Chapter 1

Invariant theory of finite groups

1.1 Introduction

In this introduction we shall make some historical remarks and give some examples[1]. Some of the basic theorems and concepts of computational algebra can be found in 19th century papers on classical invariant theory. The roots of invariant theory can be traced back to Lagrange (1773-1775) and Gauss (1801) who were interested in the problem of representing integers by quadratic binary forms and used the discriminant to distinguish between non equivalent forms.

Algebraic invariants such as the discriminant show up also in algebraic geometry when one asks for properties of geometric objects which are invariant under certain classes of transformations. For example, the geometric significance of the discriminant is that a quadratic binary form defines two distinct points on the projective line $\mathbb{P}^1(\mathbb{C})$ if and only if its discriminant is non zero.[2] People became interested in such invariant properties especially after the introduction of homogeneous coordinates by Moebius (1827) and Plucker (1830)[3]. This was a major impetus for invariant theory.

In the first decades of invariant theory (1840-1870), people were mainly concerned with the discovery of particular invariants. The major case of interest was that of forms of degree d in n variables with $SL_n(\mathbb{C})$ acting by linear substitution (see example 1.7).

In order to understand some of the most basic questions which can be studied in invariant theory we shall consider two simple examples first. They are both concerned with finite groups which will be the main object of our

[1] This section is mainly taken from [5], [6], [3]

[2] The geometric aspects of the invariant theory of binary forms is explained in [8]

[3] Find the command for umlauts

investigation. We shall assume, unless otherwise specified, that the base field k has characteristic zero, the group G is a finite group of matrices and the action of $M \in G$ over a polynomial f is

$$M \cdot f(\mathbf{x}) = f(M \cdot \mathbf{x}).$$

Note that $\mathbf{x}$ is to be thought of as a column vector. However when we write $f(\mathbf{x})$ componentwise, like in $f(x_1, x_2)$, we shall write it as a row vector for simplicity.

Example 1.1. [4] Consider the finite group

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

This is sometimes called the *Klein four group*. We consider the action

$$g \cdot f(x) = f(g \cdot (x)).$$

If a polynomial $f \in \mathbb{K}[x, y]$ is invariant under V_4 then

$$f(x, y) = f(-x, y) = f(x, -y)$$

and it is immediate to show that the converse is true. If

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

the condition $f(x, y) = f(-x, y)$ is equivalent to $a_{ij} = 0$ for i odd and the condition $f(x, y) = f(x, -y)$ is equivalent to $a_{ij} = 0$ for j odd. Thus we can write

$$f(x, y) = g(x^2, y^2)$$

for a *unique* polynomial $g(x, y) \in \mathbb{K}[x, y]$. Conversely, every polynomial of this form is invariant under V_4. This proves that

$$\mathbb{K}[x, y]^{V_4} = \mathbb{K}[x^2, y^2]$$

In the example 1.2 we shall see that if we consider invariants for other groups, even for a subgroup of V_4 itself, things may become more complicated.

[4] Taken from [3]

Example 1.2. [5] Let

$$C_2 = \left\{ \pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\} \subseteq V_2$$

Of course a polynomial which is invariant under V_4 is also invariant under C_2 but now we have more invariant polynomials, like for example xy. It is not hard to show that f is invariant under C_2 if and only if we can write

$$f(x, y) = g(x^2, y^2, xy)$$

for *at least* a polynomial $g(x, y) \in \mathbb{K}[x, y]$ and therefore

$$\mathbb{K}[x, y]^{C_2} = \mathbb{K}[x^2, y^2, xy].$$

The ring $\mathbb{K}[x^2, y^2, xy]$ however is fundamentally different from the previous example because uniqueness breaks down: a given invariant can be written in terms of x^2, y^2 and xy in more than one way. For example x^4y^2 is clearly invariant and

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

Example 1.3. Let

$$S_2 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\} \subseteq V_2$$

The group S_2 is isomorphic to C_2 but its action on $\mathbb{K}[x, y]$ is a different one. It can be shown (see example 1.4 for a general statement about symmetric groups) that

$$\mathbb{K}[x, y]^{S_2} = \mathbb{K}[x + y, xy]$$

Example 1.4. [6]. The claims of this example will be proved in the next section. Suppose the symmetric group S_n acts on $V = \mathbb{K}^n$ by

$$\sigma \cdot (x_1, x_2, \ldots, x_n) = (x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}), \quad \sigma \in S_n.$$

Let us write

$$\phi(t) := (t - x_1)(t - x_2) \cdots (t - x_n) = t^n - \sigma_1 t^{n-1} + \sigma_2 t^{n-2} - \cdots + (-1)^n \sigma_n$$

[5] Taken from [3]

[6] Taken from [6]

with $\sigma_1, \ldots, \sigma_n \in \mathbb{K}[x_1, \ldots, x_n]$ the so called *elementary symmetric polynomials.* Formulas for $\sigma_1, \ldots, \sigma_n$ are given by:

$$\begin{array}{rcl}
\sigma_1 & = & x_1 + x_2 + \cdots + x_n \\
\sigma_2 & = & x_1x_2 + x_1x_3 + \cdots + x_1x_n + x_2x_3 + \cdots + x_{n-1}x_n \\
\vdots & \vdots & \vdots \\
\sigma_r & = & \displaystyle\sum_{i_1<i_2<\ldots i_r} x_{i_1}x_{i_2}\cdots x_{i_r} \\
\vdots & \vdots & \vdots \\
\sigma_n & = & x_1x_2\cdots x_n
\end{array}$$

Claim The invariant ring of S_n in this representation is generated by the algebraically independent invariants $\sigma_1, \ldots, \sigma_n$.
This claim will be proved in the next section.

From these examples we see that given a matrix group G, invariant theory has two basic questions to answer about the ring of invariants $\mathbb{K}[x_1, \ldots, x_n]^G$

1. *Finite generation* Can we find finitely many homogeneous invariants $f_1, \ldots, f_m$ such that every invariant is a polynomial in $f_1, \ldots, f_m$?

2. *Uniqueness* In how many ways can an invariant be written in terms of $f_1, \ldots, f_m$ i.e. how to describe the ideal of relations of $f_1, \ldots, f_m$?

For finite groups acting on a ring of polynomials with coefficient in an algebraically closed field of characteristic zero we will give complete answers to both questions in Section 1.4 and describe an algorithm for finding all invariants and all relations between them.
For completing this program we need to introduce a fundamental tool for doing computations in polynomial rings, i.e. Groebner basis. However, before introducing it in Section 1.3, we shall deal with the problem of finding all polynomial invariants for the symmetric group in Section 1.4. This will give a concrete introduction to the problem of introducing a complete linear order on polynomials and to the problem of generalizing the division algorithm of one variable polynomials, which are the two problems which lie at the root of the theory of Groebner basis
Things are in general more difficult for infinite groups. The following examples deal with this more general situation

Example 1.5. [7] Suppose that $char(\mathbb{K}) = 0$. Gordan proved that the invariant rings of the 2-dimensional special linear group $SL(2, \mathbb{K})$ over $\mathbb{K}$ are always finitely generated (see [9]).

[7] Taken from [6].

Let V_d be the vector space

$$\{a_0x^d + a_1x^{d-1}y + \cdots + a_dy^d \mid a_0, a_1, \ldots, a_d \in K\}$$

of homogeneous polynomials of degree d in x and y. Such polynomials are often referred to as *binary forms.* The coordinate ring $\mathbb{K}[V_d]$ can be identified with $\mathbb{K}[a_0, a_1, \ldots, a_d]$. We can define an action of $SL(2, \mathbb{K})$ on V_d by

$$\sigma \cdot g(x, y) := g(\alpha x + \gamma y, \beta x + \delta y), \quad \sigma = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in Sl(2\mathbb{K})$$

There are many ways of constructing invariants for binary forms. One important invariant of a binary form is the *discriminant.*
For $d = 2$, one has $\mathbb{K}[V_2]^{SL_2} = \mathbb{K}[\Delta(g_2)]$, where $\Delta(g_2) = a_1^2 - 4a_0a_2$ is the well known discriminant of a quadratic polynomial $g_2 = a_0x^2 + a_1xy + a_2y^2$.
For $d = 3$, one has $\mathbb{K}[V_3]^{SL_2} = \mathbb{K}[\Delta(g_3)]$, where

$$\Delta(g_3) = a_1^2a_2^2 - 4a_0a_2^3 - 4a_1^3a_3 - 27a_0^2a_3^2 + 18a_0a_1a_2a_3$$

is the discriminant of a cubic polynomial $g_3 = a_0x^3 + a_1x^2y + a_2xy^2 + a_3y^3$.
For $d = 4$, one has $\mathbb{K}[V_4]^{SL_2} = \mathbb{K}[f_2, f_3]$, where

$$f_2 = a_0a_4 - \frac{1}{4}a_1a_3 + \frac{1}{12}a_2^2 \text{ and } f_3 = det\begin{pmatrix} a_0 & a_{12}/4 & a_2/6 \\ a_1/4 & a_2/6 & a_3/4 \\ a_2/6 & a_3/4 & a_4 \end{pmatrix}$$

The discriminant $\Delta(g_4)$ can be expressed in f_2 and f_3, namely $\Delta(g_4) = 2^8(f_2^3 - 27f_3^2)$. For $d = 5, 6, 8$, the invariant rings are also explicitly known (see [13] and [15]). See also [12]

In Example 1.5 the ring of invariants is finitely generated even if the group G was not finite. In one of his famous problems (the fourteenth) Hilbert raised the question if the ring of invariants is always [8] finitely generated, but this is not always the case, as the following construction due to Nagata proves.

Example 1.6. [9] This is the counter example of Nagata to Hilbert's fourteenth problem. Take $\mathbb{K} = \mathbb{C}$ and complex numbers $a_{i,j}$ algebraically independent over $\mathbb{Q}$ where $i = 1, 2, 3$ and $j = 1, 2, \ldots, 16$. Let $G \subseteq GL(32, \mathbb{C})$ be the group of all block diagonal matrices

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_{16} \end{pmatrix}$$

[8] check
[9] Taken from [6]

where

$$A_j = \begin{pmatrix} c_j & c_j b_j \\ 0 & c_j \end{pmatrix}$$

for $j = 1, 2, \ldots, 16$. Here the c_j and b_j are arbitrary complex numbers such that $c_1 c_2 \cdots c_{16} = 1$ and $\sum_{j=1}^{16} a_{i,j} b_j = 0$ for $i = 1, 2, 3$. Then $\mathbb{K}[x_1, \ldots, x_{32}]^G$ is not finitely generated (see [10]).

Example 1.7. [10] Let $V = \mathbb{K}^n$. The group $GL(V)$ acts on $End(V)$ by conjugation.[11]

$$\sigma \cdot A := \sigma A \sigma^{-1}, \qquad \sigma \in GL(V),\ A \in End(V).$$

The characteristic polynomial of $A \in End(V)$ is given by

$$\chi(t) := det(tI - A) = t^n - g_1 t^{n-1} + g_2 t^{n-2} - \cdots + (-1)^n g_n.$$

We view $g_1, \ldots, g_n$ as functions of A. The coefficients $g_i \in \mathbb{K}[End(V)]$ are clearly invariant under the action of $GL(V)$. Let us show that $\mathbb{K}[End(V)]^G = \mathbb{K}[g_1, \ldots, g_n]$. Consider the set of diagonal matrices

$$\mathcal{T} := \{ \begin{pmatrix} x_1 & & & \\ & x_2 & & \\ & & \ddots & \\ & & & x_4 \end{pmatrix} | x_1, x_2, \ldots, x_n \in k\}$$

The group S_n can be viewed as the subgroup of GL_n of permutation matrices. The set $\mathcal{T}$ is stable under the action of S_n. The restriction of $\chi(t)$ to $\mathcal{T}$ is S_n-invariant, in fact it is equal to

$$\phi(t) := (t - x_1)(t - x_2) \cdots (t - x_n)$$

Restricting g_i to $\mathcal{T}$ yields the elementary symmetric polynomial f_i. It follows that $g_1, \ldots, g_n$ are algebraically independent. If $h \in k[\mathbb{K}[End(V)]^{GL(V)}$, then the restriction of h to $\mathcal{T}$ is S_n-invariant. We can find a polynomial ψ such that the restriction of h to $\mathcal{T}$ is equal to $\psi(f_1, \ldots, f_n)$. Let U be the set of matrices which have distinct eigenvalues. Every matrix with distinct eigenvalues can be conjugated into $\mathcal{T}$, so $U \subseteq G \cdot \mathcal{T}$. The set U is Zariski dense since it is the complement of the Zariski closed set defined by $\Delta(\chi) = 0$. It follows that $h = \psi(g_1, \ldots, g_n)$ because $h - \psi(g_1, \ldots, g_n)$ vanishes on $G \cdot \mathcal{T} \supset U$. The trick of this example (reducing the computation of $k[\mathbb{K}[V]^G$ to the computation of $k[\mathbb{K}[W]^H$ with $W \subseteq U$ and $H \subseteq G$) works in a more general setting (see [11]).

[10] Taken from [6]

[11] $Gl(V)$ is the group of invertible linear transformations $L : V \to V$. It is isomorphic to $Gl(n, \mathbb{K})$ and an explicit isomorphism can be given by choosing a basis $\{e_1, \ldots, e_n\}$ in V. The isomorphism $L \to A = (a_{ij}) \in Gl(n, \mathbb{K})$ is given explicitly by $A(e_j) := \sum a_{ij} e_i$

1.2 Monomial ordering and Symmetric Polynomials

In this Section we shall prove the Claim in Example 1.4, i.e. we describe completely the ring of polynomial invariants under the action of the symmetric group[12]. Let us recall here the definition of the elementary symmetric polynomials in n variables.

$$\begin{aligned}
\sigma_1 &= x_1 + x_2 + \cdots + x_n \\
\sigma_2 &= x_1x_2 + x_1x_3 + \cdots + x_1x_n + x_2x_3 + \cdots + x_{n-1}x_n \\
\vdots &\ \vdots\ \vdots \\
\sigma_r &= \sum_{i_1<i_2<\ldots i_r} x_{i_1}x_{i_2}\cdots x_{i_r} \\
\vdots &\ \vdots\ \vdots \\
\sigma_n &= x_1x_2\cdots x_n
\end{aligned}$$

Exercise 1.1. *Cocoa* Define a function which returns the elementary symmetric polynomials

From the elementary symmetric polynomials we can construct other symmetric polynomials by taking polynomials in $\sigma_1, \ldots, \sigma_n$. Thus for example, for $n = 3$

$$\sigma_2^2 - \sigma_1\sigma_3 = x_1^2x_2^2 + x_1^2x_2x_3 + x_1^2x_3^2 + x_1x_2^2x_3 + x_1x_2x_3^2 + x_2^2x_3^2$$

is a symmetric polynomial. The result we want to prove here is that *all* symmetric polynomials can be *uniquely* represented in this way. To prove this we follow Gauss. We use an inductive procedure and the notion we need to carry on this induction is that of an order on monomials[13]. Ordering monomials in the polynomial ring $\mathbb{K}[x]$ is simple: one does this by degree. This ordering is implicit in the usual division algorithm for polynomials in $\mathbb{K}[x]$. However for $\mathbb{K}[x, y]$ or $\mathbb{K}[x_1, \ldots, x_n]$ it is less clear how to order monomials. We write monomials in $x_1, \ldots, x_n$ as

$$x^\alpha = x_1^{a_1} \cdots x_n^{a_n}$$

so that $\alpha = (a_1, \ldots, a_n) \in \mathbb{Z}^n_{\geq 0}$ is the vector of exponents. Then a *monomial order* is any total order $>$ on monomials with the following two properties.

[12] This section is mainly taken from [3]

[13] This section on ordering is taken by [4]

1. (Well ordering) The order $>$ is a well ordering on the set of monomials, i.e., any non empty subset of monomials has at least element under $>$.
2. (Compatibility) If $x^\alpha > x^\beta$, then $x^\alpha x^\gamma > x^\beta x^\gamma$ for any monomial x^γ.

Exercise 1.2. (Characteristic zero) A monomial order $>$ has the property $x^\alpha > 1$ whenever $x^\alpha \neq 1$.

Solution $1 > x^\alpha$ would imply $1 > x^\alpha > x^{2\alpha} > x^{3\alpha} > \dots$ by compatibility, which would contradict well ordering.

Once we have specified a monomial order $>$ on $\mathbb{K}[x_1, \dots, x_n]$, we can order the terms of a polynomial. If we write $f \in \mathbb{K}[x_1, \dots, x_n]$ as

$$f = \sum_\alpha c_\alpha x^\alpha, \quad c_\alpha \in \mathbb{K}, \tag{1.1}$$

then a *term* of f is $c_\alpha x^\alpha$ for $c_\alpha \neq 0$.

Definition 1.1. Given a monomial order $>$ on $\mathbb{K}[x_1, \dots, x_n]$, and a non zero f as in (1.1)

1. The *multidegree* of f is

$$\text{multideg}(f) = \max(\alpha \in \mathbb{Z}^n_{\geq 0} : c_\alpha \neq 0)$$

(the maximum is taken with respect to $>$)

2. The *leading coefficient* of f is

$$LC(f) = c_{\text{multideg}(f)}$$

3. The *leading monomial* of f is

$$LM(f) = x^{\text{multideg}(f)}$$

4. The *leading term* of f is

$$LT(f) = LC(f) \cdot LM(f) = max_>\{c_\alpha x^\alpha | c_\alpha \neq 0\}$$

where $max_>$ means the maximum with respect to $>$.

The leading term is sometimes called the *initial term.*
One of the simplest monomial orders is *lexicographic* order (lex for short) To define it, we order variables first, say

$$x_1 > x_2 > \cdots > x_n$$

We then define

$$ax_1^{i_1}x_2^{i_2}\cdots x_n^{i_n} > bx_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}$$

if $i_1 > j_1$, or if $i_1 = j_1$ and $i_2 > j_2$, or if $i_1 = j_1$ and $i_2 = j_2$ and $i_3 > j_3$, or ... If we list the variables differently, we get a different lex order, so that there are $n!$ possible lex orders on monomials in $\mathbb{K}[x_1, \ldots, x_n]$.

Exercise 1.3. Let $f = 4xy^2z + 4z^2 - 5x^3 + 7x^2z^2$ and let $>$ be lex order as above. Compute multidegree(f), LC(f), LM(f) and LT(f)

Solution multideg$(f) = (3, 0, 0)$; $LC(f) = -5$, $LM(f) = x^3$, $LT(f) = -5x^3$.

Exercise 1.4. Find the leading term of the polynomial $p = 12x_1^3 + 3x_1x_2x_3 - 7x_1^2x_2$ with respect to lex when $x_1 > x_2 > x_3$, when $x_2 > x_3 > x_1$ and when $x_2 > x_1 > x_3$.

Solution $12x_1^3$, $3x_1x_2x_3$ and $-7x_1^2x_2$ respectively[14].

We are now ready to prove the *fundamental theorem of symmetric polynomials.*[15]

Theorem 1.1. Every symmetric polynomial in $\mathbb{K}[x_1, \ldots, x_n]$ can be written uniquely as a polynomial in the elementary symmetric functions $\sigma_1, \ldots, \sigma_n$

Proof The proof is taken from [3]. We will use lex order with $x_1 > x_2 > \cdots > x_n$. Given a non zero symmetric polynomial $f \in \mathbb{K}[x_1, \ldots, x_n]$, let $LT(f) = ax^\alpha$. If $\alpha = (\alpha_1, \ldots, \alpha_n)$, we first claim that $\alpha_1 \geq \alpha_2 \cdots \geq \alpha_n$. To prove this, suppose that $\alpha_i < \alpha_{i+1}$ for some i. Let β be the exponent vector obtained from α by switching α_i and α_{i+1}. We shall write it as $\beta = (\ldots, \alpha_{i+1}, \alpha_i, \ldots)$. Since ax^α is a term of f, it follows that ax^β is a term of $f(\ldots, x_{i+1}, x_i, \ldots) = f$, the last equality by symmetry, and thus ax^β is a term of f. This is impossible since $\beta > \alpha$ under lex order, and our claim is proved.
Now let

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

[14] Perhaps it is better to introduce the division algorithm before the proof of the claim

[15] In general there is a first fundamental theorem which gives a set of generators for the ring of invariants and a second fundamental theorem which describes the relations (syzigies) among generators

To compute the leading term of h, first note that $LT(\sigma_r) = x_1x_2\cdots x_r$ for $1 \leq r \leq n$. Hence,

$$\begin{array}{rcl} LT(h) & = & LT(\sigma_1^{\alpha_1-\alpha_2}\sigma_2^{\alpha_2-\alpha_3}\cdots\sigma_{n-1}^{\alpha_{n-1}-\alpha_n}\sigma_n^{\alpha_n}) \\ & = & LT(\sigma_1)^{\alpha_1-\alpha_2}LT(\sigma_2)^{\alpha_2-\alpha_3}\cdots LT(\sigma_{n-1})^{\alpha_{n-1}-\alpha_n}LT(\sigma_n)^{\alpha_n} \\ & = & x_1^{\alpha_1-\alpha_2}(x_1x_2)^{\alpha_2-\alpha_3}\cdots(x_1\cdots x_n)^{\alpha_n} \\ & = & x^\alpha. \end{array}$$

It follows that f and ah have the same leading term, and thus

$$\text{multideg}(f - ah) < \text{multideg}(f)$$

whenever $f - ah \neq 0$.
Now set $f_1 = f - ah$ and note that f_1 is symmetric since f and ah are. Hence, if $f_1 \neq 0$, we can repeat the above process to form $f_2 = f_1 - a_1h_1$, where a_1 is a constant and h_1 is a product of powers of symmetric functions, defined as above. Further we know that $LT(f_2) < LT(f_1)$ when $f_2 \neq 0$. Continuing in this way, we get a sequence of polynomials $f, f_1, f_2, \ldots$ with

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

Since lex order is a well ordering, the sequence must be finite, hence $f_{t+1} = 0$ for some t and it follows easily that

$$f = ah + a_1h_1 + \cdots + a_th_t$$

which shows that f is a polynomial in the symmetric functions.
Finally we need to prove uniqueness. Suppose that we have a symmetric polynomial f which can be written

$$f = g_1(\sigma_1, \ldots, \sigma_n) = g_2(\sigma_1, \ldots, \sigma_n)$$

Hence g_1 and g_2 are polynomials in n variables, say $y_1, \ldots, y_n$. We need to prove that $g_1 = g_2 \in \mathbb{K}[y_1, \ldots, y_n]$. If we set $g = g_1 - g_2$, then $g(\sigma_1, \ldots, \sigma_n) = 0$ in $\mathbb{K}[x_1, \ldots, x_n]$ and we need to prove that $g = 0$ in $\mathbb{K}[y_1, \ldots, y_n]$. Suppose that $g \neq 0$. If we write $g = \sum_\beta a_\beta y^\beta$, then $g(\sigma_1, \ldots, \sigma_n)$ is a sum of the polynomials $g_\beta = a_\beta\sigma_1^{\beta_1}\sigma_2^{\beta_2}\cdots\sigma_n^{\beta_n}$, where $\beta = (\beta_1, \ldots, \beta_n)$. Furthermore, the argument used above to show that $LT(h) = x^\alpha$ shows that

$$LT(g_\beta) = a_\beta x_1^{\beta_1+\cdots\beta_n}x_2^{\beta_2+\cdots\beta_n}\cdots x_n^{\beta_n}.$$

It is easy to show that the map

$$(\beta_1, \ldots, \beta_n) \mapsto (\beta_1 + \cdots + \beta_n, \beta_2 + \cdots + \beta_n, \ldots, \beta_n)$$

is injective. Thus the g_β's have distinct leading terms. In particular, if we pick β such that $LT(g_\beta) > LT(g_\gamma)$ for all $\gamma \neq \beta$, then $LT(g_\beta)$ will be greater than all terms of the g_γ's. It follows that there is nothing to cancel $LT(g_\beta)$ and thus $g(\sigma_1, \dots, \sigma_n)$ cannot be zero in $\mathbb{K}[x_1, \dots, x_n]$. This contradiction completes the proof of the theorem. □

Exercise 1.5. The proof of the fundamental theorem of symmetric functions gives an algorithm for writing a symmetric polynomial in terms of $\sigma_1, \dots, \sigma_n$. Use this algorithm to express

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

as a polynomial in the elementary symmetric functions.

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

Exercise 1.6. Express $x^3 + y^3$ and $x^4 + y^4$ as a polynomial in σ_1 and σ_2.

Solution $x^3 + y^3 = \sigma_1^3 - 3\sigma_2\sigma_1$ and $\sigma_1^4 - 4\sigma_2\sigma_1^2 + 2\sigma_2^2$

It is possible to give a different algorithm[16] to express a symmetric polynomial in term of elementary symmetric functions. This algorithm is based on a procedure to divide polynomials in the ring of polynomials with multiple variables. This procedure is a fundamental tool in computational algebra and it can be used to give a complete algorithmic description of the ring of invariants for finite groups in characteristic zero. In the next section we shall describe the generalization of the division to polynomials with more variables and the fundamental tool to carry it on properly, namely *Groebner basis*.
There exists another important set of generators of the ring of symmetric functions in n variables. We define the k-th power sum

$$s_k = x_1^k + x_2^k + \cdots + x_n^k$$

Theorem 1.2. If $\mathbb{K}$ is a field containing the rational numbers $\mathbb{Q}$, then every symmetric polynomials in $\mathbb{K}[x_1, \dots, x_n]$ can be written as a polynomial in the power sums $s_1, \dots, s_n$.

Proof Because of theorem 1.1 it is enough to prove that $\sigma_1, \dots, \sigma_n$ are polynomials in $s_1, \dots, s_n$. The Newton identities state

$$s_k - \sigma_1 s_{k-1} + \cdots + (-1)^{k-1}\sigma_{k-1}s_1 + (-1)^k k\sigma_k = 0 \qquad 1 \leq k \leq n \qquad (1.2)$$

[16] How different?

For $k = 1$ $\sigma_1 = s_1$. Assume by induction that σ_k is a polynomial in $s_1, \dots, s_n$. Equation 1.1 implies

$$\sigma_k = (-1)^{k-1}\frac{1}{k}(s_k - \sigma_1 s_{k-1} + \cdots + (-1)^{k-1}\sigma_{k-1}s_1$$

1.3 The division algorithm and Groebner basis

[17]For $\mathbb{K}[x]$, the unique monomial order is given by $1 < x < x^2 < \cdots$, so that, for $f \in \mathbb{K}[x]$, its leading term $LT(f)$ is simply the term of highest degree. Thus the usual division algorithm for polynomials in $\mathbb{K}[x]$ can be stated as follows: if $f, g \in \mathbb{K}[x]$ and $g \neq 0$, then we can write f uniquely in the form

$$f = qg + r,$$

where no term of r is divisible by $LT(f)$. This might seem like a complicated way to of saying $r = 0$ or $deg(r) < deg(g)$, but it generalizes nicely to multiple variables.

The general division algorithm will divide $f \in \mathbb{K}[x_1, \dots, x_n]$ by $f_1, \dots, f_r \in \mathbb{K}[x_1, \dots, x_n]$[18]. We assume that a monomial order is given and we are looking for an expression of the form

$$f = q_1 f_1 + \cdots + q_s f_s + r, \tag{1.3}$$

where (generalizing the one variable case) the "reminder" r should satisfy

$$\text{no term of } r \text{ is divisible by any of } LT(f_1), \dots, LT(f_s) \tag{1.4}$$

We also want to minimize cancellation of leading terms among the $q_i f_i$, so that we will require

$$LT(f) \geq LT(q_i f_i), \quad 1 \leq i \leq s \tag{1.5}$$

An expression (1.3) which satisfies (1.4) and (1.5) is called a *standard expression.*

Given f and $f_1, \dots, f_s$, there is a simple algorithm for producing a standard expression. The basic idea of the algorithm is the same as in the one variable case: we want to cancel the leading term of f (w.r.t a fixed monomial order) by multiplying some f_i by an appropriate monomial and then subtract.

[17] This section is mainly taken from [4].

[18] Why dividing by r polynomials? The point is that we want to find the reminder of a polynomials in an ideal and ideals of $\mathbb{K}[x_1, \dots, x_n]$ are finitely generated ideals but in general not principal.

Example 1.8. [19] We will first divide $f = xy^2 + 1$ by $f_1 = xy + 1$ and $f_2 = y + 1$, using lex order with $x > y$. We want to employ the same scheme as for division of one-variable polynomials. The goal is to find a_1, a_2, r such that $f = a_1 f_1 + a_2 f_2 + r$ and no term of r is divisible by any of $LT(f_1)$ and $LT(f_2)$. We do it by recursion according to the following setup

$$\begin{array}{rcl} a_1 & : & \\ a_2 & : & \\ xy+1 & & \\ y+1 & & \\ & & xy^2+1 \end{array}$$

The leading terms $LT(f_1) = xy$ and $LT(f_2) = y$ both divide the leading term $LT(f) = xy^2$. Since f_1 is listed first, we will use it. Thus we divide xy^2 by xy, leaving y and then subtract $y \cdot f_1$ from f.

$$\begin{array}{rcl} a_1 & : & y \\ a_2 & : & \\ xy+1 & & \\ y+1 & & \\ & & |xy^2+1- \\ & & |xy^2+y \\ & & ------ \\ & & -y+1 \end{array}$$

Now we repeat the same process on $-y + 1$. This time we must use f_2 since $LT(f_1) = xy$ does not divide $LT(-y + 1) = -y$. We obtain

$$\begin{array}{rcl} a_1 & : & y \\ a_2 & : & -1 \\ xy+1 & & \\ y+1 & & \\ & & |-y+1 \\ & & |-y-1 \\ & & ------ \\ & & 2 \end{array}$$

[19] Taken from [3]

Since $LT(f_1)$ and $LT(f_2)$ do not divide 2, the reminder is $r = 2$ and we are done. Thus

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

Example 1.9. [20] Let us divide $f = x^2y + xy^2 + y^2$ by $f_1 = xy - 1$ and $f_2 = y^2 - 1$, using lex order with $x > y$. The first step is

$a_1 \quad : x$

$a_2 \quad :$

$r \quad :$

$xy - 1$

$y^2 - 1$

$|x^2y + xy^2 + y^2 -$

$|x^2y - x$

$- - - - - -$

$xy^2 + x + y^2$

The second step

$a_1 \quad : x + y$

$a_2 \quad :$

$r \quad :$

$xy - 1$

$y^2 - 1$

$|xy^2 + x + y^2 -$

$|xy^2 - y$

$- - - - - -$

$x + y^2 + y$

Now something new happens. The leading term of the polynomial to be divided further is not divisible by the leading terms of f_1 and f_2. But if we move this leading term in the reminder we can proceed further.

$a_1 \quad : x + y$

$a_2 \quad :$

$r \quad : x$

$xy - 1$

$y^2 - 1$

$y^2 + y$

[20] Taken from [3]

and we get

$$\begin{array}{rll} a_1 & : x+y & \\ a_2 & \quad : 1 & \\ r & \quad : x & \\ xy-1 & & \\ y^2-1 & & \\ & & y^2+y- \\ & & y^2-1 \\ & & -\,-\,-\,-\,- \\ & & y+1 \end{array}$$

Move again the leading term to the reminder you are left with one. Move again to the remainder, you are left with zero and you are done. Thus, the reminder is $x+y+1$, and we obtain

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

The algorithm works in general as follows. Start with f, a reminder variable initially set to zero and s polynomials q_i initially set to zero. Then, if $LT(f)$ is divisible by some $LT(f_i)$, we pick the smallest such i and write

$$f = (LT(f)/LT(f_i))f_i + f'.$$

The reminder is unchanged and we add $(LT(f)/LT(f_i))$ to q_i. On the other hand, if no $LT(f_i)$ divides $LT(f)$, then we add $LT(f)$ to the reminder, leave all q_i's unchanged and write

$$f = LT(f) + f'.$$

In each case, note that $LT(f) > LT(f')$. Now repeat the above process using f' and the current value of the reminder. Since $>$ is a well ordering, after finitely many steps, the process must stop. It is easy to prove that the result is a standard expression for f[21]. Although this algorithm is easy to carry out, it does not behave as well as one would like. For instance, the algorithm depends on how the polynomials $f_1, \ldots, f_s$ are ordered, and changing the order can give a different result. To illustrate this point we consider the following example. Using lex order with $x > y$ on $\mathbb{K}[x,y]$, let's

[21] See [3]

divide $f = xy^2 - x$ by $f_1 = xy+1$ and $f_2 = y^2-1$. Using the above algorithm, one easily gets

$$xy^2 - x = y \cdot (xy+1) + 0 \cdot (y^2-1) - x - y, \tag{1.6}$$

so that $-x-y$ is the reminder. But if we divide the same polynomial f using f_2, f_1, the algorithm gives

$$xy^2 - x = x \cdot (y^2-1) + 0 \cdot (xy+1) + 0, \tag{1.7}$$

where the reminder is now zero. Hence reminders are not unique.

Exercise 1.7. *Cocoa* Cocoa has an already defined function `DivAlg` to perform division of a polynomial by a list of polynomials which return its quotient and remeinder. Define a function *MyDivAlg* which implements the divisiona algoritm by scratch.

This example reveals another problem with division. If we consider the ideal generated by f_1, f_2, then (1.7) shows that dividing an element of an ideal by a basis of the ideal may fail to give a zero remainder. We now shall see that if we divide by the polynomials in a *Groebner basis*, then the shortcomings of the division algorithm disappear.
Given a monomial order $>$ on $\mathbb{K}[x_1, \ldots, x_n]$ and an ideal I, the *ideal of leading terms* $< LT(I) >$ is the ideal generated by the leading terms $LT(f)$ for $f \in I - 0$. If $I =< f_1, \ldots, f_m >$ then

$$< LT(f_1), \ldots, LT(f_s) > \subseteq < LT(I) > \tag{1.8}$$

but equality *need not occur.*

Example 1.10. If $f_1 = x^3 - 2xy$ and $f_2 = x^2y - x - 2y^2$, then $x^2 = yf_1 - xf_2 \in < f_1, f_2 >$ hence $x^2 \in LT(< f_1, f_2 >)$. However, using lex order with $x > y$, we have $LT(f_1) = x^3$ and $LT(f_2) = x^2y$, hence $x^2 \notin < LT(f_1), LT(f_2) >$

A Groebner basis occurs when we get equality in (1.8). More precisely

Definition 1.2. Given a monomial order $>$ and and an ideal $I \subseteq \mathbb{K}[x_1, \ldots, x_n]$, we say that $\{g_1, \ldots, g_t\}$ is a *Groebner basis* of I if

$$< LT(g_1), \ldots, LT(g_s) > = < LT(I) >$$

Given an ideal I, the ideal $LT(I)$ is a *monomial ideal.* These ideals have some nice properties which we consider now.

Definition 1.3. An ideal $I \subseteq \mathbb{K}[x_1, \dots, x_n]$ is a **monomial ideal** if there is a subset $A \subseteq \mathbb{Z}^n$ (possibly infinite) such that I consists of all polynomials which are finite sums of the form $\sum_{\alpha \in A} h_\alpha x^\alpha$, where $h_\alpha \in \mathbb{K}[x_1, \dots, x_n]$.

Exercise 1.8. Let $I =< x^\alpha : \alpha \in A >$ be a monomial ideal. Then a monomial x^β lies in I if and only if x^β is divisible by x^α for some $\alpha \in A$

Solution If x^β is a multiple of x^α for some $\alpha \in A$, then $x^\beta \in I$ by definition of ideal. Conversely, if $x^\beta \in I$, then $x^\beta = \sum_{i=1}^s h_i x^{\alpha(i)}$, where $h_i \in \mathbb{K}[x_1, \dots, x_n]$ and $\alpha(i) \in A$. If we expand each h_i as a linear combination of monomials, we see that every term on the right side of the equation is divisible by some $x^{\alpha(i)}$. Hence the left hand side x^β must have the same property

Monomial ideals can always be generated by a finite set of monomials.

Theorem 1.3. Dickson's Lemma. A monomial ideal $I =< x^\alpha, A \subseteq \mathbb{Z}^n >$ can be written down in the form $I =< x^{\alpha(1)}, \dots, x^{\alpha(s)} >$, where $\alpha(1), \dots, \alpha(s) \in A$. In particular, I has a finite basis.

Proof See [3], Chapter 2, section 4, Theorem 5.

Theorem 1.4. For any ideal $I \subseteq \mathbb{K}[x_1, \dots, x_n]$, $LT(I)$ is a monomial ideal.

Proof See [3], Chapter 2, section 5, Proposition 3, part (i).

Theorem 1.5. Fix a monomial order $>$ on $\mathbb{K}[x_1, \dots, x_n]$, and let $I \subseteq \mathbb{K}[x_1, \dots, x_n]$ be an ideal. Then I has a Groebner basis, and furthermore, any Groebner basis of I is a basis of I.

Proof $< LT(I) >$ is a monomial ideal by Theorem 1.4, hence it has a finite monomial basis $h_1 \dots, h_s$ by Theorem 1.3 (Dickson's Lemma). Since $< LT(I) >$ is generated by the leading terms of elements of I, expressing each h_i as a combination of leading terms shows that we can find $g_1, \dots, g_t \in I$ such that $h_1, \dots, h_s \in< LT(g_1), \dots, LT(g_t) >$. Then

$$< LT(I) >=< h_1, \dots, h_s >\subseteq< LT(g_1), \dots, LT(g_t) >\subseteq< LT(I) >,$$

so that $< LT(I) >=< LT(g_1), \dots, LT(g_t) >$. By definition $g_1, \dots, g_t$ is a Groebner basis.
It remains to be proved that $I \subseteq< g_1, \dots, g_t >$. Let $f \in I$. Divide f by $g_1, \dots, g_t$ and get $f = q_1 g_1 + \cdots + q_t g_t + r$. We want to prove that $r = 0$. If not, $r = f - (q_1 g_1 + \cdots + q_t g_t) \in I - \{0\}$, so that

$$LT(r) \in< LT(I) >=< LT(g_1), \dots, LT(g_t) > .$$

This implies that $LT(r)$ is divisible by some $LT(g_i)$ by Exercise 1.8, which is impossible by (1.4) (no term of the reminder is divisible by any of the $LT(g_i)$'s). Hence r must be zero. □

The same reasoning gives the following result

Theorem 1.6. If $g_1, \dots, g_t$ is a Groebner basis for I and $f \in \mathbb{K}[x_1, \dots, x_n]$, then $f \in I$ if and only if the remainder of f on division by $g_1, \dots, g_t$ is zero.

This tells us that once we have a Groebner basis for an ideal, we have an *algorithmic method* for deciding when a given polynomials lies in the ideal. Another important property of Groebner Basis is that the remainders are unique in the following sense.

Proposition 1.1. If $g_1, \dots, g_t$ is a Groebner basis for I and $f \in \mathbb{K}[x_1, \dots, x_n]$, then f can be written uniquely in the form

$$f = g + r$$

where $g \in I$ and no term of r is divisible by any $LT(g_i)$.

Sketch of Proof Suppose $f = g + r$ and $f = g' + r'$. Then $r - r' = g' - g$ and $r - r'$ belongs to the ideal I. Then its leading term belongs to $LT(I) =< LT(g_1), \dots, LT(g_t) >$, since $g_1, \dots, g_t$ is a Groebner basis for I. But no leading term of r and no leading term of r' is divisible by any of $LT(g_i)$. Hence $r - r' = 0$ and $g' = g$. □

Remark 1.1. Proposition 1.1 implies that the reminder on division by a Groebner basis is unique. If we let $G = \{g_1, \dots, g_t\}$ be the Groebner basis, then the reminder of f on division by G will be denoted

$$r = \overline{f}^G.$$

These reminders can be used to get unique coset representatives for elements of the quotient ring $\mathbb{K}[x_1, \dots, x_n]/I$.

These propositions are nice but in order for them to be useful we need to compute Groebner basis. Furthermore, since the definition of Groebner basis involves checking $LT(f)$ for all non zero f in the ideal, it is not clear how to prove that a given basis of an ideal is a Groebner basis. Fortunately, Buchberger provided algorithms for solving both of these problems. The key tool is the *S-polynomial* of $f_1, f_2 \in \mathbb{K}[x_1, \dots, x_n]$, which is defined to be

$$S(f_1, f_2) = \frac{x^\gamma}{LT(f_1)} f_1 - \frac{x^\gamma}{LT(f_2)} f_2$$

where $x^\gamma = \operatorname{lcm}(LM(f_1), LM(f_2))$. The basic idea of the S-polynomial is that it is the simplest combination of f_1 and f_2 which cancels leading terms.

Exercise 1.9. Let $f_1 = x^3 - 2xy$ and $f_2 = x^2y - 2y^2 - x$. Compute $S(f_1, f_2)$ with respect to the lex order with $x > y$.

Solution x^2. Since $LT(x^2)$ is divisible by neither $LT(f_1)$ nor $LT(f_2)$, we see that f_1, f_2 is not a Groebner basis of $< f_1, f_2 >$.

In general we can use S-polynomials to tell if we have a Groebner basis.

Theorem 1.7. Buchberger's criterion. A basis $G = \{g_1, \ldots, g_t\} \subseteq I$ is a Groebner Basis of I if and only if for all $i < j$, we have

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

Here, $\overline{S(g_i, g_j)}^G$ denotes the remainder of $S(g_i, g_j)$ on division by G.

Proof see [3]. □

This criterion gives an algorithm for detecting Groebner bases. Moreover it suggests how to modify a bases to turn it into a Groebner one. Namely, if $F = \{f_1, \ldots, f_j\}$ fails because $S(f_i, f_j)^F \neq 0$ for some $i < j$. then we should add this remainder to the bases and try again.

Example 1.11. Let $F = \{f_1, f_2\} = \{x^3 - 2xy, x^2y - 2y^2 - x\}$. We know that this is not a Groebner basis w.r.t lex order with $x > y$, in fact we know by Exercise 1.9 that $\overline{S(f_1, f_2)}^F = x^2 = f_3$, so that, setting $F_1 = \{f_1, f_2, f_3\}$, we compute:

$$\begin{aligned}
\overline{S(f_1, f_2)}^{F_1} &= 0 \\
\overline{S(f_1, f_3)}^{F_1} &= -2xy = f_4 \\
\overline{S(f_2, f_3)}^{F_1} &= -x - 2y^2 = f_5
\end{aligned}$$

So we do not have a Groebner basis yet. Adding the non zero remainders to F_1 we get $F_2 = \{f_1, f_2, f_3, f_4, f_5\}$, and then we compute

$$\begin{aligned}
\overline{S(f_1, f_5)}^{F_2} &= -4y^3 \\
\overline{S(f_4, f_5)}^{F_2} &= -2y^3 \\
\overline{S(f_i, f_j)}^{F_2} &= 0 \text{ all others } i < j
\end{aligned}$$

It sufficies to add $f_6 = y^3$ to F_2 giving $F_3 = \{f_1, f_2, f_3, f_4, f_5, f_6\}$. This time we get

$$\overline{S(f_i, f_j)}^{F_3} = 0 \ 1 \leq i < j \leq 6$$

so that a Groebner basis of $< x^3 - 2xy, x^2y - 2y^2 - x >$ for lex order with $x > y$ is

$$F_3 = \{< x^3 - 2xy, x^2y - x - 2y^2, x^2, -2xy, -x - 2y^2, y^3\} \qquad (1.9)$$

The general case is similar to the example just completed.

Theorem 1.8. Buchberger's algorithm. Given $\{f_1, \dots, f_s\} \subseteq \mathbb{K}[x_1, \dots, x_n]$, consider the algorithm which starts with $F = \{f_1, \dots, f_s\}$ and then repeats two steps

- (Compute Step) Compute $\overline{S(f_i, f_j)}$ for all $1 \geq i < j \geq |F|$.
- (Augment Step) Augment F by adding the nonzero $\overline{S(f_i, f_j)}^F$ until the compute step gives only zero reminders.

This algorithm always terminates and the final value of F is a Groebner basis of $< f_1, \dots, f_s >$.

This crude form of the algorithm can be enhanced in many ways, see [3] and [1]. The Groebner basis in (1.9) in unnecessarily large. A standard way to simplify a Groebner basis $G = \{g_1, \dots, g_t\}$ is the following: first replace each g_i with its reminder on division by $\{g_1, \dots, g_{i-1}, g_{i+1}, \dots, g_t\}$, then discard any remainders that are zero, and finally, for those that are left, make the coefficient of their leading terms equal to 1. This gives what is called a *reduced Groebner basis*. For example, the reduced Groebner basis associated to (1.9) is

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

In general, for a fixed monomial order, an ideal has a *unique* reduced Groebner basis.

1.4 Groebner basis and symmetric polynomials

We can show now that it is possible to use the division algorithm with respect to a suitable Groebner basis in order to express a symmetric polynomial as a polynomial in the symmetric elementary functions. The same approach can be used to express an invariant polynomial with respect to generators of more general invariant rings[22][23].

[22] Mainly taken from [3]

[23] One may wonder if also for reflection groups there exists a nattural combinatorial basis with good Groebner basis properties

Theorem 1.9. In the ring $\mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_n]$ fix a monomial order where any monomial involving one of $x_1, \ldots, x_n$ is greater than all monomials in $\mathbb{K}[y_1, \ldots, y_n]$. Let G be a Groebner basis of the ideal

$$< \sigma_1 - y_1, \ldots, \sigma_n - y_n > \subseteq \mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_n]$$

.
Given $f \in \mathbb{K}[x_1, \ldots, x_n]$, let $g = \overline{f}^G$ be the remainder of f on division by G. Then:

- f is symmetric if and only if $g \in \mathbb{K}[y_1, \ldots, y_n]$.
- If f is symmetric, then $f = g(\sigma_1, \ldots, \sigma_n)$ is the unique expression of f as a polynomial in the elementary symmetric polynomials $\sigma_1, \ldots, \sigma_n$.

Proof We have

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

To prove the first claim, suppose that $g \in \mathbb{K}[y_1 \ldots, y_n]$. Then for each i, substitute σ_i for y_i in the above formula for f. This will not affect f since it involves only $x_1, \ldots, x_n$. The crucial observation is that under this substitution, every polynomial in $< \sigma_1 - y_1, \ldots, \sigma_n - y_n >$ goes to zero. Since $g_1, \ldots, g_t$ lie in this ideal, it follows that

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

Hence f is symmetric.
Conversely, suppose that $f \in \mathbb{K}[x_1, \ldots, x_n]$ is symmetric, Then $f = g(\sigma_1, \ldots, \sigma_n)$ for some $g \in \mathbb{K}[y_1, \ldots, y_n]$. We want to show that g is the remainder of f on division by G. To prove this, first note that in $\mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_n]$, a monomial in $\sigma_1, \ldots, \sigma_n$ can be written as follows:

$$\begin{aligned} \sigma_1^{\alpha_1} \cdots \sigma_n^{\alpha_n} &= (y_1 + (\sigma_1 - y_1))^{\alpha_1} \cdots (y_n + (\sigma_n - y_n))^{\alpha_n} \\ &= y_1^{\alpha_1} \cdots y_n^{\alpha_n} + B_1 \cdot (\sigma_1 - y_1) + \cdots + B_n \cdot (\sigma_n - y_n) \end{aligned}$$

for some $B_1, \ldots, B_n \in \mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_n]$. Multiplying by an appropriate constant and adding over the exponents appearing in g, it follows that

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

where $C_1, \ldots, C_n \in \mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_n]$. Since $f = g(\sigma_1, \ldots, \sigma_n)$ we can write this as

$$f = C_1(\sigma_1 - y_1) + \cdots + C_n(\sigma_n - y_n) + g(y_1, \ldots, y_n) \qquad (1.10)$$

We want to show that g is the remainder of f on division by G.
The first step is to show that no term of g is divisible by an element of $LT(G)$. If this were not so, then there would be $g_i \in G, g_i \neq 0$, where $LT(g_i)$ divides some term of g. Hence $LT(g_i)$ would involve only $y_1, \ldots, y_n$ since $g \in \mathbb{K}[y_1, \ldots, y_n]$. By our hypothesis on the ordering, it would follow that $g_i \in \mathbb{K}[y_1, \ldots, y_n]$. Now replace every y_i with the corresponding σ_i. Since $g_i \in < \sigma_1 - y_1, \ldots, \sigma_n - y_n >$, we conclude that $g_i \mapsto 0$ under the substitution $y_i \mapsto \sigma_i$. Then $g_i \in \mathbb{K}[y_1, \ldots, y_n]$ would mean $g_i(\sigma_1, \ldots, \sigma_n) = 0$. By the uniqueness part of Theorem (1.1), this would imply $g_i = 0$, which is impossible since $g_i \neq 0$. This proves our Claim.
It follows that in (1.10), no term of g is divisible by an element of $< LT(G) >$, and since G is a Groebner basis, g is the remainder of f on division by G [24], hence we have proved the first part of the Theorem.
The second part follows immediately. □

A minor adaption of the same proof can give us something more.

Theorem 1.10. Suppose that $f_1, \ldots, f_m \in \mathbb{K}[x_1, \ldots, x_n]$ are given. Fix a monomial order in $\mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_m]$ where any monomial involving one of $x_1, \ldots, x_n$ is greater than all monomials in $\mathbb{K}[y_1, \ldots, y_n]$. Let G be a Groebner basis of the ideal

$$< y_1 - f_1, \ldots, y_m - f_m > \subseteq \mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_m].$$

Given $f \in \mathbb{K}[x_1, \ldots, x_n]$, let $g = \overline{f}^G$ be the remainder of f on division by G. Then:

- $f \in \mathbb{K}[f_1, \ldots, f_m]$ if and only if $g \in \mathbb{K}[y_1, \ldots, y_m]$.
- If $f \in \mathbb{K}[f_1, \ldots, f_m]$, then $f = g(f_1, \ldots, f_m)$ is an expression of f as a polynomial in $f_1, \ldots, f_m$.

Proof The proof follows closely the proof of Theorem 1.9. If $f = g(f_1, \ldots f_m)$, arguing as above we get

$$f = C_1(y_1 - f_1) + \cdots + C_m(y_m - f_m) + g(y_1, \ldots, y_m)$$

Now g needs not to be the remainder of f on division by G, we still need to reduce some more. Let $G' = G \cap \mathbb{K}[y_1, \ldots, y_m]$. Renumbering if necessary we can assume $G' = \{g_1, \ldots, g_s\}$. If we divide g by G' we get

$$g = B_1 g_1 + \cdots + B_s g_s + g' \qquad (1.11)$$

[24] See Proposition 1.1.

where $B_1, \ldots B_s, g' \in \mathbb{K}[y_1, \ldots, y_m]$ and hence

$$f = C_1'(y_1 - f_1) + \cdots + C_m'(y_m - f_m) + g'(y_1, \ldots, y_m)$$

This follows because in (1.11), each g_i lies in $< y_1 - f_1, \ldots, y_m - fm >$. We claim that g' is the remainder of f on division by G. This will prove that the remainder lies in $g \in \mathbb{K}[y_1, \ldots, y_m]$. Since G is a Groebner basis g' is the remainder of f on division by G provided that no term of g' is divisible by an element of $LT(G)$. To prove that g' has this property, suppose that there is $g_i \in G$ where $LT(g_i)$ divides some term of g'. Then $LT(g_i)$ involves only $y_1, \ldots, y_m$ since $g' \in \mathbb{K}[y_1, \ldots, y_m]$. By our hypothesis on the ordering, it follows that $g_i \in \mathbb{K}[y_1, \ldots, y_m]$ and hence $g_i \in G'$. Since g' is a remainder on division by G', $LT(g_i)$ cannot divide any term of g'. This contradiction shows that g' is the desired remainder. □

To use Theorem 1.9 we need to compute a Groebner basis for the ideal $< \sigma_1 - y_1, \ldots, \sigma_n - y_n >$. This is not hard when we use lex order. Given variables $u_1, \ldots, u_s$, let

$$h_i(u_1, \ldots, u_s) = \sum_{|\alpha|=i} u^\alpha$$

be the sum of *all* monomials of total degree i in $u_1, \ldots, u_s$. Then we get the following Groebner basis.

Theorem 1.11. Fix lex order on $\mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_n]$ with

$$x_1 > \cdots > x_n > y_1 > \cdots y_n$$

Then the polynomials

$$g_k = h_k(x_k, \ldots, x_n) + \sum_{i=1}^{k} (-1)^i h_{k-i}(x_k, \ldots, x_n) y_i, \qquad k = 1, \ldots, n.$$

form a Groebner basis for the ideal $< \sigma_1 - y_1, \ldots, \sigma_n - y_n >$.

Proof see [3], p.316

1.5 Generators for the Ring of Invariants

Definition 1.4. Given $f_1, \ldots, f_n \in \mathbb{K}[x_1, \ldots, x_n]$, we let $\mathbb{K}[f_1, \ldots, f_m]$ denote the subset of $\mathbb{K}[x_1, \ldots, x_n]$ consisting of all polynomial expressions in $f_1, \ldots, f_m$ with coefficients in $\mathbb{K}$.

This means that the elements $f \in \mathbb{K}[f_1, \dots, f_m]$ are those polynomials which can be written in the form

$$f = g(f_1, \dots, f_m)$$

where g is a polynomial in m variables with coefficients in k.
Elements of the ring of invariants $\mathbb{K}[x_1, \dots, x_n]^G$ are easily made by means of the Reynolds operator.

Definition 1.5. Given a finite matrix group $G \subseteq Gl(n, \daleth)$, the *Reynolds operator* of G is the map $R_G : \mathbb{K}[x_1, \dots, x_n] \to \mathbb{K}[x_1, \dots, x_n]$ defined by the formula

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

for $f(x) \in \mathbb{K}[x_1, \dots, x_n]$.

One can think of $R_G(f)$ as averaging the effect of G on f.
The following properties are easy to prove and are left as an exercise.

Proposition 1.2. 1. R_G is $\daleth$-linear in f.

2. If $f \in \mathbb{K}[x_1, \dots, x_n]$, then $R_G(f) \in \mathbb{K}[x_1, \dots, x_n]^G$

3. If $f \in \mathbb{K}[x_1, \dots, x_n]^G$, then $R_G(f) = f$

Example 1.12. Consider the cyclic matrix group $C_4 \subseteq Gl(2, k)$ of order 4 generated by

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

Obviously

$$\mathbb{K}[x, y]^{C_4} = \{f \in \mathbb{K}[x, y] : f(x, y) = f(-x, y)\}.$$

One can easily check 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 (1.2), we can compute some invariants as follows:

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

It will be shown in Theorem 1.12 that these three invariants generate in $\mathbb{K}[x, y]^{C_4}$.

It is easy to prove that, for any monomial x^α, the Reynolds operator gives a homogeneous invariant of total degree $|\alpha|$ whenever it is non zero. The following theorem of Emmy Noether shows that we can always find finitely many of these invariants that generate $\mathbb{K}[x_1, \ldots, x_n]$.

Theorem 1.12. Given a finite matrix group $G \subseteq Gl(n, \daleth)$, we have

$$\mathbb{K}[x_1, \ldots, x_n]^G = \mathbb{K}[R_G(x^\beta) : |\beta| \leq |G|]$$

In particular, $\mathbb{K}[x_1, \ldots, x_n]^G$ is generated by finitely many homogeneous invariants.

Proof Let $f = \sum_\alpha c_\alpha x^\alpha \in \mathbb{K}[x_1, \ldots, x_n]^G$. Then

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

Hence, every invariant is a linear combination of the $R_G(x^\alpha)$. It remains to be proved that $R_G(x^\alpha)$ is a polynomial in $R_G(x_\beta)$, $|\beta| \leq G$. Let

$$(x_1 + \cdots + x_n)^k = \sum_{|\alpha|=k} a_\alpha x^\alpha. \tag{1.12}$$

where c_α are the multinomials coefficients

$$c_\alpha = \binom{k}{\alpha} = \frac{k!}{\alpha_1!\alpha_2! \cdots \alpha_n!}$$

Let $\mathbf{x}$ be the column vector of variables $x_1, \ldots, x_n$ and let $\mathbf{u}$ be the column vector of variables $u_1, \ldots, u_n$. Then $(\mathbf{u} \cdot \mathbf{x})^k = \sum_{|\alpha|=k} c_\alpha \mathbf{u}^\alpha \mathbf{x}^\alpha$. If $A \in G$

$$(\mathbf{u}A \cdot \mathbf{x})^k = \sum_{|\alpha|=k} c_\alpha \mathbf{u}^\alpha (A \cdot \mathbf{x})^\alpha \tag{1.13}$$

If we sum both members of (1.13) over all A's in G, we obtain

$$\sum_{A \in G} (\mathbf{u}A \cdot \mathbf{x})^k = \sum_{|\alpha|=k} c_\alpha \mathbf{u}^\alpha \left(\sum_{A \in G} (A \cdot \mathbf{x})^\alpha \right) \tag{1.14}$$

The right hand side of (1.14) is

$$\sum_{|\alpha|=k} |G| c_\alpha \mathbf{u}^\alpha R_G(x^\alpha)$$

Note how the sum on the right encodes all $R_G(x^\alpha)$ with $|\alpha| = k$. This is why we use the variables $u_1, \ldots, u_n$: they prevent any cancellation from occurring. The left hand side of (1.14) is the k-th power sum S_k of the $|G|$ quantities $U_A = (\mathbf{u}A \cdot \mathbf{x})$. We write this as $S_k = S_k(U_A : A \in G)$. By Theorem 1.2 of Section 1.2, every symmetric function in the $|G|$ quantities U_A is a polynomial in $S_1, \ldots, S_{|G|}$. Since S_k is symmetric in the U_A, it follows that

$$S_k = F(S_1, \ldots, S_{|G|}) \tag{1.15}$$

for some polynomial F with coefficients in k. Substituting $S_h = \sum_{A \in G}(\mathbf{u}A \cdot \mathbf{x})^h$ in (1.15), we obtain

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

Expanding the right side and equating the coefficients of u^α, it follows that

$$b_\alpha R_G(x^\alpha) = \text{a polynomial in the } R_G(x^\beta), \qquad |\beta| \leq |G|.$$

Since ⅂ has characteristic zero, the coefficient $b_\alpha = |G| a_\alpha$ is not zero in ⅂, and, hence, $R_G(x^\alpha)$ has the desired form. □

1.6 Elimination theory

We begin with an example[25].

Example 1.13. To eliminate y from the system of equations

$$\begin{cases} x - y &= 0 \\ xy - 3x + 2 &= 0 \end{cases}$$

consider the polynomial consequence

$$x \cdot (x - y) + 1 \cdot (xy - 3x + 2) = x^2 - 3x + 2 = 0$$

This eliminates y and allow us to find the x values of the solutions. In terms of ideals, we can write this as

$$x^2 - 3x + 2 \in < x - y, xy - 3x + 2 > \cap \mathbb{K}[x]$$

[25] The first example should be a system of linear equations solved with Gaussian elimination

Note that $< x-y, xy-3x+2 > \cap \mathbb{K}[x]$ gives *all* possible way of eliminating y from our system of equations.
In general, given equations

$$\begin{cases} f_1 &= 0 \\ & \vdots \\ f_s &= 0 \end{cases}$$

with $f_1, \dots, f_s \in \mathbb{K}[x_1, \dots, x_n]$, we can form the ideal $I =< f_1, \dots, f_s >$ and then successively eliminate variables by considering the intersections

$$\begin{aligned} I \cap \mathbb{K}[x_2, \dots, x_n] & \qquad \text{which eliminates } x_1; \\ I \cap \mathbb{K}[x_3, \dots, x_n] & \qquad \text{which eliminates } x_1, x_2; \\ \vdots & \\ I \cap \mathbb{K}[x_n] & \qquad \text{which eliminates } x_1, x_2, \dots, x_n; \end{aligned} \tag{1.16}$$

These are called *elimination ideals* and one of the goals of elimination theory is to find elements (preferably generators) of each $I \cap \mathbb{K}[x_k, \dots, x_n]$

Amazingly, we can find basis of *all* of these ideals *simultaneously* by using a lexicographic Groebner basis. Here is the precise result.

Theorem 1.13. (Elimination theory) If $I =< f_1, \dots, f_s > \subseteq \mathbb{K}[x_1, \dots, x_n]$ is an ideal and $G = \{g_1, \dots, g_t\}$ is a Groebner basis for I for lex order with $x_1 > x_2 > \dots > x_n$, then for each k between 2 and n, the set

$$G \cap \mathbb{K}[x_k, \dots, x_n]$$

is a Groebner basis for the elimination ideal

$$I \cap \mathbb{K}[x_k, \dots, x_n]$$

Proof Given any $f \neq 0$ in $I \cap \mathbb{K}[x_k, \dots, x_n]$, we have $f \in I$, so that $LT(f)$ is divisible by $LT(g_i)$ for some $g_i \in G$ by definition of Groebner basis. Since $f \in \mathbb{K}[x_k, \dots, x_n]$, its leading term $LT(f)$ does not involve $x_1, \dots, x_{k-1}$. Hence the same is true for $LT(g_i)$ Because of our hypothesis on the order, $LT(g_i) \in \mathbb{K}[x_k, \dots, x_n]$ implies $g_i \in \mathbb{K}[x_k, \dots, x_n]$. Hence we have proved that for any $f \neq 0 \in I \cap \mathbb{K}[x_k, \dots, x_n]$, $LT(f)$ is divisible by $LT(g_i)$ for some $g_i \in G \cap \mathbb{K}[x_k, \dots, x_n]$ and we are done for the definition of Groebner basis.

Example 1.14. Suppose we want to find the minimum and maximum values of the function $f(x, y, z) = x^3 + 2xyz - z^2$ subject to the constraint equation

$g(x, y, z) = x^2 + y^2 + z^2 - 1 = 0$. This is a typical constrained min-max problem. By the method of Lagrange multipliers, we get the equations

$$\begin{cases} \nabla f &= \lambda \nabla g \\ g &= 0 \end{cases}$$

which can be written in the form

$$\begin{cases} 3x^2 + 2yz - 2x\lambda &= 0 \\ 2xz - 2y\lambda &= 0 \\ 2xy - 2z - 2z\lambda &= 0 \\ x^2 + y^2 + z^2 - 1 &= 0 \end{cases}$$

The four polynomials which appears in the left hand side of the above equations can be turned into a Groebner basis with respect to the lex order with $\lambda > x > y > z$. We obtain

$$\lambda - \frac{3}{2}x - \frac{3}{2}yz - \frac{167616}{3835}z^6 - \frac{36717}{590}z^4 - \frac{134419}{7670}z^2 = 0 \qquad (1.17)$$

$$x^2 + y^2 + z^2 - 1 = 0 \qquad (1.18)$$

$$xy - \frac{19584}{3835}z^5 + \frac{1999}{295}z^3 - \frac{6403}{3835}z = 0 \qquad (1.19)$$

$$xz + yz^2 - \frac{1152}{3835}z^5 - \frac{108}{295}z^3 + \frac{2556}{3835}z = 0 \qquad (1.20)$$

$$y^3 + yz^2 - y - \frac{9216}{3835}z^5 + \frac{906}{295}z^3 - \frac{2562}{3835}z = 0 \qquad (1.21)$$

$$y^2z - \frac{6912}{3835}z^5 + \frac{827}{295}z^3 - \frac{3839}{3835}z = 0 \qquad (1.22)$$

$$yz^3 - yz - \frac{576}{59}z^6 + \frac{1605}{118}z^4 - \frac{453}{118}z^2 = 0 \qquad (1.23)$$

$$z^7 - \frac{1763}{1152}z^5 + \frac{655}{1152}z^3 - \frac{11}{288}z = 0 \qquad (1.24)$$

The last equation involves only z, and it factors as

$$z(z^2 - 1)(z^2 - \frac{4}{9})(z^2 - \frac{11}{128}) = 0$$

which implies

$$z = 0 \quad \pm 1 \quad \pm \frac{2}{3} \quad \pm \frac{\sqrt{11}}{8\sqrt{2}}$$

This solves (1.24). (1.21), (1.22) and (1.23) involve only y and z. Thus by setting z equal to each of the values we obtained by solving (1.24), we can solve for the corresponding y. Continuing this way, we can find the values for x (and λ, which are not needed). They are

$$\begin{array}{lll}
z = 0 & y = 0 & x = \pm 1 \\
z = 0 & y = \pm 1 & x = 0 \\
z = \pm 1 & y = 0 & z = 0 \\
z = 2/3 & y = 1/3 & x = -2/3 \\
z = -2/3 & y = -1/3 & x = -2/3 \\
z = \sqrt{11}/8\sqrt{2} & y = -3\sqrt{11}/\sqrt{2} & x = -3/8 \\
z = -\sqrt{11}/8\sqrt{2} & y = 3\sqrt{11}/\sqrt{2} & x = -3/8
\end{array}$$

The code for computing the Groebner basis in example 1.14 with CoCoA is

```
Use R::=Q[l,x,y,z];
F:=x^3+2x*y*z-z^2;
G:=x^2+y^2+z^2-1;
MJ:=Jacobian([F])-l*Jacobian([G]);
ListPol:=MJ[1];
Append(ListPol,G);
Id:=Ideal(ListPol);
Set Indentation;
GB:=GBasis(Id);
GB;
```

1.7 Relations among the Generators for the Ring of Invariants

In Section 1.5 we have seen (in principle) how to find generators $F = \{f_1, \dots, f_m\}$ for the ring of polynomial invariants of a finite group, i.e. elements of $\mathbb{K}[x_1, \dots, x_n]$ such that

$$\mathbb{K}[x_1, \dots, x_n]^G = \mathbb{K}[f_1, \dots, f_m].$$

In this section we want to discuss how to produce all relations between these generators, i.e. all polynomials $h \in \mathbb{K}[y_1, \dots, y_m]$ such that $h(f_1, \dots, f_m) = 0$. If we call this set I_F, it is easy to prove that I_F is a prime ideal of $\mathbb{K}[y_1, \dots, y_m]$[26]
We call I_F the *ideal of relations* for $F = \{f_1, \dots, f_m\}$ or the *first syzygy ideal.*

[26] see [4], chap 7, par 4, proposition 1

Exercise 1.10. If $\mathbb{K}[x_1,\dots,x_n]^G = \mathbb{K}[f_1,\dots,f_m]$, let $I_F \subseteq \mathbb{K}[y_1,\dots,y_m]$ the ideal of relations. Then there is a ring isomorphism

$$\mathbb{K}[y_1,\dots,y_m]/I_F \cong \mathbb{K}[x_1,\dots,x_n]^G$$

We can compute I_F explicitly using elimination theory.

Theorem 1.14. If $\mathbb{K}[x_1,\dots,x_n]^G = \mathbb{K}[f_1,\dots,f_m]$, consider the ideal

$$J_F =< f_1 - y_1,\dots,f_m - y_m >\subseteq \mathbb{K}[x_1,\dots,x_n,y_1,\dots,y_m].$$

1. I_F is the n-th elimination ideal of J_F. Thus $I_F = J_F \cap \mathbb{K}[y_1,\dots,y_n]$

2. Fix a monomial order in $\mathbb{K}[x_1,\dots,x_n,y_1,\dots,y_m]$ where any monomial involving one of $x_1,\dots,x_n$ is greater than all monomials in $\mathbb{K}[y_1,\dots,y_m]$ and let G be a Groebner basis of J_F. Then $G \cap \mathbb{K}[y_1,\dots,y_m]$ is a Groebner basis for I_F in the monomial order induced on $\mathbb{K}[y_1,\dots,y_m]$.

Proof To relate J_F to the ideal of relations I_F, we will need the following characterization of J_F: if $p \in \mathbb{K}[x_1,\dots,x_n,y_1,\dots,y_m]$, then we claim that

$$p \in J_F \iff p(x_1,\dots,x_n,f_1,\dots,f_m) = 0 \text{ in } \mathbb{K}[x_1,\dots,x_n]. \tag{1.25}$$

One implication ($\Longrightarrow$) is obvious since the substitution $y_i \mapsto f_i$ takes all elements of $J_F =< f_1 - y_1,\dots,f_m - y_m >$ to zero. On the other hand, given $p \in \mathbb{K}[x_1,\dots,x_n,y_1,\dots,y_m]$, if we replace each y_i in p by $f_i - (f_i - y_i)$ and expand, we obtain

$$p(x_1,\dots,x_n,y_1,\dots,y_m) = p(x_1,\dots,x_n,f_1,\dots,f_m) + \\ +B_1(f_1 - y_1) + \cdots + B_m(f_m - y_m)$$

for some $B_1,\dots,B_m \in \mathbb{K}[x_1,\dots,x_n,y_1,\dots,y_m]$. In particular, if

$$p(x_1,\dots,x_n,f_1,\dots,f_m) = 0,$$

then

$$p(x_1,\dots,x_n,y_1,\dots,y_m) = B_1(f_1 - y_1) + \cdots + B_m(f_m - y_m) \in J_F.$$

This completes the proof of (1.25).
Now intersect each side of (1.25) with $\mathbb{K}[y_1,\dots,y_m]$. For $p \in \mathbb{K}[y_1,\dots,y_m]$ this proves

$$p \in J_F \cap \mathbb{K}[y_1,\dots,y_m] \iff p(f_1,\dots,f_m) = 0 \text{ in } \mathbb{K}[x_1,\dots,x_n],$$

so that $J_F \cap \mathbb{K}[y_1,\dots,y_m] = I_F$ by definition of I_F. Thus, point 1) is proved and point 2) is then an immediate consequence of the elimination theory of Section 1.6

Example 1.15. In Example 1.2 we saw that the ring of invariants of $C_2 = \{\pm I_2\} \subseteq GL(2,k)$ is given by $\mathbb{K}[x,y]^{C_2} = \mathbb{K}[x^2, y^2, xy]$. Let $F = (x^2, y^2, xy)$ and let the new variables be u, v, w. Then the ideal of relations is obtained by eliminating x, y from the equations

$$\begin{cases} u & = & x^2 \\ v & = & y^2 \\ w & = & xy \end{cases}$$

If we use lex order with $x > y > u > v > w$, than a Groebner basis for the ideal $J_F =< u - x^2, v - y^2, w - xy >$ consists of the polynomials

$$x^2 - u,\ xy - w\ xv - yw\ xw - yu\ y^2 - v,\ uv - w^2.$$

It follows from Proposition 1.14 that

$$I_F =< uv - w^2 > .$$

This says that all relations between x^2, y^2 and xy are generated by the obvious relations $x^2 \cdot y^2 = (xy)^2$. The the ring of invariants is therefore

$$\mathbb{K}[x,y]^{C_2} \cong \mathbb{K}[u,v,w]/ < uv - w^2 >$$

1.8 Hilbert series and Molien theorem

In this Section, we prove some results about Hilbert series of rings and some applications to rings of invariants[27].

Definition 1.6. For a graded vector space $V = \oplus_{d=k}^{\infty} V_d$ with V_d finite dimensional for all d we define the *Hilbert series* of V as the formal Laurent series

$$H(V,t) = \sum_{d=k}^{\infty} dim(V_d) t^d$$

In the literature, Hilbert series are sometimes called Poincaré series. In our applications, V will always be a graded algebra or a graded module.

Example 1.16. Let us compute the Hilbert series of $\mathbb{K}[x_1, \ldots, x_n]$. There are $\binom{n+d-1}{n-1}$ monomials of degree d, therefore the Hilbert series is

$$H(\mathbb{K}[x_1, \ldots, x_n], t) = \sum_{d=0}^{\infty} \binom{n+d-1}{n-1} t^d$$

This is exactly the power series expansion of $(1-t)^{-n}$.

[27] I should define the concept of graded algebra and graded moduli first, discuss only the behaviour of Hilbert series w.r.t. symmetric products, look at the examples in Sloane.

Remark 1.2. If V and W are two graded vector spaces, then the tensor product $V \otimes W$ has also a natural grading, namely

$$(V \otimes W)_d = \oplus_{d_1+d_2=d} V_{d_1} \otimes W_{d_2}$$

It is obvious from this formula that $H(V \otimes W, t) = H(V,t)H(W,t)$. Suppose that $R = \mathbb{K}[x_1, \ldots, x_n]$ and x_i has degree $d_i > 0$. Then we have $R = \mathbb{K}[x_1] \otimes \mathbb{K}[x_2] \otimes \mathbb{K}[x_n]$ as graded algebras and $H(\mathbb{K}[x_i], t) = (1 - t^{d_i})^{-1}$. It follows that

$$H(R,t) = \frac{1}{(1-t^{d_1}) \cdots (1-t^{d_n})} \tag{1.26}$$

Remark 1.3. If

$$0 \to V^{(1)} \to V^{(2)} \to \cdots \to V^{(r)} \to 0 \tag{1.27}$$

is an exact sequence of graded vector spaces (all maps respect degree) with $V_d^{(i)}$ finite dimensional for all i and d, then

$$\sum_{i=1}^{r} (-1)^i H(V^{(i)}, t) = 0$$

This is clear because the degree d part of (1.27) is exact for all d.

Theorem 1.15. (Hilbert) If $R = \oplus_{d=0}^{\infty} R_d$ is a finitely generated graded algebra over a field $\daleth = R_0$, then $H(R,t)$ is the power series of a rational function. The radius of convergence of this power series is at least one. Moreover, if $M = \oplus_{d=k}^{\infty} M_d$ is a finitely generated graded R-module, the $H(M,t)$ is the Laurent series of a rational function (which may have a pole at 0).

Proof Let $A = \mathbb{K}[x_1, \ldots, x_n]$ be the polynomial ring, graded in such a way that $\deg(x_i) = d_i > 0$. Then $H(A,t)$ is a rational function by equation (1.26) and the radius of convergence of the power series is 1 if $n > 0$ and ∞ if $n = 0$. For any integer e, we define the A-module $A(e)$ by $A(e) = \oplus_{d=-e}^{\infty} A(e)_d$ with $A(e)_d := A_{e+d}$. It is clear that $H(A(e), t) = t^{-e} H(A,t)$ is again a rational function. A module is free if it is isomorphic to a direct sum $A = \oplus_i A(i)$, hence the Hilbert series of a finitely generated free module M is a rational function. If M is a finitely generated A-module, then by Hilbert's syzygy theorem (see [7], thm 31.13), there exists a resolution

$$0 \to F^{(r)} \to F^{(r-1)} \to \cdots \to F^1 \to F^{(0)} \to M \to 0 \tag{1.28}$$

where $F^{(i)}$ is finitely generated free A-module for all i, and the sequence is exact. It follows from Remark 1.3 that

$$H(M,t) = \sum_{i=0}^{r} (-1)^i H(F^{(i)},t) \tag{1.29}$$

so $H(M,t)$ is a rational function. If M is non-negatively graded, then the same is true for all $F^{(i)}$, so the radius of convergence of $H(M,t)$ is at least 1. Let R be an arbitrary finitely generated graded algebra over $k = R_0$. Then, for some n and some $d_1, \ldots, d_n > 0$, there exists a homogeneous ideal $I \subseteq A$ such that $A/I \cong R$. Hence R is a finitely generated, not negatively graded A-module, and the claim follows. Moreover, any finitely generated graded R-module M is also a finitely generated graded A-module. □

We consider now the Hilbert series of the ring of invariants, which is also called the *Molien series*. If $T \in Gl(n,\mathbb{K})$, then T acts on $\mathbb{K}[x_1,\ldots,x_n]$ and T restricts to a linear transformation on each $\mathbb{K}[x_1,\ldots,x_n]_d$ which is finite dimensional. We shall write tr_dT for the trace of $T|\mathbb{K}[x_1,\ldots,x_n]_d$.[28]

Proposition 1.3. If $T \in Gl(n)$ then $\sum_{d=0}^{\infty}(tr_dT)t^d = det(1-tT)^{-1}$.

Proof It will be convenient for the proof to extend the base field to its algebraic closure, to have access to eigenvalues. Then we may choose a basis $\{x_1,\ldots,x_n\}$ for $\urcorner^n$ so that T is represented by a triangular matrix, with eigenvalues $\lambda_1,\ldots,\lambda_n$. Also $\prod_{i=1}^n x_i^{a_i} : (a_1,\ldots,a_n) \in \mathbb{Z}_+^n,\ \sum_i a_i = d\}$ is a basis for $\mathbb{K}[x_1,\ldots,x_n]_d$, and if that basis is ordered lexicographically then $T|\mathbb{K}[x_1,\ldots,x_n]_d$ is also represented by a triangular matrix with eigenvalues $\{\prod_{i=1}^n \lambda_i^{a_i} : (a_1,\ldots,a_n) \in \mathbb{Z}_+^n,\ \sum_i a_i = d\}$. Thus

$$(tr_dT)t^d = \sum\{\prod_{i=1}^{n}(\lambda_i t)^{a_i} : (a_1,\ldots,a_n) \in \mathbb{Z}_+^n,\ \sum_i a_i = d\}$$

and

$$\sum_{d=0}^{\infty}(tr_dT)t^d = \sum\{\prod_{i=1}^{n}(\lambda_i t)^{a_i} : (a_1,\ldots,a_n) \in \mathbb{Z}_+^n,\}$$

Now note that $(1-\lambda_i t)^{-1} = \sum_{a=0}^{\infty}(\lambda_i t)^a$, so

$$\begin{array}{rcl} \prod_{i=1}^n (1-\lambda_i t)^{-1} & = & \prod_{i=1}^n \sum_{a_i=0}^{\infty}(\lambda_i t)^{a_i} \\ & = & \sum\{\prod_{i=1}^n(\lambda_i t)^{a_i} : (a_1,\ldots,a_n) \in \mathbb{Z}_+^n,\} \\ & = & \sum_{d=0}^{\infty}(tr_dT)t^d \end{array}$$

[28] This section is taken from [2]

To complete the proof observe that

$$\prod_{i=1}^{n}(1-\lambda_i t)^{-1} = \frac{1}{det(1-tT)}$$

□

Corollary 1.1. $\sum_{d=0}^{\infty}(\dim(\mathbb{K}[x_1,\ldots,x_n]))t^d = (1-t)^{-n}$

Proof Take $T = 1$.

Proposition 1.4. Suppose W is a subspace of V and P is a projection of V onto W, i.e. $PV = W$ and $P^2 = P$. The $\dim(W) = tr(P)$

Proof Let $W' = (1-P)V$. Hence, $V = W \oplus W'$ and the restriction of P to W' is zero, so by choosing a suitable basis we can represent P by the matrix

$$\begin{pmatrix} Id_W & 0 \\ 0 & 0 \end{pmatrix} \tag{1.30}$$

and $\dim(W) = tr(P)$.
Let $G \subseteq Gl(n)$ be a finite group and let $\mathcal{I} = \mathbb{K}[x_1,\ldots,x_n]^G$ be the algebra of invariants of $\mathcal{G}$. The *Molien series* of G is the Hilbert series of $\mathcal{I}$ i.e. is the power series

$$\Phi(t) := \sum_{d=0}^{\infty}(\dim\mathcal{I}_d)t^d$$

Proposition 1.5. If $G \subseteq GL(V)$ is finite then $\dim(\mathcal{I}_d) = \frac{1}{|G|}\sum\{tr_d T : T \in G\}$

Proof Let T_d be the restriction of $T \in G$ to $\mathbb{K}[x_1,\ldots,x_n+_d$. The restriction to $\mathbb{K}[x_1,\ldots,x_n+_d$ of the Reynolds operator, $\mathcal{R}_d := \frac{1}{|G|}\sum_{T\in G} Tr(T_d)$is a projection. Hence by proposition 1.4

$$\dim\mathcal{I}_d = Tr(\mathcal{R}_d) = \frac{1}{|G|}\sum\{tr_d T : T \in G\}$$

Theorem 1.16. (Molien's theorem) If $G \subseteq GL(n,\mathbb{K})$ is finite then its Molien series is

$$\Phi(t) = \frac{\sum_{T\in G} det(1-tT)^{-1}}{|G|}$$

Proof

$$\begin{aligned}\Phi(t) &= \sum_{d=0}^{\infty} \dim(\mathcal{I}_d)t^d = \sum_{d=0}^{\infty} \frac{1}{|G|} \sum\{tr_d T : T \in G\} = \\ &= \frac{1}{|G|}\sum_{T\in G}(\sum_{d=0}^{\infty}(tr_d T)t^d) \\ &= \frac{1}{|G|}\sum_{T\in G}\frac{1}{det(1-tT)}\end{aligned}$$

Example 1.17. Let G be the dihedral group over 3 elements. The matrices

$$T_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad T_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad T_3 = \begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix}$$

are representatives of the conjugacy classes, of respective sizes 1, 3 and 2. We have

$$\det(Id - tT_1) = (1-t)^2$$
$$\det(Id - tT_2) = 1 - t^2$$
$$\det(Id - tT_3) = 1 + t + t^2$$

and thus

$$\phi(t) = \frac{1}{6}\left(\frac{1}{(1-t)^2} + \frac{3}{1-t^2} + \frac{2}{1+t+t^2}\right)$$

We recall that

$$\frac{1}{1-t} = 1 + t + t^2 + t^3 + \dots \tag{1.31}$$

By setting $t = t^2$ in (1.31) we get

$$\frac{1}{1-t^2} = 1 + t^2 + t^4 + t^6 + \dots$$

whose coefficients follows the pattern

$$1\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1\ \dots \tag{1.32}$$

By differentiating (1.31) we get

$$\frac{1}{(1-t)^2} = 1 + 2t + 3t^2 + 4t^3 + \dots$$

whose coefficients follows the pattern

$$1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ \dots \tag{1.33}$$

Notice moreover that

$$\begin{aligned}\frac{1}{1+t+t^2} &= \frac{1-t}{1-t^3} = (1-t)(1+t^3+t^6+t^9+\dots) \\ &= (1-t)+(t^3-t^4)+(t^6-t^7)+(t^9-t^{10})\end{aligned}$$

whose coefficients follows the pattern

$$1\ -1\ 0\ 1\ -1\ 0\ 1\ -1\ 0\ \dots \tag{1.34}$$

Hence, the pattern of coefficients of $\phi(t)$ is

$$1\ 0\ 1\ 1\ 1\ 1\ 2\ 1\ 2\ 2\ 2\ 2\ 3\ 2\ 3\ 3\ 3\ 3\ \dots \tag{1.35}$$

Hence we get; at degree zero, the constants; at degree one, no invariants; at degree 2, the invariant $f_2 = x^2 + y^2$, which is invariant for the whole $O(2, \mathbb{R})$; at degree 3 the invariant $f_3 = \dots$, that can be computed by applying the Reynold operator to monomial of degree 3; in degree 4, the invariant $f_4 = f_2^2$; in degree 5, the invariant $f_5 = f_2 f_3$. In degree 6 we have $f_6 = f_2^3$ and $g_6 = f_3^2$. Since f_2 and f_3 are irreducible, f_6 and g_6 are independent. Hence, by Noether theorem, $\mathbb{K}[x_1, x_2]^G$ is generated by f_2 and f_3, which are independent.

The group of the example is a Coxeter group, i.e. a finite effective subgroup of $O(\mathbb{R}^n)$ which is generated by reflections. For these groups the following facts hold[29].

Theorem 1.17. Let $\mathcal{G} \subseteq O(\mathbb{R}^n)$ be a Coxeter group. Then there exist n algebraically independent polynomials $f_1, \dots, f_n \in \mathbb{R}[x_1, \dots, x_n]$ such that

$$\mathbb{R}[x_1, \dots, x_n]^{\mathcal{G}} \cong \mathbb{R}[f_1, \dots, f_n]$$

Moreover, $\mathcal{G}$ has no invariant of degree one while it has always the obvious degree two invariant $x_1^2 + \cdots + x_n^2$.
The *basic generators* $f_1, \dots, f_n$ are not uniquely determined, but their degrees $d_1, \dots, d_n$ are. These degrees satisfy the following properties

1. $|\mathcal{G}| = \prod_{i=1}^n d_i$
2. The total number of reflections in $\mathcal{G}$ is $\sum_{i=1}^n (d_i - 1)$

We close this section with a table (Table 1.1), taken from [2], in which the degrees of the basic generators are given for each Coxeter group.

[29] We refer to [2] for their proofs.

$\mathcal{G}$	$d_1, \ldots, d_n$
$\mathcal{A}_n$	$2, 3, \ldots, n+1$
$\mathcal{B}_n$	$2, 4, \ldots, 2n$
$\mathcal{D}_n$	$2, 4, \ldots, n-2, n, n, n, n+2, \ldots, 2n-2$ (n even)
	$2, 4, \ldots, n-1, n, , n+1, \ldots, 2n-2$ (n odd)
$\mathcal{H}_2^n$	$2, n$
$\mathcal{G}_2$	$2, 6$
$\mathcal{F}_4$	$2, 6, 8, 12$
$\mathcal{I}_3$	$2, 6, 10$
$\mathcal{I}_4$	$2, 12, 20, 30$
$\mathcal{E}_6$	$2, 5, 6, 8, 9, 12$
$\mathcal{E}_7$	$2, 6, 8, 10, 12, 14, 18$
$\mathcal{E}_8$	$2, 8, 12, 14, 18, 20, 24, 30$

Table 1.1: Degree of basic generators of Coxeter groups

Appendix A

Tensor product

We shall denote by $\mathbb{K}$ a numeric field.

Definition A.1. Let $\mathbf{U}$ $\mathbf{V}$ and $\mathbf{W}$ be three dimensional vector spaces over $\mathbb{K}$. A *bilinear map* $\phi : \mathbf{U} \times \mathbf{V} \to \mathbf{W}$ is a function ϕ which is linear in u for each $v \in \mathbf{V}$ and linear in v for each $u \in \mathbf{U}$.

The set $Bil(\mathbf{U} \times \mathbf{V}, \mathbf{W})$ of bilinear maps is a vector space.

Definition A.2. A bilinear map $\phi : \mathbf{U} \times \mathbf{V} \to \mathbf{W}$ is the *tensor product* of $\mathbf{V}$ and $\mathbf{W}$ if it has the following universal property. For every vector space $\mathbf{Z}$ and every bilinear application $g : \mathbf{U} \times \mathbf{V} \to \mathbf{Z}$ there exists a unique linear map $h : \mathbf{W} \to \mathbf{Z}$ such that $g = h \circ \phi$, i.e. such that the following diagram is commutative

$$\begin{tikzcd}
\mathbf{U} \times \mathbf{V} \arrow[rr] \arrow[dr, "g"'] & & \mathbf{W} \arrow[dl, "h"] \\
& \mathbf{Z} &
\end{tikzcd}$$

.

The universal property defines the tensor product uniquely modulo natural isomorphisms. In fact, let ϕ and ϕ' two tensor products $\mathbf{U}$ and $\mathbf{V}$. Then, by the universal property of ϕ there exists one and only one linear map h such that $h \circ \phi = \phi'$ and by the universal property of ϕ' there exists one and only one linear map h' such that $h' \circ \phi' = \phi$. Hence, $h' \circ h \circ \phi = \phi$ and $h \circ h' \circ \phi' = \phi'$. By the universal property of ϕ, $h' \circ h$ coincides with the identity and by the universal property of ϕ', $h \circ h'$ coincides with the identity. We give now a concrete way to build the tensor product of two finite dimensional vector spaces.

Proposition A.1. Let $e_1, \ldots, e_n$ be a basis of $\mathbf{U}$ and let $f_1, \ldots, f_m$ be a basis of $\mathbf{V}$. Let $\mathbf{W}$ be the free vector space over the symbols $u_i \otimes v_j$. The bilinear function $\phi : \mathbf{U} \times \mathbf{V} \to \mathbf{W}$ defined by

$$\phi(\sum_{i=1}^{n} a_i e_i, \sum_{j=1}^{m} b_i f_i) := \sum_{i,j} a_i b_j u_i \otimes v_j$$

is a tensor product of $\mathbf{U}$ and $\mathbf{V}$

Proof Let $g : \mathbf{U} \times \mathbf{V} \to \mathbf{Z}$ be a bilinear function. Let us define $h : \mathbf{W} \to \mathbf{Z}$ by

$$h(\sum c_{i,j} u_i \otimes v_j) = c_{i,j} \sum g(e_i, b_j)$$

It is immediate to check that h is linear. Moreover $h \circ \phi = g$. In fact

$$\begin{array}{ccccc} (h \circ \phi)(\sum a_i e_i, \sum b_j f_j) & = & h(\phi(\sum a_i e_i, \sum) & = \\ h(\sum a_i b_j \phi(e_i, f_j)) & = & \sum a_i b_j h(u_i \otimes v_j) & = \\ \sum a_i b_j g(e_i, f_j) & = & g(\sum a_i e_i, \sum b_j f_j) \end{array} \tag{A.1}$$

If $\overline{h}$ is a linear function $\mathbf{W} \to \mathbf{Z}$ such that $\overline{h} \circ \phi = g$, then $\overline{h}(u_i \otimes w_j) = h(u_i \otimes w_j)$ and since $u_i \otimes v_j$ is a basis, $h = \overline{h}$. □

Bibliography

[1] Bayer D., Mumford D. "What can be computed in algebraic geometry" in *Computational Algebraic Geometry and Commutative Algebra* Eisenbud E. and Robbiano L. editors, Cambridge U. Press, Cambridge, 1993, 1-48.

[2] Benson C. T., Grove L. C. *Finite Reflection Groups*, Second Edition, Springer-Verlag, New York, 1970

[3] Cox D., Little J. O'Shea D., *Ideals, Varieties and Algorithms*, second edition, Springer-Verlag, New York, 1996.

[4] Cox D., *Introduction to Groebner basis* in Applications of Computational Algebraic Geometry, Proceedings of Symposia in Applied Mathematics, vol. 53, ed. Cox D. and Sturmfels B., American Mathematical Society

[5] Decker W., *Computational aspects of algebraic geometry*, Notes for the 1999 Summer School in Nordfjordeid's Sophus Lie Conference Centre.

[6] Derksen H., Kemper G., *Computational Invariant Theory* Encyclopaedia of Mathematical Sciences, Subseries Invariant Theory and Algebraic Transformation Groups, vol. 1, Springer, New York, 2002.

[7] Esembud D., *Commutative algebra with a view toward Algebraic Geometry*, Springer Verlag, New York, 1995.

[8] Enriques F. and Chisini O., *Teoria geometrica delle funzioni algebriche*

[9] Gordan P., "Bewis, dass jede Covariante und Invariante einer binaren Form eine ganze Funktion mit numerischen Coefficienten einer binaren Form eine ganze Funktion mit numerischen Coefficienten einer endlichen Anzahl solcher Formen ist" *J. Reine Angew. Math.*, **69** (1868), 323-35.

[10] Nagata M., *On the 14th problem of Hilbert*, Am. J. Math **81** (1959), 766-772.

[11] Popov V. L., Vinberg E. B., *Invariant theory*, in: N. N. Parshin, I. R. Shafarevich eds, *Algebraic Geometry IV*, Encyclopaedia of Mathematical

[12] Procesi C. *Invariant theory*, Note

[13] Shioda T. "On the graded ring of invariants of binary octavics" *Am. J. Math.* **89** (1967), 1022-1046.

[14] Sloane N. J. A. *Error correcting-codes and invariant theory: new applications of a nineteenth century tecnique*, Amer. Math. Monthly **84** (1977)

[15] Springer T. A., *Invariant Theory*, vol. 585 of Lecture Notes in Mathematics, Springer Verlag, New York, 1977.

[16] Stanley R. P., *Invariants of finite groups*, Bulletin (New Series) of the American Mathematical Society, Volume 1, Number 3, May 1979

[17] Sturmfels B., White N., *Groebner basis and invariant theory* Adv. in Math. **76** (1989), 245-259

[18] Yale P, *Geometry and symmetry*, Holden Day, San Francisco, 1968.