Invariant theory of finite groups

Mashal Mehr
Bachelor thesis
Bachelor of Mathematics and Physics
Supervisor: Prof. Martijn Kool

June 9, 2019

Utrecht University

Contents

1 Introduction

Invariant theory started in the 19th century as the study of invariant algebraic forms. Many important theorems were proved by mathematicians while they were studying invariant theory. For example, Hilbert proved his basis theorem and the Nulstellensatz while working on invariant theory. Over the years several discoveries in this field form the basis of algebraic geometry as we know it today. In this paper we will take a step back and look at invariant theory through the lens of polynomials. Through this we will give an elementary introduction to invariant theory. In particular, we do not presume prior knowledge of group thory.

We will begin by introducing polynomials in n variables and examine their properties in comparison to polynomials in one variable. Furthermore we will look at ideals and discuss how we can construct and characterize them given a certain set of polynomials. We then construct a division algorithm for multivariate polynomials and afterwards we will study Dickson's lemma. This lemma describes how we can finitely generate a monomial ideal and forms the crux of Hilbert's basis theorem. Hilbert's basis theorem is an important result as it generalizes the results of Dickson's lemma towards all ideals. Additionally we introduce Gröbner bases in the proof of this theorem. Together with the division algorithm these bases form a powerful tool to characterize polynomials of an ideal.

We move on to symmetric polynomials and begin with introducing the elementary symmetric polynomials. These polynomials form the building blocks for all symmetric polynomials and we will show this in the fundamental theorem of symmetric polynomials. Furthermore we will look at ways to determine whether a polynomials is symmetric using the division algorithm. Finally we will introduce the power sums and examine how they are related the elementary symmetric polynomials.

In the last chapter we will generalize the notion of symmetric polynomials, towards invariant polynomials of finite matrix groups. Here we will examine the sets that contain these invariant polynomials and determine how we can characterize their elements. Furthermore we will introduce the Reynolds operator. This operator allows us to generate invariant polynomials and additionally we can generate all invariant polynomials using this powerful tool.

This thesis is based on the book *Ideals, Varieties and Algorithms* by Cox, Little and O'Shea[1]. In particular I made use of chapters 1,2 and 7.

2 Polynomials, ideals and orderings

In this chapter we will discuss polynomials in n variables and some basic algebraic objects related to polynomials, namely ideals. Furthermore we will discuss orderings on polynomials in n variables and examine how this affects the division algorithm in multiple variables.

2.1 Polynomials in n variables

Before we can discuss polynomials however, we need to know what a field is. Broadly speaking a field is a set on which addition, subtraction, multiplication and division, as we know it, is defined. Common examples of fields are the real numbers $\mathbb{R}$ and complex numbers $\mathbb{C}$. On the other hand the integers $\mathbb{Z}$ are not a field as division is not properly defined. A formal definition may be found at [2, Def 1.1.1]

Now that we have defined fields we can look at polynomials. We will start with the most elementary form a polynomial can take, namely a monomial.

Definition 2.1 *A monomial in $x_1, x_2, \ldots, x_n$ is a product of the form*

$$x_1^{\alpha_1} \cdot x_2^{\alpha_2} \ldots x_n^{a_n}$$

where the exponents α_i are nonnegative. The total degree of this monomial is given by $\alpha_1 + \alpha_2 + \cdots + \alpha_n$

We usually shorten this notation by representing the exponents with the n-tuple $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_2)$. Thus the above monomial is written as.

$$x^\alpha = x_1^{\alpha_1} \cdot x_2^{\alpha_2} \ldots x_n^{a_n}$$

The total degree is then given by $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$.

In general a polynomial is not just given by a single monomial, but a linear combination of monomials with coefficients in k. To be more precise a polynomial is defined as follows.

Definition 2.2 *A polyomial f over the field k in n variables is a linear combination of monomials with coefficients in k. We will write it as follows*

$$f = \sum_\alpha a_\alpha x^\alpha, \ a_\alpha \in k$$

where we sum over a finite number of n-tuples α. The set of all these polynomials is called $k[x_1, x_2, \ldots, x_n]$.

It is easy to show that the sum and product of a polynomial is again a polynomial. Furthermore we say that the polynomial f divides a polynomial g, if there exists a polynomial $h \in k[x_1, x_2, \ldots, x_n]$ such that $g = fh$. In fact $k[x_1, x_2, \ldots, x_n]$ acts almost like a field, except for the existence of multiplicative inverse e.g. $1/x^\alpha$ is not a polynomial. This kind of a structure is called a commutative ring and thus we will call $k[x_1, x_2, \ldots, x_n]$ a polynomial ring.

For polynomials in a small number of variables we will usually leave out the indices. So for example we will use $k[x, y, z]$ for 3 variables.

Similar to how we can talk about the total degree of a monomial, we can talk about the total degree of a polynomial.

Definition 2.3 *Let $f = \sum_\alpha a_\alpha x^\alpha$ be a nonzero polynomial in $k[x_1, x_2, \ldots, x_n]$. The total degree of the polynomial deg(f) is the maximal $|\alpha|$ such that a_α is nonzero. The degree of the zero polynomial is undefined.*

Note that for a polynomial in multiple variables the total degree is not uniquely determined. Consider for example the polynomial,

$$f = -\frac{7}{2}x^1y^4 + 3x^3y^2 + y^2$$

of total degree 5. Here both the first and second term have maximal total degree. This is not possible for polynomials in one variable, as terms of the same degree can simply be added up. In chapter 2.3 we will discuss this further.

Let us introduce the affine space now

Definition 2.4 *Let n be a positive integer and k field. The n-dimensional affine space over k is defined as*

$$k^n = \{(a_1, a_2, \ldots, a_n) \quad | \quad a_1, a_2, \ldots, a_n \in k\}$$

The most common example of an affine space is the n-dimensional Euclidean space $\mathbb{R}^n$

The affine space allows us to identify a polynomial $f = \sum_\alpha a_\alpha x^\alpha$ as a function from $k^n \to k$. Evaluating the polynomial at a point $(b_1, b_2, \ldots, b_n)$ of k^n is equivalent to making the substitution $x_i \mapsto b_i$ for all $i = 1, 2, \ldots, n$. Since the coefficients a_α are elements of k, we find that $f(b_1, b_2, \ldots, b_n) \in k$ and thus the function is well-defined. This has interesting consequences for the zero function, the function such that $f(a_1, a_2, \ldots, a_n) = 0$ for all $(a_1, a_2, \ldots, a_n) \in k^n$. Namely this function is not necessary the zero polynomial, the polynomial whose coefficients are equal to zero, for all fields. Consider for example a field with only two elements $\{0, 1\}$ with the property that $1 + 1 = 0$. It is an easy exercise to confirm that this forms a field and it is usually denoted as $\mathbb{F}_2$. For this let us examine the polynomial $xy(x + y) \in \mathbb{F}_2[x, y]$. Now xy is equal to zero whenever x or y are zero, but setting them both to one implies that $x + y$ is zero. Thus our nonzero polynomial gives us the zero function on $\mathbb{F}_2^2$.

The following proposition gives us a solution to this problem, namely working on an infinite field k.

Proposition 2.5 *Let k be an in infinite field, then $f \in k[x_1, x_2, \ldots, x_n]$ is the zero polynomial if and only if $f : k^n \to k$ is the zero function.*

Proof It is obvious that the zero polynomial gives us the zero function, so we have to show that f is the zero polynomial whenever $f(a_1, a_2, \ldots, a_n)$ for all $(a_1, a_2, \ldots, a_n) \in k^n$. For this we will use induction on the number of variables n.

Let $n = 1$ and $f \in k[x]$ such that it vanishes at all points of k. Note that a nonzero polynomial in $k[x]$ of degree m has at most m distinct roots. A proof can be found in [2, Thm 3.1.8]. Since f is the zero function it follows that $f(a) = 0$ for all $a \in k$. Hence f has infinitely many distinct roots, since k is infinite. From this we can conclude that f must be the zero polynomial.

Now let the proposition hold for $n-1$ and let $f \in k[x_1, x_2, \ldots, x_n]$ such that it vanishes at all points of k^n. By factoring out x_n out of the terms we can write f as follows,

$$f = \sum_{i=0}^{m} g_i(x_1, x_2, \ldots, x_{n-1}) \cdot x_n^i$$

where m is the total degree of f and g_i are polynomials in $k[x_1, x_2, \ldots, x_{n-1}]$.

We will show that g_i is the zero polynomial for all i, which results in f being the zero polynomial. We fix $(a_1, a_2, \ldots, a_{n-1}) \in k^{n-1}$ arbitrarily and this reduces our polynomial to $f(a_1, a_2, \ldots, a_{n-1}, x_n) \in k[x_n]$. Since f is the zero function it vanishes for all $a_n \in k[x_n]$. Furthermore since the proposition holds for $n = 1$ we find that $f(a_1, a_2, \ldots, a_{n-1}, x_n)$ is the zero polynomial and its coefficients are zero. Using the formula for f we find that the coefficients for $f(a_1, a_2, \ldots, a_{n-1}, x_n)$ are the polynomials $g_i(a_1, a_2, \ldots, a_{n-1})$ for all i. Since we chose $(a_1, a_2, \ldots, a_{n-1}) \in k^{n-1}$ arbitrarily we find that every $g_i(x_1, x_2, \ldots, x_{n-1})$ vanishes for all points in k^{n-1}. Hence each $g_i(x_1, x_2, \ldots, x_{n-1})$ is the zero function on k^{n-1}. Using our induction hypothesis we can conclude that g_i is the zero polynomial in $k[x_1, x_2, \ldots, x_{n-1}]$ for all i. Every coefficient in f is thus zero and so f is the zero polynomial. □

One immediate consequence of this proposition is that two polynomials are equal when their functions are equal.

Corollary 2.6 *Let k be an infinite field and let $f, g \in k[x_1, x_2, \ldots, x_n]$. Then $f = g$ in $k[x_1, x_2, \ldots, x_n]$ if and only if $f, g : k^n \to k$ are the same function.*

Proof Let $f, g \in k[x_1, x_2, \ldots, x_n]$ have the same function on k^n. Then $f - g$ gives us the zero function and by proposition 2.5 we have that $f - g$ is the zero polynomial. Hence $f = g$ in $k[x_1, x_2, \ldots, x_n]$. The converse of the proof is trivial. □

2.2 Ideals

So far we have only looked at polynomials themselves and not any algebraic objects related to them. We will start with the most basic object.

Definition 2.7 *A subset $I \subseteq k[x_1, x_2, \ldots, x_n]$ is called an ideal if it satisfies:*

1. $0 \in I$
2. *If $f, g \in I$ then we have that $f + g \in I$*
3. *If $f \in I$ then for all $h \in k[x_1, x_2, \ldots, x_n]$ we have that $fh \in I$*

Given a collection of polynomials $\{f_1, f_2, \ldots, f_m\}$ it is possible for us to construct an ideal with these polynomials.

Definition 2.8 *Let $f_1, f_2, \ldots, f_m$ be polynomials in $k[x_1, x_2, \ldots, x_n]$. We call the set*

$$\langle f_1, f_2, \ldots, f_m \rangle = \left\{ \sum_{i=1}^{m} h_i f_i \,|\, h_1, h_2, \ldots, h_m \in k[x_1, x_2, \ldots, x_n] \right\}$$

the ideal generated by $f_1, f_2, \ldots, f_m$.

As the name implies this set forms an ideal. To see this note that $0 \in \langle f_1, f_2, \ldots, f_m \rangle$ if we set $h_i = 0$ for all i. Furthermore let $f = \sum_{i=1}^{m} p_i f_i$ and

$g = \sum_{i=1}^{m} q_i f_i$ be elements of this set. Then $f + g = \sum_{i=1}^{m}(p_i + q_i)f_i$ and $fh = \sum_{i=1}^{m}(hp_i)f_i$, for all $h \in k[x_1, x_2, \ldots, x_n]$, are again elements of $\langle f_1, f_2, \ldots, f_m\rangle$.

Let us examine a polynomial f in this ideal. Given how we defined this ideal we can write f as

$$f = h_1 f_1 + h_2 f_2 + \cdots + h_m f_m$$

thus the collection $f_1, f_2, \ldots, f_m$ divides f and in a similar way every polynomial in $\langle f_1, f_2, \ldots f_m\rangle$ can be described this way. Thus we can describe $\langle f_1, f_2, \ldots f_m\rangle$ as the ideal that contains all polynomials that can be divided by $f_1, f_2, \ldots, f_m$. If we want to determine whether a polynomial f is an element of $\langle f_1, f_2, \ldots f_m\rangle$, the natural solution would be to divide by $f_1, f_2, \ldots, f_m$. This however is not as easy as it seems and we will examine it further in chpater 2.3 and 2.4.

Another question we could ask ourselves is whether a given ideal I can always be generated by a finite set of polynomials $\{f_1, f_2, \ldots, f_m\}$. In this case we call $\{f_1, f_2, \ldots, f_m\}$ a basis of I and as it turns out every ideal has such a basis. We will discuss this further in chapter 2.5, where we prove Hilbert's basis theorem.

2.3 Monomial orderings

Division of polynomials in $k[x_1, x_2, \ldots, x_n]$ brings several issues to the table, that we would not encounter in $k[x]$. For one we can order polynomials by degree in $k[x]$. This allows us to systematically work through the division.

Let us for example divide $f = 3x^4 - 2x^2 + 1$ by $g = x^2 - 1$. If we order the polynomials by degree, we see that the leading term of g divides $3x^4 = 3x^2 \cdot x^2$. So we subtract $3x^2 g$ from f, which gives us $-5x^2 + 1$. Repeating this process until the degree of f is lower than 2, gives us $f = (3x^2 - 5)(x^2 + 1) - 4$.

This is not so simple in $k[x_1, x_2, \ldots, x_n]$ as the notion of ordering is not as straightforward as we have seen in chapter 2.1. To define an ordering on $k[x_1, x_2, \ldots, x_n]$ let us consider a monomial x^α. This monomial is represented by its exponent vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ which is an element of $\mathbb{Z}^n_{\geq 0}$. In this way we can represent every monomial with an element of $\mathbb{Z}^n_{\geq 0}$, so any ordering on $\mathbb{Z}^n_{\geq 0}$ automatically induces an ordering on our monomials. So for any ordering the statement $\alpha > \beta$, is equivalent to $x^\alpha > x^\beta$.

We want our orderings to satisfy several conditions to make sure we can work with them. Since we want to compare our monomials, we want our ordering to be a total order. For two distinct α and β in a total order we have that either $\alpha > \beta$ or $\beta > \alpha$ holds. While this might seem obvious not every ordering is a total order. Consider for example the ordering by subsets for sets. Then neither $\{\alpha\} \subseteq \{\beta\}$ nor $\{\beta\} \subseteq \{\alpha\}$ holds for distinct α and β. Another property of total orders is transitivity, that is to say that $\alpha > \gamma$ whenever we have that $\alpha > \beta$ and $\beta > \gamma$.

Furthermore we want for all $\alpha, \beta, \gamma \in \mathbb{Z}^n_{\geq 0}$, that $\alpha > \beta$ implies $\alpha + \gamma > \beta + \gamma$. Since we are working with monomials and want to divide, we want an ordering to be preserved after multiplication with another monomial. This condition arises from the fact that we get $x^\alpha \cdot x^\gamma = x^{\alpha+\gamma}$ after multiplication. So we want that $x^{\alpha+\gamma} > x^{\beta+\gamma}$ whenever $x^\alpha > x^\beta$. We will not have to worry about sums of monomials, since we simply add up the coefficients if the orders coincide.

Lastly we want our total order to be a well-order. That is to say for every nonempty subset $S \subseteq \mathbb{Z}^n_{\geq 0}$ there exists a least element $\sigma \in S$, such that for every $\alpha \in S$ not equal to σ we have $\alpha > \sigma$. The reason for this is that a well-order

has the following property that comes in handy.

Lemma 2.9 *A total order $>$ on a set X is a well-order if and only if every decreasing sequence $a_0 > a_1 > a_2 > \dots$ eventually terminates.*

Proof We first assume that $>$ is a well-order. Let $a_0 > a_1 > a_2 > \dots$ be a decreasing sequence and denote with $S = \{a_i \mid i \in \mathbb{Z}_{\geq 0}\}$ the subset that contains this sequence. This set is nonempty since $a_0 \in S$ and thus has a least element s. Let m be the smallest integer such that $a_m = s$. Then for every $m' > m$ we have that $a_{m'} > a_m$, since $a_m = s$ is the least element. This leads to a contradiction so every sequence terminates.

For the converse we will use contraposition. Assume that $>$ is not a well-order. Thus every non-empty subset has no least element and let S be one of these subsets. Choose an element $a_0 \in S$. Since this is not the least element we can find another element a_1 such that $a_0 > a_1$. Again a_1 is not a least element so we can repeat this process over and over. This gives us an infinitely strictly decreasing sequence

$$a_0 > a_1 > a_2 > \dots$$

and this concludes our proof □

We will make use of this property in algorithms, to make sure that they terminate at some point.

To recap this all we arrive at the following definition.

Definition 2.10 *A monomial ordering $>$ on $k[x_1, x_2, \dots, x_n]$ is a relation $>$ on $\mathbb{Z}^n_{\geq 0}$ satisfying the following conditions:*

1. *$>$ is a total order on $\mathbb{Z}^n_{\geq 0}$*
2. *For all $\alpha, \beta, \gamma \in \mathbb{Z}^n_{\geq 0}$ if $\alpha > \beta$ then $\alpha + \gamma > \beta + \gamma$*
3. *$>$ is a well-order on $\mathbb{Z}^n_{\geq 0}$*

The most prominent monomial ordering we will use is lexicographical ordering or lex order for short.

Definition 2.11 *Let $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ and $\beta = (\beta_1, \beta_2, \dots, \beta_n)$ be in $\mathbb{Z}^n_{\geq 0}$. We say that $\alpha >_{lex} \beta$ in lex order whenever the leftmost nonzero entry of $\beta - \alpha$ is negative. We will write $x^\alpha >_{lex} x^\beta$ whenever $\alpha > \beta$.*

We will look at some examples:

1. $(2, 0, 2) >_{lex} (1, 5, 3)$ since $\beta - \alpha = (-1, 5, 1)$
2. $(3, 2, 4) >_{lex} (3, 0, 1)$ since $\beta - \alpha = (0, -2, 3)$

In this way naturally we find that $x_1 >_{lex} x_2 >_{lex} \cdots >_{lex} x_n$ since

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

If we are working in $k[x, y, z]$ or some other ring where we use no subscripts we uphold alphabetical order. So $x >_{lex} y >_{lex} z$.

We still must check whether lex ordering satisfies the conditions of a monomial ordering.

Proposition 2.12 *Lex ordering on $\mathbb{Z}_{\geq 0}^n$ is a monomial ordering.*

Proof That lex order is a total order follows from the definition and the fact that ordering on $\mathbb{Z}_{\geq 0}$ is a well-order. Now let $\alpha, \beta, \gamma \in \mathbb{Z}_{\geq 0}^n$ such that $\alpha > \beta$. Then after adding γ to both α and β, we have that $(\beta + \gamma) - (\alpha + \gamma) = \beta - \alpha$. So the leftmost nonzero entry stays the same and thus $\alpha + \gamma > \beta + \gamma$

To show that lex ordering is a well-order we assume that there exists a strictly infinite decreasing sequence

$$\alpha_0 >_{lex} \alpha_1 >_{lex} \alpha_2 >_{lex} \dots$$

It follows that the first entry of the α_i is decreasing by definition of the lex order. Since $\mathbb{Z}_{\geq 0}$ is a well-ordering the first entry cannot form an infinite decreasing sequence. So there exists a $k \in \mathbb{Z}_{\geq 0}$ such that for all $l > k$ the first entry in α_l repeats. We can then begin at α_l for $l > k$ and repeat this procedure for the second entry. After repeating this procedure n times we thus find an $m \in \mathbb{Z}_{\geq 0}$, such that the sequence terminates at α_m. By lemma 2.9 we can conclude that $>_{lex}$ is a well-order and thus a monomial order. □

We will briefly go over some alternatives to lex order. While we will primarily use lex order in the proofs, we could just as well have used different monomial orders.

An example of such an order is graded lexicographic order, or grlex order in short. For this order we first sort the monomials by total degree and then use lex order whenever several monomials have the same degree. Similar to grlex order we have graded reverse lexicographic orderor grevlex order. Here we also first sort monomials by order, but for monomials of the same total degree we favour the rightmost smallest power.

To see how this works in practice let us consider the following polynomial $f = 2x^3 + xy^2z^2 - 4x^2z^3 - z \in k[x, y, z]$. In lex order we would sort this as

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

In grlex order however the second and third term have highest degree and thus the order would be

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

Similar to grlex order, grevlex order favours the second and third term. However $xy^2z^2 >_{grevlex} x^2z^3$, since z^2 has a lower power than z^3. So we get

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

Before we move on the division algorithm let us first introduce some terminology.

Definition 2.13 *Let $f = \sum_\alpha a_\alpha x^\alpha$ be a polynomial in $k[x_1, x_2, \dots, x_n]$. Furthermore let $>$ be a monomial order.*

1. *The multidegree of f is given by* $\text{multideg}(f) = \max\{\alpha \in \mathbb{Z}_{\geq 0}^n \mid a_\alpha \neq 0\}$.
2. *The leading monomial of f is given by* $\texttt{LM}(f) = x^{\text{multideg}(f)}$.

3. *The leading coefficient of f is given by* $\mathtt{LC}(f) = a_{\text{multideg}(f)}$.
4. *The leading term of f is given by* $\mathtt{LT}(f) = \mathtt{LC}(f) \cdot \mathtt{LM}(f)$.

For the multidegree we also have the following identities.

Lemma 2.14 *Let $f, g \in k[x_1, x_2, \ldots, x_n]$ be nonzero polynomials. Then we have*

1. $\text{multideg}(fg) = \text{multideg}(f) + \text{multideg}(g)$.
2. *If $f + g \neq 0$, then* $\text{multideg}(f + g) \leq \max(\text{multideg}(f), \text{multideg}(g))$. *Equality occurs whenever* $\text{multideg}(f) \neq \text{multideg}(g)$.

Proof For the first identity note that

$$\begin{aligned} x^{\text{multideg}(fg)} &= \mathtt{LM}(fg) = \mathtt{LM}(f) \cdot \mathtt{LM}(g) \\ &= x^{\text{multideg}(f)} \cdot x^{\text{multideg}(g)} = x^{\text{multideg}(f)+\text{multideg}(g)} \end{aligned}$$

so $\text{multideg}(fg) = \text{multideg}(f) + \text{multideg}(g)$.

For the second identity let $\text{multideg}(f) = \text{multideg}(g)$. Since their leading monomials have the same multidegree, we can simply add their coefficients which gives us

$$\mathtt{LT}(f + g) = (\mathtt{LC}(f) + \mathtt{LC}(g)) \cdot x^{\text{multideg}(f)}$$

If $\mathtt{LC}(f)$ is equal to $-\mathtt{LC}(g)$ then the leading term of $f + g$ vanishes and thus $\text{multideg}(f+g) < \text{multideg}(f) = \max(\text{multideg}(f), \text{multideg}(g))$. Otherwise we get an equality. In the case that the multidegree of f and g differ, the leading coefficient of $f + g$ is simply the leading coefficient of the polynomial whose multidegree is highest. Hence we get an equality □

2.4 A division algorithm in $k[x_1, x_2, \ldots, x_n]$

Now that we have monomial orderings we can begin to formulate a division algorithm in $k[x_1, x_2, \ldots, x_n]$. Returning to our problem raised in chapter 2.2, we want to know whether a polynomial $f \in k[x_1, x_2, \ldots, x_n]$ is an element of the ideal $\langle f_1, f_2, \ldots, f_m \rangle$. For this we want to divide by $f_1, f_2, \ldots f_m \in k[x_1, x_2, \ldots, x_n]$ and write f in the form

$$f = h_1 f_1 + h_2 f_2 + \cdots + h_m f_m + r$$

with $h_1, h_2, \ldots h_m \in k[x_1, x_2, \ldots, x_n]$ suitably chosen and $r \in k[x_1, x_2, \ldots, x_n]$ a remainder.

Similar to division in $k[x]$ we want to systematically eliminate leading terms by multiplying our f_i's with suitable monomials, such that the leading terms cancel. We will go over some examples to examine how this differs from the case in $k[x]$ and then state the algorithm.

For our first example let us divide $f = xy^2 + 3y + 1$ by $f_1 = xy + 1$ and $f_2 = y + 1$. Furthermore we will order f_1, f_2 as $F = (f_1, f_2)$ We will use lex order and let $x > y$. Here the leading term of f is given by $\mathtt{LT}(f) = xy^2$, and

is divisible by both $\texttt{LT}(f_1) = xy$ and $\texttt{LT}(f_2) = y$. Since we listed f_1 first we will divide f by $y \cdot f_1$. This then gives us

$$xy^2 + 3y + 1 - y \cdot f_1 = 2y + 1$$

The leading term is now given by $\texttt{LT}(2y + 1) = 2y$. Since we can not divide this by the leading term of f_1 we will use f_2. Thus we get that

$$2y + 1 - 2 \cdot f_2 = -1$$

Now neither $\texttt{LT}(f_1)$ nor $\texttt{LT}(f_2)$ divide -1, so our remainder is given by $r = -1$. Hence dividing f by f_1 and f_2 gives us

$$f = y \cdot (xy + 1) + 2 \cdot (y + 1) - 1$$

For another example we will look at $f = x^2y + xy^2 + y^2$ and divide it by $f_1 = xy - 1$ and $f_2 = y^2 - 1$. Again we will use lex order $x > y$ and order our division by $F = (f_1, f_2)$. The first two steps follow as before and we obtain

$$x^2y + xy^2 + y - x \cdot f_1 = xy^2 + x + y^2$$
$$xy^2 + x + y^2 - y \cdot f_1 = x + y^2 + y$$

Now both $\texttt{LT}(f_1) = xy$ and $\texttt{LT}(f_2) = y^2$ do not divide $\texttt{LT}(x + y^2 + y) = x$. In the one dimensional case the algorithm would terminate at this point, since every other term in the polynomials has lower degree than the leading term and thus cannot be divided. However in our case we find that y^2 is divisible by $\texttt{LT}(f_2) = y^2$. So we simply set x as the remainder r and continue our algorithm.

This leaves us with the polynomial $y^2 + y$. We divide by f_2 and obtain

$$y^2 + y - f_2 = y + 1$$

Which leaves us with $y + 1$. Since we cannot divide this polynomial any further we add this to our remainder $r = x$ and ultimately our remainder is given by $r = x + y + 1$. So after division we obtain

$$x^2y + xy^2 + y - x = (x + y) \cdot (xy - 1) + 1 \cdot (y^2 - 1) + x + y + 1$$

These two examples give a clear picture of how the division algorithm works. One important property of the algorithm is that the remainder is not divisible by any of the polynomials by which we are dividing.

Theorem 2.15 (The division algorithm in $k[x_1, x_2, \ldots, x_n]$) *Let $>$ be a monomial order on $\mathbb{Z}^n_{\geq 0}$ and let $F = (f_1, f_2, \ldots, f_s)$ be an ordered s-tuple of polynomials in $k[x_1, x_2, \ldots, x_n]$. Then every $f \in k[x_1, x_2, \ldots, x_n]$ can be written as*

$$f = q_1 f_1 + q_2 f_2 + \cdots + q_s f_s + r$$

where $q_i, r \in k[x_1, x_2, \ldots, x_n]$. Furthermore r is either 0 or a it is a linear combination of monomials with coefficients in k, which are not divisible by any

of $LT(f_1), LT(f_2), \ldots, LT(f_s)$. *We call* r *a remainder of* f *by division by* F. *Furthermore if* $q_i f_i \neq 0$ *then*

$$multidegree(f) \geq multidegree(q_i f_i)$$

Proof The proof and a detailed explanation of the algorithm is given in [1, p.64] □

One major way in which this division algorithm differs from the one-dimensional case, is that the remainder in $k[x_1, x_2, \ldots, x_n]$ is not uniquely determined. For this we return to our example $f = x^2y+xy^2+y^2$ with $f_1 = xy-1$ and $f_2 = y^2-1$. Once more we use lex order $x < y$, but this time we will order our polynomials as $F = (f_2, f_1)$. So we divide by f_2 first if possible. The division is then given by

$$\begin{aligned} x^2y + xy^2 + y^2 - x \cdot f_1 &= xy^2 + x + y^2 \\ xy^2 + x + y^2 - x \cdot f_2 &= 2x + y^2 \\ 2x + y^2 \to r &= 2x \\ y^2 - f_2 &= 1 \\ 1 \to r &= 2x + 1 \end{aligned}$$

where every line with an arrow denotes a step in the algorithm where the leading term cannot be divided by f_1 and f_2. Thus our polynomial f is given by

$$x^2y + xy^2 + y^2 = x \cdot (xy - 1) + (x + 1) \cdot (y^2 - 1) + 2x + 1$$

Now our remainder is different than before. Thus we see that our remainder is dependent on the way we order our division.

This problem causes some further complications when we look at finitely generated ideals. If we examine a polynomial $f \in \langle f_1, f_2, \ldots, f_m \rangle$, then we can write it as $f = q_1f_1 + q_2f_2 + \cdots + q_mf_m = q_1f_1 + q_2f_2 + \cdots + q_mf_m + 0$. So clearly a polynomial g lies in this ideal whenever its remainder by division on $f_1, f_2, \ldots, f_m$ is zero. This is however not a sufficient condition as the ordering of our division influences the remainder. This best shown through the following example:

Let $f = xy^2 - x \in k[x, y]$ and we divide this by $f_1 = xy - 1$ and $f_2 = y^2 - 1$. Using lex order $x > y$ and dividing by $F = (f_1, f_2)$ gives us

$$xy^2 - x = y \cdot (xy - 1) + 0 \cdot (y^2 - 1) + (-x + y)$$

If we reverse the order and divide by $F = (f_2, f_1)$, then we get

$$xy^2 - x = x \cdot (y^2 - 1) + 0 \cdot (xy - 1) + 0$$

Obviously we can conclude that $f \in \langle f_1, f_2 \rangle$ from the second division, but we cannot infer this from the first division.

Clearly the division algorithm in $k[x_1, x_2, \ldots, x_n]$ is not as nice as in $k[x]$. In chapter 2.6 we will look at ways to solve the problems that we have shown here. As it turns out we can resolve these issues by working in a specific basis called a Gröbner basis.

2.5 Dickson's lemma and Hilbert's basis theorem

One of the other questions we raised in chapter 2.2, was whether we could find a finite basis $\{f_1, f_2, \ldots, f_m\}$ for every ideal I in $k[x_1, x_2, \ldots, x_n]$. In this chapter we will provide the solution to this question and lay the groundwork for Gröbner bases.

To solve this question we make use of the fact that every nonzero polynomial f in $k[x_1, x_2, \ldots, x_n]$ has a unique leading term $\mathtt{LT}(f)$, once we fix a monomial order. Naturally an ideal generated by monomials, would be an appropriate place to start. So let us define monomial ideals in $k[x_1, x_2, \ldots, x_n]$.

Definition 2.16 *We call an ideal $I \subseteq k[x_1, x_2, \ldots, x_n]$ a monomial ideal, if there exists a subset $A \subseteq \mathbb{Z}^n_{\geq 0}$ such that I contains all polynomials which are finite sums of the form $\sum_{\alpha \in A} h_\alpha x^\alpha$ with $h_\alpha \in k[x_1, x_2, \ldots, x_n]$. We denote this with $I = \langle x^\alpha \,|\, \alpha \in A \rangle$.*

Note that we do not require A to be a finite subset, so I is a priori not finitely generated. An example of a monomial ideal is given by $I = \langle xy^3, x^2yz, z^5 \rangle \subseteq k[x, y, z]$

The following lemma tells us how we can characterize elements of a monomial ideal.

Lemma 2.17 *Let $I = \langle x^\alpha \,|\, \alpha \in A \rangle$ be a monomial ideal. Then a monomial x^β is an element of I if and only if there exists an $\alpha \in A$ such that x^α divides x^β.*

Proof If x^β is divisible by x^α for some $\alpha \in A$, then by definition of an ideal $x^\beta \in I$. For the converse let $x^\beta \in I$ and given by $x^\beta = \sum_{i=1}^s h_i x^{\alpha(i)}$, where $h_i \in k[x_1, x_2, \ldots, x_n]$ and $\alpha(i) \in A$. We can then write each h_i as a sum of its monomials, which gets us

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

After collecting every term of the same multidegree, the terms of the right hand side of the equation are divisible by some $x^{\alpha(i)}$. Thus the left hand side x^β must also be divisible by $x^{\alpha(i)}$. □

We can expand this idea further and show that a given polynomial f lies in a monomial ideal whenever its monomials lie in I.

Lemma 2.18 *Let I be a monomial ideal and let $f \in k[x_1, x_2, \ldots, x_n]$. Then the following are equivalent:*

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

Proof The implications 3.) ⇒ 2.) ⇒ 1.) and 2.) ⇒ 3.) are obvious. We will prove the implication 1.) ⇒ 3.).

Let $f \in I$ be given by $f = \sum_{i=0}^{s} h_i x^{\alpha(i)}$ where $h_i \in k[x_1, x_2, \ldots, x_n]$ and $\alpha(i) \in A$. If we expand a term in f in the same manner as in lemma 2.17 we get

$$h_i x^{\alpha(i)} = \sum_j c_j x^{\beta(i,j)} x^{\alpha(i)}$$

Since $x^{\alpha(i)} \in I$ divides $x^{\beta(i,j)} x^{\alpha(i)}$, we get by lemma 2.17 that the monomial $x^{\beta(i,j)} x^{\alpha(i)}$ lies in I. So f is a k-linear combination of monomials in I. This completes our proof. □

From this we can immediately see that a monomial ideal is uniquely determined by its monomials. Hence

Corollary 2.19 *Two monomial ideals are the same if and only if they contain the same monomials.*

The main result we are concerned with regards to monomial ideals in $k[x_1, x_2, \ldots, x_n]$ is to prove that they are finitely generated

Theorem 2.20 (Dickson's lemma) *Let $I = \langle x^\alpha \mid \alpha \in A \subseteq k[x_1, x_2, \ldots, x_n] \rangle$ be a monomial ideal. Then I can be written in the form $I = \langle x^{\alpha(1)}, x^{\alpha(2)}, \ldots, x^{\alpha(m)} \rangle$ where*
$\alpha(1), \alpha(2), \ldots, \alpha(m) \in A$.

Proof We will use a proof by induction on the number of variables, n. Let $n = 1$. Then I is generated by the monomials x_1^α, where $\alpha \in A \subseteq \mathbb{Z}_{\geq 0}$. Let β be the smallest element of A. Then since $\beta \leq \alpha$ for all $\alpha \in \mathbb{Z}_{\geq 0}$, we have that x_1^β divides all generators x_1^α. From this we can conlude that $I = \langle x^\beta \rangle$.

Now assume the theorem holds for $1, 2, \ldots, n-1$. To clear up the notation we will work in $k[x_1, x_2, \ldots, x_{n-1}, y]$. This way we can write monomials as $x^\alpha y^s$, where $\alpha \in \mathbb{Z}_{\geq 0}^{n-1}$ and $s \in \mathbb{Z}_{\geq 0}$.

Now let $I \subseteq k[x_1, x_2, \ldots, x_{n-1}, y]$ be a monomial ideal. We denote with J the ideal generated by the monomials x^α, with the property that $x^\alpha y^s \in I$ for some $s \geq 0$. Because J lies in $k[x_1, x_2, \ldots, x_{n-1}]$ our inductive hypothesis holds. Thus we can find finitely many monomials $x^{\alpha(i)}$ that generate J, so $J = \langle x^{\alpha(1)}, x^{\alpha(2)}, \ldots, x^{\alpha(s)} \rangle$.

By definition of J we find for all i that $x^{\alpha(i)} y^{s_i} \in I$ for some $s_i \geq 0$. Now let s be the largest among the s_i. Then for each $0 \geq l \geq n-1$ we will construct ideals $J_l \subseteq k[x_1, x_2, \ldots, x_{n-1}]$ generated by the monomials x^β, such that $x^\beta y^l \in I$. We can apply our inductive hypothesis again to find monomials $x^{\alpha_l(i)}$ such that $J_l = \langle x^{\alpha_l(1)}, x^{\alpha_l(2)}, \ldots, x^{\alpha_l(s_l)} \rangle$.

We claim that I is generated by the following monomials:

$$\begin{aligned}
\text{From } J &: x^{\alpha(1)} y^m, x^{\alpha(2)} y^m, \ldots, x^{\alpha(s)} y^m \\
\text{From } J_0 &: x^{\alpha_0(1)}, x^{\alpha_0(2)}, \ldots, x^{\alpha_0(s_0)} \\
\text{From } J_1 &: x^{\alpha_1(1)} y, x^{\alpha_1(2)} y, \ldots, x^{\alpha_1(s_1)} y \\
&\vdots \\
\text{From } J_{m-1} &: x^{\alpha_{m-1}(1)} y^{m-1}, x^{\alpha_{m-1}(2)} y^{m-1}, \ldots, x^{\alpha_{m-1}(s_{m-1})} y^{m-1}
\end{aligned}$$

The important thing to note is that every monomial in I is divisible by one of these monomials. To prove this let $x^\beta y^q \in I$. For $q \geq m$ there exists an i, such that $x^{\alpha(i)}y^m$ divides $x^\beta y^q$ by the construction of J. In the case that $q < m$ we can once more find an i, such that $x^{\alpha_q(i)}y^q$ divides $x^\beta y^q$ by the construction of J_q. Using lemma 2.17 we find that the above ideal generates an ideal with the same monomials as I and by Corollary 2.19 it follows that these are the same ideal. This proves our claim.

Now the only thing left is to prove that we can generate I using a finites set of generators in a given set of generators. We will switch back to variables in $x_1, x_2, \ldots, x_n$, so our ideal is given by $I = \langle x^\alpha \mid \alpha \in A\rangle \subseteq k[x_1, x_2, \ldots, x_n]$. Thus we need to show that I is generated by finitely many monomials x^α, where $\alpha \in A$. We have shown that $I = \langle x^{\beta(1)}, x^{\beta(2)}, \ldots, x^{\beta(s)}\rangle$ for some $x^{\beta(i)} \in I$. Since $x^{\beta(i)}$ lies in I there exists some $x^{\alpha(i)}$ with $\alpha(i) \in A$, such that $x^{\alpha(i)}$ divides $x^{\beta(i)}$ by lemma 2.17. Now let $f \in I = \langle x^{\beta(1)}, x^{\beta(2)}, \ldots, x^{\beta(s)}\rangle$ be nonzero. Then we can write it as follows

$$f = \sum_{i=0}^{s} f_i x^{\beta(i)} = \sum_{i=0}^{s} f_i x^{\gamma(i)} x^{\alpha(i)} = \sum_{i=0}^{s} (f_i x^{\gamma(i)}) x^{\alpha(i)}$$

Thus we have that $I \subseteq \langle x^{\alpha(1)}, \alpha^{\alpha(2)}, \ldots, x^{\alpha(s)}\rangle$ and we can conclude that $I = \langle x^{\alpha(1)}, \alpha^{\alpha(2)}, \ldots, x^{\alpha(s)}\rangle$. □

This theorem allows us to answer the question raised in chapter 2.4 for monomial ideals. Namely

Proposition 2.21 *Let $I = \langle x^{\alpha(1)}, x^{\alpha(2)}, \ldots, x^{\alpha(s)}\rangle$ be a monomial ideal. Then a polynomial f lies in I if and only if the remainder of f on division by $x^{\alpha(1)}, x^{\alpha(2)}, \ldots, x^{\alpha(s)}$ is zero.*

Proof Obviously if the remainder of f on division by $x^{\alpha(1)}, x^{\alpha(2)}, \ldots, x^{\alpha(s)}$ is zero, then f lies in I. So let us assume that $f \in I = \langle x^{\alpha(1)}, x^{\alpha(2)}, \ldots, x^{\alpha(s)}\rangle$. If we divide f by $x^{\alpha(1)}, x^{\alpha(2)}, \ldots, x^{\alpha(s)}$ the division algorithm gives us

$$f = h_1 x^{\alpha(1)} + h_2 x^{\alpha(2)} + \cdots + h_s x^{\alpha(s)} + r$$

where r is some remainder. By Lemma 2.18 we then have that $r \in I$ and thus there exists an $x^\alpha \in I$ that divides r by lemma 2.17. Since the $x^{\alpha(i)}$ generate I, there is some $x^{\alpha(i)}$ that divides x^α. Since r is not divisible by any $x^{\alpha(i)}$, it follows that r must be zero. □

For a general ideal I we still have to do some work, as we can only apply Dickson's lemma on monomial ideals. As we have stated before each polynomial $f \in I$ has a unique leading term $\mathtt{LT}(f)$, once we fix a monomial order. Thus we can construct a monomial using these leading terms as follows.

Definition 2.22 *Let $I \subseteq k[x_1, x_2, \ldots, x_n]$ be an ideal other than $\{0\}$ and fix a monomial ordering on $k[x_1, x_2, \ldots, x_n]$. Then:*

1. *We denote by $\mathtt{LT}(I)$ the set of leading terms of nonzero polynomials in I. So*

$$\mathtt{LT}(I) = \{cx^\alpha \mid \text{there exists } f \in I \text{ such that } \mathtt{LT}(f) = cx^\alpha\}$$

2. *We denote by* $\langle \mathtt{LT}(I) \rangle$ *the ideal generated by the elements of* $\mathtt{LT}(I)$

One might think that for a finitely generated ideal $I = \langle f_1, f_2, \ldots, f_m \rangle$, we find that $\langle \mathtt{LT}(f_1), \mathtt{LT}(f_2), \ldots, \mathtt{LT}(f_m) \rangle$ and $\langle \mathtt{LT}(I) \rangle$ are the same ideal. While by definition $\mathtt{LT}(f_i) \in \mathtt{LT}(I) \subseteq \langle \mathtt{LT}(I) \rangle$ implies that $\langle \mathtt{LT}(f_1), \mathtt{LT}(f_2), \ldots, \mathtt{LT}(f_m) \rangle \subseteq \langle \mathtt{LT}(I) \rangle$. The converse inclusion does not always hold.

Consider for example the ideal $I = \langle f_1, f_2 \rangle$ generated by $f_1 = x^3 - 2xy$ and $f_2 = x^2y + x - 2y^2$. Using lex ordering in $k[x, y]$, we find that

$$x \cdot (x^2y + x - 2y^2) - y \cdot (x^3 - 2x) = x^2$$

and so $x^2 \in I$. Its leading term $x^2 = \mathtt{LT}(x^2)$ is thus an element of $\langle \mathtt{LT}(I) \rangle$. On the other hand x^2 is not divisible by $\mathtt{LT}(f_1) = x^3$ or $\mathtt{LT}(f_2) = x^2y$, so by lemma 2.17 we find that $x^2 \notin \langle \mathtt{LT}(f_1), \mathtt{LT}(f_2) \rangle$.

We will adress this problem once we tackle Gröbner bases in the next chapter. For now we want to prove that $\langle \mathtt{LT}(I) \rangle$ is a monomial ideal. This allows us to find a finite basis for it using Dickson's lemma.

Proposition 2.23 *Let* $I \subseteq k[x_1, x_2, \ldots, x_n]$ *be an ideal different from* $\{0\}$. *Then we find that:*

1. $\langle \mathtt{LT}(I) \rangle$ *is a monomial ideal*
2. *There exists* $g_1, g_2, \ldots, g_m \in I$ *such that* $\langle \mathtt{LT}(I) \rangle = \langle \mathtt{LT}(g_1), \mathtt{LT}(g_2), \ldots, \mathtt{LT}(g_m) \rangle$

Proof 1) Let $\langle \mathtt{LM}(g) \,|\, g \in I/\{0\} \rangle$ be the monomial generated by the leading monomials of nonzero polynomials g in I. Since $\mathtt{LT}(g) = \mathtt{LC}(g) \cdot \mathtt{LM}(g)$ is an element of $\langle \mathtt{LM}(g) \,|\, g \in I/\{0\} \rangle$, we have that $\langle \mathtt{LT}(I) \rangle \subseteq \langle \mathtt{LM}(g) \,|\, g \in I/\{0\} \rangle$. Dividing by the leading coefficient gives us the reverse inclusion and thus $\langle \mathtt{LT}(I) \rangle = \langle \mathtt{LM}(g) \,|\, g \in I/\{0\} \rangle$. So $\langle \mathtt{LT}(I) \rangle$ is a monomial ideal.

2) Since $\langle \mathtt{LT}(I) \rangle$ is a monomial ideal, we can use Dickson's lemma to find $g_1, g_2, \ldots, g_m \in I$ such that $\langle I \rangle = \langle \mathtt{LM}(g_1), \mathtt{LM}(g_2), \ldots, \mathtt{LM}(g_m) \rangle$. It easily follows then that $\langle I \rangle = \langle \mathtt{LT}(g_1), \mathtt{LT}(g_2), \ldots, \mathtt{LT}(g_m) \rangle$. This completes our proof $\square$

Together with the division algorithm this allows us to find a finite basis for any ideal I.

Theorem 2.24 (Hilbert basis theorem) *Every ideal* $I \subseteq k[x_1, x_2, \ldots, x_n]$ *has a finite basis. In other words* $I = \langle g_1, g_2, \ldots, g_m \rangle$ *for some* $g_1, g_2, \ldots, g_m \in I$.

Proof If $I = \{0\}$ we simply take $\{0\}$ to be our generating set, so let I be a nontrivial ideal.

We fix a monomial ordering and let $\langle \mathtt{LT}(I) \rangle$ be the resulting ideal of leading terms. By proposition 2.23 we can find $g_1, g_2, \ldots, g_m \in I$, such that $\langle \mathtt{LT}(I) \rangle = \langle \mathtt{LT}(g_1), \mathtt{LT}(g_2), \ldots, \mathtt{LT}(g_m) \rangle$. We will show that $I = \langle g_1, g_2, \ldots, g_m \rangle$.

Obviously we have that $\langle g_1, g_2, \ldots, g_m \rangle \subseteq I$ since all $g_i \in I$. We use the division algorithm to divide f by $g_1, g_2, \ldots, g_m$ and we find that

$$f = h_1 g_1 + h_2 g_2 + \cdots + h_m g_m + r$$

where r is some remainder. Note that we can write r as

$$r = f - h_1 g_1 - h_2 g_2 - \cdots - h_m g_m \in I$$

So if $r \neq 0$, then $\mathtt{LT}(r)$ is an element of $\langle \mathtt{LT}(I) \rangle = \langle \mathtt{LT}(g_1), \mathtt{LT}(g_2), \ldots, \mathtt{LT}(g_m) \rangle$. Using lemma 2.17 we find that r is divisible by some $\mathtt{LT}(g_i)$. This leads to a contradiction since r is a remainder and hence r must be zero. This reduces our equation of f to

$$f = h_1 g_1 + h_2 g_2 + \cdots + h_m g_m$$

This is an element of $\langle g_1, g_2, \ldots, g_m \rangle$, so we can conclude that $I \subseteq \langle g_1, g_2, \ldots, g_m \rangle$. This completes our proof □

2.6 Gröbner bases

The basis $\{g_1, g_2, \ldots, g_m\}$ generated in the proof of theorem 2.24 has the additional property, namely that $\langle \mathtt{LT}(I) \rangle = \langle \mathtt{LT}(g_1), \mathtt{LT}(g_2), \ldots, \mathtt{LT}(g_m) \rangle$. Since this does not hold for every basis of an ideal, we will call the bases that have this property Gröbner bases.

Definition 2.25 *Fix a monomial order on $k[x_1, x_2, \ldots, x_n]$. We call a set $G = \{g_1, g_2, \ldots, g_m\}$ of an ideal $I \subseteq k[x_1, x_2, \ldots, x_n]$, different from $\{0\}$, a Gröbner basis if*

$$\langle LT(I) \rangle = \langle LT(g_1), LT(g_2), \ldots, LT(g_m) \rangle$$

For the ideal $\{0\}$ we define the empty set $\emptyset$ to be the Gröbner basis.

From the proof of theorem 2.24 we can immediately infer the following corollary

Corollary 2.26 *Fix a monomial order. Then every ideal $I \subseteq k[x_1, x_2, \ldots, x_n]$ has a Gröbner basis. Furthermore every Gröbner base for I is a basis of I.*

Proof Let I be a nonzero ideal. The set $G = \{g_1, g_2, \ldots, g_m\}$ constructed in the proof theorem 2.24 is a Gröbner basis of I and by the same arguments in the proof it is also a basis of I. □

The reason why Gröbner bases are so important, is that they have the following property regarding the division algorithm

Proposition 2.27 *Let $I \subseteq k[x_1, x_2, \ldots, x_n]$ be an ideal and let $G = \{g_1, g_2, \ldots, g_m\}$ be a Gröbner basis for I. For any given $f \in k[x_1, x_2, \ldots, x_n]$, there exists a unique $r \in k[x_1, x_2, \ldots, x_n]$ which satisfies the following properties:*

1. *No term of r is divisible by any of the $LT(g_1), LT(g_2), \ldots, LT(g_m)$*
2. *There exists a $g \in I$, such that $f = g + r$*

In particular r is the remainder of f by division by G. This is independent of how we list the elements of G.

Proof The division algorithm gives us that $f = h_1 g_1 + h_2 g_2 + \cdots + h_m g_m + r$ and thus r satisfies 1). Furthermore since G is a basis of I, it follows that

$g = h_1 g_1 + h_2 g_2 + \cdots + h_m g_m \in I$. This satisfies 2) and proves the existence of r.

For uniqueness assume that $f = g + r = g' + r'$ satisfies 1) and 2). By the definition of an ideal we have that $r - r' = g' - g \in I$. Since $r \neq r'$, it follows that $\mathtt{LT}(r - r') \in \langle \mathtt{LT}(I) \rangle = \langle \mathtt{LT}(g_1), \mathtt{LT}(g_2), \ldots, \mathtt{LT}(g_m) \rangle$. However using lemma 2.17 we have that $\mathtt{LT}(r - r')$ is divisible by some $\mathtt{LT}(g_i)$. This leads to a contradiction, since no terms of r, r' are divisible by $\mathtt{LT}(g_1), \mathtt{LT}(g_2), \ldots, \mathtt{LT}(g_m)$. So $r - r'$ is zero and r is unique.

The final part of the proposition follows from the uniqueness of r □

This proposition allows us to answer the question whether a polynomial $f \in k[x_1, x_2, \ldots, x_n]$ is an element of an ideal I.

Corollary 2.28 *Let $I \subseteq k[x_1, x_2, \ldots, x_n]$ be an ideal and $G = \{g_1, g_2, \ldots, g_m\}$ a Gröbner basis for it. Then a given polynomial $f \in k[x_1, x_2, \ldots, x_n]$ lies in I if and only if the remainder of f by division by G is zero.*

Proof We have shown already that $f \in I$ if its remainder is zero. For the converse note that $f = f + 0$. Since $f \in I$ we satisfy the conditions of Proposition 2.27. Hence 0 is the remainder of f on division by G. □

Thus if we want to check whether a polynomial f lies in an ideal I, we simply need to compute its remainder with respect to the Gröbner basis G of the ideal. We will denote this remainder in the following way

Definition 2.29 *Let $F = (f_1, f_2, \ldots, f_m)$ be an ordered m-tuple. We will denote the remainder of f on division by F with $\overline{f}^F$. If F is a Gröbner basis of an ideal then we can treat F as a set without an order, as $\overline{f}^F$ is uniquely determined.*

For example take $F = (x^3 z - yz, y^2 z^2 + z) \subseteq k[x, y, z]$ with lex order. Than we have that

$$\overline{x^4 y^2 z^3}^F = -xyz^2$$

since the division algorithm gives us that

$$x^4 y^2 z^3 = xy^2 z^2 \cdot (x^3 z - yz) + xyz \cdot (y^2 z^2 + z) - xyz^2$$

One concern we have not adressed so far, is that we need to find a Gröbner basis to fully utilize the division algorithm. One way to do this is by using the Buchberger algorithm. For this we make use of S-polynomials

Definition 2.30 *Let $f, g \in k[x_1, x_2, \ldots, x_n]$ be polynomials. Then:*

1. *If $\mathit{multideg}(f) = \alpha$ and $\mathit{multideg}(g) = \beta$, then we let $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n)$ where $\gamma_i = \max(\alpha_i, \beta_i)$ for all i. We call x^γ the least common multiple of $\mathit{LM}(f)$ and $\mathit{LM}(g)$. We denote this by $x^\gamma = lcm(\mathit{LM}(f), \mathit{LM}(g))$*
2. *We define the S-polynomial of f and g as follows*

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

The S-polynomials are defined in such a way that the leading terms of f and g cancel out. To see this, let us consider $f = x^2y - 2x$ and $g = 7xy^3 + y$ in $k[x, y]$ with lex order. Then γ is given by $\gamma = (2, 3)$, so

$$\begin{aligned} S(f,g) &= \frac{x^2y^3}{x^2y} \cdot f - \frac{x^2y^3}{7xy^3} \cdot g \\ &= y^2 \cdot f - \frac{1}{7}x \cdot g \\ &= -2xy^2 - \frac{1}{7}xy \end{aligned}$$

We have seen before that for an ideal $\langle f_1, f_2, \ldots, f_m \rangle$ we can find polynomials such that their leading terms are not elements of $\langle \mathtt{LT}(f_1), \mathtt{LT}(f_2), \ldots, \mathtt{LT}(f_m) \rangle$. By applying the S-polynomials repeatedly on the f_i we can expand our basis and additionally the basis we construct in this way turns out to be a Gröbner basis. We do this via the following algorithm which was developed by Buchberger.

Theorem 2.31 (Buchberger's algorithm) *Let $I = \langle f_1, f_2, \ldots, f_m \rangle \neq \{0\}$ be a polynomial ideal. Then a Gröbner basis for I can be constructed in a finite number of steps via the following algorithm:*

Input : $F = (f_1, f_2, \ldots, f_s)$
Output : a Gröbner basis $G = (g_1, g_2, \ldots, g_t)$ for I, with $F \subseteq G$

$G := F$
REPEAT
    $G' := G$
    FOR each pair $\{p, q\}, p \neq q$ in G' DO
        $r := \overline{S(p,q)}^{G'}$
        IF $r \neq 0$ THEN $G := G \cup \{r\}$
UNTIL $G = G'$
RETURN G

Proof A proof can be found in [1, p.91] □

Buchberger introduced this algorithm in his PhD thesis[3] to compute Gröbner bases, which he also introduced in his thesis. In fact Gröbner bases are named after his thesis adviser Wolfgang Gröbner.

3 Symmetric polynomials

In this chapter we will take a look at symmetric polynomials. We will look at the elementary symmetric polynomials and show that we can write all symmetric polynomials as a linear combination of the elementary polynomials. Furthermore we will consider power sums and show how they relate to the elementary

symmetric polynomials. We will find that they are related by Newton's indentity and prove this identity.

3.1 Elementary symmetric polynomials

Symmetric polynomials are something common to mathematics and we probably know some in one form of another. Consider for example the equation $x^2 + y^2 + z^2 - r^2$ that describes a sphere of radius r. Changing the order of the variables does not change the equation. Hence this polynomial is symmetric. In general a symmetric polynomial is defined as.

Definition 3.1 *We call a polynomial $f \in k[x_1, x_2, \ldots, x_n]$ symmetric if*

$$f(x_1, x_2, \ldots, x_n) = f(x_{j_1}, x_{j_2}, \ldots, x_{j_n})$$

holds for every permutation $(x_1, x_2, \ldots, x_n) \mapsto (x_{j_1}, x_{j_2}, \ldots, x_{j_n})$

The most important examples of symmetric polynomials, as we will see, are the elementary symmetric polynomials.

Definition 3.2 *Given $k[x_1, x_2, \ldots, x_n]$ a polynomial ring. The elementary symmetric polynomials $\sigma_1, \sigma_2, \ldots, \sigma_n \in k[x_1, x_2, \ldots, x_n]$ are given by.*

$$\begin{aligned}
\sigma_1 &= x_1 + x_2 + \cdots + x_n \\
\sigma_2 &= \sum_{j_1 < j_2} x_{j_1} x_{j_2} \\
&\vdots \\
\sigma_m &= \sum_{j_1 < j_2 < \cdots < j_m} x_{j_1} x_{j_2} \ldots x_{j_m} \\
&\vdots \\
\sigma_n &= x_1 x_2 \ldots x_n
\end{aligned}$$

In other words σ_m is a sum of monomials of products of m distinct variables. While these polynomials are called symmetric it is a priori not obvious that they are symmetric. To prove this we will first prove the following lemma.

Lemma 3.3 *Let $\sigma_j^{(i)}$ be the j-th elementary symmetric polynomial in $x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ for $j < n$. Furthermore let $\sigma_0 = 1$ and $\sigma_n^{(i)} = 0$. Then the following equation $\sigma_j = \sigma_j^{(i)} + x_i \sigma_{j-1}^{(i)}$ holds for all i, j*

Proof We can write σ_j as follows

$$\begin{aligned}
\sigma_j &= \sum_{k_1 < k_2 < \cdots < k_j} x_{k_1} x_{k_2} \ldots x_{k_j} \\
&= \sum_{\substack{k_1 < k_2 < \cdots < k_j \\ \forall i,\, k_i \neq j}} x_{k_1} x_{k_2} \ldots x_{k_j} + \sum_{\substack{k_1 < k_2 < \cdots < k_j \\ \exists i,\, k_i = j}} x_{k_1} x_{k_2} \ldots x_{k_j}
\end{aligned}$$

Note that all the terms in the first sum do not contain x_j and are all the products of j distinct variables. Hence the firt term is simply $\sigma_j^{(i)}$. Furthermore since every term of the second sum contains x_j, we can factor this out leaving us with

$$\begin{aligned}
&\sum_{\substack{k_1<k_2<\cdots<k_j \\ \forall i,\, k_i\neq j}} x_{k_1}x_{k_2}\dots x_{k_j} + \sum_{\substack{k_1<k_2<\cdots<k_j \\ \exists i,\, k_i=j}} x_{k_1}x_{k_2}\dots x_{k_j} \\
&= \sigma_j^{(i)} + x_j \cdot \sum_{\substack{k_1<k_2<\cdots<k_{j-1} \\ \forall i,\, k_i\neq j}} x_{k_1}x_{k_2}\dots x_{k_{j-1}} \\
&= \sigma_j^{(i)} + x_j\sigma_{j-1}^{(i)}
\end{aligned}$$

where we use the same argumentation for $\sigma_{j-1}^{(i)}$ once we factor out x_j. Hence $\sigma_j = \sigma_j^{(i)} + x_i\sigma_{j-1}^{(i)}$ □

We can now prove that elementary symmetric polynomials are indeed symmetric.

Theorem 3.4 *The elementary symmetric polynomials $\sigma_1, \sigma_2, \dots, \sigma_n$ are symmetric in $k[x_1, x_2, \dots, x_n]$*

Proof We will introduce a new variable X and consider the following polynomial

$$f(X) = (X - x_1)(X - x_2)\dots(X - x_n) \tag{1}$$

with roots in $x_1, x_2, \dots, x_n$. Using induction we claim that this is equal to

$$f(X) = X^n - \sigma_1 X^{n-1} + \cdots + (-1)^{n-1}\sigma_{n-1}X + (-1)^n\sigma_n \tag{2}$$

Obviously this equation holds for $n = 1$, so let us assume it holds for $n - 1$. Expanding equation (1) and using the induction hypothesis gives us

$$\begin{aligned}
f(X) =& (X - x_1)(X - x_2)\dots(X - x_{n-1})(X - x_n) \\
=& (X^{n-1} - \sigma_1^{(n)}X^{n-2} + \cdots + (-1)^{n-2}\sigma_{n-2}^{(n)}X + (-1)^{n-1}\sigma_{n-1}^{(n)})(X - x_n) \\
=& X^n - \sigma_1^{(n)}X^{n-1} + \cdots + (-1)^{n-2}\sigma_{n-2}^{(n)}X^2 + (-1)^{n-1}\sigma_{n-1}^{(n)}X \\
& - x_nX^{n-1} + \cdots + (-1)^{n-2}x_n\sigma_{n-3}^{(n)}X^2 + (-1)^{n-1}x_n\sigma_{n-2}^{(n)}X + (-1)^n x_n\sigma_{n-1}^{(n)} \\
=& X^n - \sigma_1X^{n-1} + \cdots + (1)^{n-1}\sigma_{n-1}X + (-1)^n\sigma_n
\end{aligned}$$

Here we used lemma 3.3 for the last equality and thus we find that the equations (1) and (2) are equivalent.

Now if we perform a permutation on the variables $x_1, x_2, \dots, x_n$ only the order of the factors in equation (1) changes, so f remains unchanged. Hence the coefficients $(-1)^m\sigma_m$ in equation (2) remain the same and thus the polynomials $\sigma_1, \sigma_2, \dots, \sigma_n$ are symmetric □

3.2 The fundamental theorem of symmetric polynomials

The elementary symmetric polynomials allow us to make several constructions, for one products and sums of the elementary symmetric polynomials are symmetric. This allows us to easily make symmetric polynomials, for example in

$k[x, y, z]$

$$\sigma_3\sigma_2 - \sigma_1^2 = x^2y^2z + x^2yz^2 + xy^2z^2 - x^2 - 2xy - xz - y^2 - 2yz - z^2$$

is symmetric. An interesting question we could aks ourselves is whether every symmetric polynomial can be expressed in terms of the elementary symmetric polynomials? Not only is this possible, but it turns out we can do this in a unique manner for every symmetric polynomial.

Theorem 3.5 (The fundamental theorem of symmetric polynomials) *Every symmetric polynomial can be written uniquely as a polynomial in the elementary symmetric polynomials*

Proof We will use lex order with $x_1 > x_2 > \cdots > x_n$. Let $f \in k[x_1, x_2, \ldots, x_n]$ be a nonzero symmetric polynomial. We will denote its leading term with $\mathtt{LT}(f) = ax^\alpha$ and let the exponent α be given by $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$. We claim that $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$ and use a proof by contradiction. Assume that $\alpha_{i+1} > \alpha_i$ for some i and let α' be the vector we get by switching α_{i+1} and α_i. We denote this by $\alpha' = (\ldots, \alpha_{i+1}, \alpha_i, \ldots)$. It follows that $ax^{\alpha'}$ is a term of $f(\ldots, \alpha_{i+1}, \alpha_i, \ldots)$, since ax^α is a term of f. Since f is symmetric we have that $f(\ldots, \alpha_{i+1}, \alpha_i, \ldots) = f$ and thus $ax^{\alpha'}$ is a term of f. By lex order $\alpha' > \alpha$ so $ax^{\alpha'}$ is the leading term of f, which leads to a contradiction. This proves our claim.

We will introduce the following polynomial

$$h = \sigma_1^{\alpha_1-\alpha_2}\sigma_2^{\alpha_2-\alpha_3}\ldots\sigma_{n-1}^{\alpha_{n-1}-\alpha_n}\sigma_n^{\alpha_n}$$

This polynomial is symmetric and it has an interesting property with regards to its leading term. For this we would like to note that $\mathtt{LT}(\sigma_m) = x_1x_2\ldots x_m$ in our lex order. Computing the leading term of h then gets us

$$\begin{aligned}
\mathtt{LT}(h) &= \mathtt{LT}(\sigma_1^{\alpha_1-\alpha_2}\sigma_2^{\alpha_2-\alpha_3}\ldots\sigma_{n-1}^{\alpha_{n-1}-\alpha_n}\sigma_n^{\alpha_n}) \\
&= \mathtt{LT}(\sigma_1)^{\alpha_1-\alpha_2}\mathtt{LT}(\sigma_2)^{\alpha_2-\alpha_3}\ldots\mathtt{LT}(\sigma_{n-1})^{\alpha_{n-1}-\alpha_n}\mathtt{LT}(\sigma_n)^{\alpha_n} \\
&= x_1^{\alpha_1-\alpha_2}(x_1x_2)^{\alpha_2-\alpha_3}\ldots(x_1x_2\ldots x_{n-1})^{\alpha_{n-1}-\alpha_n}(x_1x_2\ldots x_n)^{\alpha_n} \\
&= x_1^{\alpha_1}x_2^{\alpha_2}\ldots x_n^{\alpha_n} = x^\alpha
\end{aligned}$$

hence ah has the same leading term as f and thus we can subtract the polynomials. If this gives us the zero polynomial then we have found a decomposition in $\sigma_1, \sigma_2, \ldots, \sigma_n$, so let us assume that $f_1 = f - ah \neq 0$. Since this is again a symmetric polynomial we can find a h_1 in the same way as before and repeat this process over and over. This gives us a chain of $f, f_1, f_2, \ldots$ and since we subtract the leading term at each step we find that

$$\text{multideg}(f) > \text{multideg}(f_1) > \text{multideg}(f_2) > \ldots$$

Since the lex order is a well-ordering this chain must terminate for some m, hence $f_{m+1} = 0$. From this we get that

$$f = ah + a_1h_1 + a_2h_2 + \cdots + a_mh_m$$

so f can be written as a polynomial in the elementary symmetric polynomials.

We now have to show that can be done uniquely. So let $g_1, g_2 \in k[y_1, y_2, \ldots, y_n]$ be polynomials such that $f = g_1(\sigma_1, \sigma_2, \ldots, \sigma_n) = g_2(\sigma_1, \sigma_2, \ldots, \sigma_n)$. We need to prove that $g_1 = g_2$ in $k[y_1, y_2, \ldots, y_n]$, so we will consider the polynomial $g = g_1 - g_2$. In $k[x_1, x_2, \ldots, x_n]$ we have that $g(\sigma_1, \sigma_2, \ldots, \sigma_n) = 0$, so if we can show that $g = 0$ in $k[y_1, y_2, \ldots, y_n]$ we have proved uniqueness. Now assume that $g \neq 0$. If we examine a monomial y^α of g, then in $g(\sigma_1, \sigma_2, \ldots, \sigma_n)$ the leading term is given by $x_1^{\alpha_1+\alpha_2+\cdots+\alpha_n} x_2^{\alpha_2+\alpha_3+\cdots+\alpha_n} \ldots x_n^{\alpha_n}$ We can compute this similar to how we computed the leading term of h. Note that the map

$$(\alpha_1, \alpha_2, \ldots, \alpha_n) \mapsto (\alpha_1 + \alpha_2 + \cdots + \alpha_n, \alpha_2 + \alpha_3 + \cdots + \alpha_n, \ldots, \alpha_n)$$

is injective, so every monomial of g has a distinct leading term in $g(\sigma_1, \sigma_2, \ldots, \sigma_n)$. From this we can conclude that the leading term of the largest monomial, by lex order, cannot be cancelled, hence $g(\sigma_1, \sigma_2, \ldots, \sigma_n) \neq 0$. This contradiction completes the proof of the theorem. □

An alternative proof for this theorem may be found in [4, Lem 3.1.14]. This proof uses Galois theory and is smaller in size. However it does not provide an algorithm with which we can compute a decomposition. Let us use this algorithm on an example. Consider in $k[x, y, z]$ the polynomial

$$f = -x^3y - x^3z + x^2y^2 + x^2z^2 - xy^3 - xz^3 - y^3z + y^2z^2 - yz^3$$

The leading term is given by $-x^3y = -\mathtt{LT}(\sigma_1^2\sigma_2)$, so

$$f_1 = f + \sigma_1^2\sigma_2 = 3x^2y^2 + 5x^2yz + 3x^2z^2 + 5xy^z + 5xyz^2 + 3y^2z^2$$

Now the leading term is $3x^2y^2 = 3\mathtt{LT}(\sigma_2^2)$, which gives us

$$f_2 = f - \sigma_2^2 = -x^2yz - xy^2z - xyz^2$$

This simply leaves us to

$$f_3 = f_2 + \sigma_1\sigma_3 = 0$$

Hence we have

$$f = -\sigma_1^2\sigma_2 + \sigma_2^2 - \sigma_1\sigma_3$$

as an expression of f into the elementary symmetric polynomials.

Note that while we have used lex order in the proof and in this example, but we could have used any other monomial order as well. This can be seen in this proof [5, Thm 1.1.1] of the fundamental theorem, where the author used grlex order.

3.3 Finding symmetric polynomials

The fundamental theorem presupposes that we are working with a symmetric polynomial, but a priori we have no way of knowing whether a given polynomial is symmetric or not. As it turns out combining the division algorithm with Gröbner bases, gives us a powerful tool to check for symmetry. We do this as follows.

Proposition 3.6 *Fix a monomial order in $k[x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n]$ such that any monomial containing one of $x_1, x_2, \ldots, x_n$ is greater than all monomials in $k[y_1, y_2, \ldots, y_n]$. Let G be a Gröbner basis of $\langle \sigma_1 - y_1, \sigma_2 - y_2, \ldots, \sigma_n - y_n \rangle \subseteq k[x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n]$. Given $y \in k[x_1, x_2, \ldots, x_n]$, let $g = \overline{f}^G$ be the remainder by division on G. Then:*

1. *f is symmetric if and only if $g \in k[y_1, y_2, \ldots, y_n]$*
2. *If f is symmetric, then $f = g(\sigma_1, \sigma_2, \ldots, \sigma_n)$ is the unique expression of f as a polynomial in the elementary symmetric polynomials*

Proof Let $f \in k[x_1, x_2, \ldots, x_n]$ and $g \in k[x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n]$ its remainder on division by $G = \{g_1, g_2, \ldots, g_2\}$ as above. So

$$f = A_1 g_1 + A_2 g_2 + \cdots + A_t g_t + g$$

for some $A_1, A_2, \ldots, A_t \in k[x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n]$. We can assume $g_i \neq 0$ for all i.

Let us first assume that $g \in k[y_1, y_2, \ldots, y_n]$. We will use the substitution $y_i \mapsto \sigma_i$ for all i. Since f is a polynomial in $x_1, x_2, \ldots, x_n$ it stays unaffected. Note that every polynomial in $\langle y_1 - \sigma_1, y_2 - \sigma_2, \ldots, y_n - \sigma_n \rangle$ goes to zero under this substitution. In particular since $g_1, g_2, \ldots, g_n$ form a Gröbner basis of this ideal, they all go to zero under this substitution. This reduces our equation for f to

$$f = g(\sigma_1, \sigma_2, \ldots, \sigma_n)$$

so f is symmetric.

Now let $f \in k[x_1, x_2, \ldots, x_n]$ be symmetric, so $f = g(\sigma_1, \sigma_2, \ldots, \sigma_n)$ for some $g \in k[y_1, y_2, \ldots, y_n]$. We want to prove that this g is the remainder of f on division by G. Note that in $k[x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n]$ we can expand a monomial in $\sigma_1, \sigma_2, \ldots, \sigma_n$ as follows.

$$\begin{aligned}\sigma^{\alpha_1}\sigma^{\alpha_2}\ldots\sigma^{\alpha_n} &= (y_1 - (y_1 - \sigma_1))^{\alpha_1}(y_2 - (y_2 - \sigma_2))^{\alpha_2}\ldots(y_n - (y_n - \sigma_n))^{\alpha_n} \\ &= y_1^{\alpha_1} y_2^{\alpha_2}\ldots y_n^{\alpha_n} + B_1(y_1 - \sigma_1) + B_2(y_2 - \sigma_2) + \cdots + B_n(y_n - \sigma_n)\end{aligned}$$

for some $B_1, B_2, \ldots, B_n \in k[x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n]$. Substituting this in $g(\sigma_1, \sigma_2, \ldots, \sigma_n)$ then gives us

$$g(\sigma_1, \sigma_2, \ldots, \sigma_n) = g(y_1, y_2, \ldots, y_n) + C_1(y_1 - \sigma_1) + C_2(y_2 - \sigma_2) + \cdots + C_n(y_n - \sigma_n)$$

where $C_1, C_2, \ldots, C_n \in k[x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n]$ are suitably chosen polynomials. Using the fact that $f = g(\sigma_1, \sigma_2, \ldots, \sigma_n)$ we get that

$$f = g(y_1, y_2, \ldots, y_n) + C_1(y_1 - \sigma_1) + C_2(y_2 - \sigma_2) + \cdots + C_n(y_n - \sigma_n)$$

Now we have to show that g is the remainder of f on division by G.

For this we need to show that g is not divisible by $\mathtt{LT}(g_i)$ for all i. Assume that there is an i such that $\mathtt{LT}(g_i)$ divides g. Since g is a polynomial in $k[y_1, y_2, \ldots, y_n]$, it follows by our ordering that $\mathtt{LT}(g_i)$ contains none of $x_1, x_2, \ldots, x_n$. Thus $g_i \in k[y_1, y_2, \ldots, y_n]$. We have earlier seen that g_i goes to zero under the substitution $y_i \mapsto \sigma_i$, since $g_i \in \langle y_1 - \sigma_1, y_2 - \sigma_2, \ldots, y_n - \sigma_n \rangle$. Hence $g_i(\sigma_1, \sigma_2, \ldots, \sigma_n) = 0$ is a symmetric polynomial. By the uniqueness of theorem 3.5 it follows that $g_i = 0$. This leads to a contradiction, since $g_i \neq 0$. Hence $\mathtt{LT}(g_i)$ does not divide g for all i. Since G is a Gröbner basis proposition 2.27 we find that g is the remainder of f by division on G. This proves that the remainder is an element of $k[y_1, y_2, \ldots, y_n]$ if f is symmetric.

The second part of the proposition from the arguments above. □

3.4 Correspondence between $\sigma_i \leftrightarrow s_i$

When working with symmetric polynomials it is often helpful to work with homogeneous polynomials. These are defined as follows.

Definition 3.7 *A polynomial $f \in k[x_1, x_2, \ldots, x_n]$ is homogeneous of total degree m if every term of f has total degree m.*

The elementary symmetric polynomials are prime examples of homogeneous polynomials, as each term in σ_i has degree i. One important observation to make is that each polynomial can be written uniquely in its homogeneous polynomials. For $f \in k[x_1, x_2, \ldots, x_n]$ let the m-the homogeneous component be the polynomial that is the sum of all terms with total degree m. We denote this polynomial with f_m and then $f = \sum_m f_m$.

The following proposition establishes a link between symmetric polynomials and their homogeneous components.

Proposition 3.8 *A polynomial $f \in k[x_1, x_2, \ldots, x_n]$ is symmetric if and only if all its homogeneous components are symmetric.*

Proof Obviously a polynomial is symmetric if its homogeneous components are symmetric, so let f be a symmetric polynomial. Note that a permutation $(x_1, x_2, \ldots, x_n) \mapsto (x_{j_1}, x_{j_2}, \ldots, x_{j_n})$ takes a monomial of total degree m to a monomial with the same total degree. Since $f(x_1, x_2, \ldots, x_n) = f(x_{j_1}, x_{j_2}, \ldots, x_{j_n})$ it follows that its homogeneous components must be symmetric. □

From this proposition we can conclude that whenever we work with symmetric polynomials, we can assume that it is homogeneous.

One important group of homogeneous polynomials is the power sums

$$s_j = x_1^j + x_2^j + \cdots + x_n^j$$

Since the power sums are symmetric we can write them as a polynomial in the elementary symmetric polynomials. Furthermore it turns out that every symmetric polynomial can also be written as sum in the power sums. To prove this we make use of the following identity.

Lemma 3.9 (Newtons identity) *Let $x_1, x_2, \ldots, x_n$ be variables. Then the following holds for all $j \geq 1$*

$$s_j - \sigma_1 s_{j-1} + \cdots + (-1)^{j-1}\sigma_{j-1}s_1 + (-1)^j j\sigma_j = 0$$

Proof We will prove this using induction on n. Note that $\sigma_0 = 1$ and $\sigma_i = 0$ for all $i < 0$ or $i > n$. We first consider the case $j = 1$. The equation then reduces to $s_1 - (-1)^1 \cdot 1 \cdot \sigma_1 s = s_1 - \sigma_1 = 0$, since $s_1 = \sigma_1$ for all n. In particular the case $n = 1$ is equivalent to this case, since all σ_i vanish except for σ_1.

Thus let $j > 1$ from now on. Furthermore let the identity hold for $1, 2, \ldots, n-1$. We will write the equation as the following sum $\sum_{k=0}^{j-1}(-1)^k s_{j-k}\sigma_k + (-1)^j j\sigma_j$. Using the identity $\sigma_j = \sigma_j^{(n)} + x_n\sigma_{j-1}^{(n)}$, where the superscript (n) denotes we

omit x_n, from lemma 3.3 and the following identity $s_j = s_j^{(n)} + x_n^j$, we can write the sum as follows

$$\begin{aligned}\sum_{k=0}^{j-1}(-1)^k s_{j-k}\sigma_k + (-1)^j j\sigma_j &= \sum_{k=0}^{j-1}(-1)^k(s_{j-k}^{(n)} + x_n^{j-k})(\sigma_k^{(n)} + x_n\sigma_{k-1}^{(n)}) + (-1)^j j(\sigma_j^{(n)} + x_n\sigma_{j-1}^{(n)}) \\ &= \sum_{k=0}^{j-1}(-1)^k(s_{j-k}^{(n)}\sigma_k^{(n)} + x_n^{j-k}\sigma_k^{(n)} + x_n^{j-k+1}\sigma_{k-1}^{(n)} + x_n s_{j-k}^{(n)}\sigma_{k-1}^{(n)}) \\ &\quad + (-1)^j j(\sigma_j^{(n)} + x_n\sigma_{j-1}^{(n)})\end{aligned}$$

Note that by our induction hypothesis $\sum_{k=0}^{j-1}(-1)^k s_{j-k}^{(n)}\sigma_k^{(n)} + (-1)^j j\sigma_j^{(n)} = 0$ is the case $n-1$. This simplifies our sum to

$$\sum_{k=0}^{j-1}(-1)^k(x_n^{j-k}\sigma_k^{(n)} + x_n^{j-k+1}\sigma_{k-1}^{(n)} + x_n s_{j-k}^{(n)}\sigma_{k-1}^{(n)}) + (-1)^j j x_n\sigma_{j-1}^{(n)}$$

The second simplification we make is with regards to the first and second term of ther sum. If we compute the first term at some k and the second term at $k+1$, their sum will vanish

$$(-1)^k x_n^{j-k}\sigma_k^{(n)} + (-1)^{(k+1)}x_n^{j-(k+1)+1}\sigma_{(k+1)-1}^{(n)} = x_n^{j-k}\sigma_k^{(n)} - x_n^{j-k}\sigma_k^{(n)} = 0$$

This leaves us only with the first term for $k = j-1$ and the second term for $k = 0$. Note that $\sigma_{-1} = 0$, so the second term and third term vanish at $k = 0$. Thus we are left with the first term at $k = j-1$ which is given by $(-1)^{j-1}x_n\sigma_{j-1}^{(n)}$. This gives us the following sum

$$\begin{aligned}&\sum_{k=0}^{j-1}(-1)^k x_n s_{j-k}^{(n)}\sigma_{k-1}^n + (-1)^{j-1}x_n\sigma_{j-1}^{(n)} + (-1)^j j x_n\sigma_{j-1}^{(n)} \\ =&\sum_{k=1}^{j-1}(-1)^k x_n s_{j-k}^{(n)}\sigma_{k-1}^n + (-1)^j(j-1)x_n\sigma_{j-1}^{(n)} \\ =&x_n(-1)\left[\sum_{k=1}^{j-1}(-1)^{(k-1)}s_{j-k}^{(n)}\sigma_{k-1}^n + (-1)^{(j-1)}(j-1)\sigma_{j-1}^{(n)}\right]\end{aligned}$$

Here the sum inside the brackets is the case for $j-1$ and $n-1$, hence by our induction hypothesis this sum vanishes. With this we complete our proof □

We can now move on to prove our main theorem.

Theorem 3.10 *If k is a field containing the rational number $\mathbb{Q}$, then every symmetric polynomial in $k[x_1, x_2, \ldots, x_n]$ can be written as a polynomials in the power sums $s_1, s_2, \ldots, s_m$.*

Proof We will prove this theorem using induction. For $n = 1$ we simply have that $\sigma_1 = s_1$. We will assume the claim is true for $1, 2, \ldots, j-1$. Using Newtons identity from lemma 3.9 we get that

$$\sigma_j = (-1)^{j-1}\frac{1}{j}(s_j - \sigma_1 s_{j-1} + \cdots + (-1)^{j-1}\sigma_{j-1}s_1)$$

Here we can divide by j since $\mathbb{Q}$ is contained within our field. Then by our inductive hypothesis we have that $\sigma_1, \sigma_2, \ldots, \sigma_n$ can be written in the power sums. Thus substitung it in the above equations we find that σ_j is a polynomial in $s_1, s_2, \ldots, s_j$ □

4 Ring of invariants of finite groups

In this chapter we will generalize the notion of symmetric polynomials. We will introduce finite matrix groups and examine which polynomials stay invariant under their group action. Furthermore we will look at ways to generate invariant polynomials.

For this chapter we will assume our fields k to contain the rational numbers $\mathbb{Q}$. Such fields are said to be of characteristic zero.

4.1 Finite matrix groups

Let us examine the power sum $s_2(x_1, x_2, \ldots, x_n) = x_1^2 + x_2^2 + \cdots + x_n^2$ from chapter 3.4. We have already seen that this polynomial is symmetric, but it is also invariant under different transformations. Consider for example the linear map $x_i \mapsto -x_i$ for all i. Then clearly we have that $s_2(-x_1, -x_2, \ldots, -x_n) = s_2(x_1, x_2, \ldots, x_n)$, thus the power sum s_2 is invariant under this linear map.

Thus we will consider the following set

Definition 4.1 *Let $GL(n, k)$ be the set of invertible $n \times n$ matrices with entries in k.*

This set is called the general linear group and has a few interesting properties. Using linear algebra we can easily find that $\mathrm{GL}(n, k)$ is closed under matrix multiplication and that every matrix $A \in \mathrm{GL}(n, k)$ has an inverse which lies in $\mathrm{GL}(n, k)$. Thus the $n \times n$ identity matrix $I_n = A \cdot A^{-1}$ is also an element of $\mathrm{GL}(n, k)$.

Most importantly for every $A \in \mathrm{GL}(n, k)$ there is a corresponding invertible linear map $L_a : k^n \to k^n$ via matrix multiplication. In fact it can be shown that every linear map can be represented by a matrix A [6, p.210]. Hence $\mathrm{GL}(n, k)$ is a natural environment for us to work in.

More specifically we are interested in the following subsets

Definition 4.2 *We call a nonempty subset $G \subseteq GL(n, k)$ a finite matrix group if it is closed under matrix multiplication and finite. We call the number of elements in G the order of G and denote it with $|G|$.*

An example of a finite matrix group is the set $\{I_n\}$ that merely contains the identity matrix. For a less trivial example consider a matrix $A \in \mathrm{GL}(n, k)$ such that $A^m = I_n$ for some positive integer m. The set $C_m = \{I_n, A, A^2, \ldots, A^{m-1}\}$ generated by A is a finite matrix group. It is closed under multiplication and of order m. Namely for every $k > m$ we have that

$$A^k = A^{k-n*m} \cdot A^{n*m} = A^{k-n*m} \cdot (A^m)^n = A^{k-n*m} \in C_m$$

for some positive integer n. This group is called a cyclic group of order m.

For example one can easily see that the matrix

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathrm{GL}(2, k)$$

forms the cyclic group $C_2 = \{I_2, A\}$ of order 2.

Another important example of a finite matrix group are the permutation matrices. We have already encountered permutations when working with symmetric polynomials in chapter 4.1. Consider now for example a permutation σ which sends $(x_1, x_2, \dots, x_n) \mapsto (x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)})$, where the subscript i gets permuted to the subscript $\sigma(i)$. Thus the permutation matrix for σ is given by the matrix M_σ, which corresponds to the linear map that takes $(x_1, x_2, \dots, x_n)$ to $(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)})$. So

$$M_\sigma \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_{\sigma(1)} \\ x_{\sigma(2)} \\ \vdots \\ x_{\sigma(n)} \end{pmatrix}$$

It is easy to show that $\sigma(i)$-th column is given by the i-th column of I_n, so we leave this as an exercise for the reader.

As there are $n!$ possible ways to permute n variables, it follows that there are $n!$ permutation matrices. Furthermore the permutation matrices are closed under matrix multiplication via

$$M_\tau \cdot M_\sigma = M_{\sigma\tau}$$

where $\sigma\tau$ takes i to $\sigma(\tau(i))$. To see this we will examine how the vector $\sum_i x_i e_i$ changes under $M_\tau \cdot M_\sigma$, where $e_1, e_2, \dots, e_n$ is the standard basis of k^n. If we apply M_σ we get that $M_\sigma \cdot (\sum_i x_i e_i) = \sum_i x_{\sigma(i)} e_i$, since the i-th column of I_n is located at the $\sigma(i)$-th column of M_σ. By this same argument we get that $x_{\sigma(i)} e_i$ gets mapped to $x_{\sigma(\tau(i))} e_i$ under M_τ. Thus $M_\tau \cdot M_\sigma = M_{\sigma\tau}$ and the permutation matrices are closed under multiplication. Hence this set forms a finite matrix group in $\mathrm{GL}(n, k)$ and we will denote this group by S_n.

To end this chapter we would like to highlight some properties of finite matrix groups

Proposition 4.3 *Let $G \subseteq GL(n, k)$ be a finite matrix group. Then the following holds:*

1. *$I_n \in G$*
2. *If $A \in G$, then there exists a positive integer m such that $A^m = I_n$*
3. *If $A \in G$, then $A^{-1} \in G$*

Proof Let $A \in G$ and assume that 2) holds. Since G is closed under multiplication we have that $I_n = A^m \in G$. So 1) holds.

To prove 2), note that $\{A, A^2, \dots\} \subseteq G$ since G is closed. Furthermore we have that $A^i = A^j$ for some $i > j$, since G is finite. We can multiply both sides by A^{-j} because A is invertible. Thus we find that $A^{i-j} = I_n$ and this proves 2).

From 2) we find that $I_n = A^m = A^{m-1} \cdot A = A \cdot A^{m-1}$. So $A^{-1} = A^{m-1} \in G$, since G is closed under multiplication. This concludes our proof $\square$

4.2 Rings of invariants

We have looked at finite matrix groups and we would like to relate them to polynomials. To do this, we let $A = (a_{ij}) \in \mathrm{GL}(n, k)$ and we examine the linear map corresponding to this matrix. This maps the element $(x_1, x_2, \ldots, x_n)$ in k^n to $(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \ldots, a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n)$, which is again an element of k^n. Thus for any polynomial $f \in k[x_1, x_2, \ldots, x_n]$, we find that the polynomial

$$g(x_1, x_2, \ldots, x_n) = f(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \ldots, a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n)$$

is again a polynomial in $k[x_1, x_2, \ldots, x_n]$. We will shorten notation by denoting the column vector of $x_1, x_2, \ldots, x_n$ by $\mathbf{x}$. This reduces our equation to

$$g(\mathbf{x}) = f(A \cdot \mathbf{x})$$

We are interested in polynomials that are invariant under these maps. We define this as follows.

Definition 4.4 *Let $G \subseteq GL(n, k)$ be a finite matrix group. We say that a polynomial $f \in k[x_1, x_2, \ldots, x_n]$ is invariant under G if*

$$f(\boldsymbol{x}) = f(A \cdot \boldsymbol{x})$$

for all $A \in G$. The set that contains all polynomials invariant under G is denoted by $k[x_1, x_2, \ldots, x_n]^G$.

One example of invariant polynomials that we have already encountered, are the symmetrical polynomials. Indeed all symmetric polynomials are invariant under the permutation matrices $S_n \subseteq \mathrm{GL}(n, k)$. Thus we get that

$$k[x_1, x_2, \ldots, x_n]^{S_n} = \{\text{All symmetric polynomials in } k[x_1, x_2, \ldots, x_n]\}$$

Furthermore from theorem 3.5 we know that every symmetric polynomial can be expressed as a polynomial in the elementary symmetric polynomials with coefficients in k. Hence we get

$$k[x_1, x_2, \ldots, x_n]^{S_n} = k[\sigma_1, \sigma_2, \ldots, \sigma_n]$$

Additionally this decomposition is unique.

The question we ask ourselves is whether all invariants $k[x_1, x_2, \ldots, x_n]^G$ share these properties. One property that they all share is their algebraic structure

Proposition 4.5 *Let $G \subseteq GL(n, k)$. The set $k[x_1, x_2, \ldots, x_n]^G$ contains the constant polynomials and is closed under addition and multiplication.*

Proof Let f be a constant polynomial, thus $f(\mathbf{x}) = f$ holds for all $\mathbf{x} \in k^n$. So we get that $f(A \cdot \mathbf{x}) = f = f(\mathbf{x})$ for all $A \in G$ and hence $f \in k[x_1, x_2, \ldots, x_n]^G$.

Now let $f, g \in k[x_1, x_2, \ldots, x_n]^G$ and let $A \in G$. Then we have that

$$(f + g)(A \cdot \mathbf{x}) = f(A \cdot \mathbf{x}) + g(A \cdot \mathbf{x}) = f(x) + g(x) = (f + g)(x)$$
$$(fg)(A \cdot \mathbf{x}) = f(A \cdot \mathbf{x})g(A \cdot \mathbf{x}) = f(x)g(x) = (fg)(x)$$

Hence $f+g$ and fg are elements of $k[x_1, x_2, \ldots, x_n]^G$. This concludes our proof $\square$

Together with addition and multiplication the set $k[x_1, x_2, \ldots, x_n]^G$ thus forms a commutative ring. Since $k[x_1, x_2, \ldots, x_n]^G$ is a subset of $k[x_1, x_2, \ldots, x_n]$, we do not have to check for associativity and the other related properties as these hold in $k[x_1, x_2, \ldots, x_n]$. So in particular we call $k[x_1, x_2, \ldots, x_n]^G$ a subring of $k[x_1, x_2, \ldots, x_n]$.

One property of symmetric polynomials was that their homogeneous components were also symmetrical polynomials as shown in proposition 3.9. A similar result holds for invariant polynomials of any finite matrix group

Proposition 4.6 *Let $G \subseteq GL(n,k)$ be a finite matrix group. Then a polynomials f is invariant under G if and only if its homogeneous components are invariant under G.*

Proof We will prove this in a similar way as in the proof of proposition 3.9. Let $A = (a_{ij}) \in G$ and let $x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n}$ be a monomial of degree $m = i_1 + i_2 + \cdots + i_n$. Under matrix multiplication this monomial gets mapped to

$$(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n)^{i_1} \ldots (a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n)^{i_n} \quad (1)$$

We will use induction on n to prove that every term in $(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n)^{i_1}$ has total degree i_1.

This obviously holds for $n = 1$, so let it hold for $1, 2, \ldots, k-1$ and let $n = k$. Using the binomial theorem we get that

$$\begin{aligned} &((a_{11}x_1 + a_{12}x_2 + \cdots + a_{1(n-1)}x_{n-1}) + a_{1n}x_n)^{i_1} \\ =&\sum_{j=0}^{i_1} \binom{i_1}{j} (a_{11}x_1 + a_{12}x_2 + \cdots + a_{1(n-1)}x_{n-1})^j (a_{1n}x_n)^{i_1-j} \end{aligned}$$

By our induction hypothesis every term of $(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1(n-1)}x_{n-1})^j$ has total degree j, so every term in the sum has degree i_1. We can repeat this argument for all i_j and thus we can conclude that every term in equation (1) has total degree m.

Now let $f \in k[x_1, x_2, \ldots, x_n]^G$. Via matrix multiplication A maps every term of total degree m to a sum of terms with total degree m. Since $f(A \cdot \mathbf{x}) = f(\mathbf{x})$, it follows that the m-th homogeneous component is also invariant under G.

The converse is trivial and we conclude this proof $\square$

This proposition allows us to determine whether a polynomial is invariant, by examing its homogeneous components. This will prove to be useful, when we determine the generators of $k[x_1, x_2, \ldots, x_n]^G$ in chapter 4.3.

Another way to determine whether a polynomial is invariant under G, is to check whether it is invariant under the generators of G. We will prove this in the following lemma.

Lemma 4.7 *Let $G \subseteq GL(n,k)$ be a finite matrix group. We say that $A_1, A_2, \ldots, A_m \in G$ generates G, if every $A \in G$ can be written as*

$$A = B_1 B_2 \cdots B_s$$

where $B_i \in \{A_1, A_2, \ldots, A_m\}$ for all i. Then $f \in k[x_1, x_2, \ldots, x_n]$ lies in $k[x_1, x_2, \ldots, x_n]^G$ if and only if f is invariant under the generators of G, so

$$f(\boldsymbol{x}) = f(A_1 \cdot \boldsymbol{x}) = f(A_2 \cdot \boldsymbol{x}) = \cdots = f(A_m \cdot \boldsymbol{x})$$

Proof We will first prove that f is invariant under the product of matrices $B_1 B_2 \cdots B_s$, if it is invariant under $B_1, B_2 \ldots, B_s$. Using induction this obviously holds for $s = 1$. We assume it holds for $1, 2, \ldots, k-1$, then for $s = k$ we have that

$$f((B_1 B_2 \cdots B_k) \cdot \mathbf{x}) = f((B_1 B_2 \cdots B_{k-1}) \cdot (B_k \cdot \mathbf{x})) = f(B_k \cdot \mathbf{x}) = f(\mathbf{x})$$

Now assume that f is invariant under the generators $A_1, A_2, \ldots, A_m$ of G. Let A be an element of G. Then we can write it as $A = B_1 B_2 \cdots B_s$ with $B_i \in \{A_1, A_2, \ldots, A_m\}$. Since f is invariant under all A_i, it follows immediately that f is invariant under A and thus $f \in k[x_1, x_2, \ldots, x_n]^G$. The converse is trivial and this concludes our proof $\square$

To see how this lemma works in action we will consider the following matrix group

$$G_8 = \left\{ \begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}, \begin{pmatrix} \pm i & 0 \\ 0 & \pm i \end{pmatrix} \right\} \subseteq \mathrm{GL}(2, \mathbb{C})$$

of order 8. For the readers familiar with group theory, this group is isomorphic to $\mathbb{Z}^2 \times \mathbb{Z}^4$. It is easy to see that this group is generated by the matrices

$$\begin{pmatrix} i & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$$

Thus according to lemma 4.7 a polynomial $f \in \mathbb{C}[x, y]$ is invariant under G_8 if and only if the following holds

$$f(x, y) = f(ix, y) = f(x, iy)$$

If we write f as follows $f = \sum_{k,l} a_{kl} x^k y^l$, then first equality is given by

$$\begin{aligned} f(x, y) = f(ix, y) &\iff \sum_{k,l} a_{kl} x^k y^l = \sum_{k,l} (ix)^k y^l \\ &\iff \sum_{k,l} a_{kl} x^k y^l = \sum_{k,l} a_{kl} i^k x^k y^l \end{aligned}$$

So we get that $a_{kl} = a_{kl} i^k$. Since this only holds whenever k is a multiple of 4, it follows that a_{kl} vanishes otherwise. Thus x always has a power that divides 4. The same holds for y, when we repeat this process for $f(x, y) = f(x, iy)$. So we can find a unique polynomial $g \in \mathbb{C}[x, y]$ such that

$$f(x, y) = g(x^4, y^4)$$

Furthermore every polynomial of this form is invariant under G_8, so we get that

$$\mathbb{C}[x, y]^{G_8} = \mathbb{C}[x^4, y^4]$$

From this we can gather that every polynomial f that is invariant under G_8, can be written uniquely as a polynomial in x^4 and y^4.

Another interesting example is given by the finite matrix group

$$V_4 = \left\{ \pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \pm \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\} \subseteq \mathrm{GL}(2,k)$$

of order 4. This group is isomorphic to $\mathbb{Z}^2 \times \mathbb{Z}^2$. This group is generated by the matrices

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

so by lemma 4.7 we find that any invariant polynomial $f \in k[x,y]$ obeys the equation $f(x,y) = f(-x,-y) = f(y,x)$. For $f = \sum_{i,j} a_{ij} x^i y^j$, the first equality then implies that

$$\sum_{i,j} a_{ij} x^i y^j = \sum_{i,j} a_{ij} (-x)^i (-y)^j = \sum_{i,j} a_{ij} (-1)^{i+j} x^i y^j$$

So $a_{ij} = 0$ whenever $i+j$ is odd. It follows if f is an invariant polynomial, that every monomial $x^i y^j$ in f has the form

$$x^i y^j = \begin{cases} x^{2m} y^{2l} = (x^2)^m (y^2)^l & \text{if i,j are even} \\ x^{2m+1} y^{2l+1} = (x^2)^m (y^2)^l xy & \text{if i,j are odd} \end{cases}$$

Thus we find that f is polynomial in x^2, y^2, and xy, since all its monomial are polynomials in these invariants. So $f \in k[x^2, y^2, xy]$.

As an aside this ring of invariants differs from the rings we have seen so far, since we cannot write every invariant polynomial uniquely in terms of x^2, y^2, and xy. Take for example the polynomial x^2y^4, which is clearly invariant under the map $(x,y) \mapsto (-x,-y)$. Then we have that

$$x^2 y^4 = x^2 \cdot (y^2)^2 = (xy)^2 \cdot y^2$$

This stems from the fact that $x^2 \cdot y^2 = (xy)^2$.

Getting back to our original problem at hand, we still have the equality $f(x,y) = f(y,x)$ to examine. It follows from this equation that every invariant polynomial is also symmetric. We have already seen that our invariant polynomials lie in $k[x^2, y^2, xy]$, so we have to reduce this ring to its symmetrical polynomials. Let $f = \sum_{ijk} a_{ijk} (x^2)^i (y^2)^j (xy)^k$ be an invariant polynomial in $k[x^2, y^2, xy]$. Then we have that

$$\begin{aligned} f(x,y) = f(y,x) &\iff \sum_{i,j,k} a_{ijk} (x^2)^i (y^2)^j (xy)^k = \sum_{i,j,k} a_{ijk} (x^2)^j (y^2)^i (yx)^k \\ &\iff a_{ijk} (x^2)^i (y^2)^j = a_{jik} (x^2)^i (y^2)^j \quad \text{for all i,j} \\ &\iff a_{ijk} = a_{jik} \quad \text{for} \end{aligned}$$

In the case that $i = j$ we simply find that $(x^2)^i (y^2)^j = (x^2 y^2)^i = (xy)^{2i}$, so the monomials are polynomials in xy. On the other hand for $i < j$ we have that

$a_{ijk} = a_{jik}$, so

$$\begin{aligned}
&a_{ijk}(x^2)^i(y^2)^j(xy)^k + a_{jik}(x^2)^j(y^2)^i(xy)^k \\
=&a_{ijk}(xy)^k((x^2)^i(y^2)^j + (x^2)^j(y^2)^i) \\
=&a_{ijk}(xy)^k(x^2y^2)^i(x^2+y^2)^{j-i} \\
=&a_{ijk}(xy)^{k+2i}(x^2+y^2)^{j-1}
\end{aligned}$$

So the monomials in i and j form a polynomial in xy and $x^2 + y^2$. Thus a polynomial f that is invariant under V_4 can be expressed as a polynomial in xy and $x^2 + y^2$. This happens uniquely as there is no algebraic relation between only xy and $x^2 + y^2$. Conversely every polynomial in $k[x^2 + y^2, xy]$ is clearly invariant under V_4. So we can conclude that

$$k[x, y]^{V_4} = k[x^2 + y^2, xy]$$

4.3 Generators for rings of invariants

In this chapter we will determine how we can generate the ring of invariants $k[x_1, x_2, \ldots, x_n]^G$ for a finite matrix group $G \subseteq \mathrm{GL}(n, k)$. So far we have used that we can write certain polynomials as a polynomial in $f_1, f_2, \ldots, f_m$. To expand on this we define as follows

Definition 4.8 *Given $f_1, f_2, \ldots, f_m \in k[x_1, x_2, \ldots, x_n]$, we denote by $k[f_1, f_2, \ldots, f_m]$ the subset of $k[x_1, x_2, \ldots, x_n]$ that contains all polynomial expressions in $f_1, f_2, \ldots, f_m$ with coefficients in k.*

In other words it contains all polynomials $f \in k[x_1, x_2, \ldots, x_n]$ that can be written as

$$f = g(f_1, f_2, \ldots, f_m)$$

with g a polynomial in m variables and coefficients in k.

This set forms a subring of $k[x_1, x_2, \ldots, x_n]$ as it is closed under addition and multiplication and contains the constant polynomials. Thus $k[f_1, f_2, \ldots, f_m]$ is the subring generated by $f_1, f_2, \ldots, f_m$. One detail we have to keep in mind is that this is different from the ideal $\langle f_1, f_2, \ldots, f_m \rangle$, despite the fact that they are both generated by $f_1, f_2, \ldots, f_m$. As an example consider the ring $k[x^2]$ and the ideal $\langle x^2 \rangle$. Then $k[x^2]$ is not a subset of $\langle x^2 \rangle$, since the constant polynomials are not elements of $\langle x^2 \rangle$. Furthermore the converse inclusion does not hold as $x^3 = x^2 \cdot x \in \langle x^2 \rangle$, but there is no $m \in \mathbb{Z}_{\geq 0}$ such that $(x^2)^m = x^3$.

One important object that we make use of to determine $k[x_1, x_2, \ldots, x_n]^G$ is the Reynolds operator

Definition 4.9 *Given a finite matrix group $G \subseteq GL(n, k)$, the Reynolds operator of G is defined by the map $R_G : k[x_1, x_2, \ldots, x_n] \to k[x_1, x_2, \ldots, x_n]$ given by*

$$R_G(f)(\boldsymbol{x}) = \frac{1}{|G|} \sum_{A \in G} f(A \cdot \boldsymbol{x})$$

for $f(\boldsymbol{x}) \in k[x_1, x_2, \ldots, x_n]$.

We do not have to worry about division by $|G|$, as our field k has characteristic zero. The Reynolds operator has a few interesting properties that we will take a look at

Proposition 4.10 *Let R_G be the Reynolds operator of a finite matrix group $G \subseteq GL(n,k)$. Then R_G has the following properties:*

1. *R_G is k-linear in f*
2. *If $f \in k[x_1, x_2, \ldots, x_n]$, then $R_G(f) \in k[x_1, x_2, \ldots, x_n]^G$*
3. *If $f \in k[x_1, x_2, \ldots, x_n]^G$, then $R_G(f) = f$*

Proof For (1), let a, b be elements of k and f, g polynomials in $k[x_1, x_2, \ldots, x_n]$. Then we have that

$$\begin{aligned} R_G(af + bg)(\mathbf{x}) &= \frac{1}{|G|} \sum_{A \in G} (af + bg)(A \cdot \mathbf{x}) \\ &= \frac{1}{|G|} \left(\sum_{A \in G} af(A \cdot \mathbf{x}) + bg(A \cdot \mathbf{x}) \right) \\ &= \frac{a}{|G|} \sum_{A \in G} f(A \cdot \mathbf{x}) + \frac{b}{|G|} \sum_{B \in G} g(B \cdot \mathbf{x}) \\ &= aR_G(f)(\mathbf{x}) + bR_G(g)(\mathbf{x}) \end{aligned}$$

Thus we find that $R_G(af + bg) = aR_G(f) + bR_G(g)$, so R_G is k-linear in f

To prove (2), let $B \in G$. Then we want to prove that $R_G(f)(B\mathbf{x}) = R_G(f)(\mathbf{x})$. By the definition of the Reynolds operator we get that

$$R_G(f)(B\mathbf{x}) = \frac{1}{|G|} \sum_{A \in G} f(A \cdot B\mathbf{x}) = \frac{1}{|G|} \sum_{A \in G} f(AB \cdot \mathbf{x})$$

The important thing to note here is that for each $A \in G$ the product AB produces a unique element in G. If we assume that $AB = A'B$, then we can multiply both sides by B^{-1} and we get that $A = A'$. Since this produces $|G|$ distinct elements we can conclude that

$$G = \{AB \mid A \in G\}$$

Hence it does not matter whether we sum over $f(A \cdot \mathbf{x})$ or $f(AB \cdot \mathbf{x})$, as we produce the same terms albeit possibly in a different order. This gives us that

$$\frac{1}{|G|} \sum_{A \in G} f(AB \cdot \mathbf{x}) = \frac{1}{|G|} \sum_{A \in G} f(A \cdot \mathbf{x}) = R_G(f)(\mathbf{x})$$

Thus $R_G(f)(B\mathbf{x}) = R_G(f)(\mathbf{x})$ and we can conclude that $R_G(f) \in k[x_1, x_2, \ldots, x_n]^G$.

At last, if $f \in k[x_1, x_2, \ldots, x_n]^G$, then it simply follows that

$$R_G(f)(\mathbf{x}) = \frac{1}{|G|} \sum_{A \in G} f(A \cdot \mathbf{x}) = \frac{1}{|G|} \sum_{A \in G} f(\mathbf{x}) = f(\mathbf{x})$$

because f is invariant. This proves (3) and finishes our proof $\square$

As shown we can use the Reynolds operator to produce invariant polynomials. For an example we will consider the cyclic group $C_4 \subseteq \mathrm{GL}(2,k)$ of order 4 generated by

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

By lemma 4.7 we know that the ring of invariants contains all polynomials f of the form $f(x,y) = f(-y,x)$. It is easy to see that the Reynolds operator is given by

$$R_{C_4}(f)(x,y) = \frac{1}{4}(f(x,y) + f(-y,x) + f(-x,-y) + f(y,-x))$$

By Proposition 4.10 we can then find some invariants of $k[x,y]^{C_4}$ by

$$\begin{aligned} R_{C_4}(y^2)(x,y) &= \frac{1}{4}(y^2 + x^2 + (-y)^2 + (-x)^2) = \frac{1}{2}(x^2 + y^2) \\ R_{C_4}(xy)(x,y) &= \frac{1}{4}(xy + (-y)x + (-x)(-y) + y(-x)) = 0 \\ R_{C_4}(xy^3)(x,y) &= \frac{1}{4}(xy^3 + (-y)x^3 + (-x)(-y)^3 + y(-x)^3) = -\frac{1}{2}(x^3y - xy^3) \\ R_{C_4}(x^2y^2) &= \frac{1}{4}(x^2y^2 + (-y)^2x^2 + (-x)^2(-y)^2 + y^2(-x)^2) = x^2y^2 \end{aligned}$$

Hence we find that $x^2 + y^2, x^3y - xy^3, x^2y^2 \in k[x,y]^{C_4}$ and as we will find out these invariants generate $k[x,y]^{C_4}$.

We have shown earlier that a monomial x^α will be sent to a homogeneous polynomial of total degree $|\alpha|$, under matrix multiplication. Thus if we apply the Reynolds operator on a monomial we find that $R_G(x^\alpha)$ is a homogeneous invariant of total degree $|\alpha|$. Furthermore we can finitely generate $k[x_1, x_2, \ldots, x_n]^G$ with these invariants and we will show this through a theorem of Emmy Noether. For this we first need the following lemma

Lemma 4.11 *Let $x_1, x_2, \ldots, x_n$ be variables. Then we have that*

$$(x_1 + x_2 + \cdots + x_n)^m = \sum_{|\alpha|=m} a_\alpha x^\alpha$$

where a_α is a positive integer.

Proof Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n_{\geq 0}$ and let $|\alpha| = m$. Then we define the multinomial coefficient as

$$\binom{m}{\alpha} = \frac{m!}{\alpha_1!\alpha_2! \ldots \alpha_n!}$$

and we claim that

$$\binom{m}{\alpha} = \binom{\alpha_1}{\alpha_1}\binom{\alpha_1 + \alpha_2}{\alpha_2} \cdots \binom{\alpha_1 + \alpha_2 + \cdots + \alpha_n}{\alpha_n}$$

The right side here is composed of a product of binomial coefficients. If we write out the right hand side using the definition of the binomial coefficient we get

that

$$\begin{aligned}
&\binom{\alpha_1}{\alpha_1}\binom{\alpha_1+\alpha_2}{\alpha_2}\cdots\binom{\alpha_1+\alpha_2+\cdots+\alpha_n}{\alpha_n}\\
=&1\cdot\frac{(\alpha_1+\alpha_2)!}{\alpha_1!\alpha_2!}\cdot\frac{(\alpha_1+\alpha_2+\alpha_3)!}{(\alpha_1+\alpha_2)!\alpha_3!}\cdots\frac{(\alpha_1+\alpha_2+\cdots+\alpha_n)!}{(\alpha_1+\alpha_2+\cdots+\alpha_{n-1})!\alpha_n!}\\
=&1\cdot\frac{1}{\alpha_1!\alpha_2!}\cdot\frac{1}{\alpha_2!}\cdots\frac{(\alpha_1+\alpha_2+\cdots+\alpha_n)!}{\alpha_n!}\\
=&\frac{(\alpha_1+\alpha_2+\cdots+\alpha_n)!}{\alpha_1!\alpha_2!\ldots\alpha_n!}=\frac{m!}{\alpha_1!\alpha_2!\ldots\alpha_n!}
\end{aligned}$$

Here we make use of the fact that the numerators and the denominators cancel each other out telescopically. This proves our claim and since the binomial coefficients are positive integers, it follows that the multinomial coefficient is also a positive integer.

We now claim that $a_\alpha = \binom{m}{\alpha}$ and thus

$$(x_1+x_2+\cdots+x_n)^m=\sum_{|\alpha|=m}\binom{m}{\alpha}x^\alpha$$

We will prove this using induction on n. This obviously holds for $n=1$ and in the case of $n=2$ this is simply the binomial theorem. So assume it holds for $n=1,2,\ldots,k-1$ and let $n=k$. Then using the binomial theorem we find that

$$((x_1+x_2+\cdots+x_{k-1})+x_k)^m=\sum_{\alpha_k=0}^{m}\binom{m}{\alpha_k}x_k^{\alpha_k}(x_1+x_2+\cdots+x_{k-1})^{m-\alpha_k}$$

We can now apply our induction hypothesis which gives us

$$(x_1+x_2+\cdots+x_{k-1})^{m-\alpha_k}=\sum_{|\alpha^{(k)}|=m-\alpha_k}\binom{m-\alpha_k}{\alpha^{(k)}}x^{\alpha^{(k)}}$$

where $\alpha^{(k)}\in\mathbb{Z}_{\geq 0}^{k-1}$ denotes the exponent vector $\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_k)$ where we omit the k-th exponent. One important detail to note is that the coefficient in this sum is a multinomial coefficient. If we insert this into our equation we then get the following equation

$$\begin{aligned}
&\sum_{\alpha_k=0}^{m}\binom{m}{\alpha_k}x_k^{\alpha_k}\left(\sum_{|\alpha^{(k)}|=m-\alpha_k}\binom{m-\alpha_k}{\alpha^{(k)}}x^{\alpha^{(k)}}\right)\\
=&\sum_{\alpha_k=0}^{m}\sum_{|\alpha^{(k)}|=m-\alpha_k}\binom{m}{\alpha_k}\binom{m-\alpha_k}{\alpha^{(k)}}x_k^{\alpha_k}x^{\alpha^{(k)}}
\end{aligned}$$

Now note that $x_k^{\alpha_k}x^{\alpha^{(k)}}=x^\alpha$ and that the product of the coefficients is given by

$$\binom{m}{\alpha_k}\binom{m-\alpha_k}{\alpha^{(k)}}=\frac{m!}{\alpha_k!(m-\alpha_k)!}\frac{(m-\alpha_k)^!}{\alpha_1!\alpha_2!\ldots\alpha_{k-1}!}=\frac{m!}{\alpha_1!\alpha_2!\ldots\alpha_k!}=\binom{m}{\alpha}$$

Thus we can conclude that

$$\sum_{\alpha_k=0}^{m} \sum_{|\alpha^{(k)}|=m-\alpha_k} \binom{m}{\alpha_k} \binom{m-\alpha_k}{\alpha^{(k)}} x_k^{\alpha_k} x^{\alpha^{(k)}} = \sum_{|\alpha|=m} \binom{m}{\alpha} x^{\alpha}$$

and this completes our proof. □

We will now state the theorem

Theorem 4.12 *Given a finite matrix group $G \subseteq GL(n,k)$, let $x^{\beta_1}, x^{\beta_2}, \ldots, x^{\beta_m}$ be monomials with total degree lower than $|G|$. Then*

$$k[x_1, x_2, \ldots, x_n]^G = k[R_G(x^{\beta_1}), R_G(x^{\beta_2}), \ldots, R_G(x^{\beta_m})] = k[R_G(x^{\beta}) \,|\, |\beta| \leq |G|]$$

More specifically $k[x_1, x_2, \ldots, x_n]^G$ is finitely generated by homogeneous invariants.

Proof Let $f \in k[x_1, x_2, \ldots, x_n]^G$ be given by $f = \sum_\alpha c_\alpha x^\alpha$. By proposition 4.10 we have that

$$f = R_G(f) = R_G\left(\sum_\alpha c_\alpha x^\alpha\right) = \sum_\alpha c_\alpha R_G(x^\alpha)$$

Thus we can write every invariant polynomial as a linear combination of the $R_G(x^\alpha)$. So the only thing that is left is to prove that we can express the $R_G(x^\alpha)$ as a polynomial in the $R_G(x^\beta)$ for $|\beta| \leq |G|$.

For this we will make use of the identity shown in lemma 4.11. For this we need some notation, so let $A = (a_{ij}) \in G$ and we will denote the i-th row of A with A_i. This gives us that $A_i \cdot \mathbf{x} = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$ and for $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n_{\geq 0}$ we have that

$$(A \cdot \mathbf{x})^\alpha = (A_1 \cdot \mathbf{x})^{\alpha_1} (A_2 \cdot \mathbf{x})^{\alpha_2} \ldots (A_n \cdot \mathbf{x})^{\alpha_n}$$

The Reynolds operator is thus given by

$$R_G(x^\alpha) = \frac{1}{|G|} \sum_{A \in G} (A \cdot \mathbf{x})^\alpha$$

We will introduce new variables $u_1, u_2, \ldots, u_n$ and make the substition $x_i \mapsto u_i A_i \cdot \mathbf{x}$ in the identity of lemma 4.11. This gives us the equation

$$(u_1 A_1 \cdot \mathbf{x} + u_2 A_2 \cdot \mathbf{x} + \cdots + u_n A_n \cdot \mathbf{x})^m = \sum_{|\alpha|=m} a_\alpha (A \cdot \mathbf{x})^\alpha u^\alpha$$

We can now sum over all $A \in G$ and this gives us

$$\begin{aligned} S_m &= \sum_{A \in G} (u_1 A_1 \cdot \mathbf{x} + u_2 A_2 \cdot \mathbf{x} + \cdots + u_n A_n \cdot \mathbf{x})^m \\ &= \sum_{|\alpha|=m} a_\alpha \left(\sum_{A \in G} (A \cdot \mathbf{x})^\alpha \right) u^\alpha \\ &= \sum_{|\alpha|=m} b_\alpha R_G(x^\alpha) u^\alpha \end{aligned}$$

with $b_\alpha = a_\alpha |G|$. Here the last sum contains all $R_G(x^\alpha)$ with $|\alpha| = m$ and the variables $u_1, u_2, \ldots, u_n$ ensure that no two terms cancel eachother out.

Furthermore S_m is the m-th of the $|G|$ objects

$$U_A = u_1 A_1 \cdot \mathbf{x} + u_2 A_2 \cdot \mathbf{x} + \cdots + u_n A_n \cdot \mathbf{x}$$

which are indexed by $A \in G$. Since S_m is a power sum it is symmetric in the U_A, thus by theorem 3.10 we write it as a polynomial in $S_1, S_2, \ldots, S_{|G|}$. So

$$S_m = F(S_1, S_2, \ldots, S_{|G|})$$

where F is a polynomial with coefficients in k. We can substitute into our equation which give us

$$\sum_{|\alpha|=m} b_\alpha R_G(x^\alpha) u^\alpha = F(\sum_{|\beta|=1} b_\beta R_G(x^\beta) u^\beta, \sum_{|\beta|=2} b_\beta R_G(x^\beta) u^\beta, \ldots, \sum_{|\beta|=|G|} b_\beta R_G(x^\beta) u^\beta)$$

If we expand the right hand side and equate the coefficients of u^α, then we find that every $b_\alpha R_G(x^\alpha)$ is a polynomial in the $R_G(x^\beta)$. These have total degree less than $|G|$ and furthermore the coefficient $b_\alpha = a_\alpha |G|$ is nonzero, since our field k is of characteristic zero. Hence we can conclude our proof. □

With this theorem we find that we can finitely generate the ring of invariants. Furthermore we can find such a basis by applying the Reynolds operator on all monomials with total degree less or equal to $|G|$.

Let us return to the cyclic group $C_4 \subseteq \mathrm{GL}(2, k)$ of order 4 from before. Theorem 4.12 then tells us that we can find the ring of invariants if we compute $R_{C_4}(x^i y^j)$ for all $i + j \leq 4$. Then we get the following results:

$x^i y^j$	$R_{C_4}(x^i y^j)$	$x^i y^j$	$R_{C_4}(x^i y^j)$
x	0	xy^2	0
y	0	y^3	0
x^2	$\frac{1}{2}(x^2 + y^2)$	x^4	$\frac{1}{2}(x^4 + y^4)$
xy	0	$x^3 y$	$\frac{1}{2}(x^3 y + xy^3)$
y^2	$\frac{1}{2}(x^2 + y^2)$	$x^2 y^2$	$x^2 y^2$
x^3	0	xy^3	$-\frac{1}{2}(x^3 y + xy^3)$
$x^2 y$	0	y^4	$\frac{1}{2}(x^4 + y^4)$

Thus $k[x, y]^{C_4}$ is generated by $x^2 y^2, x^2 + y^2, x^3 y + xy^3$ and $x^4 + y^4$. However we have that $x^4 + y^4 = (x^2 + y^2)^2 - 2x^2 y^2$. So we can write our ring of invariants as

$$k[x, y]^{C_4} = k[x^2 y^2, x^2 + y^2, x^3 y + xy^3]$$

For another example consider the cyclic group $C_3 \subseteq \mathrm{GL}(2, k)$ of order 3 generated by

$$A = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$$

The Reynolds operator is then given by

$$R_{C_3} f(x, y) = \frac{1}{3}(f(x, y) + f(-y, x - y) + f(-x + y, -x))$$

and we get the following results when we compute $R_{C_3}(x^i y^j)$ for $i + j \leq 3$:

x^iy^j	$R_{C_4}(x^iy^j)$	x^iy^j	$R_{C_4}(x^iy^j)$	x^iy^j	$R_{C_4}(x^iy^j)$
x	0	xy	$\frac{1}{3}(x^2-xy+y^2)$	x^2y	$\frac{1}{3}(-x^3+3x^2y-y^3)$
y	0	y^2	$\frac{2}{3}(x^2-xy+y^2)$	xy^2	$\frac{1}{3}(-x^3+3xy^2-y^3)$
x^2	$\frac{2}{3}(x^2-xy+y^2)$	x^3	x^2y-xy^2	y^3	$-(x^2y-xy^2)$

Here we find another relation between these polynomials namely

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

Thus the ring of invariants is given by

$$k[x,y]^{C_3}=k[x^2-xy+y^2,-x^3+3xy^2-y^3,-x^3+3x^2y-y^3]$$

Given a finite matrix group G we now know that we can find a finite basis such that $k[x_1,x_2,\ldots,x_n]^G=k[f_1,f_2,\ldots,f_m]$. One question we can ask ourselves is whether we can determine if a polynomial f lies in $k[f_1,f_2,\ldots f_m]$. The following proposition answers this question in a way similar to proposition 3.6. Furthermore it also shows us a way to write f as a polynomial in $f_1,f_2,\ldots,f_m$

Proposition 4.13 *Let $f_1,f_2,\ldots,f_m\in k[x_1,x_2,\ldots,x_n]$ be given. Fix a monomial order in $k[x_1,x_2,\ldots,x_n,y_1,y_2,\ldots,y_m]$ such that any monomial containing one of $x_1,x_2,\ldots,x_n$ is greater than all monomials in $k[y_1,y_2,\ldots,y_m]$. Let G be a Gröbner basis of the ideal $\langle f_1-y_1,\ldots,f_m-y_m\rangle\subseteq k[x_1,x_2,\ldots,x_n,y_1,y_2,\ldots,y_m]$. Given $f\in k[x_1,x_2,\ldots,x_n]$ and let $g=\overline{f}^G$ be a remainder of f on division by G. Then we have that:*

1. *$f\in k[f_1,f_2,\ldots,f_m]$ if and only if $g\in k[y_1,y_2,\ldots,y_n]$*
2. *If $f\in k[f_1,f_2,\ldots,f_m]$, then $f=g(f_1,f_2,\ldots,f_m)$ is an expression of f as a polynomial in $f_1,f_2,\ldots,f_m$*

Proof Let $f\in k[x_1,x_2,\ldots,x_n]$, then its division by $G=\{g_1,g_2,\ldots,g_s\}$ is given by

$$f=A_1g_1+A_2g_2+\cdots+A_sg_s+g$$

where $A_1,A_2,\ldots,A_s,g\in k[x_1,x_2,\ldots,x_n,y_1,y_2,\ldots,y_m]$.

Let us first assume that $g\in k[y_1,y_2,\ldots,y_m]$. Then we apply the substitution $y_i\mapsto f_i$ for all i. This does not affect f as it is a polynomial in $x_1,x_2,\ldots,x_n$. Note that every polynomial in $\langle f_1-y_1,\ldots,f_m-y_m\rangle$ vanishes under this substition. Since the polynomials $g_1,g_2,\ldots,g_m$ lie in this ideal they vanish as well, which reduces our equation to $f=g(f_1,f_2,\ldots,f_m)$. So $f\in k[f_1,f_2,\ldots,f_m]$.

Conversely, assume that $f=h(f_1,f_2,\ldots,f_m)$. Similar to our proof in proposition 3.6 we rewrite our equation to

$$f=C_1(f_1-y_1)+C_2(f_2-y_2)+\cdots+C_s(f_m-y_m)+h(y_1,y_2,\ldots,y_m)$$

We need to show that h is the remainder of f on division by G. For this we consider $G'=G\cap k[y_1,y_2,\ldots,y_m]$, as this gives us the elements in G that only

contain $y_1, y_2, \ldots, y_m$. We then have that $G' = \{g_1, g_2, \ldots, g_t\}$ where we relabel if necessary and $t \leq s$. We then divide h on G' which gives us

$$h = B_1 g_1 + B_2 g_2 + \cdots + B_t g_t + g'$$

where $B_1, B_2, \ldots, B_s, g; \in k[y_1, y_2, \ldots, y_m]$. We can substitute this back into our equation which gives us

$$f = C_1'(f_1 - y_1) + C_2'(f_2 - y_2) + \cdots + C_m'(f_m - y_m) + g'$$

Since each g_i lies in $\langle f_1 - y_1, \ldots, f_m - y_m \rangle$, we can write them as polynomials in $f_1 - y_1, \ldots, f_m - y_m$. If we can prove that g' is the remainder of f on division by G, then the remainder lies in $k[y_1, y_2, \ldots, y_m]$.

By proposition 2.27 we know that g' is a remainder of f on division by G if no $\mathtt{LT}(g_i)$ divides a term of g, as G is a Gröbner basis. Assume that there is a g_i such that $\mathtt{LT}(g_i)$ divides a term of g'. Then $\mathtt{LT}(g_i)$ involves only $y_1, y_2, \ldots, y_m$, since g' lies in $k[y_1, y_2, \ldots, y_m]$. Since any monomial that contains one of $x_1, x_2, \ldots, x_n$ is greater than all monomials in $k[y_1, y_2, \ldots, y_m]$, it follows that $g_i \in k[y_1, y_2, \ldots, y_m]$. Thus g_i is also an element of G'. But g' is a remainder on division by G', so this leads to a contradiction as $\mathtt{LT}(g_i)$ cannot divide any terms of g'. Hence g' is the remainder of f by division on G.

The second part follows immediately from the above arguments and we conclude our proof $\square$

5 Conclusion

In the first chapter we have introduced polynomials in n variables and looked at their differences with regards to polynomials in one variable. The key difference was the fact that there is no canonical ordering on multivariate polynomials and this introduced several problems that we otherwise would not have to worry about. We solved this problem by introducing monomial orderings and this further allowed us to construct a division algorithm in $k[x_1, x_2, \ldots, x_n]$. By the end of the chapter we proved Hilbert's basis theorem, which is a major result as it answers the question whether every ideal is finitely generated. Furthermore we introduced Gröbner bases, which combined with the division algorithm allows us to determine whether a given polynomial lies in a ideal I.

In the second chapter we introduced symmetric polynomials. We have studied the elementary symmetric polynomials and determined whether we can express all symmetric polynomials as a polynomial in the elementary symmetric polynomials in the fundamental theorem. Furthermore, in the proof of this theorem we described an algorithm for this decomposition. Then by using division on a Gröbner basis G, we found a way to check whether a polynomial is symmetric or not. At last we discussed Newton's identity, which gave us a correspondence between symmetric polynomials and the power sums.

In the final chapter we generalized and we looked at polynomials invariant under finite matrix groups. We examined their ring of invariants by computing various examples and looked at how we could characterize them. We found that a polynomial lies in this ring, whenever it is invariant under the generators of

the matrix group. After this we examined how we could generate the ring of invariants and we introduced the Reynolds operator as a solution. We found that the Reynolds operator produced invariant polynomials. But most importantly, we found through Emmy Noether's theorem, that we could generate the ring of invariants by applying the Reynolds operator on certain monomials. To close this chapter we looked at an alternative way to determine whether a polynomial lies in the ring of invariants, similar to how we did with symmetric polynomials.

References

[1] David A. Cox, John Little, Donal O'Shea: Ideals, Varieties and Algorithms
Undergraduate Texts in Mathematics, Springer International Publishing, 2015

[2] Frits Beukers: Rings and Galois theory
Department of Mathematics, Utrecht University, 2018

[3] Bruno Buchberger, An algorithm for finding the basis elements of the residue class ring of azero dimensional polynomial ideal
Johannes Kepler University of Linz (JKU), 1965

[4] Steven H. Weintraub: Galois theory
Universitext, Springer-Verlag New York, 2006

[5] Bernd Sturmfels: Algorithms in invariant theory
Texts & Monographs in Symbolic Computation, Springer-Verlag Wien, Second edition, 2008

[6] Walter Rudin: Principles of Mathematical Analysis
Mcgraw-Hill Education - Europe, Third edition, 1976