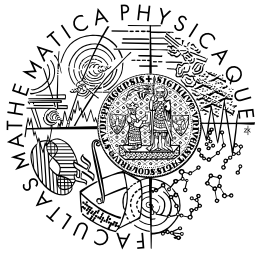

FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

BACHELOR THESIS

Martin Žurav

Invariant theory for finite groups

Department of Algebra

Supervisor of the bachelor thesis:	doc. RNDr. Jan Šťovíček, Ph.D.
Study programme:	Mathematics
Study branch:	General Mathematics

Prague 2018

I declare that I carried out this bachelor thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In ..... date ..... signature of the author

This thesis would have never come into being without the help of my supervisor doc. RNDr. Jan Šťovíček, Ph.D., whose enormous global overview and huge understanding of the given topic I deeply appreciate. Moreover, his responsible approach, patience and a lot of useful advice served as a great inspiration and made creating this text an amazing experience.

Title: Invariant theory for finite groups

Author: Martin Žurav

Department: Department of Algebra

Supervisor: doc. RNDr. Jan Šťovíček, Ph.D., Department of Algebra

Abstract: The prime goal of this thesis is to give a decent introduction to the theory of invariants for finite groups. We begin our characterisation with symmetric polynomials and their fundamental properties. In particular, we study the ring of symmetric polynomials and we prove that it is finitely generated by elementary symmetric functions. Then we deal with some of the criteria for a polynomial to be symmetric.
In the second part, we generalise these ideas for any finite subgroup of $GL(n, k)$. We define an action of a finite linear group on $k[x_1, ..., x_n]$ and consider polynomials that are invariant under such action. We show that they form a ring which is always finitely generated, as follows from the Noether's bound theorem.
At the end, we describe the ring of invariants and relations among its generators more profoundly.

Keywords: Symmetric polynomials, Group action, Invariant polynomials, Ring of invariants

Contents

Introduction

Invariant theory has played a very important role in the development of modern commutative algebra. To see why this daring statement should be possibly true it is sufficient to realise that studying a property of 'being invariant' led to the discovery of some of the most essential theorems in this field, such as Hilbert's basis theorem, Hilbert's Nullstellensatz or Lasker-Noether theorem ([5], page 1). This theory has rich history, too. Its popularity reached the top in the second half of the nineteenth and in the twentieth century with work of David Hilbert, Emmy Noether or Claude Chevalley for instance. Moreover, it is good to mention that symmetric polynomials (the most famous invariants) were objects of interest of even Gauss, who used them to give his second proof of the fundamental theorem of algebra in 1816 ([2], page 314).
This thesis should serve as an introductory text to the theory of invariants, namely to polynomials which are invariant under action of finite subgroups of $GL(n, k)$, rings of those invariants and generators of such rings. This text is inspired by *Ideals, varietes, and algorithms*, a book of David Cox, John Little and Donal O'Shea. The first three chapters are significantly based on ideas of Cox, Little and O'Shea. However, the fourth chapter, concerning relations among generators of the ring of invariants, will be analysed differently, using only algebra-theoretic arguments (such as the weak form of Nullstellensatz) instead of Groebner bases and elimination theory like it is done in *Ideals, varietes, and algorithms*, since it is not an object of study of this thesis.
Let us now start with some of the most fundamental definitions and statements which we will use throughout this whole text.

1. Preliminaries

We assume that the reader has some prior knowledge of basic algebraic terms, such as groups, rings, (algebraically closed) fields, (prime) ideals, a congruence modulo ideal (which is always an equivalence relation) and ring homomorphisms, as they are part of the standard Algebra I. and II. courses (for details see [2], [3] or [1]). Note that all rings in this text are assumed to be commutative with a unit element and each of them is not considered to be a prime ideal in itself. When talking about composition of permutations, we will use right-to-left evaluation each time.
We will start our discussion with definitions of some of the most common objects of algebraic geometry, such as multivariable polynomials, affine varieties, and coordinate rings. All these definitions will be used later in the text.

Definition 1.1. *Let $x_1, ..., x_n$ be variables (or indeterminates), $n \in \mathbb{N}$, $n \geq 1$. Then a monomial in $x_1, ..., x_n$ is a product of the form*

$$x_1{}^{\alpha_1} \cdot x_2{}^{\alpha_2} \cdot ... \cdot x_n{}^{\alpha_n},$$

where all of the exponents $\alpha_1, ..., \alpha_n$ are nonnegative integers. The total degree of such monomial is the sum $\alpha_1 + ... + \alpha_n$.

Note. Let $\alpha = (\alpha_1, ..., \alpha_n)$ be an n-tuple of nonnegative integers. Then x^α denotes the monomial $x_1{}^{\alpha_1} \cdot x_2{}^{\alpha_2} \cdot ... \cdot x_n{}^{\alpha_n}$. The *total degree* of the monomial x^α, denoted by $|\alpha|$, is naturally defined as $|\alpha| = \alpha_1 + ... + \alpha_n$. We will use this notation in order to define a *polynomial* in the next paragraph.

Definition 1.2. *Let k be a field, $n \in \mathbb{N}$, $n \geq 1$, $x_1, ..., x_n$ variables. A polynomial f in $x_1, ..., x_n$ over k is a finite k-linear combination (with coefficients in k) of monomials in $x_1, ..., x_n$. So, every polynomial can be written (uniquely up to order) as*

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

where the sum is over a finite number of n-tuples $\alpha = (\alpha_1, ..., \alpha_n)$. The total degree of f is the maximum $|\alpha|$ such that $a_\alpha \neq 0$. The set of all polynomials in $x_1, ..., x_n$ with coefficients in k (or a polynomial ring, as one can easily verify) is denoted by $k[x_1, ..., x_n]$.

Definition 1.3. *Let $0 \neq f = \sum_\alpha a_\alpha x^\alpha \in k[x_1, ..., x_n]$ and let $>$ be the lexicographical ordering of n-tuples of nonnegative integers. Then we define the multidegree of f as*

$$multideg(f) = max\{\alpha \in \mathbb{N}_0^n : a_\alpha \neq 0\},$$

where max stands for the maximum with respect to $>$. The leading coefficient of f is then

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

the leading monomial of f is

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

and, finally, let the leading term of f be

$$LT(f) = LC(f) \cdot LM(f).$$

Note. Let $f \in k[x_1, ..., x_n]$, i.e., $f = f(x_1, ..., x_n)$. Note that it is not required for all variables $x_1, ..., x_n$ to appear in f. By $f(x_1, ..., x_n)$ we only mean that all variables appearing in f are among $x_1, ..., x_n$.

Example. Let us consider polynomial ring $\mathbb{Q}[x, y, z]$. An example of a polynomial in this ring can be $f(x, y, z) = \frac{4}{5}x^5y^7z^7 - xy$. According to the above note, $g(x, y, z) = y$ is also a valid polynomial in $\mathbb{Q}[x, y, z]$.

Lemma 1.4. *Let $f \in k[x_1, ..., x_n]$, where k is infinite. Then $f(\boldsymbol{a}) = 0$ for all $\boldsymbol{a} \in k^n$ if and only if f is the zero polynomial.*

Proof. Suppose $f(\mathbf{a}) = f(a_1, ..., a_n) = 0$ for every $\mathbf{a} \in k^n$. Using induction on the number of variables, we will show that f must be the zero polynomial then.
The base case is $n = 1$, meaning $f \in k[x_1]$. Let's set $m = deg(f)$. If f is nonzero then f has at most m roots ([1], page 78). But k being infinite by assumption forces f to be the zero polynomial.
Next, assume that the statement is true for every polynomial in $k[x_1, ..., x_{n-1}]$. Surely, we can write f as

$$f = \sum_{i=0}^{m} g_i(x_1, ..., x_{n-1})x_n^i, \tag{1.1}$$

where all g_i's are polynomials in the first $n - 1$ variables and m is the greatest exponent of x_n occurring in f. By choosing any $(a_1, ..., a_{n-1})$ of k^{n-1} we get a one-variable polynomial $f(a_1, ..., a_{n-1}, x_n) \in k[x_n]$. In this situation, the assumption on f and the base case give us that $f(a_1, ..., a_{n-1}, x_n)$ is the zero polynomial in $k[x_n]$. But then (1.1) implies that $g_i(a_1, ..., a_{n-1}) = 0$ for all $i \in \{0, ..., m\}$. Moreover, all g_i's vanish on every point of k^{n-1}, as follows from the arbitrariness of choice of $(a_1, ..., a_{n-1})$. We can now use our inductive hypothesis to get that g_i is the zero polynomial for each i. It follows that f itself must be the zero polynomial.
The converse is trivial. □

We are now going to define the fundamental object of algebraic geometry.

Definition 1.5. *Let k be a field and let $S \subseteq k[x_1, ..., x_n]$. Consider the set*

$$\boldsymbol{V}(S) = \Big\{(a_1, ..., a_n) \in k^n : f(a_1, ..., a_n) = 0 \text{ for all } f \in S\Big\}.$$

Then $\boldsymbol{V}(S)$ is called the affine variety defined by S. We say that a nonempty affine variety V is irreducible if for any two affine varieties $V_1, V_2 \subseteq k^n$ an expression $V = V_1 \cup V_2$ implies $V = V_1$ or $V = V_2$.

Note. Therefore, an affine variety is actually a set of all common solutions of $f(x_1, ..., x_n) = 0$, as f varies over S.

Example. Let $k = \mathbb{R}$. It is obvious that $\mathbf{V}(x^2 + y^2 - 4)$ in $\mathbb{R}^2$ is equal to the circle of radius 2 centred at the origin. However, $\mathbf{V}(x^2 + y^2 - 4, y - x)$ in $\mathbb{R}^2$ equals only to a set of two points $\Big\{[\sqrt{2},\sqrt{2}],[-\sqrt{2},-\sqrt{2}]\Big\}$ and $\mathbf{V}(x^2 + y^2 - 4, y - x, y - x^2)$ is empty.

Definition 1.6. *Let $V \subseteq k^n$ be an affine variety. Then we define*

$$\boldsymbol{I}(V) = \Big\{f \in k[x_1, ..., x_n]; f(a_1, ..., a_n) = 0 \text{ for all } (a_1, ..., a_n) \in V\Big\}.$$

Lemma 1.7. *$\boldsymbol{I}(V)$ is an ideal of $k[x_1, ..., x_n]$ for every affine variety $V \subseteq k^n$.*

Proof. We need to check that $\mathbf{I}(V)$ contains the zero element and whether it is closed under addition and multiplication by elements of $k[x_1, ..., x_n]$. The zero polynomial vanishes on every point of k^n. In particular, it vanishes on V, because $V \subseteq k^n$. Therefore, $0 \in \mathbf{I}(V)$. Now, let $f, g \in \mathbf{I}(V)$, $h \in k[x_1, ..., x_n]$ and $(a_1, ..., a_n) \in V$. Then

$$(f + g)(a_1, ..., a_n) = f(a_1, ..., a_n) + g(a_1, ..., a_n) = 0 + 0 = 0.$$

Hence, $f + g \in \mathbf{I}(V)$. Similarly,

$$(h \cdot f)(a_1, ..., a_n) = h(a_1, ..., a_n) \cdot f(a_1, ..., a_n) = h(a_1, ..., a_n) \cdot 0 = 0.$$

Thus, $h \cdot f \in \mathbf{I}(V)$, what completes the proof. □

Lemma 1.8. *Let V be an affine variety in k^n. Then V is irreducible if and only if $\boldsymbol{I}(V)$ is prime.*

Proof. See ([3], page 7). □

Definition 1.9. *Let $V \subseteq k^m$ and $W \subseteq k^n$ be affine varieties. A mapping $\varphi : V \to W$ is said to be a polynomial mapping if there exist polynomials $f_1, ..., f_n \in k[x_1, ..., x_m]$ such that the following condition is satisfied:*

$$\varphi(a_1, ..., a_m) = \Big(f_1(a_1, ..., a_m), ..., f_n(a_1, ..., a_m)\Big)$$

for all $(a_1, ..., a_m) \in V$. If so, then we say that φ is represented by $(f_1, ..., f_n)$.

Note. In case of $W = k$, the definition above says that for a map $\varphi : V \to k$ to be a polynomial mapping it means there exists a polynomial $f \in k[x_1, ..., x_m]$ representing φ.

Note. Now, we are going to explore the situation that two polynomials represent the same polynomial function. So, for instance, let $V = \mathbf{V}(x^2 - 1, x + y, z^5 - 3) \subseteq \mathbb{R}^3$ and $f = x^2 - yz^2 + z^3$. Then f represents a polynomial function $\varphi : V \to \mathbb{R}$, $\varphi(x, y, z) = f(x, y, z)$. Now, consider

$$g = x^2 - yz^2 + z^3 + A \cdot (x^2 - 1) + B \cdot (x + y) + C \cdot (z^5 - 3),$$

where A, B, C are arbitrary polynomials in $\mathbb{R}[x, y, z]$. Then for every point $\mathbf{a} = (a_1, a_2, a_3) \in V$ we have

$$g(\mathbf{a}) = f(\mathbf{a}) + A(\mathbf{a}) \cdot 0 + B(\mathbf{a}) \cdot 0 + C(\mathbf{a}) \cdot 0 = f(\mathbf{a}),$$

because $x^2 - 1, x + y$ and $z^5 - 3$ all vanish on every point of V by the definition of V. It means that g represents the same polynomial function φ as f. To sum up, adding any polynomial $h \in \mathbf{I}(V)$ to f does not change the value of f at any point of V. This leads to the following lemma.

Lemma 1.10. *Let $V \subseteq k^n$ be an affine variety and $f, g \in k[x_1, ..., x_n]$. Then f and g represent the same polynomial function on V if and only if $f - g \in \boldsymbol{I}(V)$. More generally, $(f_1, ..., f_m)$ and $(g_1, ..., g_m)$ represent the same polynomial mapping from V to k^m if and only if $f_i - g_i \in \boldsymbol{I}(V)$ for each $i \in \{1, ..., m\}$.*

Proof. We will only prove the first part of the statement as the second part follows immediately from the first one. So, let f and g represent the same polynomial function on V. Then $(f-g)(\mathbf{a}) = f(\mathbf{a}) - g(\mathbf{a}) = 0$ at every point $\mathbf{a} \in V$. Thus, $f-g \in \mathbf{I}(V)$ by the definition of $\mathbf{I}(V)$. Conversely, let $f-g \in \mathbf{I}(V)$. By definition, $0 = (f-g)(\mathbf{a}) = f(\mathbf{a}) - g(\mathbf{a})$ for all $\mathbf{a} \in V$. It follows that f and g represent the same polynomial function on V. □

Corollary. There is a one-to-one correspondence between the distinct polynomial functions $\varphi : V \to k$ and the equivalence classes of polynomials under congruence modulo $\mathbf{I}(V)$.

Definition 1.11. *Let $V \subseteq k^n$ be an affine variety. The set of all polynomial functions $\varphi : V \to k$ will be denoted by $k[V]$.*

For the purposes of this text it is necessary to define a structure of a commutative ring on two sets, namely on $k[V]$ and on $k[x_1, ..., x_n]/I$.

Note. Consider $\varphi, \psi \in k[V]$ and $\mathbf{a} \in V$. Then we can define addition and multiplication on $k[V]$ as follows:

$$(\varphi + \psi)(\mathbf{a}) = \varphi(\mathbf{a}) + \psi(\mathbf{a}),$$

$$(\varphi.\psi)(\mathbf{a}) = \varphi(\mathbf{a}).\psi(\mathbf{a}),$$

$$(-\varphi)(\mathbf{a}) = -\varphi(\mathbf{a}).$$

In addition, $\varphi + 0 = \varphi$ and $\varphi.1 = \varphi$ where 0 is the zero function and 1 denotes the function identically equal to 1. Now, let $f_1, f_2 \in k[x_1, ..., x_n]$ be representatives of φ and $g_1, g_2 \in k[x_1, ..., x_n]$ be representatives of ψ. Then, for $\mathbf{a} \in V$ we have

$$f_1(\mathbf{a}) = \varphi(\mathbf{a}) = f_2(\mathbf{a}), g_1(\mathbf{a}) = \psi(\mathbf{a}) = g_2(\mathbf{a}),$$

$$(f_1 + g_1)(\mathbf{a}) = f_1(\mathbf{a}) + g_1(\mathbf{a}) = \varphi(\mathbf{a}) + \psi(\mathbf{a}) = f_2(\mathbf{a}) + g_2(\mathbf{a}) = (f_2 + g_2)(\mathbf{a}),$$

$$(f_1.g_1)(\mathbf{a}) = f_1(\mathbf{a}).g_1(\mathbf{a}) = \varphi(\mathbf{a}).\psi(\mathbf{a}) = f_2(\mathbf{a}).g_2(\mathbf{a}) = (f_2.g_2)(\mathbf{a}),$$

$$(-f_1)(\mathbf{a}) = (-\varphi)(\mathbf{a}) = -\varphi(\mathbf{a}) = -(f_2)(\mathbf{a}) = (-f_2)(\mathbf{a}).$$

Hence, the operations are well-defined and by the equations above we have just defined a commutative-ring structure on $k[V]$. All the axioms of a commutative ring hold because they are satisfied in $k[x_1, ..., x_n]$.

Definition 1.12. *Let $V \subseteq k^n$ be an affine variety. Then the coordinate ring of V is the ring $k[V]$.*

Note. Let I be an ideal of $k[x_1, ..., x_n]$ and let $[f], [g] \in k[x_1, ..., x_n]/I$. We define the basic ring operations on $k[x_1, ..., x_n]/I$ as follows:

$$[f] + [g] = [f + g], [f].[g] = [f.g],$$

$$-[f] = [-f],$$

$$1 = [1], 0 = [0] = I.$$

These operations are well-defined and the proof can be found in [2], page 220. Therefore, we have a commutative-ring structure on $k[x_1, ..., x_n]/I$, too.

The last corollary can be specified a bit more.

Theorem 1.13. *Let $V \subseteq k^n$ be an affine variety and set $F : k[x_1..., x_n]/\boldsymbol{I}(V) \to k[V]$ such that $F\Big([f]\Big) = \varphi$, where f represents the polynomial function φ. Then F is a ring isomorphism.*

Proof. F is well-defined by **Lemma 1.10**, so we need to show that F is a bijective homomorphism. By definition, every φ of $k[V]$ is represented by some polynomial f of $k[x_1, ..., x_n]$. Hence, F is onto. Now, suppose $F\Big([f]\Big) = F\Big([g]\Big)$. Then *Corollary* 1 of **Lemma 1.10** implies that $[f] = [g]$ in $k[x_1, ..., x_n]/\mathbf{I}(V)$, what proves that F is injective. Finally, F must respect the identity element, addition and multiplication. So, take $[f], [g] \in k[x_1, ..., x_n]/\mathbf{I}(V)$ arbitrarily. By definition, we already know that $[f] + [g] = [f + g]$ and $[f].[g] = [f.g]$. In addition, $f + g$ (or $f.g$) represents $\varphi + \psi$ (or $\varphi.\psi$), whenever f represents φ and g represents ψ. Then we have

$$F\Big([f] + [g]\Big) = F\Big([f + g]\Big) = \varphi + \psi = F\Big([f]\Big) + F\Big([g]\Big),$$

$$F\Big([f].[g]\Big) = F\Big([f.g]\Big) = \varphi.\psi = F\Big([f]\Big).F\Big([g]\Big),$$

$$F\Big([1]\Big) = id_{k[V]}.$$

We see that F is a homomorphism and the proof is complete. □

Theorem 1.14. ***(The weak Nullstellensatz):*** *Let k be an algebraically closed field. Then $\boldsymbol{V}(I)$ is nonempty for every proper ideal $I \subseteq k[x_1, ..., x_n]$.*

Proof. See [2], page 168 or [3], page 10. □

2. Symmetric polynomials

We are now ready to introduce some theory of symmetric polynomials. They appear more naturally than one may think. For instance, consider monic polynomial $f = x^3 + bx^2 + cx + d \in k[x]$. Let α, β, and γ be the roots of f in the algebraic closure $\overline{k}$ of k. In $\overline{k}[x]$ we then have

$$x^3 + bx^2 + cx + d = (x - \alpha)(x - \beta)(x - \gamma).$$

Expanding the right-hand side gives

$$x^3 + bx^2 + cx + d = x^3 - (\alpha + \beta + \gamma)x^2 + (\alpha\beta + \alpha\gamma + \beta\gamma)x - \alpha\beta\gamma.$$

Therefore, the coefficients of f are polynomials in its roots α, β and γ. Furthermore, any change in ordering of roots does not change f itself, so after expanding we always get the same right-hand side. This means that the coefficients of f remain unchanged after permuting its roots. This leads to the following definition.

Definition 2.1. *Let $f \in k[x_1, ..., x_n]$ be a polynomial. Then f is said to be symmetric (or S_n-invariant), if f remains unchanged after we permute its variables. Formally,*

$$f(x_{\pi(1)}, ..., x_{\pi(n)}) = f(x_1, ..., x_n)$$

for every permutation $\pi \in S_n$.

Example. Let $f, g, h \in k[x, y]$, such that $f = x + y$, $g = x^2 + 4xy + y^2$, and $h = x - y$. Then definitely f and g are symmetric while h is not whenever the characteristics of k is different from 2 $\big($if $char(k)$=2 then $h = f$ and, hence, h is symmetric$\big)$. Turning into n-variable case, another example of a symmetric polynomial is $F = \prod_{1 \leq i < j \leq n}(x_i - x_j)^2$.

Note. When talking about symmetric polynomials, we need to be precise and specify the ring of polynomials we work with. We have already seen that the field we are working over matters. However, other problems may possibly occur, even with k fixed. For example, it might be the case that a polynomial is symmetric in $k[X_1]$, but is not symmetric in $k[X_2]$, where X_1, X_2 are some sets of variables and $X_1 \subset X_2$. As an example consider the polynomial f from the above example. Then f is symmetric in $k[x, y]$, but f is not symmetric in $k[x, y, z]$.

Note. It follows immediately that a sum and a product of symmetric polynomials is again a symmetric polynomial and constant polynomials are symmetric, too. Hence, **the set of all symmetric polynomials forms a subring of the corresponding ring of polynomials**. We will generalise this fact in the next chapter. However, we can already describe generators of the ring of symmetric polynomials. To do so, we will need the following crucial definition.

Definition 2.2. *Let $k[x_1, ..., x_n]$ be a given polynomial ring. Then for $i \in \mathbb{N}_0$ we define a polynomial $\sigma_i \in k[x_1, ..., x_n]$ as follows:*

$$\sigma_0 = 1,$$

$$\sigma_1 = x_1 + ... + x_n,$$

$$\vdots$$

$$\sigma_l = \sum_{1 \le i_1 < i_2 < \ldots < i_l \le n} x_{i_1} x_{i_2} \cdots x_{i_l},$$

$$\vdots$$

$$\sigma_n = x_1 x_2 \cdots x_n,$$

$$\sigma_m = 0 \text{ for } m > n,$$

where σ_l *is a sum of* $\binom{n}{l}$ *terms.*

Note. By generalising the case of $f = x^3 + bx^2 + cx + d$ from above we see that $\sigma_1, ..., \sigma_n$ all appear (up to a sign) as coefficients of every monic polynomial $f \in k[x]$ of degree n if we name the roots of f by $x_1, x_2, ..., x_n$. Formally, $\sigma_1, ..., \sigma_n$ satisfy the relation

$$\prod_{i=1}^{n}(x - x_i) = x^n - \sigma_1 x^{n-1} + \sigma_2 x^{n-2} - ... + (-1)^{n-1}\sigma_{n-1}x + (-1)^n \sigma_n$$

in $k[x_1, ..., x_n, x]$. It is easy to see that $\sigma_1, ..., \sigma_n$ are all symmetric polynomials. Such $\sigma_1, ..., \sigma_n$ are called *elementary symmetric functions* (or *polynomials*).

There are certainly more places where one can (rather unexpectedly) find symmetric polynomials in mathematics. Just to mention one more, realise that the inequality of algebraic and geometric means (AM-GM inequality) can be formulated as

$$\frac{\sigma_1}{n} \ge \sqrt[n]{\sigma_n},$$

where x_i's are nonnegative reals.

The following famous theorem solves our problem of generators of the ring of symmetric polynomials.

Theorem 2.3. ***(The fundamental theorem of symmetric polynomials)****: Let* $k[x_1, ..., x_n]$ *be given. Then every symmetric polynomial in* $k[x_1, ..., x_n]$ *can be written uniquely as a polynomial in the elementary symmetric functions* $\sigma_1, ..., \sigma_n$*.*

Proof. At first, let f be a symmetric polynomial in $k[x_1, ..., x_n]$, order $\mathbb{N}_0^n$ lexicographically and let $LT(f) = a_0 x^\alpha$, where $\alpha = (\alpha_1, ..., \alpha_n)$ as before. Observe that $\alpha_1 \ge ... \ge \alpha_n$: Suppose, for the sake of contradiction, that there is an index i such that $\alpha_i < \alpha_{i+1}$. Let $\beta = (..., \alpha_{i+1}, \alpha_i, ...)$, where β is the vector of exponents which we get from α by swapping α_i and α_{i+1}. It follows that $\beta > \alpha$. But then $a_0 x^\beta$ is a term of $f(..., x_{i+1}, x_i, ...)$, because $a_0 x^\alpha$ is a term of $f(x_1, ..., x_n)$. However, f is symmetric by assumption, what forces $f(..., x_{i+1}, x_i, ...) = f(x_1, ..., x_n)$. This implies that $a_0 x^\beta$ is a term of f, and that is a contradiction with the maximality of α.
We will now remove the leading term from f. For this purpose, let

$$g_0 = \sigma_1^{\alpha_1 - \alpha_2} \sigma_2^{\alpha_2 - \alpha_3} \cdots \sigma_{n-1}^{\alpha_{n-1} - \alpha_n} \sigma_n^{\alpha_n}.$$

Clearly, for each $l \in \{1, ..., n\}$ it is true that $LT(\sigma_l) = x_1 x_2 \cdots x_l$. This implies that

$$LT(g_0) = LT(\sigma_1)^{\alpha_1 - \alpha_2} LT(\sigma_2)^{\alpha_2 - \alpha_3} \cdots LT(\sigma_n)^{\alpha_n}$$

$$= x_1^{\alpha_1-\alpha_2}(x_1x_2)^{\alpha_2-\alpha_3}\cdots(x_1\cdots x_n)^{\alpha_n} = x_1^{\alpha_1}x_2^{\alpha_2}\cdots x_n^{\alpha_n} = x^{\alpha}.$$

Therefore, f and g_0 have the same leading monomial. Consequently, f and a_0g_0 have the same leading term, i.e., either $f - a_0g_0 = 0$ or, at least,

$$multideg(f - a_0g_0) < multideg(f).$$

We know that both f and a_0g_0 are symmetric. Thus, $f_1 = f - a_0g_0$ is symmetric. If f_1 is the zero polynomial then we are done. If not, we can eliminate the leading term of f_1 in the same way we did it for f in order to get $f_2 = f_1 - a_1g_1$, where $a_1 \in k$ and g_1 is again a product of σ_i's to the appropriate powers. Once more, either $f_2 = 0$ and the process is over, or at least, the multidegree of f_2 is strictly less than the one of f_1. In case of $f_2 \neq 0$ we can repeat this procedure as many times as needed and end up with a polynomial sequence $f, f_1, f_2, f_3, ...$ such that

$$multideg(f) > multideg(f_1) > multideg(f_2) > \cdots .$$

It is important to realise that the lexicographic order is a well-ordering on $\mathbb{N}_0^n$. Thus, the sequence must be finite. However, our procedure can come to an end if and only if there exists some index $l \in \mathbb{N}$ such that $f_{l+1} = 0$. But then

$$f = \sum_{i=0}^{l} a_ig_i$$

proves the existence of an expression of f as a polynomial in σ_i's.
It remains to prove the uniqueness. Suppose there are two polynomial expressions of f in $\sigma_1, ..., \sigma_n$, namely

$$f = h_1(\sigma_1, ..., \sigma_n) = h_2(\sigma_1, ..., \sigma_n),$$

where h_1 and h_2 are some polynomials in $k[y_1, ..., y_n]$. We want $h_1 = h_2$.
Let $h = h_1 - h_2$. Then the equation above says that $h(\sigma_1, ..., \sigma_n) = 0$ in $k[x_1, ..., x_n]$. We will show that $h = 0$ in $k[y_1, ..., y_n]$, what will consequently imply the desired uniqueness. For the sake of contradiction suppose $h \neq 0$. Let $h = \sum_\beta a_\beta y^\beta$, $\beta = (\beta_1, ..., \beta_n)$. By substituting σ_i's into h we have $h(\sigma_1, ..., \sigma_n) = \sum_\beta h_\beta$, where $h_\beta = a_\beta\sigma_1^{\beta_1}\sigma_2^{\beta_2}\cdots\sigma_n^{\beta_n}$. It is straightforward to see that

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

Now, consider $\varphi : \mathbb{N}_0^n \to \mathbb{N}_0^n$, such that

$$\varphi\Big((\beta_1, ..., \beta_n)\Big) = (\beta_1 + \cdots + \beta_n, \beta_2 + \cdots + \beta_n, ..., \beta_n).$$

Observe that φ is injective: Suppose $\varphi(\beta) = \varphi(\gamma)$. Then $\beta_n = \gamma_n$ because of the last coordinate. But then $\beta_{n-1} = \gamma_{n-1}$ because of the last but one coordinate, and so on. It follows that $\beta = \gamma$.
We see that h_β's have pairwise different leading terms. Since there are only finitely many of them and, in particular, the ordering is linear, we can choose β such that $LT(h_\beta) > LT(h_\gamma)$ $\Big($meaning $multideg(h_\beta) > multideg(h_\gamma)\Big)$ for all $\gamma \neq \beta$. It follows that $LT(h_\beta)$ is strictly greater than any term of any h_γ for $\gamma \neq \beta$. But this exactly means that the term $LT(h_\beta)$ cannot be cancelled in $h(\sigma_1, ..., \sigma_n)$, leaving $h(\sigma_1, ..., \sigma_n) \neq 0$ in $k[x_1, ..., x_n]$ – a contradiction.
The proof is now complete. □

The proof of the above theorem is useful in finding the desired polynomial expression itself:

Exercise. Let $f = (x^2+y^2)(x^2+z^2)(y^2+z^2) \in k[x,y,z]$. Then

$$f = x^4y^2 + x^4z^2 + x^2y^4 + 2x^2y^2z^2 + x^2z^4 + y^4z^2 + y^2z^4.$$

Since f is obviously symmetric, it can be expressed in terms of σ_1, σ_2 and σ_3. The leading term of f is $x^4y^2 = LT(\sigma_1^2\sigma_2^2)$. Then

$$f_1 = f - \sigma_1^2\sigma_2^2$$
$$= -2x^4yz - 2x^3y^3 - 8x^3y^2z - 8x^3yz^2 - 2x^3z^3 - 8x^2y^3z - 13x^2y^2z^2$$
$$-8x^2yz^3 - 2xy^4z - 8xy^3z^2 - 8xy^2z^3 - 2xyz^4 - 2y^3z^3.$$

In this situation, the leading term of f_1 is $-2x^4yz = -2LT(\sigma_1^3\sigma_3)$. We have

$$f_2 = f_1 + 2\sigma_1^3\sigma_3$$
$$= -2x^3y^3 - 2x^3y^2z - 2x^3yz^2 - 2x^3z^3 - 2x^2y^3z - x^2y^2z^2 - 2x^2yz^3$$
$$-2xy^3z^2 - 2xy^2z^3 - 2y^3z^3.$$

Now, the leading term of f_2 is $-2x^3y^3 = -2LT(\sigma_2^3)$. Similarly,

$$f_3 = f_2 + 2\sigma_2^3$$
$$= 4x^3y^2z + 4x^3yz^2 + 4x^2y^3z + 11x^2y^2z^2 + 4x^2yz^3 + 4xy^3z^2 + 4xy^2z^3.$$

This time, the leading term of f_3 is $4x^3y^2z = 4LT(\sigma_1\sigma_2\sigma_3)$. But then

$$f_4 = f_3 - 4\sigma_1\sigma_2\sigma_3 = -x^2y^2z^2.$$

Hence, it is obvious that

$$f_4 + \sigma_3^2 = 0.$$

Finally, it follows that

$$f = \sigma_1^2\sigma_2^2 - 2\sigma_1^3\sigma_3 - 2\sigma_2^3 + 4\sigma_1\sigma_2\sigma_3 - \sigma_3^2.$$

Exercise. Suppose we have the following system of equations in $k[x,y]$:

$$x + y = a,$$
$$x^2 + xy + y^2 = b,$$
$$x^3 + x^2y^2 + y^3 = c.$$

We want to find the relation between a, b and c, assuming the solution of the system exists. We have

$$x + y = a = \sigma_1$$
$$x^2 + xy + y^2 = b = (x+y)^2 - xy = \sigma_1^2 - \sigma_2$$
$$\Longrightarrow a^2 - \sigma_2 = b \Longrightarrow a^2 - b = \sigma_2$$
$$x^3 + x^2y^2 + y^3 = c = (x+y)^3 + x^2y^2 - 3x^2y - 3xy^2$$
$$= \sigma_1^3 + \sigma_2^2 - 3\sigma_1\sigma_2.$$

But then

$$c = a^3 + (a^2-b)^2 - 3a(a^2-b).$$

It might be useful to look for some sufficient conditions for a polynomial to be S_n-invariant. We will prove a few of them and show some other properties of symmetric polynomials in the rest of this chapter.

Definition 2.4. *Let $f \in k[x_1, ..., x_n]$. Then f is homogeneous of total degree l, if every monomial appearing in f has total degree l.*

Example. The i-th elementary symmetric polynomial σ_i is homogeneous of total degree i.

Note. Let $f \in k[x_1, ..., x_n]$. Then f can be written uniquely (up to order) as a sum of homogeneous polynomials, because $f = \sum_l f_l$, where f_l is the sum of all terms of f of total degree l. Such f_l is said to be an l-th *homogeneous component* of f. Hence, we receive another criterion for a polynomial to be symmetric.

Lemma 2.5. *Let $f \in k[x_1, ..., x_n]$. Then f is symmetric if and only if all homogeneous components of f are symmetric.*

Proof. Suppose f is symmetric, $\pi \in S_n$ and let $x_{\pi(1)}, ..., x_{\pi(n)}$ be a permutation of variables. It is obvious that π does not change the total degree of the terms of f, i.e., if a is a term of f of total degree l, then, after permutation of variables, $\tilde{a}$ is again a term of total degree l. Therefore, the symmetricity of f implies that all homogeneous components of f must be symmetric, too. Conversely, let all homogeneous components f_l of f be symmetric. Since $f = \sum_l f_l$ and all f_l's are symmetric, f must be symmetric as well. □

Exercise. Let $f \in k[x_1, ..., x_n]$. If

$$f(x_1, x_2, x_3, ..., x_n) = f(x_2, x_1, x_3, ..., x_n) = f(x_2, x_3, ..., x_n, x_1),$$

then f is symmetric. Hence, in order to verify whether f is symmetric or not, it is sufficient to check two specific permutations instead of $n!$.

Proof. We will show that $\varphi = (1\ 2)$ and $\psi = (1\ 2\ ...\ n)$ generate S_n. Then the statement will be an immediate consequence. To do so, we will use a fact, that every permutation can be written as a product of disjoint cycles. Furthermore, every cycle $(x_1\ x_2\ ...\ x_l)$ can be written as $(x_1\ x_2)(x_2\ x_3)...(x_{l-1}\ x_l)$. It follows that S_n is generated by transpositions. We will use this to prove that $\{\varphi, \psi\}$ is also a generating set for S_n. It is useful to observe that for $i \in \{1, ..., n-2\}$ we have

$$\alpha_i = (1\ 2\ ...\ n)^i(1\ 2)(1\ 2\ ...\ n)^{-i} = (i+1\ i+2).$$

Indeed, $(1\ 2\ ...\ n)^{-1} = (n\ ...\ 2\ 1)$, so $(1\ 2\ ...\ n)^{-i} = (1\ 2\ ...\ n)^{n-i}$ sends an element l to an element $l-i$ mod n whenever $l \neq i$ and it sends i to n. Similarly, $(1\ 2\ ...\ n)^i$ sends $n-i$ to n and any other element l to $l+i$ mod n. If $l = i+1$, then $\alpha_i(l) = \alpha_i(i+1) = i+2$. If $l = i+2$, then $\alpha_i(l) = \alpha_i(i+2) = i+1$. For all other l's we have $\alpha_i(l) = (l)$. Hence, $\alpha_i = (i+1\ i+2) = (i+2\ i+1)$. To complete the proof, it is sufficient to show that $(1\ 2), ..., (n-1\ n)$ generate all transpositions, i.e., we want to prove that any transposition $(a\ b)$ can be written as a product of α_i's and $(1\ 2)$, where $a, b \in \{1, ..., n\}, a \neq b$.
Without loss of generality, we can assume $a < b$, since $(a\ b) = (b\ a)$. We will use induction on $l = b - a$. The base case of $l = 1$ is trivial, because then

$(a\ b) = (i\ i+1)$, what is either equal to φ or α_i for some $i \in \{1, ..., n-2\}$. Anyway, it is an element of our set. Now, suppose that $l = b - a = m$ for some $m \in \{2, ..., n-1\}$ and that the statement is true for every $l < m$. In this situation, we can write

$$(a\ b) = (a\ a+1)(a+1\ b)(a\ a+1).$$

However, $(a\ a+1)$ belongs to our set and the difference of b and $a+1$ is strictly less than m. So, we can use the inductive assumption to receive that $(a\ b)$ is really of the desired form. □

The following definition describes another important type of symmetric polynomials.

Definition 2.6. *Let $l \in \mathbb{N}$. Then we define a polynomial $s_l \in k[x_1, ..., x_n]$ as*

$$s_l = x_1^l + ... + x_n^l.$$

Such s_l is obviously symmetric and is called the l-th power sum.

Lemma 2.7. ***(The Newton identities)**:*
Let $k[x_1, ..., x_n]$ be given, $m \in \mathbb{N}$. Then we have

1. *If $m \geq n$ then*

$$\sum_{i=0}^{m}(-1)^i \sigma_i s_{m-i} = \sum_{i=0}^{n}(-1)^i \sigma_i s_{m-i} = 0,$$

   where $s_0 = n \cdot 1$.

2. *If $m \leq n$ then*

$$\sum_{i=0}^{m-1}(-1)^i \sigma_i s_{m-i} + (-1)^m m\sigma_m = 0.$$

Proof. For proof see ([4], pages 1-3). □

Corollary. Let $f \in k[x_1, ..., x_n]$ be symmetric, where k is a field of characteristic zero. Then f can be written as a polynomial in the power sums $s_1, ..., s_n$.

Proof. The fundamental theorem of symmetric polynomials states that every symmetric polynomial can be written as a polynomial in σ_i's, $i \in \{1, ..., n\}$. Therefore, we only need to prove that each such σ_i can be written as a polynomial in s_i's. We will use induction on l to show that σ_l has the desired property. The base case of l being equal to 1 is trivial, because $\sigma_1 = s_1$.
Now, suppose the statement holds for $\sigma_1, \sigma_2, ..., \sigma_{l-1}$. Using the second Newton identity we receive

$$\sigma_l = \frac{(-1)^{l+1}}{l}\sum_{i=0}^{l-1}(-1)^i \sigma_i s_{l-i}.$$

Note that division by l is well-defined because of the assumption on k. Our inductive hypothesis now completes the proof. □

Note. Consider working in $k[x_1, ..., x_n]$, where $char(k) \neq 2$. An intuitive generalisation of S_n-invariant polynomials might be the one concerning A_n-invariant polynomials, i.e., polynomials which remain unchanged after every even permutation, but may possibly change after any odd permutation. In fact, the set of all A_n-invariant polynomials can be divided into three disjoint (up to the zero polynomial) parts:

a) polynomials, which are S_n-invariant as well. Those are exactly the symmetric polynomials.

b) polynomials, which are invariant under even permutations, but change their sign under an odd permutation, i.e.,

$$f(x_{\pi(1)}, ..., x_{\pi(n)}) = sgn(\pi) \cdot f(x_1, ..., x_n),$$

where π is a permutation on n elements. Such polynomials are called *alternating* and a typical representative is the *Vandermonde* polynomial $V_n = \prod_{1 \leq i < j \leq n}(x_i - x_j)$, the determinant of the Vandermonde matrix. Note the importance of the second powers of terms of F from the first example of this chapter. In fact, the set of all alternating polynomials forms an R-bimodule, where R is the ring of symmetric polynomials, since a sum of two alternating polynomials is again an alternating polynomial, and a product of a symmetric polynomial and an alternating polynomial is alternating, too. However, the alternating polynomials do not form an R-algebra, since a product of two alternating polynomials is symmetric. Note that if $char(k) = 2$, then $\{\text{alternating polynomials}\} = \{\text{symmetric polynomials}\}$.

c) polynomials, which are A_n-invariant, but change dramatically after an odd permutation. For instance, consider $f(x, y, z) = x^2y + xz^2 + y^2z \in k[x, y, z]$. Then it is easy to see that $f(x, y, z) = f(y, z, x) = f(z, x, y)$, which proves that f is A_3-invariant. However, $f(y, x, z) = y^2x + yz^2 + x^2z \neq \pm f(x, y, z)$.

3. Finite matrix groups, rings of invariants and their generators

We will now begin to discuss the general theory of invariants. Therefore, we need to define finite matrix groups and action of a group on polynomials. From now on, **we will always assume k to be of characteristic zero**.

Note. The set of all invertible $n \times n$ matrices with entries in a field k is denoted by $GL(n, k)$. A product of two invertible matrices is invertible, an inverse of any invertible matrix is also invertible and finally, I_n is an invertible matrix, too. Therefore, $GL(n, k)$ with matrix multiplication, inversion and identity element I_n, forms a multiplicative group, so-called *general linear group*.

Definition 3.1. *A finite subset $G \subseteq GL(n, k)$ is called a finite matrix group if it is nonempty and closed under matrix multiplication. The order of G (denoted by $|G|$) is defined as the number of elements of G.*

Note. From the definition above it is by no means obvious that every nonempty and finite subset of $GL(n, k)$ closed under multiplication is actually a group, because we did not verify all the group axioms. However, the following lemma proves that the definition above makes sense and it also gives us some basic properties of finite matrix groups. Even before, we can realise that multiplication is associative in G, since it is associative in $GL(n, k)$.

Lemma 3.2. *Let $G \subseteq GL(n, k)$ be a finite matrix group. Then*

1. *$I_n \in G$.*
2. *If $A \in G$, then $A^m = I_n$ for some positive integer m.*
3. *If $A \in G$, then $A^{-1} \in G$.*

Proof. Let $A \in G$ and consider $X = \{A, A^2, A^3, ...\}$. Then $X \subseteq G$, because G is closed under multiplication. Since G is finite, there must be some $i, j \in \mathbb{N}, i > j$, such that $A^i = A^j$. Since $G \subseteq GL(n, k)$, A is invertible, so we can multiply this equation by A^{-j} in order to get $A^{i-j} = I_n$. This proves the second part of the statement and tells us that $I_n = A^{i-j} = A^{i-j-1}A = AA^{i-j-1}$, which implies that $A^{-1} = A^{i-j-1}$ whenever $i > j + 1$, or if $i = j + 1$, then $A^{i-j} = A^1 = A = I_n$. However, either $A^{i-j-1} \in G$, because $A^{i-j-1} \in X \subseteq G$, or in the second case, $A^{-1} = I_n^{-1} = I_n = A \in X \subseteq G$. Hence, the statement $A^{-1} \in G$ holds. For the first part it is sufficient to realise that $I_n = A^{i-j} \in X \subseteq G$. □

The following two examples should illustrate the definition properly.

Example. Linear algebra gives us a one-to-one correspondence between permutations of S_n and $n \times n$ permutation matrices over k (denoted by P_n) – for a chosen $\pi \in S_n$ we can create M_π by permuting columns of the identity matrix according to π. In addition, it tells us that

$$M_\pi \cdot M_\varphi = M_{\pi \circ \varphi}, \tag{3.1}$$

what implies that P_n is closed under multiplication. It has $n!$ elements and thus is a finite matrix group. We can also define a group isomorphism between these two groups. Let $f : S_n \to P_n$, such that

$$f(\pi) = M_\pi.$$

Then it is trivial to see that f is both one-to-one and onto and (3.1) implies it is also a homomorphism. Therefore, we will often use symbol S_n instead of P_n when talking about permutation matrices.

Example. Let $A \in GL(2,k)$, $char(k) \neq 2$, such that $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and let G be the cyclic group generated by A. Then

$$G = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\}.$$

Obviously, G is a finite subset of $GL(2,k)$ and it is closed under matrix multiplication. Therefore, G is a finite matrix group.

Note. For the rest of this section let $\mathbf{x}$ denote the column vector of the variables $x_1, ..., x_n$, i.e.,

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

So, from now on, $f(\mathbf{x})$ is an abbreviation for $f(x_1, ..., x_n)$.

Definition 3.3. *Let $G \subseteq GL(n,k)$ be a finite matrix group, $f(\boldsymbol{x}) \in k[x_1, ..., x_n]$. Then $f(\boldsymbol{x})$ is invariant under G if*

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

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

Definition 3.4. *Let $f_1, ..., f_m \in k[x_1, ..., x_n]$. Then the subset of $k[x_1, ..., x_n]$ consisting of all polynomial expressions in $f_1, ..., f_m$ with coefficients in k is denoted by $k[f_1, ..., f_m]$,i.e.,*

$$k[f_1, ..., f_m] = \Big\{ f \in k[x_1, ..., x_n] \Big| f = g(f_1, ..., f_m); g \in k[y_1, ..., y_m] \Big\}.$$

Note. It is easy to prove that $k[f_1, ..., f_m]$ is a subring of $k[x_1, ..., x_n]$: it is trivially closed under addition and multiplication and it also contains every constant polynomial; in particular, it contains the zero and the identity element. We say that such subring is *generated* by $f_1, ..., f_m$ over k.

Example. Consider the group $S_n \subseteq GL(n,k)$ of permutation matrices. Then the theory of S_n-invariant polynomials from above gives us that $k[x_1, ..., x_n]^{S_n} = \Big\{$symmetric polynomials in $k[x_1, ..., x_n]\Big\}$. Moreover, the fundamental theorem of symmetric polynomials proves

$$k[x_1, ..., x_n]^{S_n} = k[\sigma_1, ..., \sigma_n].$$

Lemma 3.5. *Let $G \subseteq GL(n,k)$ be a finite matrix group. Then the set of invariants $k[x_1,...,x_n]^G$ is a subring of $k[x_1,...,x_n]$.*

Proof. The proof is very straightforward and uses only the definition of invariance. □

We can now generalise **Lemma 2.5** of **Chapter 2**.

Theorem 3.6. *Let $G \subseteq GL(n,k)$ be a finite matrix group, $f \in k[x_1,...,x_n]$. Then f is invariant under G if and only if all of its homogeneous components are.*

Proof. Let $f \in k[x_1,...,x_n]$ be invariant under G and choose $A = (a_{ij}) \in G$. We already know that f can be written as $f = \sum_l f_l$, where f_l is the l-th homogeneous component of f. Then

$$f(\mathbf{x}) = f(A \cdot \mathbf{x}) = \sum_l f_l(A \cdot \mathbf{x}) = \sum_l \sum_{i_1+...+i_n=l} f_{i_1,...,i_n}(A \cdot \mathbf{x}), \tag{3.2}$$

where $f_{i_1,...,i_n}(\mathbf{x}) = c_{i_1,...,i_n}.{x_1}^{i_1} \cdot ... \cdot {x_n}^{i_n}$, and where i_j's are nonnegative integers and $0 \neq c_{i_1,...,i_n} \in k$. By the definition of matrix multiplication,

$$(3.2) = \sum_l \sum_{i_1+...+i_n=l} c_{i_1,...,i_n}.(a_{11}x_1 + ... + a_{1n}x_n)^{i_1} \cdot ... \cdot (a_{n1}x_1 + ... + a_{nn}x_n)^{i_n}.$$

It is obvious that every monomial in $(a_{j1}x_1 + ... + a_{jn}x_n)^{i_j}$ has total degree equal to i_j. Hence, $c_{i_1,...,i_n}.(a_{11}x_1+...+a_{1n}x_n)^{i_1}\cdot...\cdot(a_{n1}x_1+...+a_{nn}x_n)^{i_n}$ is homogeneous of total degree l. It follows that $f_l(\mathbf{x})$ is the l-th homogeneous component of $f(\mathbf{x})$ if and only if $f_l(A \cdot \mathbf{x})$ is the l-th homogeneous component of $f(A \cdot \mathbf{x})$. So, if f is invariant under G, then it must be invariant componentwise, i.e., its homogeneous components must be invariant under G as well. The converse is now trivial. □

Lemma 3.7. *Let $G \subseteq GL(n,k)$ be a finite matrix group and suppose there exist $m \in \mathbb{N}$ and $A_1,...,A_m \in G$ such that every $A \in G$ can be written in the form*

$$A = B_1 B_2 \cdots B_l,$$

where $B_i \in \{A_1,...,A_m\}$ for every $i \in \{1,...,l\}$, and where $l \in \mathbb{N}$ (i.e., $A_1,...,A_m$ generate G). Then f is invariant under G if and only if

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

Proof. The proof uses simple induction. For details see [2], page 325. □

Note. Note how this lemma generalises the last *Exercise* 2 of **Chapter 2**. All it says is that it is always sufficient to check the invariance just on the generators.

We have already seen that $k[x_1,...,x_n]^{S_n}$ is finitely generated by $\sigma_1,...,\sigma_n$. It is natural to ask what we can say about $k[x_1,...,x_n]^G$ in general. Is it always the case that $k[x_1,...,x_n]^G$ is finitely generated? Well, the answer is yes and we shall prove it through the Noether's degree bound theorem. The theorem essentially uses the concept of the Reynolds operator.

Definition 3.8. *Let $G \subseteq GL(n,k)$ be a finite matrix group. Then a map $R_G : k[x_1,...,x_n] \to k[x_1,...,x_n]$ defined by the formula*

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

for $f(\boldsymbol{x}) \in k[x_1,...,x_n]$ is called the Reynolds operator of G.

Note. Division by $|G|$ is well-defined, because we assume k to be of characteristic zero.

We can now state some of the most important properties of R_G.

Lemma 3.9. *Let $G \subseteq GL(n,k)$ be a finite matrix group and let R_G be the Reynolds operator of G. Then:*

1. *R_G is k-linear in f.*
2. *If $f \in k[x_1,...,x_n]$, then $R_G(f) \in k[x_1,...,x_n]^G$.*
3. *If $f \in k[x_1,...,x_n]^G$, then $R_G(f) = f$.*
4. *If $f \in k[x_1,...,x_n]^G$ and $g \in k[x_1,...,x_n]$, then $R_G(f\cdot g) = f\cdot R_G(g)$.*
5. *R_G is a surjective mapping from $k[x_1,...,x_n]$ to $k[x_1,...,x_n]^G$.*
6. *$R_G \circ R_G = R_G$.*

Proof. 1. This is an immediate consequence of the definition.
2. Let $B \in G$. Then we have

$$R_G(f)(B\cdot \mathbf{x}) = \frac{1}{|G|}\sum_{A\in G} f\Big(A\cdot(B\cdot \mathbf{x})\Big) = \frac{1}{|G|}\sum_{A\in G} f\Big((AB)\cdot \mathbf{x}\Big). \tag{3.3}$$

We can write G as $\{A_1,...,A_{|G|}\}$. However, $A_iB \neq A_jB$, whenever $i \neq j$ (otherwise, we could multiply the equation by B^{-1} to get $A_i = A_j$ for some i and j, what is a contradiction). Hence, $\{A_1B,...,A_{|G|}B\}$ has $|G|$ distinct elements and, since G is closed under multiplication, it is a subset of G. It follows that the set is equal to G, so we can write $G = \{AB | A \in G\}$. So, the set of all $f\Big((AB)\cdot\mathbf{x}\Big)$'s equals the set of all $f(A\cdot\mathbf{x})$'s. Thus,

$$(3.3) = \frac{1}{|G|}\sum_{A\in G} f(A\cdot\mathbf{x}) = R_G(f)(\mathbf{x}).$$

$$\Longrightarrow R_G(f)(B\cdot\mathbf{x}) = R_G(f)(\mathbf{x})$$

As $B \in G$ was chosen arbitrarily, this proves *2.*
To prove *3.*, let f be invariant under G. Then

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

For *4.*, take any $g \in k[x_1,...,x_n]$ and realise that

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

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

5. $k[x_1,...,x_n]^G$ is the codomain of R_G by the second part. For a given $f \in k[x_1,...,x_n]^G$, f itself is a preimage of f by *3.*
Finally, $R_G(f)$ is an element of $k[x_1,...,x_n]^G$ by *2.* Therefore, we can use *3.* to obtain $R_G\Big(R_G(f)\Big) = R_G(f)$. □

Example. Consider a matrix group G from the *Example* 3 preceding **Definition 3.3**. Then

$$R_G(xy)(\mathbf{x}) = \frac{1}{4}\Big(xy + (-y)x + (-x)(-y) + y(-x)\Big) = 0.$$

This shows that the Reynolds operator of f might be zero, although the input polynomial f was not.

Note. Let $f \in k[x_1,...,x_n]$ be a monomial of total degree l, i.e, $f(\mathbf{x}) = {x_1}^{i_1}\cdot....\cdot{x_n}^{i_n}$, $i_1 + ... + i_n = l$ and let $A \in GL(n,k)$. Then

$$f(A\cdot\mathbf{x}) = (a_{11}x_1 + ... + a_{1n}x_n)^{i_1}\cdot ... \cdot (a_{n1}x_1 + ... + a_{nn}x_n)^{i_n}.$$

In the proof of **Theorem 3.6** we found out that $f(A\cdot\mathbf{x})$ is homogeneous of total degree l. It follows that for any finite matrix group G if $R_G(f)$ is nonzero, then it is a homogeneous invariant of total degree l.

The following useful and very important theorem, which Emmy Noether was the first to prove, tells us that the ring of invariants is always finitely generated. Furthermore, it also gives us an explicit algorithm for finding the generators.

Theorem 3.10. *(**The Noether's bound**) Let $G \subseteq GL(n,k)$ be a finite matrix group. Then*

$$k[x_1,...,x_n]^G = k\Big[R_G(x^\beta) : |\beta| \le |G|\Big].$$

In particular, $k[x_1,...,x_n]^G$ is generated by finitely many homogeneous invariants.

Proof. Let $f = \sum_\gamma c_\gamma x^\gamma \in k[x_1,...,x_n]^G$. By using **Lemma 3.9**, we get

$$f \overset{3.}{=} R_G(f) = R_G\left(\sum_\gamma c_\gamma x^\gamma\right) \overset{1.}{=} \sum_\gamma c_\gamma R_G(x^\gamma).$$

The equation above tells us that every G-invariant polynomial is a linear combination of $R_G(x^\alpha)$ with coefficients in k. Thus, we just need to show that for every $\alpha = (\alpha_1,...,\alpha_n) \in \mathbb{N}_0^n$, the homogeneous invariant $R_G(x^\alpha)$ can be expressed as a polynomial in $R_G(x^\beta)$'s, where $|\beta| \le |G|$.
To do this, we will take the same path as Emmy Noether did before: we will not fix any α and then prove that for this particular α we can find an appropriate polynomial for $R_G(x^\alpha)$. Instead, we will fix a natural number l, and we will take a look at all $R_G(x^\alpha)$'s with $|\alpha| = l$ and then try to prove the statement for all such α's at once. We will show that a certain k-linear combination of these is actually equal to the l-th power sum, which is a symmetric polynomial. Such information is going to be really helpful, since we already have a solid apparatus of working with symmetric polynomials, especially the last *Corollary* 2 of the previous chapter. Using this particular result, we will show that every power sum

S_l is a polynomial in $S_1, ..., S_{|G|}$. By that time we will have already proved that for every $1 \leq i \leq |G|$, S_i is a polynomial in $R_G(x^\beta)$'s, $|\beta| = i$, altogether with some auxiliary variables, and the theorem will then become an easy consequence. To do this formally, firstly fix $l \in \mathbb{N}$ and set

$$(x_1 + \cdots + x_n)^l = \sum_{|\alpha|=l} a_\alpha x^\alpha. \tag{3.4}$$

Note, that every monomial of total degree l is contained in the sum (and no else monomial is), since a_α's are all positive integers by the multinomial theorem.
We will use the following notation: for any $A = (a_{ij}) \in G$, set $A_i = (a_{i1}, ..., a_{in})$. Hence, $A_i \cdot \mathbf{x} = a_{i1}x_1 + ... + a_{in}x_n$. Moreover, for every $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}_0^n$ let

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

Using this notation we receive

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

$$= \frac{1}{|G|} \sum_{A \in G} (A_1 \cdot \mathbf{x})^{\alpha_1} \cdot ... \cdot (A_n \cdot \mathbf{x})^{\alpha_n} = \frac{1}{|G|} \sum_{A \in G} (A \cdot \mathbf{x})^\alpha.$$

As we have mentioned already, we want to prove the statement for all α's, such that $|\alpha| = l$, at once. Thus, we must not allow any particular α to cancel out during our calculations. For this purpose, we will use new variables $u_1, ..., u_n$. If we now substitute $u_i(A_i \cdot \mathbf{x})$ for x_i in (3.4), then for any $A \in G$ we receive

$$\Big(u_1(A_1 \cdot \mathbf{x}) + ... + u_n(A_n \cdot \mathbf{x}) \Big)^l = \sum_{|\alpha|=l} a_\alpha \Big(u_1(A_1 \cdot \mathbf{x}) \Big)^{\alpha_1} \cdot ... \cdot \Big(u_n(A_n \cdot \mathbf{x}) \Big)^{\alpha_n}$$

$$= \sum_{|\alpha|=l} a_\alpha (A \cdot \mathbf{x})^\alpha u^\alpha,$$

where $u^\alpha = u_1^{\alpha_1} \cdot ... \cdot u_n^{\alpha_n}$. Summing over all $A \in G$ we have

$$\sum_{A \in G} \Big(u_1(A_1 \cdot \mathbf{x}) + ... + u_n(A_n \cdot \mathbf{x}) \Big)^l = \sum_{A \in G} \Big(\sum_{|\alpha|=l} a_\alpha (A \cdot \mathbf{x})^\alpha u^\alpha \Big)$$

$$= \sum_{|\alpha|=l} a_\alpha \Big(\sum_{A \in G} (A \cdot \mathbf{x})^\alpha \Big) u^\alpha = \sum_{|\alpha|=l} a_\alpha |G| R_G(x^\alpha) u^\alpha. \tag{3.5}$$

Note the usefulness of our new variables: the sum on the right-hand side now contains every $R_G(x^\alpha)$ with $|\alpha| = l$.
There is even one more set of new variables we need. Let us introduce a new variable U_A for each $A \in G$, so that $U_A = u_1(A_1 \cdot \mathbf{x}) + ... + u_n(A_n \cdot \mathbf{x})$. In this situation, the left-hand side of (3.5) is equal to the l-th power sum S_l in $k[U_A; A \in G]$, i.e.,

$$S_l = \sum_{|\alpha|=l} a_\alpha |G| R_G(x^\alpha) u^\alpha.$$

Since l was chosen arbitrarily, $S_i = \sum_{|\alpha|=i} a_\alpha |G| R_G(x^\alpha) u^\alpha$ for every $i \in \mathbb{N}$. Moreover, S_l is symmetric in $k[U_A; A \in G]$ and the corollary mentioned above implies

that it can be written as a polynomial in $S_1, ..., S_{|G|}$, i.e., $S_l = F(S_1, ..., S_{|G|})$ for some $F \in k[y_1, ..., y_{|G|}]$.

$$\implies \sum_{|\alpha|=l} a_\alpha |G| R_G(x^\alpha) u^\alpha = F\Big(\sum_{|\beta|=1} a_\beta |G| R_G(x^\beta) u^\beta, ..., \sum_{|\beta|=|G|} a_\beta |G| R_G(x^\beta) u^\beta \Big).$$

Now, the need for variables $u_1, ..., u_n$ is even more obvious. It may possibly be the case that some terms in the sum in (3.5) may not be expressible in suitable $R_G(x^\beta)$'s, while the whole sum is. At the end, however, we would get the desired polynomial expression for the whole sum, but not for any particular $R_G(x^\alpha)$. Therefore, we would be able to say nothing about the existence of such expression for that one specific $R_G(x^\alpha)$. But since we used the new variables, by comparing the coefficients of u^α on both sides, we see that for each particular α with $|\alpha| = l$, $a_\alpha |G| R_G(x^\alpha)$ is equal to some polynomial in $R_G(x^\beta)$'s, $|\beta| \leq |G|$. By assumption, k is of zero characteristic, which forces $a_\alpha |G|$ to be nonzero. But then we can divide the equation to receive that $R_G(x^\alpha)$ itself is a polynomial in $R_G(x^\beta)$'s, $|\beta| \leq |G|$. The theorem is now proven. □

4. Relations among generators and the geometry of orbits

The last theorem of the previous chapter states that for $k[x_1, ..., x_n]$ and any finite subgroup $G \subseteq GL(n, k)$ there exist homogeneous invariant polynomials $f_1, ..., f_m$ such that

$$k[x_1, ..., x_n]^G = k[f_1, ..., f_m].$$

The main goal of this chapter is to study algebraic relations of these generators, and, consequently, give a more precise characterisation of the ring of invariants. As we already mentioned in the **Introduction**, we will take a rather different path in proving the final theorems than Cox, Little, and O'Shea did in [2].

Definition 4.1. *Let $F = (f_1, ..., f_m)$. Then we define a set I_F as*

$$I_F = \Big\{h \in k[y_1, ..., y_m] : h(f_1, ..., f_m) = 0 \text{ in } k[x_1, ..., x_n]\Big\}.$$

Here come some of the most important properties of I_F.

Lemma 4.2. *Let $k[x_1, ..., x_n]^G = k[f_1, ..., f_m]$, $F = (f_1, ..., f_m)$ and let $I_F \subseteq k[y_1, ..., y_m]$ be as above. Then*

1. *I_F is a proper ideal of $k[y_1, ..., y_m]$ (so-called ideal of relations for F). Furthermore, it is a prime ideal.*

2. *Let $f \in k[x_1, ..., x_n]^G$ and let $f = g(f_1, ..., f_m)$ be one representation of f in terms of $f_1, ..., f_m$. Then all such representations are given by*

$$f = g(f_1, ..., f_m) + h(f_1, ..., f_m),$$

where h varies over I_F.

Proof. 1. At first, we have to prove that I_F is an ideal. Well, $0 \in k[y_1, ..., y_m]$ belongs to I_F, since $0(f_1, ..., f_m) = 0$ in $k[x_1, ..., x_n]$. Next, take any $h_1, h_2 \in I_F$ and $g \in k[y_1, ..., y_m]$. Then:

$$h_1(f_1, ..., f_m) = h_2(f_1, ..., f_m) = 0$$

$$\implies (h_1 + h_2)(f_1, ..., f_m) = h_1(f_1, ..., f_m) + h_2(f_1, ..., f_m) = 0 + 0 = 0$$

$$\&$$

$$\implies (g \cdot h_1)(f_1, ..., f_m) = g(f_1, ..., f_m) \cdot h_1(f_1, ..., f_m) = g(f_1, ..., f_m) \cdot 0 = 0.$$

It is a proper ideal, since $1(f_1, ..., f_m) = 1 \neq 0$, therefore, 1 (and also every nonzero constant polynomial) is not an element of I_F.
Now, suppose $f \cdot g \in I_F$ for $f, g \in k[y_1, ..., y_m]$. Then

$$0 = (f \cdot g)(f_1, ..., f_m) = f(f_1, ..., f_m) \cdot g(f_1, ..., f_m).$$

By assumption, k is a field, and hence, $k[x_1, ..., x_n]$ is a domain. So, $f(f_1, ..., f_m) = 0$ or $g(f_1, ..., f_m) = 0$, meaning $f \in I_F$ or $g \in I_F$. This proves part *1.*

2. Fix $g \in k[y_1, ..., y_m]$ such that $g(f_1, ..., f_m) = f$ and choose $h \in I_F$ arbitrarily. Then

$$g(f_1, ..., f_m) + h(f_1, ..., f_m) = g(f_1, ..., f_m) + 0 = f,$$

so every $g(f_1, ..., f_m) + h(f_1, ..., f_m)$ is a representation of f. Conversely, suppose $f = \overline{g}(f_1, ..., f_m)$ for some $\overline{g} \in k[y_1, ..., y_m]$. We want to show that

$$\overline{g}(f_1, ..., f_m) = g(f_1, ..., f_m) + h(f_1, ..., f_m)$$

for some $h \in I_F$. We have

$$f = \overline{g}(f_1, ..., f_m) = g(f_1, ..., f_m) + \Big(\overline{g}(f_1, ..., f_m) - g(f_1, ..., f_m)\Big)$$

$$= g(f_1, ..., f_m) + (\overline{g} - g)(f_1, ..., f_m).$$

It suffices to prove that $(\overline{g} - g)(f_1, ..., f_m) = 0$ in $k[x_1, ..., x_n]$. However,

$$(\overline{g} - g)(f_1, ..., f_m) = \overline{g}(f_1, ..., f_m) - g(f_1, ..., f_m) = f - f = 0. \quad \square$$

The following theorem gives a nice characterisation of $k[x_1, ..., x_n]^G$ as a quotient of $k[y_1, ..., y_m]$ modulo I_F.

Theorem 4.3. *Let $k[x_1, ..., x_n]^G = k[f_1, ..., f_m]$ and let $F = (f_1, ..., f_m)$. Then $k[y_1, ..., y_m]/I_F$ and $k[x_1, ..., x_n]^G$ are ring-isomorphic.*

Proof. Consider $\varphi : k[y_1, ..., y_m] \to k[x_1, ..., x_n]^G$ such that

$$\varphi(g) = g(f_1, ..., f_m).$$

Now, φ is a homomorphism, so-called evaluation homomorphism. The kernel of φ is exactly I_F. Moreover, $k[x_1, ..., x_n]^G = k[f_1, ..., f_m]$ by assumption, meaning every invariant is a polynomial in $f_1, ..., f_m$. Consequently, φ is onto and the first isomorphism theorem now yields the result. $\square$

The ideal of relations I_F is closely related to another important term in the invariant theory.

Definition 4.4. *Let $k[x_1, ..., x_n]^G = k[f_1, ..., f_m]$ and set $F = (f_1, ..., f_m)$. Then we define V_F to be the affine variety of I_F, i.e.,*

$$V_F = \boldsymbol{V}(I_F) \subseteq k^m.$$

Theorem 4.5. *Let $F = (f_1, ..., f_m), k[f_1, ..., f_m] = k[x_1, ..., x_n]^G$. Consider $\varphi : k^n \to k^m$ given by the formula*

$$\varphi(\boldsymbol{a}) = \Big(f_1(\boldsymbol{a}), ..., f_m(\boldsymbol{a})\Big),$$

where $\boldsymbol{a} = (a_1, ..., a_n) \in k^n$. Then:

1. *V_F is the smallest affine variety in k^m containing the image of φ.*
2. *$\boldsymbol{I}\Big(\varphi(k^n)\Big) = I_F = \boldsymbol{I}(V_F)$.*
3. *V_F is an irreducible variety.*

4. $k[V_F] \cong k[x_1, ..., x_n]^G$.

Proof. At first, we will show that

$$I_F = \mathbf{I}\Big(\varphi(k^n)\Big). \tag{4.1}$$

$\subseteq$: Let $h \in I_F$ and take any $\mathbf{P} \in \varphi(k^n)$. Hence, there must be some $\mathbf{a} \in k^n$ such that $\mathbf{P} = \Big(f_1(\mathbf{a}), ..., f_m(\mathbf{a})\Big)$. Observe that

$$h(\mathbf{P}) = h\Big(f_1(\mathbf{a}), ..., f_m(\mathbf{a})\Big) = 0$$

by assumption on h.
$\supseteq$: Let $h \in \mathbf{I}\Big(\varphi(k^n)\Big)$ and set $g = h(f_1, ..., f_m) \in k[x_1, ..., x_n]$. But then $g(\mathbf{a}) = 0$ for all $\mathbf{a} \in k^n$. We assume k to be of characteristic zero, and hence, infinite. **Lemma 1.4** now forces g to be the zero polynomial, meaning $h \in I_F$.
It follows that

$$V_F = \mathbf{V}(I_F) = \mathbf{V}\Big(\mathbf{I}\Big(\varphi(k^n)\Big)\Big). \tag{4.2}$$

However, $\mathbf{I}\Big(\varphi(k^n)\Big)$ is the greatest possible set of polynomials vanishing on the image of φ. It is an immediate consequence that $V_F = \mathbf{V}\Big(\mathbf{I}\Big(\varphi(k^n)\Big)\Big)$ is the smallest affine variety containing the whole image of φ, since the mapping $\mathbf{V}$: {ideals} $\to$ {affine varieties}; $I \mapsto \mathbf{V}(I)$, is inclusion reversing ([3], page 4 and 6). Furthermore,

$$\mathbf{I}(V_F) \overset{(4.2)}{=} \mathbf{I}\Big(\mathbf{V}\Big(\mathbf{I}\Big(\varphi(k^n)\Big)\Big)\Big) \overset{3,6}{=} \mathbf{I}\Big(\varphi(k^n)\Big) \overset{(4.1)}{=} I_F.$$

So, we have already proved *1.* and *2.* The statement *3.* follows from **Lemma 1.8** and the first part of **Lemma 4.2**. Similarly, *4.* follows immediately from *2.*, **Theorem 1.13** and **Theorem 4.3**. □

Definition 4.6. *Let $G \subseteq GL(n, k)$ be a finite matrix group and let $\boldsymbol{a}$ be an element of k^n. The $G-$orbit of $\boldsymbol{a}$ is the set $G \cdot \boldsymbol{a} = \{A \cdot \boldsymbol{a}; A \in G\}$. The set of all $G-$orbits in k^n is denoted by k^n/G and we call it the orbit space.*

Note. We can define an equivalence relation $\sim_G$ on k^n as follows:

$$\mathbf{x} \sim_G \mathbf{y} \iff (\exists A \in G) : (\mathbf{x} = A \cdot \mathbf{y}).$$

It is very straightforward to check that $\sim_G$ is reflexive, symmetric and transitive. It is obvious that the equivalence classes of $\sim_G$ are exactly the G-orbits, what means that the set of equivalence classes is just k^n/G. It is a well-known fact that two classes are either the same or disjoint.

Theorem 4.7. *Let $k[x_1, ..., x_n]^G = k[f_1, ..., f_m]$, $F = (f_1, ..., f_m)$, $\boldsymbol{a} \in k^n$, where k is **algebraically closed**. Then*

1. *The polynomial mapping $\varphi : k^n \to V_F$ defined in the previous theorem is surjective.*

2. *The map sending the $G-$orbit $G\cdot \boldsymbol{a}\subseteq k^n$ to the point $\varphi(\boldsymbol{a})\in V_F$ induces a one-to-one correspondence between k^n/G and V_F.*

Proof. 1. We have $\varphi : k^n \to V_F; \varphi(\mathbf{a}) = \Big(f_1(\mathbf{a}), ..., f_m(\mathbf{a})\Big)$ for $\mathbf{a}\in k^n$. Now, choose any $\mathbf{P} = (p_1, ..., p_m) \in \mathbf{V}(I_F)$. We want to find some element in k^n, whose image is $\mathbf{P}$. What it basically means is that we are trying to solve the following system of equations:

$$f_{1,\mathbf{P}} = f_1(x_1, ..., x_n) - p_1 = 0,$$

$$\vdots$$

$$f_{m,\mathbf{P}} = f_m(x_1, ..., x_n) - p_m = 0.$$

Note that each $f_{i,\mathbf{P}} \in k[x_1, ..., x_n]$ is a G-invariant polynomial, since it is equal to the sum of two G-invariant polynomials, what is an element of $k[x_1, ..., x_n]^G$. Now, observe that the system has a solution if and only if

$$\mathbf{V}(f_{1,\mathbf{P}}, ..., f_{m,\mathbf{P}}) \neq \emptyset.$$

So, suppose, for the sake of contradiction, that there is no such solution. Then $\mathbf{V}(f_{1,\mathbf{P}}, ..., f_{m,\mathbf{P}}) \overset{3.4}{=} \mathbf{V}\Big((f_{1,\mathbf{P}}, ..., f_{m,\mathbf{P}})\Big)$ is empty. Now, **the Weak Nullstellensatz** (**Theorem 1.14**) tells us that

$$(f_{1,\mathbf{P}}, ..., f_{m,\mathbf{P}}) = k[x_1, ..., x_n].$$

Particularly, there exist some $c_1, ..., c_m \in k[x_1, ..., x_n]$ such that

$$c_1 f_{1,\mathbf{P}} + ... + c_m f_{m,\mathbf{P}} = 1.$$

Using the Reynolds operator on both sides of the above equation we have

$$R_G\left(\sum_{i=1}^{m} c_i f_{i,\mathbf{P}}\right) = R_G(1)$$

$$\overset{\text{L. 3.9}}{\Longrightarrow} \sum_{i=1}^{m} R_G(c_i) f_{i,\mathbf{P}} = \sum_{i=1}^{m} R_G(c_i)(f_i - p_i) = 1$$

in $k[x_1, ..., x_n]^G$. If we use the isomorphism from **Theorem 4.3**, we can take a look at this equation in $k[y_1, ..., y_m]/I_F$ to receive

$$u(y_1, ..., y_m) = \textstyle\sum_{i=1}^{m} \widehat{R_G(c_i)}(y_i - p_i) = 1 \bmod I_F.$$

Now, what is $u(\mathbf{P})$ equal to? Well, on one hand, $u(\mathbf{P}) = 1 \bmod I_F$, since u is constantly equal to 1 mod I_F. On the other hand, however, $(y_i - p_i)(\mathbf{P}) = 0$ for each i, and thus, $u(\mathbf{P})$ must be 0 mod I_F. Hence, we get a contradiction, since $0 \neq 1 \bmod I_F$, meaning φ is surjective.

2. Let $\overline{\varphi} : k^n/G \to V_F$ be the mapping sending the G-orbit $G\cdot\mathbf{a}$ to $\varphi(\mathbf{a}) = \Big(f_1(\mathbf{a}), ..., f_m(\mathbf{a})\Big)$. We need to show that $\overline{\varphi}$ is a well-defined bijection.

Note that each f_i is a G-invariant, and thus, $f_i(A\cdot\mathbf{x}) = f_i(\mathbf{x})$ for every $A\in G$. Now, take any $\mathbf{b}\in G\cdot\mathbf{a}$. It means that there is some $A\in G$, such that $\mathbf{b} = A\cdot\mathbf{a}$. But then $\overline{\varphi}$ is well-defined, because

$$\overline{\varphi}(G\cdot\mathbf{b}) = \Big(f_1(\mathbf{b}), ..., f_m(\mathbf{b})\Big) = \Big(f_1(A\cdot\mathbf{a}), ..., f_m(A\cdot\mathbf{a})\Big)$$

$$= \Big(f_1(\mathbf{a}), ..., f_m(\mathbf{a})\Big) = \overline{\varphi}(G \cdot \mathbf{a}).$$

By part *1.*, φ is surjective, i.e., for every $\mathbf{v} \in V_F$ there is some $\mathbf{a} \in k^n$, such that $\varphi(\mathbf{a}) = \mathbf{v}$. But then $\overline{\varphi}(G \cdot \mathbf{a}) = \mathbf{v}$, implying $\overline{\varphi}$ is surjective, too.
Finally, we need to prove that $\overline{\varphi}$ is injective. So, take two distinct orbits $G \cdot \mathbf{a}$ and $G \cdot \mathbf{b}$. By the note preceding this theorem, $G \cdot \mathbf{a}$ and $G \cdot \mathbf{b}$ are disjoint. We want to show that $\overline{\varphi}(G \cdot \mathbf{a}) \neq \overline{\varphi}(G \cdot \mathbf{b})$. In order to do so, we will find a polynomial g of $k[x_1, ..., x_n]^G$ such that $g(\mathbf{a}) \neq g(\mathbf{b})$. We will then express g in terms of $f_1, ..., f_m$ and show that $f_i(\mathbf{a}) \neq f_i(\mathbf{b})$ for some i and the theorem will follow.
Consider the set $S = G \cdot \mathbf{a} \cup G \cdot \mathbf{b} \setminus \{\mathbf{a}\}$. Then S is finite. Moreover, S is an affine variety by [3], page 5. Therefore, $S = \mathbf{V}(M)$ for some $M \subseteq k[x_1, ..., x_n]$. Note that there must exist a polynomial $f \in M$, such that $f(\mathbf{a}) \neq 0$, since we left $\mathbf{a}$ outside of S. Now, for every $A \in G$ we see that $f(A \cdot \mathbf{b}) = 0$ and $f(A \cdot \mathbf{a}) = f(\mathbf{a}) \neq 0$, whenever $A \cdot \mathbf{a} = \mathbf{a}$ and $f(A \cdot \mathbf{a}) = 0$ otherwise.
Set $g = R_G(f)$. But then $g(\mathbf{b}) = R_G(f)(\mathbf{b})$ is a sum of zeros, and, hence, is zero itself. On the other hand, $g(\mathbf{a}) = \frac{c}{|G|} f(\mathbf{a})$, where c denotes the cardinality of the set of those elements A of G, such that $A \cdot \mathbf{a} = \mathbf{a}$. Since k is assumed to be of characteristic 0, we see that $0 \neq g(\mathbf{a}) \neq g(\mathbf{b}) = 0$.
The polynomial g belongs to $k[x_1, ..., x_n]^G$ by **Lemma 3.9** and can therefore be written as a polynomial h in $f_1, ..., f_m$. Furthermore, we have already showed that $h(f_1, ..., f_m)(\mathbf{a}) \neq h(f_1, ..., f_m)(\mathbf{b})$. It follows that there must be some index i, such that $f_i(\mathbf{a}) \neq f_i(\mathbf{b})$. Consequently, $\overline{\varphi}(G \cdot \mathbf{a}) \neq \overline{\varphi}(G \cdot \mathbf{b})$, what proves the injectivity of $\overline{\varphi}$. □

Note. By definition, k^n/G is just a set. However, in case of k being algebraically closed, the last theorem enables us to define a structure of an affine variety on k^n/G in the following way:
There is a bijection $\overline{\varphi}$ between k^n/G and V_F by part *2.* and $\varphi : k^n \to V_F$ is a surjective polynomial mapping by *1.* Therefore, we have

$$k^n \overset{\varphi}{\twoheadrightarrow} V_F \overset{\overline{\varphi}^{-1}}{\hookrightarrow} k^n/G;$$

$$\mathbf{a} \mapsto \varphi(\mathbf{a}) \mapsto G \cdot \mathbf{a}.$$

Altogether, if we identify k^n/G with V_F, then there exists a surjective polynomial mapping $\tilde{\varphi}$, such that

$$\tilde{\varphi} : k^n \twoheadrightarrow k^n/G; \tilde{\varphi}(\mathbf{a}) = G \cdot \mathbf{a}.$$

So, if $k = \overline{k}$, then not only k^n/G has a structure of an affine variety, but this structure also has a property that there exists a surjective polynomial mapping $\tilde{\varphi}$ from k^n to k^n/G. Moreover, the structure of an affine variety on k^n/G is uniquely determined up to isomorphism (see [2], page 346).
Finally, using **Theorem 4.5** and **Theorem 4.7**, observe that

$$k[k^n/G] \cong k[V_F] \cong k[x_1, ..., x_n]^G,$$

i.e., the coordinate ring of k^n/G is just the ring of G-invariant polynomials. This final result now should not be very surprising, since if we take any $f \in k[x_1, ..., x_n]^G$ and $\mathbf{a} \in k^n$, then $f(\mathbf{a}) = f(A \cdot \mathbf{a})$ for all $A \in G$. This means that f has always the same value on whole G-orbits and, thus, f defines a polynomial function on k^n/G.

Conclusion

In this thesis, we studied polynomials with the property of being invariant under given group action. At the beginning, we focused on S_n-invariant polynomials. The fundamental theorem of symmetric polynomials gave us the proof that the ring $k[x_1, ..., x_n]^{S_n}$ is finitely generated by elementary symmetric functions and it also provided us with the algorithm of finding an expression of a symmetric polynomial in terms of $\sigma_1, ..., \sigma_n$. Another necessary and sufficient conditions for a polynomial to be symmetric were given.
Next, we defined how a finite linear matrix group acts on the vector of variables and then we examined invariance in general. We tried to generalise some of the facts about symmetric polynomials. We were particularly interested in answering a question concerning whether or not it is always the case that $k[x_1, ..., x_n]^G$ is finitely generated, what we proved to be true.
When we had proved that $k[x_1, ..., x_n]^G = k[f_1, ..., f_m]$, we discussed relations among these generators, using the ideal of relations I_F and its variety. Immediately, another important characterisations of the ring of invariants came up, namely that it is isomorphic to the quotient ring $k[y_1, ..., y_m]/I_F$ and $k\Big[\mathbf{V}(I_F)\Big]$.

Bibliography

[1] A. Clark. *Elements of Abstract Algebra*. Dover edition. Dover, USA, 1984. ISBN 0486647250,9780486647258.

[2] D. Cox, Little J., and O'Shea D. *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra*. Second Edition. Springer-Verlag New York, Inc., New York, 1996. ISBN 0-387-94680-2.

[3] W. Fulton. *Algebraic curves - An Introduction to Algebraic Geometry*. Third edition. 2008.

[4] D. Kalman. A Matrix Proof of Newton's Identities. *Mathematics Magazine*, 73(4):313–315, October 2000.

[5] M. D. Neusel and L. Smith. *Invariant Theory of Finite Groups*. Second edition. American Mathematical Society, USA, 2010. ISBN 978-0-8218-4981-1.