1 \in \{1, 2, \ldots, t\}$ such that $x^{\beta(j_1)} | x^{\alpha(1)}$. But since $x^{\beta(j_1)} \in I$, there exists $\ell_1 \in \{1, 2, \ldots, s\}$ such that $x^{\alpha(\ell_1)} | x^{\beta(j_1)}$. Consequently, $x^{\alpha(\ell_1)} | x^{\alpha(1)}$, but due to the construction of the basis for I, $\ell_1 = 1$ and $x^{\beta(j_1)} = x^{\alpha(1)}$.

By repeating this argument, we find that if $i \in \{1, 2, \ldots, s\}$ then there exists $j_i \in \{1, 2, \ldots, t\}$ such that $x^{\beta(j_i)} = x^{\alpha(i)}$. So,
$\{\alpha(1), \ldots, \alpha(s)\} \subseteq \{\beta(1), \ldots, \beta(u)\}$.

By interchanging the roles of the two bases for I and repeating the above argument, we also have $\{\beta(1), \ldots, \beta(u)\} \subseteq \{\alpha(1), \ldots, \alpha(s)\}$. Hence, $\{\beta(1), \ldots, \beta(u)\} = \{\alpha(1), \ldots, \alpha(s)\}$ and so, the basis $x^{\alpha(1)}, \ldots, x^{\alpha(s)}$ of I is unique. □

The unique basis of a monomial ideal in Theorem 2.3.6 is called a **minimal basis**.

2.4 The Hilbert basis theorem and Gröbner bases

Recall that Dickson's lemma (see Theorem 2.3.3) is a special case of the Hilbert basis theorem, which states that $k[x_1, \ldots, x_n]$ is a Noetherian ring — every ideal of $k[x_1, \ldots, x_n]$ is finitely generated. We will begin this section by using Dickson's lemma to prove the Hilbert basis theorem. The idea is to use the fact that with a monomial ordering on $k[x_1, \ldots, x_n]$, every polynomial $f \in k[x_1, \ldots, x_n]$ has a unique leading term $LT(f)$.

Definition 2.4.1. Let $I \subseteq k[x_1, \ldots, x_n]$ be a non-zero ideal and fix a monomial ordering on $k[x_1, \ldots, x_n]$. Define

$$LT(I) = \{cx^\alpha \mid \text{There exists } f \in I - \{0\} \text{ such that } LT(f) = cx^\alpha\}$$

We denote by $(LT(I))$ the ideal generated by the elements of $LT(I)$.

It is tempting to claim that if $I = (f_1, \ldots, f_s)$ then $(LT(I)) = (LT(f_1), \ldots, LT(f_s))$. Since $LT(f_i) \in LT(I)$ where $i \in \{1, 2, \ldots, s\}$, $LT(f_i) \in LT(I) \subseteq (LT(I))$. So, $(LT(f_1), \ldots, LT(f_s)) \subseteq (LT(I))$. But, the ideal $(LT(I))$ can be strictly larger than $(LT(f_1), \ldots, LT(f_s))$, as the following example demonstrates.

Example 2.4.2. We work in the polynomial ring $k[x, y]$. Let $I = (x^3 - 2xy, x^2y - 2y^2 + x)$. We impose the grlex ordering on $k[x, y]$, with $x >_{grlex} y$. Then,

$$x(x^2y - 2y^2 + x) - y(x^3 - 2xy) = x^2 \in I.$$

So, $x^2 = LT(x^2) \in (LT(I))$. However, x^2 is not divisible by either $LT(x^3 - 2xy) = x^3$ or $LT(x^2y - 2y^2 + x) = x^2y$. By Lemma 2.3.1, $x^2 \notin (LT(x^3 - 2xy), LT(x^2y - 2y^2 + x))$. Therefore,

$$(LT(I)) \neq (LT(x^3 - 2xy), LT(x^2y - 2y^2 + x)).$$

The point of defining the ideal $(LT(I))$ from an ideal $I \subseteq k[x_1, \ldots, x_n]$ is that $(LT(I))$ is a monomial ideal, providing the perfect setup for Theorem 2.3.3.

Lemma 2.4.1. *Let $I \subseteq k[x_1, \ldots, x_n]$ be a non-zero ideal. Then, $(LT(I))$ is a monomial ideal and there exists $g_1, \ldots, g_t \in I$ such that $(LT(I)) = (LT(g_1), \ldots, LT(g_t))$*

Proof. Assume that $I \subseteq k[x_1, \ldots, x_n]$ is a non-zero ideal.

To show: (a) The ideal $(LT(I))$ is a monomial ideal.

(b) There exists $g_1, \ldots, g_t \in I$ such that $(LT(I)) = (LT(g_1), \ldots, LT(g_t))$.

(a) For arbitrary $g \in I - \{0\}$, the leading monomials $LM(g)$ generate the monomial ideal $(LM(g) \mid g \in I - \{0\})$. Since $LM(g)$ and $LT(g)$ differ by multiplication by a constant, we find that

$$(LM(g) \mid g \in I - \{0\}) = (LT(g) \mid g \in I - \{0\}) = (LT(I)).$$

Hence, $(LT(I))$ is a monomial ideal.

(b) From part (a), $(LT(I))$ is a monomial ideal generated by the monomials $LM(g)$ for $g \in I - \{0\}$. By Dickson's lemma (see Theorem 2.3.3), there exists $g_1, \ldots, g_t \in I - \{0\}$ such that $(LT(I)) = (LM(g_1), \ldots, LM(g_t))$.

Again, since $LM(g)$ and $LT(g)$ differ by multiplication by a constant, we find that

$$(LT(I)) = (LM(g_1), \ldots, LM(g_t)) = (LT(g_1), \ldots, LT(g_t)).$$

□

Lemma 2.4.1 and Theorem 2.2.1 are the key ingredients to the proof of the Hilbert basis theorem, which we finally prove below.

Theorem 2.4.2 (Hilbert basis theorem)**.** *Let $n \in \mathbb{Z}_{>0}$ and $I \subseteq k[x_1, \ldots, x_n]$ be an ideal of $k[x_1, \ldots, x_n]$. Then, I is finitely generated.*

Proof. Assume that $n \in \mathbb{Z}_{>0}$ and $I \subseteq k[x_1, \ldots, x_n]$ is an ideal of $k[x_1, \ldots, x_n]$. Suppose that we have a monomial ordering $>$ on $k[x_1, \ldots, x_n]$. If I is the zero ideal then I is generated by the single element $0 \in k$.

So, assume that $I \neq 0$. Since we have a monomial ordering on $k[x_1, \ldots, x_n]$, we can construct an ideal of leading terms $(LT(I))$. Using Lemma 2.4.1, there exists $g_1, \ldots, g_t \in I$ such that $(LT(I)) = (LT(g_1), \ldots, LT(g_t))$. Notice that we have the containment of ideal $(g_1, \ldots, g_t) \subseteq I$ by construction.

To show: (a) $I \subseteq (g_1, \ldots, g_t)$.

(a) Assume that $f \in I$. By Theorem 2.2.1, there exists $q_1, \ldots, q_t, r \in k[x_1, \ldots, x_n]$ such that

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

To show: (aa) $r = 0$.

(aa) Suppose for the sake of contradiction that $r \neq 0$. By Theorem 2.2.1, r is a k-linear combination of monomials, none of which are divisible by any of the leading terms $LT(g_1), \ldots, LT(g_t)$. Observe that

$$r = f - q_1 g_1 - \cdots - q_t g_t \in I.$$

This means that $LT(r) \in (LT(I)) = (LT(g_1), \ldots, LT(g_t))$. By Lemma 2.3.1, there exists $j \in \{1, 2, \ldots, t\}$ such that $LT(g_j) | LT(r)$. But this contradicts the assumption that r is a k-linear combination of monomials, none of which are divisible by any of the leading terms $LT(g_1), \ldots, LT(g_t)$. So, $r = 0$.

(a) Since $r = 0$, $f \in (g_1, \ldots, g_t)$ and $I \subseteq (g_1, \ldots, g_t)$.

Consequently, $I = (g_1, \ldots, g_t)$ and the ideal I is finitely generated. □

With Theorem 2.4.2, we have completely solved the ideal description problem for $k[x_1, \ldots, x_n]$. The proof of Theorem 2.4.2 hinged on the fact that there exists $g_1, \ldots, g_t \in I$ such that $(LT(I)) = (LT(g_1), \ldots, LT(g_t))$. The generating set $\{g_1, \ldots, g_t\}$ is clearly special and deserves its own definition.

Definition 2.4.3. Let $k[x_1, \ldots, x_n]$ be equipped with a monomial order. Let $I \subseteq k[x_1, \ldots, x_n]$ be a non-zero ideal. We say that a finite subset $G = \{g_1, \ldots, g_t\}$ of I is a **Gröbner basis** if

$$(LT(I)) = (LT(g_1), \ldots, LT(g_t)).$$

Note that we define the empty set $\emptyset$ to be the Gröbner basis of the zero ideal 0.

By Lemma 2.3.1, a finite subset $G = \{g_1, \ldots, g_t\}$ of I is a Gröbner basis if and only if the leading term of any non-zero element of I is divisible by some $LT(g_i)$.

A consequence of Theorem 2.4.2 is that every ideal $I \subseteq k[x_1, \ldots, x_n]$ has a Gröbner basis and any Gröbner basis is a basis (generating set) for I. What is not clear at the moment is *how* to construct a Gröbner basis. This will be discussed in the next section where we investigate *Buchberger's algorithm.*

As mentioned previously, Theorem 2.4.2 can be reformulated to the statement that $k[x_1, \ldots, x_n]$ is a Noetherian ring. Let us make this translation precise.

Definition 2.4.4. Let R be a commutative ring. We say that R is **Noetherian** if for any ascending chain of ideals in R

$$I_1 \subseteq I_2 \subseteq I_3 \subseteq \cdots \subseteq I_n \subseteq \ldots$$

there exists $k \in \mathbb{Z}_{>0}$ such that if $\ell \in \mathbb{Z}_{\geq k}$ then $I_\ell = I_k$.

Theorem 2.4.3. *Let R be a ring. Then, the following are equivalent:*

(a) R is a Noetherian ring.

(b) Every ideal of R is finitely generated.

Proof. First, assume that R is a Noetherian ring so that R satisfies the ascending chain condition. We will prove this by contrapositive. So, suppose that there exists an ideal J of R which is not finitely generated. Pick $r_1 \in J$. Since J is not finitely generated, $r_1R \neq J$ as ideals in R and subsequently, there exists $r_2 \in J \backslash r_1R$. Since J is not finitely generated, $r_1R + r_2R \neq J$ and so, there exists $r_3 \in J \backslash (r_1R + r_2R)$.

By continuing this argument, we create an infinite ascending chain of ideals

$$r_1R \subset r_1R + r_2R \subset \cdots \subset \sum_{i=1}^{n} r_nR \subset \ldots$$

where each inclusion is strict. This demonstrates that R is not Noetherian.

Conversely, assume that every ideal of R is finitely generated. Suppose that we have the following ascending chain of ideals

$$I_1 \subseteq I_2 \subseteq \cdots \subseteq I_k \subseteq \ldots \tag{2.1}$$

Observe that the set $\bigcup_{i\geq 1} I_i$ is an ideal of R. This is because if $j_1, j_2 \in \bigcup_{i\geq 1} I_i$ then there exists $k \in \mathbb{Z}_{>0}$ such that $j_1, j_2 \in I_k$, due to equation (2.1). Hence, by assumption, $\bigcup_{i\geq 1} I_i$ must be finitely generated. Suppose that $r_1, r_2, \ldots, r_m \in R$ such that

$$r_1R + r_2R + \cdots + r_mR = \bigcup_{i\geq 1} I_i$$

as ideals. Since $r_1, r_2, \ldots, r_m \in \bigcup_{i\geq 1} I_i$, there exists $t \in \mathbb{Z}_{>0}$ as a result of equation (2.1) such that $r_1, r_2, \ldots, r_m \in I_t$. Hence, $\bigcup_{i\geq 1} I_i \subseteq I_t$. But by definition, $I_t \subseteq \bigcup_{i\geq 1} I_i$ and consequently, $I_t = \bigcup_{i\geq 1} I_i$. So, for all $s \geq t$,

$$I_t \subseteq I_s \subseteq \bigcup_{i\geq 1} I_i = I_t.$$

Therefore, $I_t = I_s$ and the ascending chain in equation (2.1) terminates. □

By Theorem 2.4.3 and the Hilbert basis theorem in Theorem 2.4.2, $k[x_1, \ldots, x_n]$ is a Noetherian ring as required. We will now give an example of a Gröbner basis.

Example 2.4.5. Here is a well-known example of a Gröbner basis. Let $A = (a_{ij}) \in M_{m\times n}(\mathbb{R})$ be a matrix in row echelon form and $J \subseteq \mathbb{R}[x_1, \ldots, x_n]$ be an ideal generated by the polynomials

$$\sum_{j=1}^{n} a_{ij}x_j$$

for $i \in \{1, 2, \ldots, m\}$. We claim that the generators form a Gröbner basis for J, with respect to a specific lexicographic order on $\mathbb{R}[x_1, \ldots, x_n]$.

Supose that the j^{th} column of A corresponds to the variable x_j for $j \in \{1, 2, \ldots, n\}$. On the polynomial ring $\mathbb{R}[x_1, \ldots, x_n]$, we impose the lexicographic order with $x_1 >_{lex} x_2 >_{lex} \cdots >_{lex} x_n$. The leading term of a generator $\sum_{j=1}^{n} a_{ij}x_j$ is therefore,

$$LT(\sum_{j=1}^{n} a_{ij}x_j) = \min_{\ell \in \{1,2,\ldots,n\}} \{x_\ell \mid a_{i\ell} \neq 0\}.$$

Now suppose that $f \in J - \{0\}$. Then, there exists polynomials $p_1, \ldots, p_n \in \mathbb{R}[x_1, \ldots, x_n]$ such that

$$f = p_1\big(\sum_{j=1}^{n} a_{1j}x_j\big) + \cdots + p_n\big(\sum_{j=1}^{n} a_{nj}x_j\big).$$

Since A is in row echelon form, if $k, \ell \in \{1, 2, \ldots, n\}$ are distinct then $LT(\sum_{j=1}^{n} a_{kj}x_j) > LT(\sum_{j=1}^{n} a_{\ell j}x_j)$. So, the leading term of f is

$$LT(f) = LT(p_1)LT(\sum_{j=1}^{n} a_{1j}x_j)$$

which is divisible by $LT(\sum_{j=1}^{n} a_{1j}x_j)$. Hence, the generators for J

$$\{\sum_{j=1}^{n} a_{\ell j}x_j \mid \ell \in \{1, 2, \ldots, n\}\}$$

form a Gröbner basis for J.

Before we investigate the properties of Gröbner bases further, we will discuss a second, geometric consequence of Theorem 2.4.2. We are able to talk about the affine variety defined by an ideal $I \subseteq k[x_1, \ldots, x_n]$.

Definition 2.4.6. Let $n \in \mathbb{Z}_{>0}$ and $I \subseteq k[x_1, \ldots, x_n]$ be an ideal. Define $V(I)$ to be the set

$$V(I) = \{(a_1, \ldots, a_n) \in k^n \mid \text{If } f \in I \text{ then } f(a_1, \ldots, a_n) = 0\}.$$

Generally a non-zero ideal $I \subseteq k[x_1, \ldots, x_n]$ contains infinitely many polynomials. However, Theorem 2.4.2 tells us that the set $V(I)$ can still be defined by a finite set of polynomial equations because I is finitely generated.

Lemma 2.4.4. *Let $n \in \mathbb{Z}_{>0}$ and $I \subseteq k[x_1, \ldots, x_n]$ be an ideal. Then the set $V(I)$ is an affine variety. In particular, if $I = (f_1, \ldots, f_s)$ then $V(I) = V(f_1, \ldots, f_s)$.*

Proof. Assume that $n \in \mathbb{Z}_{>0}$ and $I \subseteq k[x_1, \ldots, x_n]$ is an ideal. Assume that

$$V(I) = \{(a_1, \ldots, a_n) \in k^n \mid \text{If } f \in I \text{ then } f(a_1, \ldots, a_n) = 0\}.$$

By Theorem 2.4.2, there exists a finite set of polynomials $\{f_1, \ldots, f_s\}$ in $k[x_1, \ldots, x_n]$ such that $I = (f_1, \ldots, f_s)$. We claim that $V(I) = V(f_1, \ldots, f_s)$.

To show: (a) $V(f_1, \ldots, f_s) \subseteq V(I)$.

(b) $V(I) \subseteq V(f_1, \ldots, f_s)$.

(a) Assume that $(b_1, \ldots, b_n) \in V(f_1, \ldots, f_s)$ so that if $i \in \{1, 2, \ldots, s\}$ then $f_i(b_1, \ldots, b_n) = 0$. Now, assume that $g \in I$. Then, there exists polynomials $p_1, \ldots, p_s \in k[x_1, \ldots, x_n]$ such that $g = p_1 f_1 + \cdots + p_s f_s$. Note that by assumption,

$$g(b_1, \ldots, b_n) = \sum_{i=1}^{n} p_i(b_1, \ldots, b_n)(0) = 0.$$

Therefore, $(b_1, \ldots, b_n) \in V(I)$ and $V(f_1, \ldots, f_s) \subseteq V(I)$.

(b) Assume that $(c_1, \ldots, c_n) \in V(I)$. If $h \in I$ then $h(c_1, \ldots, c_n) = 0$. Since $f_1, \ldots, f_s \in I$, $f_i(c_1, \ldots, c_n) = 0$ for $i \in \{1, 2, \ldots, s\}$. Therefore, $(c_1, \ldots, c_n) \in V(f_1, \ldots, f_s)$ and $V(I) \subseteq V(f_1, \ldots, f_s)$.

Combining parts (a) and (b) together, we find that $V(I) = V(f_1, \ldots, f_s)$ as required. □

2.5 Properties of Gröbner bases

In order to reinforce why Gröbner bases are important and useful, we will prove some important properties Gröbner bases satisfy. Firstly, if we divide a polynomial $f \in k[x_1, \ldots, x_n]$ by a set of polynomials $\{f_1, \ldots, f_s\}$ which form a Gröbner basis then the remainder we obtain is unique.

Theorem 2.5.1. *Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal and $G = \{g_1, \ldots, g_t\}$ be a Gröbner basis for I. If $f \in k[x_1, \ldots, x_n]$ then there exists a* ***unique*** *$r \in k[x_1, \ldots, x_n]$ such that no term of r is divisible by any of $LT(g_1), \ldots, LT(g_t)$ and there exists $g \in I$ such that $f = g + r$.*

Proof. Assume that $I \subseteq k[x_1, \ldots, x_n]$ is an ideal and $G = \{g_1, \ldots, g_t\}$ is a Gröbner basis for I. Assume that $f \in k[x_1, \ldots, x_n]$. By Theorem 2.2.1, there exists polynomials $q_1, \ldots, q_t, r \in k[x_1, \ldots, x_n]$ such that

$$f = q_1 g_1 + \cdots + q_t g_t + r$$

where either $r = 0$ or no term of r divides $LT(g_1), \ldots, LT(g_t)$. Note that $q_1 g_1 + \cdots + q_t g_t \in I$. Hence, it remains to show that $r \in k[x_1, \ldots, x_n]$ is unique.

To show: (a) $r \in k[x_1, \ldots, x_n]$ is unique.

(a) Assume that $f = g + r = g' + r'$, where $g, g' \in I$ and $r, r' \in k[x_1, \ldots, x_n]$ are polynomial with no terms dividing $LT(g_1), \ldots, LT(g_t)$. Suppose for the sake of contradiction that $r \neq r'$. Then, $r - r' = g - g' \in I$ and the leading term

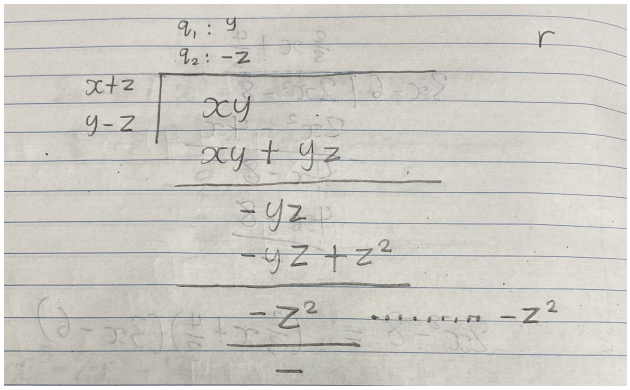
$$LT(r - r') \in (LT(I)) = (LT(g_1), \ldots, LT(g_t)).$$

By Lemma 2.3.1, there exists $i \in \{1, 2, \ldots, t\}$ such that $LT(g_i) | LT(r - r')$. This contradicts the assumption that the terms of r and r' are divisible by none of the $LT(g_1), \ldots, LT(g_t)$. So, $r = r'$ and the remainder r must be unique. □

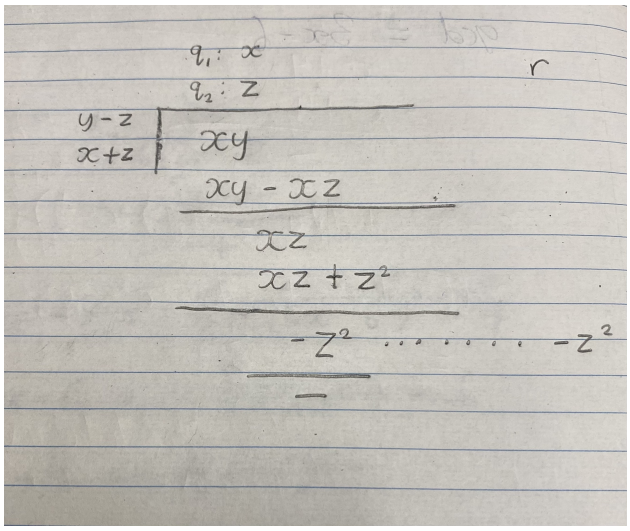
Note that as a consequence of Theorem 2.5.1, the remainder $r \in k[x_1, \ldots, x_n]$ we obtain after dividing a polynomial f by a Gröbner basis $\{g_1, \ldots, g_t\}$ remains unique even if we switch the order of the generators $g_1, \ldots, g_t$.

Note that the uniqueness in Theorem 2.5.1 only applies to the remainder. The quotient polynomials $q_1, \ldots, q_t$ in the division algorithm can still change if we switch the order of the generators $g_1, \ldots, g_t$. Let us peruse an example of this observation.

Example 2.5.1. We will work in the polynomial ring $k[x, y, z]$. We impose the lexicographic order on $k[x, y, z]$, declaring that $x >_{lex} y >_{lex} z$. The set $G = \{x + z, y - z\}$ forms a Gröbner basis. We will divide xy by $x + z$ and $y - z$ in two different ways. Firstly, we have



Next, we switch the order of the generators for G and then divide xy again. We obtain



Observe that the remainder stays the same, in accordance with Theorem 2.5.1. However, the polynomials q_1, q_2 in both computations differ.

As a consequence of Theorem 2.5.1, we have

Theorem 2.5.2. *Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal and $G = \{g_1, \ldots, g_t\}$ be a Gröbner basis for I. Let $f \in k[x_1, \ldots, x_n]$. Then, $f \in I$ if and only if when we divide f by $g_1, \ldots, g_t$ using the polynomial division algorithm, the remainder r is zero.*

Proof. Assume that $I \subseteq k[x_1, \ldots, x_n]$ is an ideal and $G = \{g_1, \ldots, g_t\}$ is a Gröbner basis for I.

To show: (a) If $f \in I$ then the remainder is zero when f is divided by $g_1, \ldots, g_t$.

(b) If the remainder is zero when f is divided by $g_1, \ldots, g_t$ then $f \in I$.

(a) Assume that $f \in I$. Then, there exists polynomials $q_1, \ldots, q_t \in k[x_1, \ldots, x_n]$ such that $f = q_1 g_1 + \cdots + q_t g_t$. Note that $f \in I$ satisfies $f = f + 0$. By Theorem 2.5.1, we find that if we divide f by $g_1, \ldots, g_t$ then 0 must be the remainder.

(b) Assume that if we divide f by $g_1, \ldots, g_t$ then the remainder is zero. Then, there exists $q_1, \ldots, q_t \in k[x_1, \ldots, x_n]$ such that $f = q_1 g_1 + \cdots + q_t g_t$. Consequently, $f \in I$. □

Theorem 2.5.2 actually provides us with a solution to the ideal membership problem in $k[x_1, \ldots, x_n]$, but this is under the assumption that given an ideal $I \subseteq k[x_1, \ldots, x_n]$, we can construct a Gröbner basis for I using an algorithm. As we have mentioned before, this will be dealt with in a separate section.

Definition 2.5.2. Let $f \in k[x_1, \ldots, x_n]$ and $F = (f_1, \ldots, f_s)$ be an ordered s-tuple consisting of polynomials $f_i \in k[x_1, \ldots, x_n]$. Define $\overline{f}^F$ to be the remainder when we divide f by the polynomials in F (beginning from left to right).

If F happens to be a Gröbner basis then we can regard F as a set since changing the order of the elements in a Gröbner basis does not affect the remainder by Theorem 2.5.1.

Suppose that $I = (f_1, \ldots, f_s) \subseteq k[x_1, \ldots, x_n]$. The main obstruction to $\{f_1, \ldots, f_s\}$ being a Gröbner basis is that there could be polynomial combinations of $f_1, \ldots, f_s$ whose leading terms are not in $(LT(I))$. One way this can occur is if the leading terms in a suitable combination cancel, leaving the smaller terms behind (with respect to a monomial ordering on $k[x_1, \ldots, x_n]$). The purpose of S-polynomials is to study this type of cancellation.

Definition 2.5.3. Let $f, g \in k[x_1, \ldots, x_n]$ be non-zero polynomials. Suppose that $multideg(f) = \alpha$ and $multideg(g) = \beta$. Let $\gamma = (\gamma_1, \ldots, \gamma_n) \in (\mathbb{Z}_{\geq 0})^n$ where for $i \in \{1, 2, \ldots, n\}$,

$$\gamma_i = \max(\alpha_i, \beta_i).$$

The monomial x^γ is called the **least common multiple** of $LM(f)$ and $LM(g)$. It is denoted by

$$x^\gamma = \text{lcm}(LM(f), LM(g)).$$

The **S-polynomial** of f and g is the combination

$$S(f,g) = \frac{x^\gamma}{LT(f)}f - \frac{x^\gamma}{LT(g)}g.$$

As explained in [CLO15], a S-polynomial is designed to produce cancellation of the leading terms. The next important lemma demonstrates that every cancellation of leading terms arising from polynomials of the same multidegree originates from the cancellation which is artificially produced by S-polynomials.

Lemma 2.5.3. *Let $s \in \mathbb{Z}_{>0}$ and suppose that we have a monomial ordering on $k[x_1, \ldots, x_n]$. For $i \in \{1, 2, \ldots, s\}$, let $p_i \in k[x_1, \ldots, x_n]$ with $multideg(p_i) = \delta$. If $multideg(\sum_{i=1}^s p_i) < \delta$ then $\sum_{i=1}^s p_i$ is a k-linear combination of the S-polynomials $S(p_j, p_l)$ for $j, l \in \{1, 2, \ldots, s\}$. Also, $multideg(S(p_j, p_l)) < \delta$.*

Proof. Assume that for $i \in \{1, 2, \ldots, s\}$, we have $p_i \in k[x_1, \ldots, x_n]$ where $multideg(p_i) = \delta$. Assume that $multideg(\sum_{i=1}^s p_i) < \delta$.

For $i \in \{1, 2, \ldots, s\}$, let $d_i = LC(p_i)$ so that $LT(p_i) = d_i x^\delta$. Since $multideg(\sum_{i=1}^s p_i) < \delta$, $\sum_{i=1}^s d_i = 0$. By definition, the S-polynomial of p_i and p_j is

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

Using the S-polynomials, we compute directly that

$$\begin{aligned}\sum_{i=1}^{s-1} d_i S(p_i, p_s) &= \sum_{i=1}^{s-1} d_i\big(\frac{1}{d_i}p_i - \frac{1}{d_s}p_s\big)\\ &= \sum_{i=1}^{s-1}(p_i - \frac{d_i}{d_s}p_s)\\ &= (p_1 + \cdots + p_{s-1}) - \frac{d_1 + \cdots + d_{s-1}}{d_s}p_s\\ &= (p_1 + \cdots + p_{s-1}) + \frac{d_s}{d_s}p_s\\ &= \sum_{i=1}^{s} p_i.\end{aligned}$$

Hence $\sum_{i=1}^{s} p_i$ can be expressed as a k-linear combination of S-polynomials $S(p_i, p_s)$. It is obvious from the expression that $S(p_i, p_s)$ has multidegree less than δ, which completes the proof. □

Buchberger's criterion uses S-polynomials to tell us when a basis is a Gröbner basis.

Theorem 2.5.4 (Buchberger's criterion)**.** *Let $I \subseteq k[x_1, \ldots, x_n]$. Let $G = \{g_1, \ldots, g_t\}$ be a basis for I so that $I = (g_1, \ldots, g_t)$. Then, G is a Gröbner basis of I if and only if for distinct pairs $i, j \in \{1, 2, \ldots, t\}$, the remainder on division of $S(g_i, g_j)$ by G (listed in some order) is zero.*

Proof. Assume that $I \subseteq k[x_1, \ldots, x_n]$ is an ideal. Assume that $G = \{g_1, \ldots, g_t\}$ is a basis for I so that $I = (g_1, \ldots, g_t)$.

To show: (a) If G is a Gröbner basis for I then for distinct pairs $i, j \in \{1, 2, \ldots, t\}$, the remainder on division of $S(g_i, g_j)$ by G (listed in some order) is zero.

(b) If for distinct pairs $i, j \in \{1, 2, \ldots, t\}$, the remainder on division of $S(g_i, g_j)$ by G (listed in some order) is zero then G is a Gröbner basis for I.

(a) Assume that G is a Gröbner basis for I. Assume that $i, j \in \{1, 2, \ldots, t\}$ are distinct. Then, the S-polynomial $S(g_i, g_j) \in I$. By Theorem 2.5.2, we find that if we divide $S(g_i, g_j)$ by $g_1, \ldots, g_t$ (in any order) then the remainder is zero.

(b) Assume that $i, j \in \{1, 2, \ldots, t\}$ are distinct and that if $S(g_i, g_j)$ is divided by $g_1, \ldots, g_t$ in some order then the remainder is zero. Assume that $f \in I - \{0\}$

To show: (ba) $LT(f) \in (LT(g_1), \ldots, LT(g_t))$.

(ba) Since $f \in I - \{0\}$, there exists polynomials $h_i \in k[x_1, \ldots, x_n]$ (with not all of them zero) such that

$$f = h_1 g_1 + h_2 g_2 + \cdots + h_t g_t.$$

We know that the multidegree satisfies

$$multideg(f) \leq \max\{multideg(h_i g_i) \mid h_i g_i \neq 0\}.$$

The idea of the proof is that there are multiple representations of f as a sum $\sum_{i=1}^{t} h_i g_i$. We pick a representation such that the quantity

$$\delta = \max\{multideg(h_i g_i) \mid h_i g_i \neq 0\}$$

is minimal. Since any monomial ordering on $k[x_1, \ldots, x_n]$ has the well-ordering property, the minimum δ does exist. Of course, we still have $multideg(f) \leq \delta$. We now divide the proof into two cases:

Case 1: $multideg(f) = \delta$.

If $multideg(f) = \delta$ then $multideg(f) = multideg(h_i g_i)$ for some $i \in \{1, 2, \ldots, t\}$. So, $LT(g_i) | LT(f)$. Consequently, $LT(f) \in (LT(g_1), \ldots, LT(g_t))$. So, $(LT(I)) = (LT(g_1), \ldots, LT(g_t))$ and consequently, G is a Gröbner basis for I.

Case 2: $multideg(f) < \delta$

Suppose for the sake of contradiction that $multideg(f) < \delta$. Write $f = \sum_{i=1}^{t} h_i g_i$, where the quantity δ is minimal. The idea is to break up this sum so that

$$f = \sum_{multideg(h_i g_i) = \delta} h_i g_i \; + \sum_{multideg(h_i g_i) < \delta} h_i g_i.$$

By breaking up the first sum even further, we have

$$f = \sum_{multideg(h_i g_i)=\delta} LT(h_i)g_i + \sum_{multideg(h_i g_i)=\delta} (h_i - LT(h_i))g_i + \sum_{multideg(h_i g_i)<\delta} h_i g_i.$$

Note that the monomials appearing in the second and third sums above all have multidegree less than δ. Since $multideg(f) < \delta$, the first sum must also have multidegree less than δ.

We recognise that the first sum

$$\sum_{multideg(h_i g_i)=\delta} LT(h_i)g_i$$

is a sum of polynomials $p_i = LT(h_i)g_i$ with multidegree δ and the multidegree of the sum is less than δ. By Lemma 2.5.3, the sum is a k-linear combination of the S-polynomials $S(p_i, p_j)$. By a direct computation, we have

$$\begin{aligned}
S(p_i, p_j) &= S(LT(h_i)g_i, LT(h_j)g_j) \\
&= \frac{x^\delta}{LT(h_i g_i)} LT(h_i)g_i - \frac{x^\delta}{LT(h_j g_j)} LT(h_j)g_j \\
&= \frac{x^\delta}{LT(h_i)LT(g_i)} LT(h_i)g_i - \frac{x^\delta}{LT(h_j)LT(g_j)} LT(h_j)g_j \\
&= x^\delta \left(\frac{1}{LT(g_i)} g_i - \frac{1}{LT(g_j)} g_j\right) \\
&= x^{\delta - \gamma_{ij}} \left(\frac{x^{\gamma_{ij}}}{LT(g_i)} g_i - \frac{x^{\gamma_{ij}}}{LT(g_j)} g_j\right) \\
&= x^{\delta - \gamma_{ij}} S(g_i, g_j)
\end{aligned}$$

where $x^{\gamma_{ij}} = \text{lcm}(LM(g_i), LM(g_j))$. Now take one of the S-polynomials $S(g_i, g_j)$. By assumption, we can apply the polynomial division algorithm to divide $S(g_i, g_j)$ by $g_1, \ldots, g_t$ to obtain

$$S(g_i, g_j) = \sum_{l=1}^{t} A_l g_l$$

where $A_l \in k[x_1, \ldots, x_n]$ and if $A_l g_l \neq 0$ then $multideg(A_l g_l) \leq multideg(S(g_i, g_j))$. Multiply both sides of the equation by $x^{\delta - \gamma_{ij}}$ to obtain

$$S(p_i, p_j) = x^{\delta - \gamma_{ij}} S(g_i, g_j) = \sum_{l=1}^{t} B_l g_l$$

where $B_l = x^{\delta - \gamma_{ij}} A_l$. If $B_l g_l \neq 0$ then

$$multideg(B_l g_l) \leq multideg(x^{\delta - \gamma_{ij}} S(g_i, g_j)) < \delta$$

because $LT(S(g_i, g_j)) < \text{lcm}(LM(g_i), LM(g_j)) = x^{\gamma_{ij}}$.

It follows that the first sum in the expansion of f can be written as

$$\sum_{multideg(h_i g_i) = \delta} LT(h_i) g_i = \sum_{l=1}^{t} \tilde{B}_l g_l$$

where if $\tilde{B}_l g_l \neq 0$ then $multideg(\tilde{B}_l g_l) < \delta$. So,

$$f = \sum_{l=1}^{t} \tilde{B}_l g_l \; + \sum_{multideg(h_i g_i) = \delta} (h_i - LT(h_i)) g_i \; + \sum_{multideg(h_i g_i) < \delta} h_i g_i.$$

is a linear combination of $g_1, \ldots, g_t$ where all the terms have multidegree less than δ. However, this contradicts the assumption that δ was minimal amongst all representations of f as a linear combination of $g_1, \ldots, g_t$.

Therefore $multideg(f) = \delta$ and the second case can never occur. This completes the proof. □

Unsurprisingly, Buchberger's criterion in Theorem 2.5.4 forms the basis for the Buchberger algorithm in the next section. Buchberger's algorithm is sometimes called the *S-pair criterion*.

Example 2.5.4. In this example, we will use Buchberger's criterion. We will work in the polynomial ring $\mathbb{R}[x, y, z]$. Let $I = (y - x^2, z - x^3)$. We claim that $G = \{y - x^2, z - x^3\}$ is a Gröbner basis, where we impose lexicographic order on $\mathbb{R}[x, y, z]$ with $y > z > x$.

To see that the claim is true, note that there is only one S-polynomial to consider by Theorem 2.5.4. It is

$$S(y - x^2, z - x^3) = \frac{yz}{y}(y - x^2) - \frac{yz}{z}(z - x^3) = -zx^2 + yx^3.$$

By applying the polynomial division algorithm, we obtain

$$S(y-x^2, z-x^3) = -zx^2 + yx^3 = x^3(y-x^2) + (-x^2)(z-x^3) + 0.$$

Since the remainder is zero, Theorem 2.5.4 tells us that G is a Gröbner basis for I.

We now claim that G is not a Gröbner basis for I if we impose lexicographic order on $\mathbb{R}[x, y, z]$ with $x > y > z$. We compute the S-polynomial as

$$S(y-x^2, z-x^3) = \frac{x^3}{-x^2}(y-x^2) - \frac{x^3}{-x^3}(z-x^3) = -xy + x^3 + z - x^3 = -xy + z.$$

But this time, if we apply the division algorithm, we obtain

$$-xy + z = 0(y-x^2) + 0(z-x^3) + (-xy+z).$$

The remainder is not zero. So, G is not a Gröbner basis with respect to the lexicographic order with $x > y > z$.

2.6 Buchberger's algorithm

Suppose that $I = (f_1, \ldots, f_s)$ is an ideal in the polynomial ring $k[x_1, \ldots, x_n]$. Buchberger's algorithm provides us with a method of converting the basis $F = \{f_1, \ldots, f_s\}$ to a Gröbner basis. The idea behind this is that by Theorem 2.5.4, we compute the S-polynomials $S(f_i, f_j)$ for distinct $i, j \in \{1, 2, \ldots s\}$. Then, we divide them by $f_1, \ldots, f_s$, using the polynomial division algorithm. If the remainder

$$\overline{S(f_i, f_j)}^F = 0$$

then we add $S(f_i, f_j)$ to F and repeat the process. In other words, we keep adding S-polynomials to F until it becomes a Gröbner basis for I.

We describe Buchberger's algorithm in detail below:

1. Suppose that $I = (f_1, \ldots, f_s)$ is an ideal in the polynomial ring $k[x_1, \ldots, x_n]$. Suppose that $F = \{f_1, \ldots, f_s\}$. Set $F = G$.

2. For distinct $i, j \in \{1, 2, \ldots, s\}$ where $i < j$, compute the S-polynomials $S(f_i, f_j)$.

3. Use the polynomial division algorithm to divide $S(f_i, f_j)$ by the polynomials in G.

4. If there exists a distinct pair (i, j) with $i < j$ such that $\overline{S(f_i, f_j)}^G = 0$, add the S-polynomial $S(f_i, f_j)$ to G and repeat the algorithm from step 1. Otherwise, G is a Gröbner basis by Theorem 2.5.4 and the algorithm terminates.

As with the other algorithms described in these notes, we must justify why Buchberger's algorithm works.

1. Does $G \subseteq I$ hold at every step of the algorithm?

   In step 1 of Buchberger's algorithm, $G = F \subseteq I$. In step 4, we enlarge G by adding the remainder $\overline{S(f_i, f_j)}^G$ to G. If $G \subseteq I$ then $S(f_i, f_j) \in I$ and by the division algorithm, $\overline{S(f_i, f_j)}^G \in I$. Therefore, $G \cup \{\overline{S(f_i, f_j)}^G\} \subseteq I$. So, $G \subseteq I$ at every step of Buchberger's algorithm.

2. Does G remain a basis for I?

   Since $F \subseteq I$ is a basis for I and $F \subseteq G \subseteq I$ at every step of Buchberger's algorithm, we deduce that G must be a basis for I.

3. Does Buchberger's algorithm terminate?

   Suppose that G' is the basis we started with in step 1 and G is the union of G' with the non-zero remainders of S-polynomials of elements of G'. Since $G' \subseteq G$, the ideals

   $$(LT(G')) \subseteq (LT(G)).$$

   Moreover, if $G' \neq G$ then $(LT(G')) \subset (LT(G))$. To see why this is the case, suppose that $r = \overline{S(f_i, f_j)}^{G'} \neq 0$ is a remainder which is adjoined to G. By Theorem 2.2.1, the leading term $LT(r)$ cannot be divisible by any of the leading terms of elements in G'. By Lemma 2.3.1, $LT(r) \notin (LT(G'))$. But by assumption, $LT(r) \in (LT(G))$. Hence, $(LT(G')) \subset (LT(G))$.

   Consequently, as we loop through the algorithm, we create a strictly ascending chain of ideals:

$$(LT(G')) \subset (LT(G_1)) \subset \cdots \subset (LT(G_k)) \subset \ldots$$

By Theorem 2.4.2, $k[x_1, \ldots, x_n]$ is a Noetherian ring. Thus, there exists $k \in \mathbb{Z}_{>0}$ such that if $m \in \mathbb{Z}_{\geq k}$ then $(LT(G_m)) = (LT(G_k))$. By the contrapositive of the previous claim, we deduce that $G_m = G_k = G$. Thus, Buchberger's algorithm must terminate.

It is mentioned in [CLO15, §7] that Buchberger's algorithm given above is not a practical method of doing the computation. There are several refinements of the algorithm given in [CLO15, §10], but we will not discuss them here. Observe that once a remainder $\overline{S(p,q)}^{G'} = 0$, the remainder will stay zero even if we add extra elements to G'. Hence, if we add new generators f_j to G' one at a time, the only remainders that need to be checked are $\overline{S(f_i, f_j)}^{G'}$ for $i \in \{1, 2, \ldots, j-1\}$.

Often, we find that with Buchberger's algorithm, the Gröbner bases constructed have unnecessary generators. We can eliminate them with the following lemma:

Lemma 2.6.1. *Let G be a Gröbner basis of an ideal $I \subseteq k[x_1, \ldots, x_n]$. Let $p \in G$ be a polynomial such that $LT(p) \in (LT(G - \{p\}))$. Then, $G - \{p\}$ is also a Gröbner basis for I.*

Proof. Assume that G is a Gröbner basis for the ideal $I \subseteq k[x_1, \ldots, x_n]$. Assume that $p \in G$ is a polynomial such that $LT(p) \in (LT(G - \{p\}))$. Since G is a Gröbner basis for I, $(LT(G)) = (LT(I))$. Since $LT(p) \in (LT(G - \{p\}))$, $(LT(G - \{p\})) = (LT(G)) = (LT(I))$, which means that $G - \{p\}$ is a Gröbner basis for I. □

Definition 2.6.1. Let I be an ideal of $k[x_1, \ldots, x_n]$ and $G = \{g_1, \ldots, g_t\}$ be a Gröbner basis for I.

We say that G is a **minimal Gröbner basis** if for $i \in \{1, 2, \ldots, t\}$, the leading constants $LC(g_i) = 1$ and $LT(g_i) \notin (LT(G - \{g_i\}))$.

We say that G is a **reduced Gröbner basis** if for $i \in \{1, 2, \ldots, t\}$, the leading constants $LC(g_i) = 1$ and no monomial in the expression for $g_i \in G$ lies in $(LT(G - \{g_i\}))$.

By definition, a reduced Gröbner basis is a minimal Gröbner basis. The point of reduced Gröbner bases is that they are unique.

Theorem 2.6.2. *Let $I \subseteq k[x_1, \ldots, x_n]$ be a non-zero ideal. For a given monomial ordering on $k[x_1, \ldots, x_n]$, I has a reduced Gröbner basis. Moreover, the reduced Gröbner basis is unique.*

Proof. Assume that $I \subseteq k[x_1, \ldots, x_n]$ is a non-zero ideal. Assume that we have a monomial ordering on $k[x_1, \ldots, x_n]$. By definition of a Gröbner basis, all minimal Gröbner bases for I must have the same leading terms.

Let G be a minimal Gröbner basis for I. We say that $g \in G$ is *fully reduced* for G if no monomial in the expression for g lies in $(LT(G - \{g\}))$. If $g \in G$ is fully reduced for G then g must be fully reduced for any other minimal Gröbner basis G' such that $g \in G'$ because the leading terms of G and G' are the same.

Now let $h \in G$ be arbitrary and set $h' = \overline{h}^{G-\{h\}}$. Let $G' = (G - \{h\}) \cup \{h'\}$.

To show: (a) G' is a minimal Gröbner basis for I.

(a) First, we observe that $LT(h') = LT(h)$. To see why this is the case, if we divide h by the polynomials in $G - \{h\}$ then the remainder $h' = \overline{h}^{G-\{h\}}$ must have $LT(h)$ as one of its terms because $LT(h) \notin (LT(G - \{h\}))$ and thus, $LT(h)$ is not divisible by any of the leading terms of $G - \{h\}$.

Subsequently, $LT(h) = LT(h')$ and $(LT(G')) = (LT(G))$.

Notice that by construction, G' is a minimal Gröbner basis for I and $h' = \overline{h}^{G-\{h\}}$ is fully reduced for G' by construction of G' and Theorem 2.2.1.

By repeating the above process for every $h \in G$, we obtain a minimal Gröbner basis G^* for I such that every element of G^* is fully reduced for G^*. Despite the fact that the Gröbner basis changes every time we apply the process, a fully reduced element stays fully reduced because the leading terms never change throughout the entire process. Hence, G^* is a reduced Gröbner basis.

We will now prove uniqueness. Suppose that G and G^* are both reduced Gröbner bases for I. Then, G and G^* are both minimal Gröbner bases and hence, $LT(G) = LT(G')$ as sets of leading terms. This means that if $g \in G$ then there exists $g' \in G'$ such that $LT(g) = LT(g')$.

To show: (b) $g = g'$.

(b) Consider the difference $g - g'$. Then, $g - g' \in I$ and since G is a Gröbner basis, the remainder $\overline{g - g'}^G = 0$ by Theorem 2.5.2. Since $LT(g) = LT(g')$, these terms must cancel each other out in the difference $g - g'$. The remaining terms in $g - g'$ are divisible by none of the leading terms $LT(G) = LT(G')$ because G and G' are reduced Gröbner bases.

So, $\overline{g - g'}^G = g - g'$ and consequently, $g - g' = 0$.

Therefore, a reduced Gröbner basis for I must be unique. □

A major consequence of Theorem 2.6.2 is that we can now tell when two sets of polynomials generate the same ideal. Suppose that $F = \{f_1, \ldots, f_s\}$ and $G = \{g_1, \ldots, g_t\}$ are two different sets of polynomials in $k[x_1, \ldots, x_n]$. Then, we compute reduced Gröbner bases for the ideals $(f_1, \ldots, f_s)$ and $(g_1, \ldots, g_t)$. If the reduced Gröbner bases for both ideals are the same then by the uniqueness of a reduced Gröbner basis in Theorem 2.6.2, $(f_1, \ldots, f_s) = (g_1, \ldots, g_t)$.

We will end with an involved example on how to use Buchberger's algorithm to compute a Gröbner basis and then, we will convert it to a reduced Gröbner basis.

Example 2.6.2. Let $I = (x^2 - y, x^3 - x) \subseteq \mathbb{Q}[x, y]$. We will compute a Gröbner basis for I using Buchberger's algorithm. We will use the lexicographic order on $\mathbb{Q}[x, y]$ with $x >_{lex} y$.

Let $G = \{x^2 - y, x^3 - x\}$. We commence by computing the S-polynomial

$$S(x^2 - y, x^3 - x) = x(x^2 - y) - (x^3 - x) = -xy + x.$$

Notice that the leading term $LT(-xy + x)$ cannot be divided by either $LT(x^2 - y) = x^2$ or $LT(x^3 - x) = x^3$. Therefore, $\overline{-xy + x}^G = -xy + x$ and consequently, we add the polynomial $-xy + x$ to G.

Now we have two more S-polynomials to compute. These are

$$S(x^2 - y, -xy + x) = y(x^2 - y) + x(-xy + x) = x^2 - y^2$$

and

$$S(x^3 - x, -xy + x) = y(x^3 - x) + x^2(-xy + x) = x^3 - xy.$$

Note that $\overline{x^3 - xy}^G = 0$ because $x^3 - xy = x(x^2 - y)$. So, we can discard the polynomial $x^3 - xy$ from our computations. On the other hand,

$$\overline{x^2 - y^2}^G = -y^2 + y.$$

So, we add $x^2 - y^2$ to the set G and then repeat Buchberger's algorithm.

Continuing in this fashion, we find that a Gröbner basis for I is

$$G = \{x^2 - y, x^3 - x, -xy + x, x^2 - y^2, y^2 - y, -xy + y^3\}.$$

Let us now find a reduced Gröbner basis for I. First, we want all the leading coefficients of each generator to be 1. So, we multiply each generator in G by an appropriate constant so that

$$G = \{x^2 - y, x^3 - x, xy - x, x^2 - y^2, y^2 - y, xy - y^3\}.$$

Next, we remove any redundant generators, in accordance with Lemma 2.6.1. For instance, $x^2 - y^2$ is redundant because

$$LT(x^2 - y^2) = x^2 \in (LT(G - \{x^2 - y^2\})) = (x^2, x^3, xy, y^2, xy).$$

Hence, we remove $x^2 - y^2$ from G so that

$$G = \{x^2 - y, x^3 - x, xy - x, y^2 - y, xy - y^3\}$$

The other redundant generators in G are $x^3 - x$ and $xy - y^3$. We are left with $G = \{x^2 - y, xy - x, y^2 - y\}$. Straight from the definition, we check that G is a reduced Gröbner basis for I.

As a quick check, we can use Theorem 2.5.2 to show that $x^2 - y, x^3 - x \in (x^2 - y, xy - x, y^2 - y)$. We already know that $x^2 - y \in (x^2 - y, xy - x, y^2 - y)$. By the polynomial division algorithm, we find that

$$x^3 - x = x(x^2 - y) + 1(xy - x).$$

Since the remainder is zero, we deduce that $x^3 - x \in (x^2 - y, xy - x, y^2 - y)$. So, $I \subseteq (x^2 - y, xy - x, y^2 - y)$ as expected.

Chapter 3

An introduction to elimination theory

3.1 The elimination and extension theorems

Similarly to [CLO15, Chapter 3, §1], we will commence with an example in order to establish what elimination theory is all about.

Example 3.1.1. Suppose that we want to solve the system of equations

$$\begin{aligned} x^2 + y + z &= 1, \\ x + y^2 + z &= 1, \\ x + y + z^2 &= 1. \end{aligned}$$

The idea is to consider the ideal

$$I = (x^2 + y + z - 1, x + y^2 + z - 1, x + y + z^2 - 1)$$

generated by these equations. We impose lexicographic order on $k[x, y, z]$ with $x >_{lex} y >_{lex} z$. By using Buchberger's algorithm, we find that a Gröbner basis for I is

$$G = \{x + y + z^2 - 1, y^2 - y - z^2 + z, 2yz^2 + z^4 - z^2, z^6 - 4z^4 + 4z^3 - z^2\}.$$

Hence, the original equations are satisfied if and only if the generators of G are equal to zero. In particular, we have the equation

$$z^6 - 4z^4 + 4z^3 - z^2 = z^2(z - 1)^2(z^2 + 2z - 1) = 0$$

which has solutions $z = 0, 1, -1 \pm \sqrt{2}$. If we substitute these values into the equations

$$2yz^2 + z^4 - z^2 = 0 \qquad \text{and} \qquad y^2 - y - z^2 + z = 0$$

then we can solve for the values of y. Once we obtain the values of y, we substitute them into the equation $x + y + z^2 - 1$ in order to obtain the values of x. The five solutions turn out to be

$$(x, y, z) = (1, 0, 0), (0, 1, 0), (0, 0, 1), (-1 \pm \sqrt{2}, -1 \pm \sqrt{2}, -1 \pm \sqrt{2}).$$

In the above example, we accentuate two key steps which allowed us to solve the system of equations. Firstly, we have the elimination step. We were able to obtain an equation where the LHS was a polynomial in z and the RHS was 0. This allowed us to isolate the variable z and solve for it first.

Secondly, we have the extension step. Once we solved for z, we were able to extend the solutions to solve for y and then solve for x. The key concept behind elimination theory is that these two steps can be generally executed.

We will first explain why the elimination step works. We require the following definition.

Definition 3.1.2. Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal. For $\ell \in \{1, 2, \ldots, n\}$, the ℓ^{th} **elimination ideal**, denoted by I_ℓ, is defined as the ideal of $k[x_{\ell+1}, \ldots, x_n]$

$$I_\ell = I \cap k[x_{\ell+1}, \ldots, x_n].$$

If $I = (f_1, \ldots, f_s)$ then the elimination ideal I_ℓ consists of all consequence of $f_1 = \cdots = f_s = 0$ which eliminate the variables $x_1, \ldots, x_\ell$. It is easy to check that I_ℓ is an ideal of $k[x_{\ell+1}, \ldots, x_n]$. We define $I_0 = I$ as the zeroth elimination ideal.

So, if we want to solve the simultaneous equations $f_1 = \cdots = f_s = 0$ then eliminating the variables $x_1, \ldots, x_\ell$ is the same as computing non-zero polynomials in the ℓ^{th} elimination ideal. The elimination theorem tells us that with a Gröbner basis, this is very easy to do.

Theorem 3.1.1 (Elimination theorem)**.** *Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal and let G be a Gröbner basis of I. Impose the lexicographic order on $k[x_1, \ldots, x_n]$ where $x_1 > x_2 > \cdots > x_n$. If $\ell \in \{0, 1, \ldots, n\}$ then the set*

$$G_\ell = G \cap k[x_{\ell+1}, \ldots, x_n]$$

is a Gröbner basis for the ℓ^{th} elimination ideal I_ℓ.

Proof. Assume that $I \subseteq k[x_1, \ldots, x_n]$ is an ideal and that G is a Gröbner basis for I. Assume that $\ell \in \{0, 1, \ldots, n\}$ and that $G_\ell = G \cap k[x_{\ell+1}, \ldots, x_n]$. We claim that $(LT(G_\ell)) = (LT(I_\ell))$. Since $G_\ell \subseteq I$, $(LT(G_\ell)) \subseteq (LT(I_\ell))$.

To show: (a) $(LT(I_\ell)) \subseteq (LT(G_\ell))$.

(a) Assume that $f \in I_\ell$. We want to show that there exists $g \in G_\ell$ such that $LT(g)|LT(f)$. Since $f \in I$, there exists $g' \in G$ such that $LT(g')|LT(f)$. Recall that we imposed lexicographic order on $k[x_1, \ldots, x_n]$ with $x_1 >_{lex} \cdots >_{lex} x_n$. Since $LT(f) \in k[x_{\ell+1}, \ldots, x_n]$, the lexicographic order ensures that $LT(g') \in k[x_{\ell+1}, \ldots, x_n]$ and subsequently, that $g' \in k[x_{\ell+1}, \ldots, x_n]$. Therefore, $g' \in G_\ell$ satisfies $LT(g')|LT(f)$. So, $f \in (LT(G_\ell))$ and $(LT(I_\ell)) \subseteq (LT(G_\ell))$.

So, $(LT(I_\ell)) = (LT(G_\ell))$, which means that $G_\ell = G \cap k[x_{\ell+1}, \ldots, x_n]$ is a Gröbner basis for the elimination ideal I_ℓ. □

By Theorem 3.1.1, the elimination step can be carried out easily, provided that we have a Gröbner basis (with the lexicographic order). Even with Buchberger's algorithm, constructing such a Gröbner basis is easier said than done. Sometimes, the Gröbner bases can be quite unwieldy to work with (see [CLO15, Chapter 2, §10, Exercise 13]). Versions of the elimination theorem exist which use more efficient monomial orderings than the lexicographic order.

Next, we will discuss the extension step. Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal and

$$V(I) = \{(a_1, \ldots, a_n) \in k^n \mid \text{If } f \in I \text{ then } f(a_1, \ldots, a_n) = 0\}$$

be the affine variety associated to I. The idea behind the extension step is to build a solution to our system of equations by considering one variable at a time. Let $\ell \in \{1, 2, \ldots, n\}$. A point $(a_{\ell+1}, \ldots, a_n) \in V(I_\ell)$ is called a *partial solution* to the original system of equations.

To extend $(a_{\ell+1}, \ldots, a_n)$ to a complete solution, we must add an extra variable to the solution, which amounts to finding a_l such that

$$(a_\ell, a_{\ell+1}, \ldots, a_n) \in V(I_{\ell-1}).$$

If $I_{\ell-1} = (g_1, \ldots, g_r)$ in $k[x_\ell, \ldots, x_n]$ then we want to find solutions $x_\ell = a_\ell$ of the equations

$$g_1(x_\ell, a_{\ell+1}, \ldots, a_n) = \cdots = g_r(x_\ell, a_{\ell+1}, \ldots, a_n) = 0.$$

Here, we run into a few problems. It might be the case that the polynomials $g_1, \ldots, g_r$ do not have a common root, which means that some partial solutions might not extend to complete solutions.

Example 3.1.3. Here is a simple example from [CLO15]. Suppose that we want to solve the system of equations

$$\begin{aligned} xy &= 1, \\ xz &= 1. \end{aligned}$$

Let $I = (xy - 1, xz - 1) \in k[x, y, z]$. A Gröbner basis for the first elimination ideal I_1 is $G_1 = \{y - z\}$. Hence, the partial solutions to the system of equations are given by (a, a) for $a \in k$. These partial solutions extend to a complete solution $(x, y, z) = (a^{-1}, a, a)$, except for the partial solution $(y, z) = (0, 0)$.

The extension of partial solutions is also sensitive to the field we are working over.

Example 3.1.4. Here is another example from [CLO15]. Suppose that we have the equations

$$\begin{aligned} x^2 &= y, \\ x^2 &= z. \end{aligned}$$

The partial solutions are $(y, z) = (a, a)$ for $a \in k$. These partial solutions extend to complete solutions, provided that $k = \mathbb{C}$. However, if $k = \mathbb{R}$ then only the solutions with $a \geq 0$ extend to complete solutions.

The above example suggests that the extension theorem should, at the very least, be stated for an algebraically closed field.

Let us provide the necessary context for the extension theorem. Suppose that we eliminate just the first variable x_1. We want to know if a partial solution $(a_2, \ldots, a_n) \in V(I_1)$ can be extended to a complete solution $(a_1, a_2, \ldots, a_n) \in V(I)$. The extension theorem tells us when this can be done.

Theorem 3.1.2 (Extension theorem)**.** *Let $I = (f_1, \ldots, f_s) \subseteq \mathbb{C}[x_1, \ldots, x_n]$ be an ideal and let I_1 be the first elimination ideal of I. If $i \in \{1, 2, \ldots, s\}$ then write f_i in the form*

$$f_i = c_i(x_2, \ldots, x_n)x_1^{N_i} + (\textit{terms where } x_1 \textit{ has degree less than } N_i).$$

Here $N_i \in \mathbb{Z}_{>0}$ and $c_i \in \mathbb{C}[x_2, \ldots, x_n]$ are non-zero polynomials. Let $(a_2, \ldots, a_n) \in V(I_1)$ be a partial solution. If $(a_2, \ldots, a_n) \notin V(c_1, \ldots, c_s)$ then there exists $a_1 \in \mathbb{C}$ such that $(a_1, a_2, \ldots, a_n) \in V(I)$.

We will defer the proof of Theorem 3.1.2 to a later, separate section. The proof we will give uses Gröbner bases. For now, we will address the conditions of Theorem 3.1.2 and demonstrate how it can be applied to solve systems of equations.

We have already established from Example 3.1.4 why Theorem 3.1.2 is stated for the algebraically closed field $\mathbb{C}$, rather than $\mathbb{R}$ or $\mathbb{Q}$. Next, we will talk about why we insist that $(a_2, \ldots, a_n) \notin V(c_1, \ldots, c_s)$ in Theorem 3.1.2.

The point is that in Theorem 3.1.2, if $i \in \{1, 2, \ldots, s\}$ then c_i is the leading coefficient of the polynomial f_i with respect to the variable x_1. So, the condition $(a_2, \ldots, a_n) \notin V(c_1, \ldots, c_s)$ tells us that if we substitute $(x_2, \ldots, x_n) = (a_2, \ldots, a_n)$ into the polynomials $c_1, \ldots, c_s \in k[x_2, \ldots, x_n]$ then $c_2, \ldots, c_s$ cannot vanish simultaneously.

To illustrate this point, let us return to Example 3.1.3. The only partial solution we could not extend to a complete solution was $(y, z) = (0, 0)$. If we substitute $(y, z) = (0, 0)$ into the leading coefficients y and z (with respect to the variable x) then they both vanish. This establishes that if the leading coefficients $c_1, \ldots, c_s$ vanish simultaneously on a partial solution then Theorem 3.1.2 may fail.

Finally, let us explain why the extension theorem can be used when eliminating any number of variables.

Theorem 3.1.3. *Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal and $\ell \in \{1, 2, \ldots, n\}$. Let $I_\ell \subseteq k[x_{\ell+1}, \ldots, x_n]$ be the ℓ^{th} elimination ideal. Then, $I_{\ell+1}$ is the first elimination ideal of I_ℓ.*

Proof. Assume that $I \subseteq k[x_1, \ldots, x_n]$ is an ideal and $\ell \in \{1, 2, \ldots, n\}$. Recall that the ℓ^{th} elimination ideal satisfies

$$I_\ell = I \cap k[x_{\ell+1}, \ldots, x_n].$$

Then,

$$\begin{aligned} I_{\ell+1} &= I \cap k[x_{\ell+2}, \ldots, x_n] \\ &= (I \cap k[x_{\ell+1}, \ldots, x_n]) \cap k[x_{\ell+2}, \ldots, x_n] \\ &= I_\ell \cap k[x_{\ell+2}, \ldots, x_n]. \end{aligned}$$

Hence, $I_{\ell+1}$ is the first elimination ideal of I_ℓ. □

We will see how Theorem 3.1.3 allows us to use the extension theorem to extend partial solutions for any number of variables by considering an example.

Example 3.1.5. We will work in the polynomial ring $\mathbb{C}[x, y, z]$. Suppose that we want to solve the system of equations

$$\begin{aligned} x^2 + y^2 + z^2 &= 1, \\ xyz &= 1. \end{aligned}$$

Let $I = (x^2 + y^2 + z^2 - 1, xyz - 1)$. A Gröbner basis for I with respect to the lexicographic order on $\mathbb{C}[x, y, z]$ (with $x > y > z$) is

$$G = \{y^4z^2 + y^2z^4 - y^2z^2 + 1, x + y^3z + yz^3 - yz\}.$$

Let $g_1 = y^4z^2 + y^2z^4 - y^2z^2 + 1$ and $g_2 = x + y^3z + yz^3 - yz$. By Theorem 3.1.1, the elimination ideals I_1 and I_2 are

$$I_1 = (g_1) \qquad \text{and} \qquad I_2 = 0.$$

Note that $V(I_2) = V(0) = \mathbb{C}$. So, every $c \in \mathbb{C}$ is a partial solution to the above system of equations.

By Theorem 3.1.3, I_2 is the first elimination ideal of I_1. We write g_1 in the prescribed form of Theorem 3.1.2:

$$g_1 = z^2(y^4) + (y^2z^4 - y^2z^2 + 1).$$

The leading coefficient of g_1 with respect to the variable y is z^2, which only vanishes if $z = 0$. By Theorem 3.1.2, if $c \in \mathbb{C} - \{0\}$ then there exists $b \in \mathbb{C}$ such that $(b, c) \in V(I_1)$ and $(y, z) = (b, c)$ is a partial solution to our original system of equations.

Now, I_1 is the first elimination ideal of I. Let us again verify that Theorem 3.1.2 applies here. First, we write $x^2 + y^2 + z^2 - 1$ and $xyz - 1$ in the forms required by Theorem 3.1.2:

$$x^2 + y^2 + z^2 - 1 = 1(x^2) + (y^2 + z^2 - 1)$$

and

$$xyz - 1 = yz(x) - 1.$$

The leading coefficients of $x^2 + y^2 + z^2 - 1$ and $xyz - 1$ with respect to the variable x are 1 and yz respectively. Note that $V(1, yz) = 0$ since the first leading coefficient is constant. Since the partial solution
$(b, c) \in \mathbb{C} \times (\mathbb{C} - \{0\})$, $(b, c) \notin V(1, yz) = 0$. By Theorem 3.1.2, there exists $a \in \mathbb{C}$ such that $(a, b, c) \in V(I)$.

Therefore, all partial solutions $z = c \in \mathbb{C} - \{0\}$ extend to complete solutions $(x, y, z) = (a, b, c)$ of the original system of equations, which are points in the affine variety $V(I)$.

Example 3.1.5 also tells us that it is particularly easy to use the extension theorem when one of the leading coefficients involved is a constant. We state this as a corollary.

Corollary 3.1.4. *Let $I = (f_1, \ldots, f_s) \subseteq \mathbb{C}[x_1, \ldots, x_n]$ be an ideal and let I_1 be the first elimination ideal of I. Suppose that there exists $i \in \{1, 2, \ldots, n\}$ such that*

$$f_i = c_i x_1^{N_i} + (\textit{terms where } x_1 \textit{ has degree less than } N_i)$$

where $N_i \in \mathbb{Z}_{>0}$ and $c_i \in \mathbb{C}$. Let $(a_2, \ldots, a_n) \in V(I_1)$ be a partial solution. Then, there exists $a_1 \in \mathbb{C}$ such that $(a_1, a_2, \ldots, a_n) \in V(I)$.

Together, the elimination and extension theorems (Theorem 3.1.1 and Theorem 3.1.2 respectively) show that if we have a system of equations and we work in polynomial ring $\mathbb{C}[x_1, \ldots, x_n]$ with lexicographic order (and $x_1 > x_2 > \cdots > x_n$) then we can prove that under specific circumstances, partial solutions extend to complete solutions. Of course, constructing the complete solutions is an entirely different question to demonstrating the existence of complete solutions. Most of time, numerical approximation methods (Newton-Rhapson for instance) are needed to compute the necessary roots required for the method to work.

3.2 Geometric interpretation of elimination

We will work over the field $\mathbb{C}$. In this section, we will highlight geometric interpretations of the elimination theorem (Theorem 3.1.1) and the

extension theorem (Theorem 3.1.2). Geometrically, the elimination theorem corresponds to projecting an affine variety to a lower dimensional subspace.

In order to make this precise, assume that $V = V(f_1, \ldots, f_s) \subseteq \mathbb{C}^n$ is an affine variety. For $\ell \in \{1, 2, \ldots, n-1\}$, define the projection map

$$\begin{array}{rccc} \pi_\ell : & \mathbb{C}^n & \to & \mathbb{C}^{n-\ell} \\ & (a_1, \ldots, a_n) & \mapsto & (a_{\ell+1}, \ldots, a_n) \end{array}$$

We want to relate the image $\pi_\ell(V)$ to the ℓ^{th} elimination ideal I_ℓ.

Lemma 3.2.1. *Let $I = (f_1, \ldots, f_s) \subseteq \mathbb{C}[x_1, \ldots, x_n]$ be an ideal. Let $V = V(f_1, \ldots, f_s)$ be the affine variety associated with I. Let $\ell \in \{1, 2, \ldots, n-1\}$. Then, in $\mathbb{C}^{n-\ell}$,*

$$\pi_\ell(V) \subseteq V(I_\ell).$$

Proof. Assume that $I = (f_1, \ldots, f_s) \subseteq \mathbb{C}[x_1, \ldots, x_n]$ is an ideal and $V = V(f_1, \ldots, f_s)$ is the affine variety associated with I. Assume that $\ell \in \{1, 2, \ldots, n-1\}$.

Assume that $(a_{\ell+1}, \ldots, a_n) \in \pi_\ell(V)$. Then, there exists $(a_1, \ldots, a_n) \in \mathbb{C}^n$ such that if $i \in \{1, 2, \ldots, s\}$ then

$$f_i(a_1, a_2, \ldots, a_n) = 0$$

and $\pi_\ell(a_1, \ldots, a_n) = (a_{\ell+1}, \ldots, a_n)$. Now assume that $p \in I_\ell = I \cap k[x_{\ell+1}, \ldots, x_n]$. Since $p \in I = (f_1, \ldots, f_s)$, there exist polynomials $q_1, \ldots, q_s \in k[x_1, \ldots, x_n]$ such that

$$p(x_{\ell+1}, \ldots, x_n) = \sum_{i=1}^{s} q_i(x_1, \ldots, x_n) f_i(x_1, \ldots, x_n).$$

Substituting $(x_1, x_2, \ldots, x_n) = (a_1, a_2, \ldots, a_n)$ into the above equation, we find that

$$p(a_{\ell+1}, \ldots, a_n) = \sum_{i=1}^{s} q_i(a_1, \ldots, a_n) f_i(a_1, \ldots, a_n) = 0.$$

Therefore, $(a_{\ell+1}, \ldots, a_n) \in V(I_\ell)$ and consequently, $\pi_\ell(V) \subseteq V(I_\ell)$. □

Similarly to the previous section, points of the affine variety $V(I_\ell)$ are called *partial solutions*. By 3.2.1, we can write the image $\pi_\ell(V)$ as

$$\pi_\ell(V) = \left\{ (a_{\ell+1}, \ldots, a_n) \in V(I_\ell) \;\middle|\; \begin{array}{l} \text{There exists } a_1, \ldots, a_\ell \in \mathbb{C} \\ \text{such that } (a_1, \ldots, a_n) \in V \end{array} \right\} \tag{3.1}$$

The points of $\pi_\ell(V)$ are partial solutions which extend to complete solutions. To illustrate this point, let us use Example 3.1.3 again.

Recall that in Example 3.1.3, we want to solve the system of equations

$$\begin{aligned} xy &= 1, \\ xz &= 1. \end{aligned}$$

Again, let $I = (xy - 1, xz - 1)$. The partial solutions are given by the affine variety

$$V(I_1) = \{(y, z) \in \mathbb{C}^2 \mid y = z\}.$$

The image of the affine variety $V = V(xy - 1, xz - 1)$ under the projection map $\pi_1 : \mathbb{C}^3 \to \mathbb{C}^2$ which projects to the latter two coordinates is

$$\pi_1(V) = \{(a, a) \in \mathbb{C}^2 \mid a \neq 0\}.$$

Notice that $\pi_1(V)$ cannot be an affine variety because $(0, 0) \notin \pi_1(V)$. This leads us to the geometric interpretation of the extension theorem, which in this context, tells us how close the set of extendable partial solutions $\pi_1(V)$ is to being an affine variety.

Theorem 3.2.2 (Geometric extension theorem)**.** *Let $I = (f_1, \ldots, f_s) \subseteq \mathbb{C}[x_1, \ldots, x_n]$ be an ideal. Let $V = V(f_1, \ldots, f_s) \subseteq \mathbb{C}^n$ be the affine variety associated to I. For $i \in \{1, 2, \ldots, s\}$, let $c_i \in \mathbb{C}[x_2, \ldots, x_n]$ be the leading coefficient of f_i with respect to the variable x_1 (see Theorem 3.1.2). Let I_1 be the first elimination ideal of I. Then, in $\mathbb{C}^{n-1}$,*

$$V(I_1) = \pi_1(V) \cup (V(c_1, \ldots, c_s) \cap V(I_1))$$

where $\pi_1 : \mathbb{C}^n \to \mathbb{C}^{n-1}$ is the projection map onto the last $n - 1$ coordinates.

Proof. Assume that $I \subseteq \mathbb{C}[x_1, \ldots, x_n]$, $V \subseteq \mathbb{C}^n$ and I_1 are defined as above.

To show: (a) $V(I_1) \subseteq \pi_1(V) \cup (V(c_1, \ldots, c_s) \cap V(I_1))$.

(b) $\pi_1(V) \cup (V(c_1, \ldots, c_s) \cap V(I_1)) \subseteq V(I_1)$.

(a) Recall that the points in $V(I_1)$ are partial solutions to the system of equations $f_1 = \cdots = f_s = 0$. Pick any point $a \in V(I_1)$. Then, a either extends to a complete solution or it does not. Hence, there are two cases to consider.

Case 1: Suppose that $a \in V(I_1)$ such that a extends to a complete solution to the system of equations $f_1 = \cdots = f_s = 0$. By equation (3.1), this means that $a \in \pi_1(V)$.

Case 2: On the contrary, suppose that $a \in V(I_1)$ does not extend to a complete solution. By the contrapositive statement of Theorem 3.1.2, this means that $a \in V(I_1) \cap V(c_1, \ldots, c_s)$.

By combining the two cases together, we find that
$V(I_1) \subseteq \pi_1(V) \cup (V(c_1, \ldots, c_s) \cap V(I_1))$.

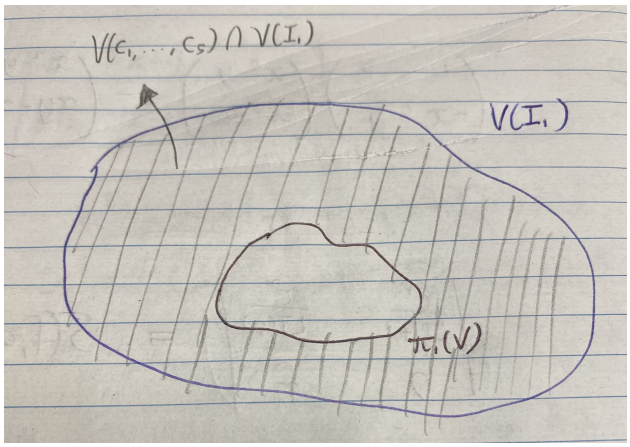
(b) Observe that $V(I_1) \cap V(c_1, \ldots, c_s) \subseteq V(I_1)$ and by Lemma 3.2.1, $\pi_1(V) \subseteq V(I_1)$. Therefore, $\pi_1(V) \cup (V(c_1, \ldots, c_s) \cap V(I_1)) \subseteq V(I_1)$.

By combining parts (a) and (b), we deduce that

$$\pi_1(V) \cup (V(c_1, \ldots, c_s) \cap V(I_1)) = V(I_1)$$

as required. □

As stated in [CLO15], Theorem 3.2.2 tells us that $\pi_1(V)$ fills up the affine variety $V(I_1)$, except possibly for a part which lies in $V(c_1, \ldots, c_s)$. What is not clear in Theorem 3.2.2 is how big $V(c_1, \ldots, c_s)$ is. If we take Example 3.1.3 again, the affine variety $V(c_1, \ldots, c_s)$ is in this case, $\{(0, 0)\}$, which is quite small.



The closure theorem gives us more information about the size of $\pi_\ell(V)$ relative to $V(I_\ell)$.

Theorem 3.2.3 (Closure theorem)**.** *Let $I = (f_1, \ldots, f_s) \subseteq \mathbb{C}[x_1, \ldots, x_n]$ be an ideal. Let $V = V(f_1, \ldots, f_s) \subseteq \mathbb{C}^n$ be the affine variety associated to I. Let $\ell \in \{1, 2, \ldots, n-1\}$ and I_ℓ be the ℓ^{th} elimination ideal of I. Let π_ℓ denote the projection map*

$$\begin{array}{rccc} \pi_\ell : & \mathbb{C}^n & \to & \mathbb{C}^{n-\ell} \\ & (a_1, \ldots, a_n) & \mapsto & (a_{\ell+1}, \ldots, a_n) \end{array}$$

Then, $V(I_\ell)$ is the smallest affine variety containing $\pi_\ell(V) \subseteq \mathbb{C}^{n-\ell}$. Moreover, if $V \neq \emptyset$ then there exists an affine variety $W \subset V(I_\ell)$ such that $V(I_\ell) \backslash W \subseteq \pi_\ell(V)$.

The proof of Theorem 3.2.3 is provided in [CLO15, §4]. The closure theorem gives us a precise structure for the image $\pi_\ell(V)$. There exists affine varieties $Z_i \subseteq W_i \subseteq \mathbb{C}^{n-\ell}$ for $i \in \{1, 2, \ldots, m\}$ such that

$$\pi_\ell(V) = \bigcup_{i=1}^{m} (W_i \backslash Z_i).$$

The closure theorem is stated for the field $\mathbb{C}$, but just like the extension theorem, it holds over any algebraically closed field.

3.3 A proof of the extension theorem

We begin with the necessary terminology.

Definition 3.3.1. Let $f \in k[x_1, \ldots, x_n]$ be non-zero. Write f in the form

$$f = c_f(x_2, \ldots, x_n) x_1^N + (\text{terms where } x_1 \text{ has degree less than } N).$$

where $N \in \mathbb{Z}_{\geq 0}$ and $c_f \in k[x_2, \ldots, x_n]$ is non-zero. Define $\deg(f, x_1) = N$. If $f = 0$ then we define $c_f = 0$.

Definition 3.3.2. Fix a monomial order on $k[x_1, \ldots, x_n]$. Let $G = \{g_1, \ldots, g_t\} \subseteq k[x_1, \ldots, x_n]$. We say that f has a **standard representation** with respect to G if there exists polynomials $A_1, \ldots, A_t \in k[x_1, \ldots, x_n]$ such that

$$f = A_1 g_1 + \cdots + A_t g_t.$$

In the standard representation above if $A_i g_i \neq 0$ then we must have $multideg(f) \geq multideg(A_i g_i)$.

We require the following lemma regarding the quantity $\deg(f, x_i)$ and the polynomial c_f.

Lemma 3.3.1. *Suppose that we impose lexicographic order on $k[x_1, \ldots, x_n]$ with $x_1 > \cdots > x_n$. Let $f = \sum_{j=1}^{t} A_j g_j$ be a standard representation.*

1. *If $A_j g_j \neq 0$ then $\deg(f, x_1) \geq \deg(A_j g_j, x_1)$.*
2. *If $N = \deg(f, x_1)$ then*

$$c_f = \sum_{\deg(A_j g_j, x_1) = N} c_{A_j} c_{g_j}.$$

Proof. Assume that $f = \sum_{j=1}^{t} A_j g_j$ is a standard representation and that we have the lexicographic order on $k[x_1, \ldots, x_n]$. First, observe that if $A_j g_j \neq 0$ for some $j \in \{1, 2, \ldots, t\}$ then

$$\deg(A_j g_j, x_1) \leq \max_{j \in \{1,2,\ldots,n\}} \deg(A_j g_j, x_1) = \deg(f, x_1).$$

Next, assume that $N = \deg(f, x_1)$. The polynomial $c_f \in k[x_2, \ldots, x_n]$ satisfies

$$f = c_f(x_2, \ldots, x_n) x_1^N + (\text{terms where } x_1 \text{ has degree less than } N).$$

Since $f = \sum_{i=1}^{t} A_i g_i$, we collect all the terms on the RHS which has degree N with respect to x_1. The result is

$$\sum_{\deg(A_j g_j, x_1) = N} c_{A_j g_j} = \sum_{\deg(A_j g_j, x_1) = N} c_{A_j} c_{g_j}.$$

Therefore, by comparing terms with degree N (with respect to x_1), we find that

$$f = \sum_{\deg(A_j g_j, x_1) = N} c_{A_j} c_{g_j}.$$

□

The proof of the extension theorem hinges on the following theorem, which tells us how Gröbner bases with the lexicographic order interact with certain partial solutions.

Theorem 3.3.2. *Let $I \subseteq k[x_1, \ldots, x_n]$, where we impose the lexicographic order on $k[x_1, \ldots, x_n]$ with $x_1 > \cdots > x_n$. Let $G = \{g_1, \ldots, g_t\}$ be a Gröbner basis for I. For $j \in \{1, 2, \ldots, t\}$, let $c_j \in k[x_2, \ldots, x_n]$ and $N_j \in \mathbb{Z}_{\geq 0}$ be such that*

$$g_j = c_j(x_2, \ldots, x_n)x_1^{N_j} + (\textit{terms where } x_1 \textit{ has degree less than } N_j).$$

Assume that $\boldsymbol{a} = (a_2, \ldots, a_n) \in V(I_1)$ is a partial solution such that $\boldsymbol{a} \notin V(c_1, \ldots, c_t)$. Then, there exists $g_o \in G$ such that

1. *The leading coefficient $c_o \in k[x_2, \ldots, x_n]$ satisfies $c_o(\boldsymbol{a}) \neq 0$,*
2. *g_o has minimal degree (with respect to x_1) among all elements $g_j \in G$ with $c_j(\boldsymbol{a}) \neq 0$,*
3. *$\deg(g_o(x_1, \boldsymbol{a})) > 0$,*
4. *If $g_o(a_1, \boldsymbol{a}) = 0$ for some $a_1 \in k$ then $(a_1, \boldsymbol{a}) \in V(I)$,*
5. *$\{f(x_1, \boldsymbol{a}) \mid f \in I\} = (g_o(x_1, \boldsymbol{a}))$ as ideals in $k[x_1]$.*

Proof. Assume that $I \subseteq k[x_1, \ldots, x_n]$ is an ideal and $G = \{g_1, \ldots, g_t\}$ is a Gröbner basis for I. Consider the set

$$\{g_j \in G \mid c_j(\mathbf{a}) \neq 0\}$$

where $c_j \in k[x_2, \ldots, x_n]$ is the leading coefficient of g_j with respect to the variable x_1, $\mathbf{a} = (a_2, \ldots, a_n) \in V(I_1)$ and $\mathbf{a} \notin V(c_1, \ldots, c_t)$.

Since the set $\{g_j \in G \mid c_j(\mathbf{a}) \neq 0\}$ is finite, we can always choose a polynomial $g_o \in G$ from this set such that $\deg(g_o, x_1)$ is minimal amongst all the polynomials in the set $\{g_j \in G \mid c_j(\mathbf{a}) \neq 0\}$. By construction, the leading coefficient c_o must satisfy $c_o(\mathbf{a}) \neq 0$.

To show: (a) $\deg(g_o(x_1, \mathbf{a})) > 0$.

(b) $\{f(x_1, \mathbf{a}) \mid f \in I\} = (g_o(x_1, \mathbf{a}))$.

(c) If $g_o(a_1, \mathbf{a}) = 0$ for some $a_1 \in k$ then $(a_1, \mathbf{a}) \in V(I)$.

(a) Suppose for the sake of contradiction that $\deg(g_o(x_1, \mathbf{a})) = 0$. Since $c_o(\mathbf{a}) \neq 0$, $\deg(g_o, x_1) = 0$. Hence, $g_o \in I_1$ and $c_o = g_o$. Since $\mathbf{a} \in V(I_1)$,

$$c_o(\mathbf{a}) = g_o(\mathbf{a}) = 0$$

which contradicts the assumption that $c_o(\mathbf{a}) \neq 0$. So, $\deg(g_o(x_1, \mathbf{a})) > 0$.

(b) Consider the evaulation map

$$\begin{array}{rccc} ev_{\mathbf{a}}: & k[x_1, \ldots, x_n] & \to & k[x_1] \\ & f(x_1, x_2, \ldots, x_n) & \mapsto & f(x_1, \mathbf{a}). \end{array}$$

This is a ring homomorphism. The image of I under $ev_{\mathbf{a}}$ is the ideal

$$ev_{\mathbf{a}}(I) = (g_1(x_1, \mathbf{a}), \ldots, g_t(x_1, \mathbf{a})) \subseteq k[x_1].$$

Notice that $ev_{\mathbf{a}}(I) = \{f(x_1, \mathbf{a}) \mid f \in I\}$. Hence, it suffices to show that if $j \in \{1, 2, \ldots, t\}$ then $g_j(x_1, \mathbf{a}) \in (g_o(x_1, \mathbf{a}))$.

To show: (ba) If $j \in \{1, 2, \ldots, t\}$ then $g_j(x_1, \mathbf{a}) \in (g_o(x_1, \mathbf{a}))$.

(ba) Assume that $j \in \{1, 2, \ldots, t\}$. We proceed by induction on the degree $\deg(g_j, x_1)$. For the base case, assume that $\deg(g_j, x_1) < \deg(g_o, x_1)$. By construction of g_o, this means that either $c_j(\mathbf{a}) = 0$, which means that g_j drops x_1-degree when evaluated at $\mathbf{a}$, or g_j vanishes identically when evaluated at $\mathbf{a}$.

Let $d_o = \deg(g_o, \mathbf{a})$. Suppose for the sake of contradiction that there exists $g_j \in G$ such that $\deg(g_j, x_1) < d_o$ and $g_j(x_1, \mathbf{a}) \neq 0$. Among all the g_j, pick a polynomial $g_b \in G$ which minimises the decrease in x_1-degree when evaluated at $\mathbf{a}$. Set $\delta = \deg(g_b, x_1) - \deg(g_b(x_1, \mathbf{a}))$ so that the x_1-degree of g_b drops by δ when we evaluate at $\mathbf{a}$.

For clarity, we will introduce even more notation. Set $d_b = \deg(g_b, x_1)$ so that $\deg(g_b(x_1, \mathbf{a})) = d_b - \delta$. Define the polynomial $S \in I$ by

$$S = c_0 x_1^{d_o - d_b} g_b - c_b g_o = c_0 x_1^{d_o - d_b}(c_b x_1^{d_b} + \ldots) - c_b(c_o x_1^{d_o} + \ldots).$$

Since the leading terms in $c_o x_1^{d_o - d_b}$ and $c_b g_o$ cancel in the subtraction above, $\deg(S, x_1) < d_o$. We aim to derive a contradiction by computing $\deg(S(x_1, a))$ in two different ways:

Method 1: We first proceed via the direct route — evaluate $S(x_1, \ldots, x_n)$ at $\mathbf{a}$ and then compute the degree. We obtain

$$S(x_1, \mathbf{a}) = c_0(\mathbf{a}) x_1^{d_o - d_b} g_b(x_1, \mathbf{a}) - c_b(\mathbf{a}) g_o(x_1, \mathbf{a}) = c_0(\mathbf{a}) x_1^{d_o - d_b} g_b(x_1, \mathbf{a}).$$

This is because $c_b(\mathbf{a}) = 0$ (the degree of g_b drops when evaluated at $\mathbf{a}$). Recall that by definition of g_o, $c_o(\mathbf{a}) \neq 0$ and

$$\deg(S(x_1, \mathbf{a})) = d_o - d_b + \deg(g_b(x_1, \mathbf{a})) = d_o - d_b + d_b - \delta = d_o - \delta.$$

Method 2: Consider a standard representation $S = \sum_{j=1}^{t} B_j g_j$. Such a standard representation exists because $S \in I$ and $\{g_1, \ldots, g_t\}$ is a Gröbner basis for I. By the first part of Lemma 3.3.1, we have for $B_j \neq 0$,

$$\deg(B_j, x_1) + \deg(g_j, x_1) = \deg(B_j g_j, x_1) \leq \deg(S, x_1) < d_o.$$

Notice that by the inequality above, the g_j which appear satisfy $\deg(g_j, x_1) < d_o = \deg(g_o, x_1)$. Therefore, either $g_j(x_1, \mathbf{a}) = 0$ or the x_1-degree of g_j drops by at least δ upon evaluation at $\mathbf{a}$. Therefore,

$$\deg(B_j(x_1, \mathbf{a})) + \deg(g_j(x_1, \mathbf{a})) \leq \deg(B_j, x_1) + \deg(g_j, x_1) - \delta < d_0 - \delta.$$

Here we implicitly used the fact that $\mathbf{a} \notin V(c_1, \ldots, c_t)$ so that at the very least, one of the g_j must drop in x_1-degree by at least δ.

Consequently,

$$\deg(S(x_1, \mathbf{a})) \leq \max(\deg(B_j(x_1, \mathbf{a})) + \deg(g_j(x_1, \mathbf{a})) < d_0 - \delta.$$

However, in Method 1 we proved that $\deg(S(x_1, \mathbf{a})) = d_o - \delta$. This gives our desired contradiction. Hence, if $g_j \in G$ then either $d(g_j, x_1) \geq d_o$ or $g_j(x_1, \mathbf{a}) = 0$. The first case will be addressed in the inductive step. The second case tells us that $g_j(x_1, \mathbf{a}) \in (g_o(x_1, \mathbf{a}))$, which proves the base case.

For the inductive hypothesis, fix $d \geq d_o$ and assume that $g_j(x_1, \mathbf{a}) \in (g_0(x_1, \mathbf{a}))$ for $g_j \in G$ with $\deg(g_j, x_1) < d$. Now assume that $g_l \in G$ with $\deg(g_l, x_1) = d$. Define the polynomial $T \in I$ by

$$T = c_o g_l - c_l x_1^{d-d_o} g_o = c_o(c_j x_1^d + \ldots) - c_j x_1^{d-d_o}(c_o x_1^{d_0} + \ldots).$$

Note that $\deg(T, x_1) < d$. Since $T \in I$, we can write it as a standard representation $T = \sum_{k=1}^{t} C_k g_k$. Arguing as in the base case, we deduce that if $B_k \neq 0$ then $\deg(g_k, x_1) < d$. By the inductive hypothesis, this means that if $B_k \neq 0$ then $g_k(x_1, \mathbf{a}) \in (g_o(x_1, \mathbf{a}))$. Now write

$$c_o g_l = T + c_l x_1^{d-d_o} g_o = c_l x_1^{d-d_o} g_o + \sum_{k=1}^{t} C_k g_k.$$

Then,

$$c_o(\mathbf{a})g_l(x_1,\mathbf{a}) = c_l(\mathbf{a})x_1^{d-d_o}g_o(x_1,\mathbf{a}) + \sum_{k=1}^{t} B_k(x_1,\mathbf{a})g_k(x_1,\mathbf{a})$$

The RHS is an element of the ideal $(g_o(x_1,\mathbf{a}))$. Since $c_o(\mathbf{a}) \neq 0$ by construction, we find that $g_l(x_1,\mathbf{a}) \in (g_o(x_1,\mathbf{a}))$. This finally completes the induction step.

By mathematical induction, we deduce that if $g_j \in G$ then $g_j(x_1,\mathbf{a}) \in (g_o(x_1,\mathbf{a}))$. So,

$$(g_o(x_1,\mathbf{a})) = (g_1(x_1,\mathbf{a}),\ldots,g_t(x_1,\mathbf{a})) = \{f(x_1,\mathbf{a}) \mid f \in I\}.$$

(c) Assume that there exists $a_1 \in k$ such that $g_o(a_1,\mathbf{a}) = 0$. By part (b),

$$(g_o(x_1,\mathbf{a})) = \{f(x_1,\mathbf{a}) \mid f \in I\}.$$

Therefore, if $g_o(a_1,\mathbf{a}) = 0$ then $f(a_1,\mathbf{a}) = 0$ and $(a_1,\mathbf{a}) \in V(I)$ as required. □

Now we are ready to prove the extension theorem. We will do it over an algebraically closed field k rather than just $\mathbb{C}$. We restate the extension theorem below for clarity.

Theorem 3.3.3 (Extension theorem V2)**.** *Let k be an algebraically closed field and $I = (f_1,\ldots,f_s) \subseteq k[x_1,\ldots,x_n]$ be an ideal and let I_1 be the first elimination ideal of I. If $i \in \{1,2,\ldots,s\}$ then write f_i in the form*

$$f_i = c_i(x_2,\ldots,x_n)x_1^{N_i} + (\textit{terms where } x_1 \textit{ has degree less than } N_i).$$

Here $N_i \in \mathbb{Z}_{>0}$ and $c_i \in k[x_2,\ldots,x_n]$ are non-zero polynomials. Let $(a_2,\ldots,a_n) \in V(I_1)$ be a partial solution. If $(a_2,\ldots,a_n) \notin V(c_1,\ldots,c_s)$ then there exists $a_1 \in k$ such that $(a_1,a_2,\ldots,a_n) \in V(I)$.

Proof. Assume that $I \subseteq k[x_1,\ldots,x_n]$ is an ideal and $G = \{g_1,\ldots,g_t\}$ is a Gröbner basis for I with respect to the lexicographic order on $k[x_1,\ldots,x_n]$ with $x_1 > \cdots > x_n$. Set $\mathbf{a} = (a_2,\ldots,a_n)$.

To show: (a) There exists $g_j \in G$ such that $c_{g_j}(\mathbf{a}) \neq 0$.

(a) Since $\mathbf{a} \notin V(c_1,\ldots,c_s)$, there exists $i \in \{1,2,\ldots,t\}$ such that $c_i(\mathbf{a}) \neq 0$. Write the polynomial $f_i \in I$ with its standard representation

$$f_i = \sum_{j=1}^{t} A_j g_j.$$

Since $c_i = c_{f_i}$ and $N_i = \deg(f_i, x_1)$, we can invoke the second part of Lemma 3.3.1 to deduce that

$$c_i = \sum_{\deg(A_j g_j, x_1) = N_i} c_{A_j} c_{g_j}.$$

Since $c_i(\mathbf{a}) \neq 0$, there exists a c_{g_j} appearing in the LHS such that $c_{g_j}(\mathbf{a}) \neq 0$.

Now we can apply Theorem 3.3.2 to obtain $g_o \in G$ with $\deg(g_o(x_1, \mathbf{a})) > 0$. Since k is algebraically closed, there exists $a_1 \in k$ such that $g_o(a_1, \mathbf{a}) = 0$. As yet another consequence of Theorem 3.3.2, $(a_1, \mathbf{a}) \in V(I)$. This proves the extension theorem for an algebraically closed field. $\square$

Chapter 4

The Hilbert Nullstellensatz

4.1 The weak Nullstellensatz

Recall that if $V \subseteq k^n$ is an affine variety then we can construct the ideal

$$I(V) = \{f \in k[x_1, \ldots, x_n] \mid \text{If } a \in V \text{ then } f(a) = 0\}.$$

Conversely, if $I \subseteq k[x_1, \ldots, x_n]$ then we can define the set

$$V(I) = \{a \in k^n \mid \text{If } f \in I \text{ then } f(a) = 0\}.$$

By the Hilbert basis theorem (see Theorem 2.4.2), there exists $f_1, \ldots, f_s \in k[x_1, \ldots, x_n]$ such that $I = (f_1, \ldots, f_s)$. Consequently,

$$V(I) = \{a \in k^n \mid \text{If } i \in \{1, 2, \ldots, s\} \text{ then } f_i(a) = 0\}$$

is an affine variety. Thus, we have a map

$$\begin{array}{ccc} \{\text{Ideals}\} & \leftrightarrow & \{\text{Affine varieties}\} \\ I & \mapsto & V(I) \\ I(V) & \leftarrow\!\shortmid & V \end{array}$$

We know that the above map is not one-to-one. For instance, if $I_1 = (1 + x^2)$ and $I_2 = (1 + x^2 + x^4)$ are ideals in $\mathbb{R}[x]$ then $V(I_1) = V(I_2) = \emptyset$.

The above example shows that it is possible to have different ideals represent the empty variety. However, if we work in an algebraically closed field then this problem goes away for the single variable case. To see why this is the case, recall that $k[x]$ is a PID. So, if $I \subseteq k[x]$ is an ideal then

$I = (f)$ for some $f \in k[x]$.

The affine variety $V(I)$ is the set of roots of f. Assuming that k is algebraically closed, we deduce that every non-constant polynomial in $k[x]$ has a root. Therefore, if $V(I) = \emptyset$ then $f \in k - \{0\}$. Consequently, $I = (f) = (1) = k[x]$. So, $I = k[x]$ is the only ideal of $k[x]$ such that $V(I) = \emptyset$.

The weak Nullstellensatz asserts that this also happens in the multivariable case. Note that in the proof we give, we will use a particular result in [CLO15] without proof.

Theorem 4.1.1 (Weak Nullstellensatz)**.** *Let k be an algebraically closed field and let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal such that $V(I) = \emptyset$. Then, $I = k[x_1, \ldots, x_n]$.*

Proof. We will prove the contrapositive of the statement. Assume that $I \subset k[x_1, \ldots, x_n]$ is an ideal of $k[x_1, \ldots, x_n]$ which is not equal to $k[x_1, \ldots, x_n]$ itself. Assume that k is an algebraically closed field.

Before we proceed, let us establish some terminology. Given $a \in k$ and $f \in k[x_1, \ldots, x_n]$, let $\overline{f} = f(x_1, \ldots, x_{n-1}, a) \in k[x_1, \ldots, x_{n-1}]$. Define the ideal

$$I_{x_n=a} = \{\overline{f} \mid f \in I\}.$$

For clarity, $I_{x_n=a}$ is an ideal of the polynomial ring $k[x_1, \ldots, x_{n-1}]$.

Now we claim that there exists $a \in k$ such that $I_{x_n=a} \subset k[x_1, \ldots, x_{n-1}]$. Before we prove the claim, note that if we prove the claim then we can iteratively use the claim to generate elements $a_1, \ldots, a_n \in k$ such that the ideal

$$I_{x_n=a_n,\ldots,x_1=a_1} \subset k.$$

Since k is a field, $I_{x_n=a_n,\ldots,x_1=a_1} = 0$ and consequently, $(a_1, \ldots, a_n) \in V(I)$. So, $V(I) \neq \emptyset$, which is the statement we want to prove.

To show: (a) There exists $a \in k$ such that $I_{x_n=a} \subset k[x_1, \ldots, x_{n-1}]$.

(a) There are two cases to consider.

Case 1: $I \cap k[x_n] \neq \{0\}$.

If $I \cap k[x_n] \neq \{0\}$ then let $f \in I \cap k[x_n]$ be non-zero. Note that f must be non-constant as otherwise, $1 \in I \cap k[x_n] \subseteq I$ which contradicts the assumption that $I \neq k[x_1, \ldots, x_n]$.

Now since k is algebraically closed, we can write

$$f = c\prod_{i=1}^{r}(x_n - b_i)^{m_i}$$

where $c, b_1, \ldots, b_r \in k$ and $c \neq 0$. Suppose for the sake of contradiction that $I_{x_n=b_i} = k[x_1, \ldots, x_{n-1}]$ for $i \in \{1, 2, \ldots, r\}$. If $i \in \{1, 2, \ldots, r\}$ then there exists $B_i \in I$ such that $B_i(x_1, \ldots, x_{n-1}, b_i) = 1$. A quick computation reveals that

$$1 = B_i(x_1, \ldots, x_{n-1}, b_i) = B_i(x_1, \ldots, x_{n-1}, x_n - (x_n - b_i)) = B_i + A_i(x_n - b_i)$$

for some $A_i \in k[x_1, \ldots, x_n]$. Since this holds for $i \in \{1, 2, \ldots, r\}$, we deduce that

$$1 = \prod_{i=1}^{r}(A_i(x_n - b_i) + B_i)^{m_i} = A\prod_{i=1}^{r}(x_n - b_i)^{m_i} + B$$

where $A = \prod_{i=1}^{r} A_i^{m_i}$ and $B \in I$. But,

$$1 = A\prod_{i=1}^{r}(x_n - b_i)^{m_i} + B = Ac^{-1}f + B \in I.$$

So, $I = k[x_1, \ldots, x_n]$, which contradicts the assumption that $I \neq k[x_1, \ldots, x_n]$. Therefore, there exists $b_i \in k$ such that $I_{x_n=b_i} \neq k[x_1, \ldots, x_{n-1}]$. This proves the claim in this particular case.

Case 2: $I \cap k[x_n] = \{0\}$.

Assume that $I \cap k[x_n] = \{0\}$. Let $G = \{g_1, \ldots, g_t\}$ be a Gröbner basis for I where we impose lexicographic order on $k[x_1, \ldots, x_n]$ with $x_1 > \cdots > x_n$. Write

$$g_i = c_i(x_n)x^{\alpha_i} + (\text{Terms with degree less than } x^{\alpha_i})$$

where $c_i(x_n) \in k[x_n]$ is non-zero and x^{α_i} is a monomial in $x_1, \ldots, x_{n-1}$.

Next, pick $a \in k$ such that if $i \in \{1, 2, \ldots, t\}$ then $c_i(a) \neq 0$. This is possible because algebraically closed fields such as k are infinite. The ideal $I_{x_n=a} \subseteq k[x_1, \ldots, x_{n-1}]$ has basis given by

$$\{\overline{g}_1, \overline{g}_2, \ldots, \overline{g}_t\}$$

where $\overline{g}_i = g_i(x_1, \ldots, x_{n-1}, a)$. If we substitute $x_n = a$ into our previous expression for g_i, we find that $LT(\overline{g}_i) = c_i(a)x^{\alpha_i}$ because $c_i(a) \neq 0$.

We claim that if $i \in \{1, 2, \ldots, t\}$ then $LT(\overline{g}_i)$ is non-constant. Suppose for the sake of contradiction that $LT(\overline{g}_i) = c_i(a)x^{\alpha_i}$ is constant. Then, $x^{\alpha_i} = 1$ and $g_i = c_i \in I \cap k[x_n] = 0$. This contradicts the fact that $c_i(x_n) \neq 0$. So, $LT(\overline{g}_i)$ is non-constant.

To show: (aa) $\{\overline{g}_1, \ldots, \overline{g}_t\}$ forms a Gröbner basis for $I_{x_n=a}$.

(aa) Assume that $g_i, g_j \in G$ and compute the S-polynomial

$$S(g_i, g_j) = c_j(x_n)x^{\gamma-\alpha_i}g_i - c_i(x_n)x^{\gamma-\alpha_j}g_j.$$

In the above expression, the subtraction $\gamma - \alpha_i$ is done entrywise. Also, $x^\gamma = \text{lcm}(x^{\alpha_i}, x^{\alpha_j})$. By construction of $S(g_i, g_j)$, $x^\gamma > LT(S)$. This can be seen clearly once we write

$$S(g_i, g_j) = c_j(x_n)x^{\gamma-\alpha_i}(c_i(x_n)x^{\alpha_i} + \ldots) - c_i(x_n)x^{\gamma-\alpha_j}(c_j(x_n)x^{\alpha_j} + \ldots).$$

Since $S(g_i, g_j) \in I$, we can express it using the standard representation

$$S(g_i, g_j) = \sum_{l=1}^{t} A_l g_l.$$

By substituting $x_n = a$, we find that

$$c_j(a)x^{\gamma-\alpha_i}\overline{g}_i - c_i(a)x^{\gamma-\alpha_j}\overline{g}_j = \overline{S(g_i, g_j)} = \sum_{l=1}^{t} \overline{A}_l \overline{g}_l.$$

Now $LT(\overline{g}_i) = c_i(a)x^{\alpha_i}$ which means that $\overline{S(g_i, g_j)}$ is the S-polynomial $S(\overline{g}_i, \overline{g}_j)$, up to the non-zero constant $c_i(a)c_j(a)$. If $A_l g_l \neq 0$ then

$$x^\gamma > LT(S) \geq LT(A_l g_l).$$

Therefore, if $\overline{A}_l \overline{g}_l \neq 0$ then

$$x^\gamma > LT(\overline{A}_l \overline{g}_l).$$

Since $x^\gamma = \text{lcm}(LM(\overline{g}_i), LM(\overline{g}_j))$, $S(\overline{g}_i, \overline{g}_j)$ has a lcm representation for $i, j \in \{1, 2, \ldots, t\}$ (see [CLO15, Chapter 3, §9, Definition 5]). So, $\{\overline{g}_1, \ldots, \overline{g}_t\}$ is a Gröbner basis for $I_{x_n=a}$. This uses [CLO15, Chapter 2, §9, Theorem 6].

(a) Since $\{\overline{g}_1, \ldots, \overline{g}_t\}$ is a Gröbner basis for $I_{x_n=a}$, $1 \notin I_{x_n=a}$ because 1 is not divisible by any of the non-constant leading terms $LT(\overline{g}_i)$. Therefore, $I_{x_n=a} \neq k[x_1, \ldots, x_{n-1}]$, which proves the claim in part (a) and gives the desired result. □

In the statement of the weak Nullstellensatz in Theorem 4.1.1, set $k = \mathbb{C}$. In this sense, the weak Nullstellensatz can be thought of as a generalisation of the fundamental theorem of algebra to multivariable polynomials because it (informally) states that **every system of polynomials which generates an ideal strictly smaller than $\mathbb{C}[x_1, \ldots, x_n]$ must have a common zero in $\mathbb{C}^n$.**

Before we proceed to Hilbert's Nullstellensatz, let us describe an application of Theorem 4.1.1. The *consistency problem* asks whether a system of polynomial equations $f_1 = \cdots = f_s = 0$, where $f_1, \ldots, f_s \in \mathbb{C}[x_1, \ldots, x_n]$, has a common solution in $\mathbb{C}^n$.

By Theorem 4.1.1, the polynomials do not have a common root in $\mathbb{C}^n$ if and only if $V(f_1, \ldots, f_s) = \emptyset$ if and only if $1 \in (f_1, \ldots, f_s)$. It is easy to verify that $\{1\}$ is the unique reduced basis of the ideal $(1) = k[x_1, \ldots, x_n]$. Thus, we devise the following algorithm to solve the consistency problem.

1. Suppose that $f_1, \ldots, f_s \in \mathbb{C}[x_1, \ldots, x_n]$ and we want to determine if the equations $f_1 = f_2 = \cdots = f_s = 0$ have a common solution in $\mathbb{C}^n$.
2. Let $I = (f_1, \ldots, f_s)$. Compute the reduced Gröbner basis G for I.
3. If $G = \{1\}$ then there is no common solution in $\mathbb{C}^n$. If $G \neq \{1\}$ then there exists $(a_1, \ldots, a_n) \in V(f_1, \ldots, f_s)$ such that

$$f_1(a_1, \ldots, a_n) = \cdots = f_s(a_1, \ldots, a_n) = 0.$$

Actually, the above algorithm is valid over any algebraically closed field because Theorem 4.1.1 holds for an algebraically closed field.

Example 4.1.1. Let $J = (x^2 + y^2 - 1, y - 1) \subseteq \mathbb{C}[x, y]$ be an ideal. We impose lexicographic order on $\mathbb{C}[x, y]$ with $x >_{lex} y$. We want to use the algorithm described previous to determine whether the equations

$$x^2 + y^2 - 1 = 0 \qquad \text{and} \qquad y - 1 = 0$$

have a common solution in $\mathbb{C}^2$. The first step is to compute a reduced Gröbner basis for J. Let $G = \{x^2 + y^2 - 1, y - 1\}$

We begin with Buchberger's algorithm. We compute the S-polynomial

$$S(x^2 + y^2 - 1, y - 1) = y(x^2 + y^2 - 1) - x^2(y - 1) = x^2 + y^3 - y.$$

If we divide $x^2 + y^3 - y$ by the polynomials in G, we obtain the remainder

$$\overline{x^2 + y^3 - y}^G = 0.$$

Thus, Buchberger's algorithm terminates and we deduce that $G = \{x^2 + y^2 - 1, y - 1\}$ is a Gröbner basis for J. In particular, it is a minimal Gröbner basis.

However, G is not a reduced Gröbner basis because the term $y^2 \in (LT(y - 1)) = (y)$. Following the proof of Theorem 2.6.2, we compute the remainder

$$\overline{x^2 + y^2 - 1}^{y-1} = x^2.$$

By Theorem 2.6.2, $G' = \{x^2, y - 1\}$ is a minimal Gröbner basis for J. We can check directly from the definition that G' is the desired reduced Gröbner basis for J.

Now since $G' = \{x^2, y - 1\} \neq \{1\}$, the weak Nullstellensatz (see Theorem 4.1.1) tells us that the equations

$$x^2 + y^2 - 1 = 0 \qquad \text{and} \qquad y - 1 = 0$$

have a common solution, which is $(x, y) = (0, 1) \in \mathbb{C}^2$.

With this information, we will now answer the question in [CLO15, Chapter 4, §1, Exercise 2] by finding a polynomial $f \in I(V(J))$ such that $f \notin J$. First, we have

$$V(J) = \{(x, y) \in \mathbb{C}^2 \mid x^2 + y^2 - 1 = 0, y - 1 = 0\} = \{(0, 1)\}$$

and

$$I(V(J)) = \{f \in \mathbb{C}[x, y] \mid f(0, 1) = 0\}.$$

Observe that the polynomial $xy \in I(V(J))$. We claim that $xy \notin J$. If we divide xy by the polynomials in the reduced Gröbner basis $G' = \{x^2, y - 1\}$, we obtain the remainder

$$\overline{xy}^{G'} = x.$$

Since $\overline{xy}^{G'} \neq 0$, we can invoke Theorem 2.5.2 to deduce that $xy \notin J$ as required.

4.2 A proof of Hilbert's Nullstellensatz

Even though we have the weak Nullstellensatz in Theorem 4.1.1, the correspondence between ideals and affine variety is still not one-to-one, even if we work over an algebraically closed field such as $\mathbb{C}$. For instance, the affine varieties $V(x) = V(x^2) = \{0\} \subseteq \mathbb{C}$. As a multivariable example, $V(x^2, y) = V(x, y) = \{(0, 0)\} \subseteq \mathbb{C}^2$. The issue here is that a power of a polynomial vanishes on the same set as the original polynomial. That is, if $f \in k[x_1, \ldots, x_n]$ and $m \in \mathbb{Z}_{>0}$ then $V(f) = V(f^m)$.

The Hilbert Nullstellensatz states that over an algebraically closed field k, the above reason is the **only reason** why different ideals can give the same variety.

Theorem 4.2.1 (Hilbert's Nullstellensatz)**.** *Let k be an algebraically closed field and $f, f_1, \ldots, f_s \in k[x_1, \ldots, x_n]$. Then, $f \in I(V(f_1, \ldots, f_s))$ if and only if there exists $m \in \mathbb{Z}_{>0}$ such that*

$$f^m \in (f_1, \ldots, f_s).$$

Proof. Assume that k is an algebraically closed field and $f, f_1, \ldots, f_s \in k[x_1, \ldots, x_n]$.

To show: (a) If there exists $m \in \mathbb{Z}_{>0}$ such that $f^m \in (f_1, \ldots, f_s)$ then $f \in I(V(f_1, \ldots, f_s))$.

(b) If $f \in I(V(f_1, \ldots, f_s))$ then there exists $m \in \mathbb{Z}_{>0}$ such that $f^m \in (f_1, \ldots, f_s)$.

(a) Assume that there exists $m \in \mathbb{Z}_{>0}$ such that $f^m \in (f_1, \ldots, f_s)$. Then, f^m must vanish on $V(f_1, \ldots, f_s)$. Consequently, f must also vanish on $V(f_1, \ldots, f_s)$ and $f \in I(V(f_1, \ldots, f_s))$.

(b) Assume that $f \in I(V(f_1, \ldots, f_s))$. Consider the ideal

$$I' = (f_1, \ldots, f_s, 1 - yf) \subseteq k[x_1, \ldots, x_n, y].$$

To show: (ba) $V(I') = \emptyset$.

(ba) Let $(a_1, \ldots, a_{n+1}) \in k^{n+1}$. There are two cases to consider here.

Case 1: Suppose that $(a_1, \ldots, a_n) \in k^n$ is a common zero of $f_1, \ldots, f_s$. Since $f \in I(V(f_1, \ldots, f_s))$, $f(a_1, \ldots, a_n) = 0$. So,

$$1 - yf = 1 - a_{n+1} f(a_1, \ldots, a_n) = 1 \neq 0.$$

Consequently, $(a_1, \ldots, a_{n+1}) \notin V(I')$.

Case 2: Now suppose that $(a_1, \ldots, a_n)$ is not a common zero of $f_1, \ldots, f_s$. Then, there exists $i \in \{1, 2, \ldots, s\}$ such that $f_i(a_1, \ldots, a_n) \neq 0$. By considering f_i as a polynomial in $x_1, \ldots, x_n, y$, we have $f_i(a_1, \ldots, a_n, a_{n+1}) \neq 0$. Therefore, $(a_1, \ldots, a_{n+1}) \notin V(I')$.

We combine the two cases above to show that if $(a_1, \ldots, a_{n+1}) \in k^{n+1}$ then $(a_1, \ldots, a_{n+1}) \notin V(I')$. Therefore, $V(I') = \emptyset$ as claimed.

(b) Since $V(I') = \emptyset$, the weak Nullstellensatz (see Theorem 4.1.1) tells us that $1 \in I'$. There exists $p_1, \ldots, p_s, q \in k[x_1, \ldots, x_n, y]$ such that

$$1 = \Big(\sum_{i=1}^{s} p_i(x_1, \ldots, x_n, y) f_i \Big) + q(x_1, \ldots, x_n, y)(1 - yf).$$

Let $y = 1/f(x_1, \ldots, x_n)$. Then,

$$1 = \sum_{i=1}^{s} p_i(x_1, \ldots, x_n, 1/f) f_i.$$

Think of the RHS as a rational function, whose denominators involve powers of f. Choose $m \in \mathbb{Z}_{>0}$ sufficiently large so that multiplying both sides by f^m clears denominators on the RHS. So,

$$f^m = \sum_{i=1}^{s} A_i f_i$$

for some polynomials $A_i \in k[x_1, \ldots, x_n]$. Therefore, $f^m \in (f_1, \ldots, f_s)$ as required. □

4.3 Radical ideals and the strong Nullstellensatz

One question remains unanswered by the Hilbert Nullstellensatz in Theorem 4.2.1 — what type(s) of ideals are ideals of an affine variety? The key observation lies with the following lemma.

Lemma 4.3.1. *Let $V \subseteq k^n$ be an affine variety and $m \in \mathbb{Z}_{>0}$. If $f^m \in I(V)$ then $f \in I(V)$.*

Proof. Assume that $V \subseteq k^n$ is an affine variety and $m \in \mathbb{Z}_{>0}$. Assume that $f^m \in I(V)$. If $a \in V$ then $(f(a))^m = 0$, which holds if and only if $f(a) = 0$. Hence, $f \in I(V)$. □

The property of an ideal of an affine variety given in Lemma 4.3.1 is important enough to warrant its own definition.

Definition 4.3.1. Let R be a commutative ring and $I \subseteq R$ be an ideal. We say that I is **radical** if the following statement is satisfied: If $m \in \mathbb{Z}_{>0}$, $f \in R$ and $f^m \in I$ then $f \in I$.

Lemma 4.3.1 tells us that the ideal of an affine variety $I(V)$ is a radical ideal. Theorem 4.2.1 tells us that an arbitrary ideal J fails to equal $I(V(J))$ if J contains integer powers f^m of a polynomial $f \notin I$ — that is, if J fails to be a radical ideal. This statement hints at a strong connection between affine varieties and radical ideals. To elucidate this, we require the following definition.

Definition 4.3.2. Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal. The **radical** of I, denoted by $\sqrt{I}$, is the set

$$\sqrt{I} = \{f \in k[x_1, \ldots, x_n] \mid \text{There exists } m \in \mathbb{Z}_{>0} \text{ such that } f^m \in I\}.$$

The radical of ideals is still an ideal, as the following lemma suggests.

Lemma 4.3.2. *Let $I \subseteq k[x_1, \ldots, x_n]$. Then, $\sqrt{I}$ is an ideal in $k[x_1, \ldots, x_n]$ such that $I \subseteq \sqrt{I}$. Moreover, $\sqrt{I}$ is a radical ideal.*

Proof. Assume that $I \subseteq k[x_1, \ldots, x_n]$ is an ideal and $\sqrt{I}$ is the radical of I.

To show: (a) $I \subseteq \sqrt{I}$.

(b) $\sqrt{I}$ is an ideal of $k[x_1, \ldots, x_n]$.

(c) $\sqrt{I}$ is a radical ideal.

(a) Assume that $f \in I$. Then, $f \in \sqrt{I}$ because $f = f^1 \in I$. Therefore, $I \subseteq \sqrt{I}$.

(b) First, assume that $f, g \in \sqrt{I}$. Then, there exists $m, n \in \mathbb{Z}_{>0}$ such that $f^m \in I$ and $g^n \in I$. Consider the polynomial

$$(f+g)^{m+n-1} = \sum_{\ell=0}^{m+n-1} \binom{m+n-1}{\ell} f^\ell g^{m+n-1-\ell}.$$

If $\ell \in \{1, 2, \ldots, m-1\}$ then $f^\ell g^{m+n-1-\ell} \in I$ because $g^n \in I$. If $\ell \in \{m, m+1, \ldots, m+n-1\}$ then $f^\ell g^{m+n-1-\ell} \in I$ because $f^m \in I$. Hence, $(f+g)^{m+n-1} \in I$ and consequently, $f + g \in \sqrt{I}$.

Now assume that $h \in k[x_1, \ldots, x_n]$. Since $f^m \in I$, $(hf)^m = h^m f^m \in I$. Therefore, $hf \in \sqrt{I}$. So, $\sqrt{I}$ is an ideal of $k[x_1, \ldots, x_n]$.

(c) Assume that $p \in k[x_1, \ldots, x_n]$, $m \in \mathbb{Z}_{>0}$ and $p^m \in \sqrt{I}$. We want to show that $p \in \sqrt{I}$. Since $p^m \in \sqrt{I}$, there exists $n \in \mathbb{Z}_{>0}$ such that $(p^m)^n = p^{mn} \in I$. Since $mn \in \mathbb{Z}_{>0}$ satisfies $p^{mn} \in I$, we deduce that $p \in \sqrt{I}$. So, $\sqrt{I}$ must be a radical ideal. □

Now we can restate Theorem 4.2.1 in terms of ideals. The following theorem is often referred to as the **strong Nullstellensatz**.

Theorem 4.3.3 (Strong Nullstellensatz)**.** *Let k be an algebraically closed field and $J \subseteq k[x_1, \ldots, x_n]$ be an ideal. Then,*

$$I(V(J)) = \sqrt{J}.$$

Proof. Assume that k is an algebraically closed field and $J \subseteq k[x_1, \ldots, x_n]$ is an ideal.

To show: (a) $\sqrt{J} \subseteq I(V(J))$.

(b) $I(V(J)) \subseteq \sqrt{J}$.

(a) Assume that $f \in \sqrt{J}$. Then, there exists $m \in \mathbb{Z}_{>0}$ such that $f^m \in J$. So, $f^m \in I(V(J))$ and by Lemma 4.3.1, $f \in I(V(J))$. So, $\sqrt{J} \subseteq I(V(J))$.

(b) Assume that $g \in I(V(J))$. By Theorem 4.2.1, there exists $m \in \mathbb{Z}_{>0}$ such that $g^m \in J$. Therefore, $g \in \sqrt{J}$ and $I(V(J)) \subseteq \sqrt{J}$.

By combining parts (a) and (b), we deduce that

$$I(V(J)) = \sqrt{J}$$

as required. □

Bibliography

[Cha22] Y. Chan. *Wedge product matrices and applications*, University of Melbourne, 2022.

[CLO15] D. Cox, J. Little, D. O'Shea. *Ideals, Varieties and Algorithms An Introduction to Computational Algebraic Geometry and Commutative Algebra*, 4th edition, Springer International Publishing, Switzerland, 2015, ISBN: 978-3-319-16720-6, MR3330490.

[Mur22] D. Murfet. *MAG Lecture 5 Dickson's Lemma*, July 13th 2022, Available at: http://www.therisingsea.org/notes/mag/MAG1-5.pdf

[Ram22] A. Ram. *Lecture 2, 2 March 2022: Macdonald polynomials*, March 2nd 2022, Available at: http://math.soimeme.org/ arun-ram/Teaching/2022MacPolys/Lect2MacPolys2022.pdf