

INVARIANT THEORY OF FINITE GROUPS

THOMAS FANGEL

1. INTRODUCTION: SYMMETRIC POLYNOMIALS

Symmetric polynomials arise when we study the roots of a polynomial in one variable. Consider for example $f = x^3 + bx^2 + cx + d$ and let $\alpha_1, \alpha_2, \alpha_3$ be its roots. Then:

$$\begin{aligned}x^3 + bx^2 + cx + d &= (x - \alpha_1)(x - \alpha_2)(x - \alpha_3) \\ &= x^3 - (\alpha_1 + \alpha_2 + \alpha_3)x^2 + (\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3)x - \alpha_1\alpha_2\alpha_3\end{aligned}$$

And by equating coefficients we get:

$$b = -(\alpha_1 + \alpha_2 + \alpha_3), \quad c = \alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3, \quad d = \alpha_1\alpha_2\alpha_3,$$

i.e. the coefficients are polynomials in the roots, and further more these polynomials are symmetric, that is they are invariant under permutation of the $\alpha_1, \alpha_2, \alpha_3$. This motivates the following:

Definition 1.1. A polynomial $f \in k[x_1, \dots, x_n]$ is *symmetric* if $f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = f(x_1, \dots, x_n)$ for any $\sigma \in \Sigma_n$.

In this definition and in the rest of the paper k is a field. Again following the observation above let f be the polynomial in one variable X given by

$$f(X) = (X - x_1)(X - x_2) \cdots (X - x_n)$$

If we expand the righthand side we get:

$$f(X) = X^n - \sigma_1 X^{n-1} + \sigma_2 X^{n-2} + \cdots + (-1)^{n-1} \sigma_{n-1} X + (-1)^n \sigma_n,$$

where $\sigma_i \in k[x_1, \dots, x_n]$. It is not hard to write out the σ_i 's:

$$\begin{aligned}\sigma_1 &= x_1 + \cdots + x_n \\ \sigma_2 &= x_1x_2 + x_1x_3 + \cdots + x_{n-1}x_n\end{aligned}$$

In general the pattern is: for each i , we choose i of the brackets from which we do not take X as a factor and sum over all the possible ways of doing that. In general we get:

$$\sigma_r = \sum_{1 \leq i_1 < i_2 < \cdots < i_r \leq n} x_{i_1} x_{i_2} \cdots x_{i_r}$$

i.e. the sum of all monomials of order r where each x_i appears with degree 1. Since x_1, \dots, x_n are the roots of f we expect the σ_i 's to be symmetric which is of course the case: if we permute the x_i 's f does not change, and therefore the coefficients σ_i does not change.

Definition 1.2. $\sigma_i, 1 \leq i \leq n$, is the *elementary symmetric function* of degree i . When we want to stress the fact that we have n variables we write σ_i^n .

It is obvious that by forming polynomials in $\sigma_1, \dots, \sigma_n$ we get new symmetric polynomials. Surprisingly this gives us all symmetric polynomials. In other words we are now able to answer the question proposed by Niels in the beginning of the course:

Theorem 1.3. Any symmetric polynomial $f \in k[x_1, \dots, x_n]$ can be written in a unique way as a polynomial $g(\sigma_1, \dots, \sigma_n)$ in the elementary symmetric functions $\sigma_1, \dots, \sigma_n$.

The proof given here is in the spirit of the course constructive.

Proof. We use lex order with $x_1 > x_2 > \dots > x_n$. Let $f \in k[x_1, \dots, x_n]$ be non-zero and symmetric, and let $\text{LT}(f) = ax^\alpha$. If $\alpha = (\alpha_1, \dots, \alpha_n)$ we claim that $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n$. Suppose not. Then $\alpha_i < \alpha_{i+1}$ for some i . Now let $\beta = (\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \alpha_i, \alpha_{i+2}, \dots, \alpha_n)$. Since ax^α is a term of f , ax^β must be a term of $f(x_1, \dots, x_{i+1}, x_i, \dots, x_n) = f(x_1, \dots, x_n)$ since f is symmetric and thus ax^β is a term of f . By construction $\beta > \alpha$, but this is impossible since ax^α was the leading term of f .

We may therefore define a polynomial h as

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

Note that $\text{LT}(\sigma_r) = x_1 \cdots x_r$ for $1 \leq r \leq n$, hence by the calculational rules for leading terms we get:

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

So $\text{LT}(f) = \text{LT}(ah)$ and thus

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

when $f - ah \neq 0$. Now set $f_1 = f - ah$ and note that f_1 is symmetric since f and ah are. If $f_1 \neq 0$ we can repeat the process and form $f_2 = f_1 - a_1 h_1$ where a_1 is a constant and h_1 is a product of $\sigma_1, \dots, \sigma_n$ to various powers. Furthermore $\text{LT}(f_2) < \text{LT}(f_1)$ whenever $f_2 \neq 0$. Continuing we get a sequence of polynomials f, f_1, f_2, \dots with

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

Since lex order is a well-ordering the sequence must terminate at some point. But this is only possible when $f_{t+1} = 0$ for some t . Keeping track of what we have done we then see that

$$f = ah + a_1 h_1 + \dots + a_t h_t$$

which shows that f is a polynomial in the elementary symmetric functions.

Now for the uniqueness. Suppose a symmetric polynomial f can be written in two ways:

$$f = g_1(\sigma_1, \dots, \sigma_n) = g_2(\sigma_1, \dots, \sigma_n), \quad g_1, g_2 \in k[y_1, \dots, y_n]$$

If we let $g = g_1 - g_2$ then $g(\sigma_1, \dots, \sigma_n) = f - f = 0$ in $k[x_1, \dots, x_n]$. We want to show that $g = 0$ in $k[y_1, \dots, y_n]$. Write $g = \sum_{\beta} a_{\beta} y^{\beta} \Rightarrow g(\sigma_1, \dots, \sigma_n) = \sum_{\beta} a_{\beta} \sigma_1^{\beta_1} \cdots \sigma_n^{\beta_n} = \sum_{\beta} a_{\beta} g_{\beta}$, where $\beta = (\beta_1, \dots, \beta_n)$. As above we get:

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

The map

$$(\beta_1, \dots, \beta_n) \mapsto (\beta_1 + \dots + \beta_n, \dots, \beta_n)$$

is injective (given by the invertible matrix with 1's on and above the diagonal and zeros under) so the g_{β} 's have distinct leading terms. Now choose g_{β} among the summands in $g(\sigma_1, \dots, \sigma_n)$ with maximal leading term i.e. $\text{LT}(g_{\beta}) > \text{LT}(g_{\gamma})$ for any other summand g_{γ} . Then there is nothing to cancel out $\text{LT}(g_{\beta})$ and $g(\sigma_1, \dots, \sigma_n)$ cannot be zero in $k[x_1, \dots, x_n]$. This is a contradiction. \square

Since the proof gives us an algorithm for writing a symmetric polynomial as a polynomial in $\sigma_1, \dots, \sigma_n$ we better give an example:

Example 1.4. Let $f = x^2y + x^2z + xy^2 + y^2z + xz^2 + yz^2$. We find $\text{LT}(f) = x^2y = \text{LT}(\sigma_1\sigma_2)$. We calculate $\sigma_1\sigma_2$:

$$\begin{aligned}\sigma_1\sigma_2 &= (x + y + z)(xy + xz + yz) \\ &= x^2y + x^2z + xy^2 + y^2z + xz^2 + yz^2 + 3xyz\end{aligned}$$

And then $f_1 = f - \sigma_1\sigma_2 = -3xyz = -3\sigma_3$. So we are done and we get:

$$f = \sigma_1\sigma_2 - 3\sigma_3$$

Actually we need not do as above. Gröbner bases gives us a general method for checking if a polynomial is symmetric and if it is, how to write it in $\sigma_1, \dots, \sigma_n$:

Proposition 1.5. In $k[x_1, \dots, x_n, y_1, \dots, y_n]$ fix a monomial order where any monomial involving any of x_i, \dots, x_n is greater than any monomial involving only the y_i 's. Let G be a Gröbner basis for the ideal $\langle \sigma_1 - y_1, \dots, \sigma_n - y_n \rangle$. Given $f \in k[x_1, \dots, x_n]$, let $g = \bar{f}^G$ be the remainder on division by G . Then

- (i) f is symmetric iff $g \in k[y_1, \dots, y_n]$.
- (ii) If f is symmetric then $f = g(\sigma_1, \dots, \sigma_n)$ is the unique expression of f as a polynomial in $\sigma_1, \dots, \sigma_n$.

Proof. Let $G = \{g_1, \dots, g_m\}$. We can assume that all $g_i \neq 0$. Otherwise just remove those that are zero.

First suppose the remainder g on division of f by G is in $k[y_1, \dots, y_n]$. Then:

$$f = A_1g_1 + \dots + A_mg_m + g$$

for some $A_1, \dots, A_m \in k[x_1, \dots, x_n, y_1, \dots, y_n]$. If we substitute σ_i for y_i all the g_i 's go to zero since they are in the ideal $\langle \sigma_1 - y_1, \dots, \sigma_n - y_n \rangle$ and f is not changed since $f \in k[x_1, \dots, x_n]$. Therefore we get $f = g(\sigma_1, \dots, \sigma_n)$ i.e. f is symmetric.

Now suppose $f \in k[x_1, \dots, x_n]$ is symmetric. Then $f = g(\sigma_1, \dots, \sigma_n)$ for some $g \in k[y_1, \dots, y_n]$. We want to show that g is the remainder on division of f by G . First note that any g_i involves some x_j . If not $g_i(\sigma_1, \dots, \sigma_n) = 0$ in $k[x_1, \dots, x_n]$ since g_i is in the ideal $\langle \sigma_1 - y_1, \dots, \sigma_n - y_n \rangle$. But the uniqueness part of Theorem 1.3 then implies that $g_i = 0$ in $k[y_1, \dots, y_n]$ contradicting our assumption. Therefore no $\text{LT}(g_i)$ divide g by the choice of the monomial order. By Proposition 1 of chapter 2, §6 of [1] it only remains to show that $f = h + g$ for some $h \in \langle \sigma_1 - y_1, \dots, \sigma_n - y_n \rangle$. Observe that in $k[x_1, \dots, x_n, y_1, \dots, y_n]$ a monomial in $\sigma_1, \dots, \sigma_n$ can be written

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

where $B_i \in k[x_1, \dots, x_n, y_1, \dots, y_n]$. We have written out the brackets and collected the terms involving only y_1, \dots, y_n and observed that all other terms involve some $(\sigma_i - y_i)$. By collecting terms and adding coefficients this implies that

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

But then

$$f = C_1(\sigma_1 - y_1) + \dots + C_n(\sigma_n - y_n) + g(y_1, \dots, y_n)$$

and we are done since the first n terms are in the ideal.

Now part (ii) follows immediately from the above. \square

One could argue, that a drawback of the above proposition is, that we have to calculate a Gröbner basis for the ideal $\langle \sigma_1 - y_1, \dots, \sigma_n - y_n \rangle$. However this is taken care of by the

following proposition which we state without proof. Given variables u_1, \dots, u_s , let

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

be the sum of all monomials of total degree i in u_1, \dots, u_s . Then we have:

Proposition 1.6. Fix lex order on $k[x_1, \dots, x_n, y_1, \dots, y_n]$ with $x_1 > \dots > x_n > y_1 > \dots > y_n$. Then the polynomials

$$g_k = h_k(x_1, \dots, x_n) + \sum_{i=1}^k (-1)^i h_{k-i}(x_1, \dots, x_n) y_i \quad , k = 1, \dots, n$$

form a Gröbner basis for the ideal $\langle \sigma_1 - y_1, \dots, \sigma_n - y_n \rangle$.

2. FINITE MATRIX GROUPS AND THE RING OF INVARIANTS

For the rest of the paper we assume that our field k is of characteristic 0.

Now we want to construct a general setup of invariant polynomials for finite matrix groups. It turns out that the symmetric polynomials are just an example of this.

Definition 2.1. Denote by $\text{GL}_n(k)$ the group of invertible $n \times n$ matrices with entries in k . A (finite) matrix group G is a (finite) subgroup of $\text{GL}_n(k)$.

Consider the polynomial ring $k[x_1, \dots, x_n]$. We may view the n variables x_1, \dots, x_n as a vector $\mathbf{x} = (x_1, \dots, x_n)$. If $G \leq \text{GL}_n(k)$, $A \in G$ and $f(\mathbf{x})$ is a polynomial in $k[x_1, \dots, x_n]$ we may consider the polynomial $g \in k[x_1, \dots, x_n]$ given by:

$$g(\mathbf{x}) = f(A^{-1}\mathbf{x})$$

This actually introduces a group action of a matrix group on $k[x_1, \dots, x_n]$:

Definition 2.2. Let $G \leq \text{GL}_n(k)$ be a matrix group. We have an action of G on $k[x_1, \dots, x_n]$ given by:

$$(A \cdot f)(\mathbf{x}) = f(A^{-1}\mathbf{x})$$

for $A \in G, f \in k[x_1, \dots, x_n]$.

Remark 2.3. It is left to the reader to check, that this is actually a group action. In [1] $(A \cdot f)(\mathbf{x})$ is defined to be $f(A\mathbf{x})$ but this is not an action since $(AB) \cdot f = B \cdot (A \cdot f)$. This is the reason for the A^{-1} . Actually the action has a lot more structure: For an $A \in G$ we can think of $A \cdot$ as a map from $k[x_1, \dots, x_n]$ to itself and actually $A \cdot \in \text{Aut}(k[x_1, \dots, x_n])$, the automorphisms of $k[x_1, \dots, x_n]$, since we have the following properties that are almost trivial:

$$\begin{aligned} A \cdot (f + g) &= A \cdot f + A \cdot g \\ A \cdot (fg) &= (A \cdot f)(A \cdot g) \\ A^{-1} \cdot (A \cdot f) &= (A^{-1}A) \cdot f = I \cdot f = f \end{aligned}$$

This observation will prove its value below.

What we are going to study is the set of polynomials that are left invariant under this action.

Definition 2.4. Let $G \leq \text{GL}_n(k)$ be a matrix group. A polynomial $f \in k[x_1, \dots, x_n]$ is invariant under G if $A \cdot f = f$ for all $A \in G$. The set of all invariant polynomials is denoted $k[x_1, \dots, x_n]^G$.

Example 2.5. With this new notion we can now describe the symmetric polynomials as the fixed set of some matrix group. Let e_i denote the standard i 'th basis vector of k^n . Define $\phi : \Sigma_n \rightarrow \text{GL}_n(k)$ to be the map:

$$\phi(\sigma) = \begin{pmatrix} | & | & & | \\ e_{\sigma(1)} & e_{\sigma(2)} & \cdots & e_{\sigma(n)} \\ | & | & & | \end{pmatrix}$$

So ϕ permutes the columns of the identity matrix. It is not hard to see, that ϕ is a group homomorphism since $\phi(\sigma\tau) = \phi(\sigma)\phi(\tau)$, and of course ϕ is injective since only $\phi(\text{Id}) = I$. This means that $\phi : \Sigma_n \rightarrow \phi(\Sigma_n)$ is an isomorphism. We call $\phi(\Sigma_n) = S_n$, the group of matrix permutations. For convenience we set $M_\sigma = \phi(\sigma)$. Now let $\mathbf{x} \in k^n$, then

$$M_\sigma \mathbf{x} = (x_{\sigma(1)}, \dots, x_{\sigma(n)})$$

that is M_σ permutes the coordinates of \mathbf{x} . Now we see that the symmetric polynomials is nothing but the set of invariants of $k[x_1, \dots, x_n]$ under the action of S_n . The statement in Theorem 1.3 can now be written as

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

i.e. every invariant can be written as a polynomial in $\sigma_1, \dots, \sigma_n$ and furthermore this representation is unique.

This example suggests that we should pursue two questions:

- Is any set of invariants of a finite matrix group finitely generated?
- If it is, is the representation of invariants in the generators unique?

To answer these questions we need some basic observations and definitions.

Proposition 2.6. Let $G \leq \text{GL}_n(k)$ be a matrix group. Then the set $k[x_1, \dots, x_n]^G$ is a subring of $k[x_1, \dots, x_n]$.

Proof. This follows from the remark above that $A \in \text{Aut}(k[x_1, \dots, x_n])$. Because if $A \cdot f = f$, $A \cdot g = g$ then it follows that $A \cdot (f + g) = f + g$ and $A \cdot (fg) = fg$, and of course all the constant polynomials are invariant. \square

Definition 2.7. Given a set $F = \{f_1, \dots, f_m\} \subset k[x_1, \dots, x_n]$ we introduce the *evaluation homomorphism* $\text{Ev}_F : k[y_1, \dots, y_m] \rightarrow k[x_1, \dots, x_n]$ given by

$$\text{Ev}_F(g) = g(f_1, \dots, f_m)$$

That this is a homomorphism is obvious. The image of Ev_F is the set of all polynomial expressions in f_1, \dots, f_m : $\text{Im}(\text{Ev}_F) = \text{Ev}_F(k[y_1, \dots, y_m]) = k[f_1, \dots, f_m]$ and is of course a subring of $k[x_1, \dots, x_n]$ generated by F ; actually it is the smallest subring of $k[x_1, \dots, x_n]$ containing F . The reason for introducing this homomorphism is, that if a set of invariants of a given matrix group is generated by a set F then it is the image of Ev_F .

Definition 2.8. Given $G \leq \text{GL}_n(k)$, $|G| < \infty$, the *Reynolds operator* of G is the map $R_G : k[x_1, \dots, x_n] \rightarrow k[x_1, \dots, x_n]$ given by:

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

for $f \in k[x_1, \dots, x_n]$.

One may think of R_G as averaging the effect of G on f . Note that we use the fact that k has characteristic zero by dividing by $|G|1_k \neq 0$. It is almost obvious that R_G has some more structure than just being a map (otherwise we would not bother introducing it!). Naively one could think that R_G is a homomorphism, but this is not so since R_G does not preserve the

multiplicative structure of $k[x_1, \dots, x_n]$: $R_G(fg)$ is not in general equal to $R_G(f)R_G(g)$. But the additive structure is clearly preserved. The following proposition reveals the properties of R_G .

Proposition 2.9. *Let R_G be the Reynolds operator of a finite matrix group G . Then:*

- (i) R_G is $k[x_1, \dots, x_n]^G$ -linear in f , i.e. $R_G(f_1g_1 + f_2g_2) = f_1R_G(g_1) + f_2R_G(g_2)$ for $f_i \in k[x_1, \dots, x_n]^G$, $g_i \in k[x_1, \dots, x_n]$, $i = 1, 2$.
- (ii) If $f \in k[x_1, \dots, x_n]$ then $R_G(f) \in k[x_1, \dots, x_n]^G$.
- (iii) R_G acts as the identity on $k[x_1, \dots, x_n]^G$.

Proof. The additivity in (i) is trivial, it is the linearity in $k[x_1, \dots, x_n]^G$ that is non-trivial. Let $f \in k[x_1, \dots, x_n]^G$, $g \in k[x_1, \dots, x_n]$:

$$\begin{aligned} R_G(fg) &= \frac{1}{|G|1_k} \sum_{A \in G} A \cdot (fg) \\ &= \frac{1}{|G|1_k} \sum_{A \in G} (A \cdot f)(A \cdot g) \\ &= \frac{1}{|G|1_k} \sum_{A \in G} f(A \cdot g) \\ &= f \frac{1}{|G|1_k} \sum_{A \in G} A \cdot g = f R_G(g) \end{aligned}$$

Now let $f \in k[x_1, \dots, x_n]$ and $B \in G$:

$$\begin{aligned} B \cdot R_G(f) &= B \cdot \left(\frac{1}{|G|1_k} \sum_{A \in G} A \cdot f \right) \\ &= \frac{1}{|G|1_k} \sum_{A \in G} B \cdot (A \cdot f) \\ &= \frac{1}{|G|1_k} \sum_{A \in G} (BA) \cdot f \\ &= \frac{1}{|G|1_k} \sum_{A' \in G} A' \cdot f \\ &= R_G(f) \end{aligned}$$

since multiplication by B is a bijective map of G into itself, so the sum is the same just permuted in some way. This proves (ii). (iii) is actually a consequence of (i). Just let $g = 1_{k[x_1, \dots, x_n]}$ and note that $R_G(1_{k[x_1, \dots, x_n]}) = 1_{k[x_1, \dots, x_n]}$. \square

The proposition tells us how to create invariants, and actually (ii) and (iii) shows that R_G maps $k[x_1, \dots, x_n]$ onto $k[x_1, \dots, x_n]^G$. Note that $k[x_1, \dots, x_n]$ carries a natural $k[x_1, \dots, x_n]^G$ -module structure. The 3 statements in the proposition can then be rephrased as the single statement that R_G is a $k[x_1, \dots, x_n]^G$ -module homomorphism from $k[x_1, \dots, x_n]$ onto $k[x_1, \dots, x_n]^G$.

Definition 2.10. A polynomial $f \in k[x_1, \dots, x_n]$ is *homogeneous of degree m* if all terms in f has degree m , or equivalently if $f(ax) = a^m f(x)$ for any $a \in k$.

For a monomial x^α it is not hard to see, that $A \cdot x^\alpha$ is (when different from zero) a homogeneous polynomial of degree $|\alpha|$. This is because the "coordinates" of $A^{-1}x$ involve only terms in x_1, \dots, x_n of degree 1. All terms of $(A^{-1}x)^\alpha$ will then have degree $|\alpha|$. This again implies that $R_G(x^\alpha)$ is a homogeneous invariant of degree $|\alpha|$ when different from zero, and therefore $R_G(h)$ is zero or homogeneous of the same total degree as h when h is homogeneous. The following theorem due to Emily Noether answers the first of our questions.

Theorem 2.11. *Given $G \leq \text{GL}_n(k)$, $|G| < \infty$, we have*

$$k[x_1, \dots, x_n]^G = k[\{R_G(x^\alpha) : |\alpha| \leq |G|\}]$$

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

A proof of this theorem can be found in [1]. Besides the fact that the theorem gives an upper bound for the number of generators it is constructive: it gives us a set of generators for the invariant set. But there are some drawbacks: first of all this set is probably far too big i.e. there might be a much smaller set of generators. And secondly the number of monomials for which we have to compute the Reynolds operator increases rapidly. The number of monomials with degree $\leq |G|$ is $\binom{n+|G|}{|G|}$. For example consider the symmetric polynomials in 3 variables $k[x, y, z]^{S_3}$. $|S_3| = 3! = 6$ and $\binom{3+6}{6} = \binom{9}{6} = 84$ but we already know that $k[x, y, z]^{S_3}$ is generated by only 3 elements, namely the three elementary symmetric functions in three variables.

Therefore Noether's theorem is not an efficient way for calculating generators. Fortunately there are improvements of Noether's theorem. Molien's Theorem, which enables us to guess the number of linearly independent homogeneous invariants of a certain degree in advance, makes it possible to find more efficient methods of calculating generators.

Totally against the spirit of the course we will state Hilbert's nonconstructive version of the theorem above and give a nonconstructive proof:

Theorem 2.12. *Given $G \leq \text{GL}_n(k)$, $|G| < \infty$, then $k[x_1, \dots, x_n]^G$ is generated by finitely many homogeneous invariants.*

Proof. Let $I \subseteq k[x_1, \dots, x_n]$ be the ideal generated by all homogeneous invariants of positive (≥ 1) total degree. By Hilbert's Basis Theorem this ideal is generated by a finite number of polynomials g_1, \dots, g_r . But since $g_j \in I$, $g_j = \sum_{i=1}^{m_j} h_{ji} f_{ji}$ for some $h_{ji} \in k[x_1, \dots, x_n]$ and homogeneous invariants f_{ji} . If we collect all the homogeneous invariants, $f_{11}, \dots, f_{r m_r}$, in these expressions of g_1, \dots, g_r , then these homogeneous invariants will generate I since they generate a generating set. So we may assume that $I = \langle f_1, \dots, f_m \rangle$, where each f_i is a homogeneous invariant.

The strategy is to show that $k[x_1, \dots, x_n]^G = k[f_1, \dots, f_m]$. Since f_i is an invariant the inclusion $k[x_1, \dots, x_n]^G \supseteq k[f_1, \dots, f_m]$ is trivial since $A \cdot$ is a homomorphism:

$$A \cdot g(f_1, \dots, f_m) = g(A \cdot f_1, \dots, A \cdot f_m) = g(f_1, \dots, f_m)$$

Now suppose $k[x_1, \dots, x_n]^G \not\subseteq k[f_1, \dots, f_m]$. Choose a $f' \in k[x_1, \dots, x_n]^G$ with positive total degree such that $f' \notin k[f_1, \dots, f_m]$. Observe that a polynomial is invariant if and only if all of its homogeneous components are invariant. All homogeneous components of f' are therefore invariant and at least one of them is not in $k[f_1, \dots, f_m]$. Among all homogeneous invariants not in $k[f_1, \dots, f_m]$ choose one with minimal total degree $l \geq 1$. This is possible since the constant polynomials are contained in both sets. Call it f . Since f is homogeneous and invariant $f \in I$. Then $f = \sum_{i=1}^m h_i f_i$.

Suppose that for some i , $h_i f_i \neq 0$ is not homogeneous of degree l . Then h_i cannot be homogeneous since f_i is. The terms of $h_i f_i$ which are not of degree l must cancel with other $h_j f_j$ or parts of these that are not of degree l . We can therefore assume that all $h_i f_i$ are zero or homogeneous of degree l . This implies that h_i must be zero or homogeneous of degree strictly less than l since f_i has at least degree 1. By (i) of Proposition 2.9 we get

$$f = R_G(f) = \sum R_G(h_i) f_i$$

and by the observation following Definition 2.10 $R_G(h_i)$ is zero or a homogeneous invariant of the same degree as h_i . By choice of l this implies that $R_G(h_i) \in k[f_1, \dots, f_m]$ for all i . But then $f = \sum R_G(h_i) f_i \in k[f_1, \dots, f_m]$ contradicting the choice of f . This proves the theorem. \square

Now that we know, that $k[x_1, \dots, x_n]^G$ for a finite matrix group is finitely generated, one can ask how we represent an invariant in terms of the generators. As in the case of symmetric polynomials we use Gröbner bases to solve this problem. Without proof we state the following more general version of Proposition 1.5. The proof is very much the same as in the case of symmetric polynomials.

Proposition 2.13. *Suppose $f_1, \dots, f_m \in k[x_1, \dots, x_n]$ are given. Fix a monomial order in $k[x_1, \dots, x_n, y_1, \dots, y_m]$ where a monomial involving any x_i is greater than a monomial involving only the y_i 's. Let G be a Gröbner basis of $\langle f_1 - y_1, \dots, f_m - y_m \rangle$. Given $f \in k[x_1, \dots, x_n]$ let $g = \bar{f}^G$ be the remainder on division by G . Then*

- (i) $f \in k[f_1, \dots, f_m]$ if and only if $g \in k[y_1, \dots, y_m]$.
- (ii) if $f \in k[f_1, \dots, f_m]$ then $f = g(f_1, \dots, f_m)$ is an expression of f as a polynomial in f_1, \dots, f_m .

Note that (ii) says nothing about uniqueness of the expression. This leads us to the next section.

3. RELATIONS AMONG GENERATORS

We have seen, that for a finite matrix group there is a finite set of generators for the invariant set of polynomials: $k[x_1, \dots, x_n]^G = k[f_1, \dots, f_m]$. This solves the first question of the last section. The second question was that of unique representation of an invariant in terms of the generators. Suppose $f \in k[x_1, \dots, x_n]^G$, $f = g_1(f_1, \dots, f_m) = g_2(f_1, \dots, f_m)$, then $h(f_1, \dots, f_m) = 0$, where $h = g_1 - g_2$. And if $h \in k[y_1, \dots, y_m]$, $h \neq 0$ is such that $h(f_1, \dots, f_m) = 0$ then any element in $k[f_1, \dots, f_m]$ can be written as both $g(f_1, \dots, f_m)$ and $g'(f_1, \dots, f_m)$, where $g' = g + h$. It follows, that uniqueness fails if there is a nonzero polynomial $h \in k[y_1, \dots, y_m]$ such that $h(f_1, \dots, f_m) = 0$. Such a polynomial is called a non-trivial algebraic relation among f_1, \dots, f_m . If we let $F = \{f_1, \dots, f_m\}$ then

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

records all algebraic relations among f_1, \dots, f_m . The following proposition justifies that I_F is called the *ideal of relations*:

Proposition 3.1. *If $k[x_1, \dots, x_n]^G = k[f_1, \dots, f_m]$ let $I_F \subset k[y_1, \dots, y_m]$ be as above. Then:*

- (i) I_F is a prime ideal of $k[y_1, \dots, y_m]$.
- (ii) Suppose $f \in k[x_1, \dots, x_n]^G$ and $f = g(f_1, \dots, f_m)$ is one representation of f . Then any other representation of f is given by

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

where $h \in I_F$.

Proof. It is easy to see that I_F is an ideal. Now suppose $g_1, g_2 \notin I_F$. Then $g_i(f_1, \dots, f_m) \neq 0$, $i = 1, 2$. This implies that $g_1 g_2 \neq 0$ i.e. $g_1 g_2 \notin I_F$ since $k[x_1, \dots, x_n]$ is an integral domain, i.e. it has no non-trivial zero-divisors. This shows that I_F is a prime ideal.

Now suppose $g_2(f_1, \dots, f_m)$ is another representation of f . Then $h = g_2 - g \in I_F$ and

$$f = g_2(f_1, \dots, f_m) = g_2(f_1, \dots, f_m) - g(f_1, \dots, f_m) + g(f_1, \dots, f_m) = g(f_1, \dots, f_m) + h(f_1, \dots, f_m)$$

□

In the last section we introduced, when given a set $F = \{f_1, \dots, f_m\} \subset k[x_1, \dots, x_n]$, the evaluation homomorphism $\text{Ev}_F : k[y_1, \dots, y_m] \rightarrow k[x_1, \dots, x_n]$. If F is the set of generators of $k[x_1, \dots, x_n]^G$ of a finite matrix group then we see that $\text{Im}(\text{Ev}_F) = k[f_1, \dots, f_m] = k[x_1, \dots, x_n]^G$. The kernel of Ev_F is all the polynomials mapped to zero in $k[x_1, \dots, x_n]$; these

are precisely the polynomials in the ideal of relations. From the fundamental homomorphism theorem we then get this nice description of the invariant ring:

Proposition 3.2. *If $k[x_1, \dots, x_n]^G = k[f_1, \dots, f_m]$ let I_F be the ideal of relations. Then there is an isomorphism:*

$$k[y_1, \dots, y_m]/I_F \simeq k[x_1, \dots, x_n]^G$$

The ideal of relations can be found using elimination theory:

Proposition 3.3. *If $k[x_1, \dots, x_n]^G = k[f_1, \dots, f_m]$ consider the ideal*

$$J_F = \langle f_1 - y_1, \dots, f_m - y_m \rangle \subset k[x_1, \dots, x_n, y_1, \dots, y_m]$$

- (i) I_F is the n th elimination ideal of J_F , $I_F = J_F \cap k[y_1, \dots, y_m]$.
- (ii) Fix a monomial order in $k[x_1, \dots, x_n, y_1, \dots, y_m]$ as in Proposition 2.13 and let G be a Gröbner basis of J_F . Then $G \cap k[y_1, \dots, y_m]$ is a Gröbner basis for I_F in the induced monomial order on $k[y_1, \dots, y_m]$.

Proof. Observe that for $p \in k[x_1, \dots, x_n, y_1, \dots, y_m]$ we have:

$$p \in J_F \Leftrightarrow p(x_1, \dots, f_1, \dots, f_m) = 0$$

The implication to the right is trivial. The other uses the fact that:

$$\begin{aligned} p(x_1, \dots, x_n, y_1, \dots, y_m) &= p(x_1, \dots, x_n, f_1 - (f_1 - y_1), \dots, f_m - (f_m - y_m)) \\ &= p(x_1, \dots, x_n, f_1, \dots, f_m) + A_1(f_1 - y_1) + \dots + A_m(f_m - y_m) \end{aligned}$$

for some $A_i \in k[x_1, \dots, x_n, y_1, \dots, y_m]$. The last equality is obtained as in the proof of Proposition 1.5. If $p(x_1, \dots, x_n, f_1, \dots, f_m) = 0$ then $p \in \langle f_1 - y_1, \dots, f_m - y_m \rangle$.

Using this we see, that if $p \in J_F \cup k[y_1, \dots, y_m]$ then $p(f_1, \dots, f_m) = 0$ i.e. $p \in I_F$. The second part now follows from elimination theory, more precisely from the exercise 5 of chapter 3, §1 generalized version of Theorem 2 of the same section in [1]. \square

This solves the uniqueness problem. But if $I_F \neq 0$ is there then in some other sense a "unique" way of representing an invariant? Gröbner bases solve this. Given a Gröbner basis for I_F with respect to some monomial order on $k[y_1, \dots, y_m]$, and $g \in k[y_1, \dots, y_m]$ let \bar{g}^G be the remainder on division by G . We have seen in chapter 5 of [1] that the remainders \bar{g}^G uniquely represent elements of $k[y_1, \dots, y_m]/I_F$, so when we have chosen a monomial order there is in some sense a unique way of writing an invariant in terms of the generators.

Note that Proposition 3.3 gives us a way of testing whether a set of polynomials are algebraic independent. We simply calculate the ideal of relations and see if it is trivial or not.

4. THE GEOMETRY OF ORBITS

So far we have studied purely algebraic properties of invariant theory. Now for the geometric aspects.

If $G \leq \text{GL}_n(k)$, $|G| < \infty$, and $k[x_1, \dots, x_n]^G = k[f_1, \dots, f_m]$ let I_F be the ideal of relations. Denote by V_F the affine variety

$$V_F = V(I_F) \subset k^m$$

We have the following.

Proposition 4.1. *In the situation as above we have:*

- (i) V_F is the smallest variety containing the parametrization:

$$\begin{aligned} y_1 &= f_1(x_1, \dots, x_n) \\ &\vdots \\ y_m &= f_m(x_1, \dots, x_n) \end{aligned}$$

- (ii) $I_F = I(V_F)$, that is I_F is the ideal of all polynomials vanishing on V_F .
- (iii) V_F is an irreducible variety.
- (iv) Let $k[V_F]$ be the coordinate ring of V_F . Then there is an isomorphism:

$$k[V_F] \simeq k[x_1, \dots, x_n]^G$$

In the proof we use a lot of our previously deduced results on affine varieties.

Proof. First note that the stated parametrization in (i) is actually a parametrization since for any $g \in I_F$, $g(f_1, \dots, f_m) = 0$ in $k[x_1, \dots, x_n] \Rightarrow g(y_1, \dots, y_m) = 0$ with y_i as in (i). From Proposition 3.3 we know that I_F is the n th elimination ideal of $J_F = \langle f_1 - y_1, \dots, f_m - y_m \rangle$. The statement in (i) now follows from the Polynomial Implizitation Theorem (Theorem 1, chapter 3 §3 of [1]).

For (ii) note that $I_F \subset I(V(I_F)) = I(V_F)$ so one inclusion is trivial. Now suppose $h \in I(V_F)$. Given any $(a_1, \dots, a_n) \in k^n$ (i) implies that

$$(f_1(a_1, \dots, a_n), \dots, f_m(a_1, \dots, a_n)) \in V_F$$

And since h vanishes on V_F we get that

$$h(f_1(a_1, \dots, a_n), \dots, f_m(a_1, \dots, a_n)) = 0$$

Since k has characteristic 0 and hence is infinite only the zero polynomial vanishes on all of k^n . Therefore $h(f_1, \dots, f_m) = 0 \in k[x_1, \dots, x_n]$ and thus $h \in I_F$.

From (ii) and Proposition 3.1 we get that $I(V_F) = I_F$ is a prime ideal and by proposition 4 of chapter 5, §1 of [1] V_F is irreducible.

Also in chapter 5 of [1] we saw that

$$k[V_F] \simeq k[y_1, \dots, y_m]/I(V_F) = k[y_1, \dots, y_m]/I_F \simeq k[x_1, \dots, x_n]^G$$

where the last isomorphism is by Proposition 3.2. This proves (iv). □

It is clear that a generating set of $k[x_1, \dots, x_n]^G$ is by no means unique. So what happens to V_F if we change the generating set? By (iv) of the above proposition we expect them to be isomorphic. From Theorem 9, chapter 5, §4 we know that two varieties are isomorphic if and only if there is an isomorphism between their coordinate rings, which is the identity on the constant functions. It is obvious that the isomorphism in (iv) above is the identity on constant functions since the map is defined by evaluation on f_1, \dots, f_m , and constant functions are the same in $k[x_1, \dots, x_n]$ and $k[y_1, \dots, y_m]$. Suppose now that F' is another generating set. Then $k[V_F] \simeq k[x_1, \dots, x_n]^G \simeq k[V_{F'}]$ and the isomorphisms are the identity on constant functions, i.e. V_F is isomorphic to $V_{F'}$. We have proved:

Corollary 4.2. *Suppose $k[x_1, \dots, x_n]^G = k[f_1, \dots, f_m] = k[f'_1, \dots, f'_{m'}]$ then the varieties $V_F \subset k^m$ and $V_{F'} \subset k^{m'}$ are isomorphic.*

Finally we want to relate the orbit space of G to the variety V_F .

We started by letting a finite matrix group G act on $k[x_1, \dots, x_n]$, but we could also let it act on the space k^n by setting $A \cdot \mathbf{a} = A\mathbf{a}$ for $A \in G$, $\mathbf{a} = (a_1, \dots, a_n) \in k^n$.

Definition 4.3. The orbit $G \cdot \mathbf{a}$ of $\mathbf{a} \in k^n$ is the set

$$G \cdot \mathbf{a} = \{A \cdot \mathbf{a} : A \in G\}$$

The set of all G -orbits in k^n is denoted k^n/G and is called the orbit space.

It is not hard to verify that

$$\mathbf{a} \sim \mathbf{b} \Leftrightarrow \mathbf{b} = A\mathbf{a} \text{ for some } A \in G$$

defines an equivalence relation on k^n . This shows that the orbits of \mathbf{a} and \mathbf{a}' are either equal or disjoint.

Our last theorem shows that k^n/G has the structure of an affine variety in the sense that there is a bijective map from k^n/G to V_F :

Theorem 4.4. *Let $G \leq \mathrm{GL}_n(k)$, $|G| < \infty$, and k be an algebraically closed field. Suppose that $k[x_1, \dots, x_n]^G = k[f_1, \dots, f_m]$. Then*

(i) *The polynomial mapping $F : k^n \rightarrow V_F$ defined by*

$$F(a_1, \dots, a_n) = (f_1(a_1, \dots, a_n), \dots, f_m(a_1, \dots, a_n))$$

is surjective i.e. it covers all of V_F .

(ii) *The map sending the G -orbit $G \cdot \mathbf{a}$ to the point $F(\mathbf{a}) \in V_F$ is a bijective map from k^n/G to V_F .*

Note that the map in (ii) is welldefined since f_i is invariant, and hence it takes on the same value on all points of a G -orbit $G \cdot \mathbf{a}$. The proof is rather technical so we leave it out here. It can be found in [1]. Note that Corollary 4.2 implies that the variety structure of the orbit space is unique up to isomorphism. Observe that we assume that our field is algebraically closed. This is because we use the Extension Theorem to prove the surjectivity in (i). Since I_F is the n th elimination ideal of J_F a point \mathbf{b} on V_F is a partial solution to an equation. The Extension Theorem is then used to extend this to a solution and produce a point \mathbf{a} such that $F(\mathbf{a}) = \mathbf{b}$. See [1] for the details.

REFERENCES

- [1] Cox, Little, O'Shea (1997), *Ideals, Varieties and Algorithms*, Second edition, Undergraduate Texts in Mathematics, Springer-Verlag New York.
- [2] Sturmfels, Bernd (1993), *Algorithms in Invariant Theory*, Springer-Verlag Wien New York.