
Chapter 4

Permutation Representations

In this chapter we study permutation representations and their invariants. Because of their combinatorial nature there is a lot one can prove about them without any machinery from abstract algebra. This is what we exploit in this chapter. In the first section we define permutation representations. Then in Section 4.2 we prove that every symmetric polynomial can be written as a polynomial in the elementary symmetric polynomials. In Section 4.3 we generalize this result to arbitrary permutation representations.

4.1. Permutation Representations

The symmetric group on n letters Σ_n acts on the set $\{1, \dots, n\}$ by permutations. Let $V = \mathbb{F}^n$ be the n -dimensional vector space over a field \mathbb{F} . Let $\mathbf{e}_1, \dots, \mathbf{e}_n$ be a basis. Then Σ_n also acts on the set $\mathcal{S} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ by permutation.

$$\Sigma_n \times \mathcal{S} \longrightarrow \mathcal{S}, (\sigma, \mathbf{e}_i) \mapsto \mathbf{e}_{\sigma(i)}$$

for all $i = 1, \dots, n$. Thus Σ_n acts on V by permuting the basis vectors. Let $\mathbf{v} = (v_1, \dots, v_n)$; then

$$\Sigma_n \times V \longrightarrow V, (\sigma, \mathbf{v}) \mapsto (v_{\sigma(1)}, \dots, v_{\sigma(n)}).$$

In the terminology introduced in Chapter 1, this means that we have rediscovered the defining representation of Σ_n :

$$\rho : \Sigma_n \hookrightarrow \mathrm{GL}(n, \mathbb{F})$$

afforded by the matrices

$$\rho(1i) = \begin{bmatrix} 0 & & & & 1 & & & & \\ & 1 & & & & & & & \\ & & \ddots & & & & & & \\ & & & 1 & & & & & \\ 1 & & & & 0 & & & & \\ & & & & & 1 & & & \\ & & & & & & \ddots & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{bmatrix}, \quad i = 2, \dots, n,$$

where the 1's in the first row and column appear in the i th place. Since Σ_n can be generated by the transpositions $(1i)$ for $i = 2, \dots, n$, it is enough to give these images. Note that by construction every matrix in $\rho(\Sigma_n)$ has exactly one 1 in each row and column. Matrices with this property are called **permutation matrices**. Let G be an arbitrary group. We call a representation

$$\rho : G \longrightarrow \mathrm{GL}(n, \mathbb{F})$$

a **permutation representation** if for each element $g \in G$ its image $\rho(g) \in \mathrm{GL}(n, \mathbb{F})$ is a permutation matrix. Identifying the symmetric group Σ_n with its image in $\mathrm{GL}(n, \mathbb{F})$ in its defining representation allows us to rephrase this as

$$\rho(g) \in \Sigma_n \quad \forall g \in G.$$

Example 4.1. Take the cyclic group of order 4, $\mathbb{Z}/4$, with generator $3 = -1$. Then

$$\rho : \mathbb{Z}/4 \longrightarrow \mathrm{GL}(4, \mathbb{F}), \quad 3 \mapsto \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

defines a faithful permutation representation of $\mathbb{Z}/4$.

Indeed, every finite group G has a permutation representation because G can be viewed as a subgroup of some symmetric group. This can be seen as follows. Let $|G| = n$. Then enumerate the elements of G , so that we have $G = \{g_1, \dots, g_n\}$. Take an arbitrary element $g \in G$. Then for every $i = 1, \dots, n$ there is a $g(i) \in \{1, \dots, n\}$ such that

$$gg_i = g_{g(i)}.$$

Moreover,

$$gg_i = gg_j \iff i = j.$$

Thus G acts by permutation on its underlying set $\{g_1, \dots, g_n\}$. We identify the g_i 's with the basis vectors of an n -dimensional vector space V , then group multiplication leads to a permutation of the basis vectors. The representation of G that is obtained in this way is called the **regular representation of G** . It is probably the most important representation of G since it encodes all other representations in a way we will not explain here, since this would lead us too far off the path. However, this is an interesting topic and you should definitely learn about this!

The representation in Example 4.1 is the regular representation of $\mathbb{Z}/4$: The group $\mathbb{Z}/4$ consists of $0, 1, 2, 3$. We identify the group elements with basis elements of a vector space

$$\{0, \dots, 3\} \longrightarrow \{\mathbf{e}_1, \dots, \mathbf{e}_4\}, \quad i \mapsto \mathbf{e}_{i+1}.$$

The group is generated by $-1 = 3$. Since $(i) - 1 = i - 1$, this translates into

$$\mathbf{e}_{i+1} \overset{+3}{\rightsquigarrow} \mathbf{e}_i$$

for the basis vectors. So, the group permutes cyclicly the basis vectors. The matrix above describes precisely this: a cyclic permutation of the basis vectors

$$\mathbf{e}_4 \rightsquigarrow \mathbf{e}_3 \rightsquigarrow \mathbf{e}_2 \rightsquigarrow \mathbf{e}_1 \rightsquigarrow \mathbf{e}_4.$$

Thus the above Example 4.1 is the regular representation of $\mathbb{Z}/4$.

Example 4.2 (Regular Representation of \mathbb{Z}/n). Take the cyclic group of order n . Then

$$\rho : \mathbb{Z}/n \longrightarrow \mathrm{GL}(n, \mathbb{F}), \quad n-1 \mapsto \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ 1 & & & & 0 \end{bmatrix}$$

defines a faithful permutation representation of \mathbb{Z}/n . This matrix describes a cyclic permutation of the basis vectors

$$\mathbf{e}_n \rightsquigarrow \mathbf{e}_{n-1} \rightsquigarrow \cdots \rightsquigarrow \mathbf{e}_2 \rightsquigarrow \mathbf{e}_1 \rightsquigarrow \mathbf{e}_n.$$

Thus this is the regular representation of \mathbb{Z}/n .

Before we proceed, let us give another example:

Example 4.3 (Regular Representation of Σ_3). The symmetric group in three letters has six elements. We enumerate them in the following way:

$$g_1 = (1), g_2 = (12), g_3 = (13), g_4 = (23), g_5 = (123), \text{ and } g_6 = (132).$$

Then the regular representation ρ of Σ_3 is given by the matrices

$$\rho(g_1) = \mathbf{I}, \quad \rho(g_2) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix},$$

$$\rho(g_3) = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}, \quad \rho(g_4) = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\rho(g_5) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \text{ and } \rho(g_6) = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Remark 4.4. It does not matter in which order we enumerate our group elements to obtain the regular representation, because a change of the order translates into a permutation of the basis. Thus we receive an equivalent representation. This justifies speaking about *the* regular representation.

Note that the degree of the regular representation always coincides with the order of the group. Note also that not every permutation representation is the regular representation.

Example 4.5. The defining representation of Σ_n is, of course, a permutation representation of degree n . However, it is not the regular representation: this has degree $n!$.

4.2. Newton, Waring, and Gauss

In this section we look at the polynomial invariants of the defining representation of the symmetric group on n letters Σ_n .

Consider the ring $\mathbb{F}[V] = \mathbb{F}[x_1, \dots, x_n]$ of polynomials in n variables. We call the sum

$$s_1 = x_1 + \cdots + x_n \in \mathbb{F}[V]$$

the **first elementary symmetric polynomial** in x_1, \dots, x_n . More generally we define the **i th elementary symmetric polynomial** as

$$s_i = \sum_I \mathbf{x}^I,$$

where the sum runs over all exponent sequences $I = (i_1, \dots, i_n)$ consisting of i 1's and zeros otherwise. These are also sometimes called **Viète polynomials**. For example

$$s_2 = x_1x_2 + x_1x_3 + \cdots + x_1x_n + x_2x_3 + \cdots + x_{n-1}x_n$$

and

$$s_n = x_1 \cdots x_n.$$

As we saw in the previous section the defining representation of Σ_n induces an action of Σ_n on the n -dimensional vector space $V = \mathbb{F}^n$ by permuting the basis elements $\mathbf{e}_1, \dots, \mathbf{e}_n$. For all $\sigma \in \Sigma_n$ we have

$$\sigma(\mathbf{e}_i) = \mathbf{e}_{\sigma(i)} \quad \forall i = 1, \dots, n,$$

and by linear extension

$$\sigma(\mathbf{v}) = \sigma(v_1, \dots, v_n) = (v_{\sigma(1)}, \dots, v_{\sigma(n)}) \quad \forall \mathbf{v} \in V.$$

Thus the induced action on the dual space V^* also permutes the basis elements x_1, \dots, x_n . For $\sigma \in \Sigma_n$ we have

$$\begin{aligned} \sigma(x_i(\mathbf{v})) &= x_i(\sigma^{-1}(\mathbf{v})) = x_i(v_{\sigma^{-1}(1)}, \dots, v_{\sigma^{-1}(n)}) \\ &= v_{\sigma^{-1}(i)} = x_{\sigma^{-1}(i)}(\mathbf{v}) \end{aligned}$$

for all $i = 1, \dots, n$. Hence for a monomial $x_1^{i_1} \cdots x_n^{i_n} \in \mathbb{F}[V]$ we obtain

$$\sigma(x_1^{i_1} \cdots x_n^{i_n}) = x_{\sigma^{-1}(1)}^{i_1} \cdots x_{\sigma^{-1}(n)}^{i_n},$$

another monomial of the same degree with permuted indices. Extending this linearly to all polynomials gives for $f = \sum a_I \mathbf{x}^I$

$$\sigma(f) = \sum a_I \sigma(\mathbf{x}^I) = \sum a_I \sigma(x_1^{i_1} \cdots x_n^{i_n}) = \sum a_I x_{\sigma^{-1}(1)}^{i_1} \cdots x_{\sigma^{-1}(n)}^{i_n}.$$

We apply an element $\sigma \in \Sigma_n$ to an elementary symmetric polynomial

$$\begin{aligned} \sigma(s_i) &= \sigma(x_1 x_2 \cdots x_i + x_1 x_2 \cdots x_{i-1} x_{i+1} + \cdots + x_{n-i+1} \cdots x_n) \\ &= x_{\sigma^{-1}(1)} x_{\sigma^{-1}(2)} \cdots x_{\sigma^{-1}(i)} + \cdots + x_{\sigma^{-1}(n-i+1)} \cdots x_{\sigma^{-1}(n)}. \end{aligned}$$

Thus the only thing that has changed is the order of summation, i.e., $\sigma(s_i) = s_i$ is a polynomial invariant under Σ_n . An arbitrary polynomial invariant under Σ_n is called **symmetric**.

Example 4.6 (Symmetric Power Sums). The symmetric power sums are defined by

$$p_i = x_1^i + \cdots + x_n^i.$$

They are symmetric because any element in Σ_n permutes the summands.

The following formulae go back to Sir Isaac Newton [1646-1723].

Proposition 4.7 (Newton's Formulae). *The symmetric power sums and elementary symmetric functions satisfy the following relations:*

$$\begin{aligned} s_1 &= p_1, \\ 2s_2 &= p_1s_1 - p_2, \\ 3s_3 &= p_1s_2 - p_2s_1 + p_3, \\ &\dots \\ is_i &= \sum_{k=1}^i (-1)^{k-1} p_k s_{i-k}, \end{aligned}$$

where we set $s_0 = 1$.

Proof. We consider the following polynomial in an indeterminate t and with coefficients from $\mathbb{F}[V]$:

$$(*) \quad \Omega(t) = \prod_{i=1}^n (1 - x_i t) \in \mathbb{F}[V][t].$$

Its logarithmic derivative with respect to t is

$$\begin{aligned} -\frac{\frac{\partial}{\partial t}\Omega(t)}{\Omega(t)} &= \frac{1}{\Omega(t)} \sum_{i=1}^n x_i \left(\prod_{j=1, j \neq i}^n (1 - x_j t) \right) \\ &= \sum_{i=1}^n \frac{x_i}{1 - x_i t} = p_1 + p_2 t + p_3 t^2 + \dots, \end{aligned}$$

where the last equality holds because

$$\frac{x_i}{1 - x_i t} = x_i + x_i^2 t + x_i^3 t^2 + x_i^4 t^3 + \dots.$$

Thus we obtain

$$(1) \quad \frac{\partial}{\partial t}\Omega(t) = -\Omega(t)(p_1 + p_2 t + p_3 t^2 + \dots).$$

On the other hand, we can expand the product in equation (*) and get

$$\Omega(t) = \sum_{i=0}^n (-1)^i s_i t^i.$$

Therefore

$$(2) \quad \frac{\partial}{\partial t}\Omega(t) = \sum_{i=0}^n (-1)^i i s_i t^{i-1},$$

and we obtain the result by equating coefficients of (1) and (2). \square

These formulae express *implicitly* the symmetric power sums in terms of the elementary symmetric functions. A little bit later Edward Waring [1734-1798] gave a formula expressing the p_i 's *explicitly* in terms of the elementary symmetric functions. More generally, he gave formulae expressing *any* symmetric polynomial in terms of the elementary symmetric ones. This gives the elementary symmetric functions their significance: any symmetric function is a polynomial in the s_i 's. We want to prove this result, but instead of giving Waring's (very long) formulae, we present the algorithmic proof due to Carl Friedrich Gauss [1777-1855]. For this we need some preliminary steps.

Example 4.8. Consider $\rho : \mathbb{Z}/2 \longrightarrow \mathrm{GL}(4, \mathbb{F})$ with

$$\rho(1) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

The dual space V^* is the disjoint union of $\{\mathbf{0}\}$ and the orbits of the following nine types:

$$\begin{aligned} o[\lambda x_1] &= \{\lambda x_1, \lambda x_2\}, \\ o[\lambda x_3] &= \{\lambda x_3, \lambda x_4\}, \\ o[\lambda x_1 + \mu x_2] &= \{\lambda x_1 + \mu x_2, \lambda x_2 + \mu x_1\}, \\ o[\lambda x_1 + \mu x_3] &= \{\lambda x_1 + \mu x_3, \lambda x_2 + \mu x_4\}, \\ o[\lambda x_1 + \mu x_4] &= \{\lambda x_1 + \mu x_4, \lambda x_2 + \mu x_3\}, \\ o[\lambda x_3 + \mu x_4] &= \{\lambda x_3 + \mu x_4, \lambda x_4 + \mu x_3\}, \\ o[\lambda x_1 + \mu x_2 + \nu x_3] &= \{\lambda x_1 + \mu x_2 + \nu x_3, \lambda x_2 + \mu x_1 + \nu x_4\}, \\ o[\lambda x_1 + \mu x_3 + \nu x_4] &= \{\lambda x_1 + \mu x_3 + \nu x_4, \lambda x_2 + \mu x_4 + \nu x_3\}, \end{aligned}$$

and

$$\begin{aligned} o[\lambda x_1 + \mu x_2 + \nu x_3 + \omega x_4] &= \{\lambda x_1 + \mu x_2 + \nu x_3 + \omega x_4, \\ &\quad \lambda x_2 + \mu x_1 + \nu x_4 + \omega x_3\}, \end{aligned}$$

for nonzero field elements λ, μ, ν , and ω . In particular, the set of basis elements consists of two orbits:

$$\{x_1, \dots, x_4\} = \{x_1, x_2\} \cup \{x_3, x_4\}.$$

Thus the polynomials $x_1 + x_2$ and $x_3 + x_4$ are also invariant.

The two polynomials we constructed in the preceding example are **orbit sums**. They are what you think they are: Take any monomial $\mathbf{x}^I \in \mathbb{F}[V]$. Then $g\mathbf{x}^I \in \mathbb{F}[V]$ for $g \in G$ is also a monomial, since we are dealing with permutation representations. Thus the orbit of \mathbf{x}^I is the set of monomials

$$o[\mathbf{x}^I] = \{g\mathbf{x}^I \mid g \in G\}.$$

The orbit sum is the sum of all orbit elements

$$o(\mathbf{x}^I) = \sum g\mathbf{x}^I.$$

It is, by construction, an invariant polynomial. This is our first method to construct invariant polynomials. As we see below, in the case of permutation invariants this suffices to find all invariant polynomials.

Before we proceed note that in general

$$o(\mathbf{x}^I) = \sum_{g \in G/\text{Stab}_G(\mathbf{x}^I)} g\mathbf{x}^I \neq \sum_{g \in G} g\mathbf{x}^I$$

as we illustrate with the next example.

Example 4.9. Reconsider Example 4.8. The orbit of x_1x_2 consists of one element, because it is invariant. Thus $o(x_1x_2) = x_1x_2$. However,

$$\sum_{g \in \mathbb{Z}/2} g(x_1x_2) = 2x_1x_2.$$

The elementary symmetric polynomials are nothing but orbit sums (under the action of the full symmetric group)

$$s_i = o(x_1 \cdots x_i) \quad \forall i = 1, \dots, n.$$

Indeed in general we have the following result.

Proposition 4.10. *Permutation invariants are sums of orbit sums.*

Proof. Let $\rho : G \hookrightarrow \mathrm{GL}(n, \mathbb{F})$ be a permutation representation. Let $f \in \mathbb{F}[V]^G$ be a homogeneous invariant of degree d . Without loss of generality we assume that f is not the sum of two other invariant polynomials. We need to show that f is the orbit sum of one of its terms. Set

$$f = \sum_I a_I \mathbf{x}^I, \quad \text{with } a_I \neq 0.$$

Observe that

$$g(\mathbf{x}^I) = g(x_1^{i_1} \cdots x_n^{i_n}) = x_{g^{-1}(1)}^{i_1} \cdots x_{g^{-1}(n)}^{i_n} = x_1^{i_{g^{-1}(1)}} \cdots x_n^{i_{g^{-1}(n)}}.$$

Thus the action of G can be interpreted as a permutation of the exponent sequence:

$$g(\mathbf{x}^I) = \mathbf{x}^{g^{-1}(I)}.$$

Thus for any $g \in G$ we have that

$$gf = \sum_I a_I g(\mathbf{x}^I) = \sum_I a_I \mathbf{x}^{g^{-1}(I)} = \sum_I a_{g(I)} \mathbf{x}^I.$$

Therefore we obtain

$$0 = f - gf = \sum_I a_I \mathbf{x}^I - \sum_I a_{g(I)} \mathbf{x}^I.$$

Since the set of monomials of degree d forms a basis of the vector space of all polynomials of degree d , they are linearly independent. Moreover, the coefficients $a_I \neq 0$ by assumption. Thus every term $a_I \mathbf{x}^I$ in the first sum is cancelled by the term $a_{g(I)} \mathbf{x}^I$, i.e.,

$$a_I = a_{g(I)}$$

as claimed. \square

Example 4.11. We come back to Example 4.8 and consider monomials of degree two. The second elementary symmetric polynomial is

$$\begin{aligned} s_2 &= x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4 \\ &= (x_1x_2) + (x_1x_3 + x_2x_4) + (x_1x_4 + x_2x_3) + (x_3x_4). \end{aligned}$$

We set the brackets around expressions that are $\mathbb{Z}/2$ -orbit sums. We see that s_2 is the sum of four $\mathbb{Z}/2$ -invariants.

Recall from Exercise 3 of Chapter 3 that the set of homogeneous invariants of degree d , $\mathbb{F}[V]_{(d)}^G$, forms an \mathbb{F} -vector space, $d \in \mathbb{N}_0$.

Proposition 4.12. *The orbit sums of any given degree d form a basis of the vector space $\mathbb{F}[V]_{(d)}^G$.*

Proof. In the preceding Proposition 4.10 we have seen that every invariant polynomial is a sum of orbit sums. Thus the orbit sums of degree d form a generating set of $\mathbb{F}[V]_{(d)}^G$. To show that they are linearly independent, assume that there is a linear relation

$$(*) \quad a_1 o_1 + \cdots + a_n o_n = 0,$$

where o_1, \dots, o_n are orbit sums of degree d and $a_1, \dots, a_n \in \mathbb{F}$. Expand $(*)$ as a sum of distinct monomials of degree d . Since those are linearly independent and the orbits are disjoint, the coefficients a_1, \dots, a_n vanish as desired. \square

In other words, we have proven that rings of permutation invariants are generated by orbit sums of monomials. With this at hand we can proceed to Gauss' proof.

We say a monomial \mathbf{x}^I is smaller in **lexicographic order** than \mathbf{x}^J , written as

$$\mathbf{x}^I <_{\text{lex}} \mathbf{x}^J,$$

if the first nonzero entry in the difference of the exponent sequences

$$J - I = (j_1 - i_1, \dots, j_n - i_n)$$

is positive.

Theorem 4.13 (Ring of Invariants of the Defining Representation of Σ_n). *Any polynomial invariant under the defining representation of the symmetric group on n letters can be written as a polynomial in the elementary symmetric polynomials.*

Proof. Let f be a homogeneous symmetric polynomial. If f is divisible by s_n , then

$$f = s_n f'$$

for some symmetric polynomial f' . Thus without loss of generality we assume that s_n does not divide f .

We want to proceed by induction on the lexicographic order. If $f \in \mathbb{F}$, then there is nothing to show. Thus our induction starts. Let

ax^I , $a \in \mathbb{F} \setminus 0$, be the leading (i.e., highest) term of f with respect to the lexicographic order. We claim that

$$i_n \leq i_{n-1} \leq \cdots \leq i_2 \leq i_1.$$

Assume to the contrary that there is an index $k \in \{1, \dots, n\}$ such that $i_k < i_{k+1}$. Then for $\tau = (k \ k+1) \in \Sigma_n$, the term

$$\tau(m) = \tau(ax_1^{i_1} \cdots x_n^{i_n}) = ax_1^{i_1} \cdots x_k^{i_{k+1}} x_{k+1}^{i_k} \cdots x_n^{i_n}$$

has higher lexicographic order. Since this also appears in f by Proposition 4.12, we obtain a contradiction.

Next observe that the symmetric polynomial

$$as_1^{i_1-i_2} s_2^{i_2-i_3} \cdots s_n^{i_n}$$

has the same leading term as f . Thus

$$f - as_1^{i_1-i_2} s_2^{i_2-i_3} \cdots s_n^{i_n}$$

is a symmetric polynomial with lower highest term. So we are done by induction. \square

Thus we can write our ring of polynomial invariants as

$$\mathbb{F}[V]^{\Sigma_n} = \mathbb{F}[s_1, \dots, s_n].$$

A set of polynomials is called **algebraically independent** if there are no algebraic relations between them.

Example 4.14. There are no algebraic relations between the variables $x_1, \dots, x_n \in \mathbb{F}[V]$. For otherwise there would be a polynomial expression

$$f(x_1, \dots, x_n) \equiv 0.$$

But this means that f is the zero polynomial. There is also no algebraic relation between x^2 and y^2 in $\mathbb{F}[x, y]$.

Indeed there are no algebraic relations among the elementary symmetric functions.

Proposition 4.15. *The elementary symmetric polynomials are algebraically independent.*

Proof. Suppose there is an algebraic relation between the elementary symmetric functions. Let $f(X_1, \dots, X_n) \in \mathbb{F}[X_1, \dots, X_n]$ be a polynomial such that

$$f(s_1, \dots, s_n) \equiv 0.$$

We have to show that f is the zero polynomial. Rewrite $f(s_1, \dots, s_n)$ as a polynomial in the x_i 's. Choose the term $a\mathbf{X}^I$ of f of highest lexicographic order. Thus $f(s_1, \dots, s_n)$, regarded as a polynomial in the x_i 's, has leading term

$$ax_1^{i_1+\dots+i_n} x_2^{i_2+\dots+i_n} \dots x_n^{i_n},$$

which, by assumption, is zero. Thus $a = 0$ as desired. \square

Remark 4.16. The issue here is that this means that every symmetric polynomial can be *uniquely* written as a polynomial in the s_1, \dots, s_n : Let $p \in \mathbb{F}[V]^{\Sigma_n}$ be an arbitrary invariant. Assume that p can be written in two ways as a polynomial in the elementary symmetric polynomials. Then

$$p = f_1(s_1, \dots, s_n) = f_2(s_1, \dots, s_n)$$

for some $f_1, f_2 \in \mathbb{F}[X_1, \dots, X_n]$. Thus

$$0 = f_1(s_1, \dots, s_n) - f_2(s_1, \dots, s_n),$$

which in turn means that $0 = f_1 - f_2$ by the preceding Proposition 4.15. Thus f_1 and f_2 are the same polynomials.

Remark 4.17. The generators of a ring of invariants are in general not algebraically independent; cf. Exercise 13 in this chapter.

4.3. Göbel's Bound

In this section we want to find a complete set of generators for an arbitrary permutation representation

$$\rho : G \hookrightarrow \mathrm{GL}(n, \mathbb{F}).$$

This result was proven by Manfred Göbel in his PhD thesis.¹ The proof we present here is based on a reworking of the original proof

¹Manfred Göbel, *Computing Bases for Permutation-Invariant Polynomials*, PhD thesis, University of Tübingen, Germany, 1996.

by Nelson Killius, who was at that time a first year graduate student visiting Northwestern University.

We note that $\rho(G) \leq \Sigma_n \leq \text{GL}(n, \mathbb{F})$ (where we identified the symmetric group with its image under its defining representation). In particular, this means that every polynomial that is invariant under the full symmetric group is invariant under the smaller group G as well.

$$\mathbb{F}[V]^{\Sigma_n} = \mathbb{F}[s_1, \dots, s_n] \hookrightarrow \mathbb{F}[V]^G \hookrightarrow \mathbb{F}[V].$$

Thus the elementary symmetric polynomials are present in every ring of permutation invariants. In other words, they form a **universal set of invariants** for the class of permutation invariants.

We know already that $\mathbb{F}[V]^G$ is generated by orbit sums by Proposition 4.12. Our next goal is to describe more precisely the set of orbit sums we actually need. For that we introduce more terminology.

Take an exponent sequence $I = (i_1, \dots, i_n)$ with $i_1, \dots, i_n \in \mathbb{N}_0$, and rearrange it in weakly decreasing order. We obtain an n -tuple $(\lambda_1(I), \dots, \lambda_n(I))$ that is a permutation of the original (i_1, \dots, i_n) . The underlying ordered set

$$\lambda(I) = (\lambda_1(I) \geq \lambda_2(I) \geq \dots \geq \lambda_n(I))$$

is called the **associated partition of I** .

We call a monomial \mathbf{x}^I **special** if the associated partition $\lambda(I)$ satisfies

- (i) $\lambda_i(I) - \lambda_{i+1}(I) \leq 1$ for all $i = 1, \dots, n-1$, and
- (ii) $\lambda_n(I) = 0$.

Example 4.18. The monomial $x_1^2 x_2^3 x_3^1 \in \mathbb{F}[x_1, x_2, x_3]$ is not special, because the associated partition $(3 \geq 2 \geq 1)$ does not satisfy property (ii). However, considered as an element in $\mathbb{F}[x_1, \dots, x_4]$ it is special. The associated partition is $(3 \geq 2 \geq 1 \geq 0)$. The monomial $x_1^1 x_2^4 x_3^1$ is not special in $\mathbb{F}[x_1, x_2, x_3]$ nor in $\mathbb{F}[x_1, \dots, x_4]$. The respective associated partitions are $(4 \geq 1 \geq 1)$ and $(4 \geq 1 \geq 1 \geq 0)$. In both cases the first condition for special monomials is not satisfied: $4 - 1 > 1$, i.e., there is a “gap”.

Since the n -tuples of the associated partitions are permutations of the original exponent sequence, two exponent sequences J and K

are permutations of each other if and only if the associated partitions are equal: $\lambda(J) = \lambda(K)$.

We want to prove that the ring of invariants of a permutation representation is generated as an algebra by the orbit sums of special monomials and the elementary symmetric polynomial² $s_n = x_1 \cdots x_n$. Thus we need to show that any permutation invariant can be rewritten as a polynomial in the orbit sums of special monomials and s_n .

The proof relies on the same idea as Gauss' proof for the Σ_n -invariants: we give an algorithm that rewrites an arbitrary invariant in the desired form. For this algorithm to work (i.e., to terminate) we need to choose a suitable order on the set of monomials.

Let $I = (i_1, \dots, i_n)$ and $J = (j_1, \dots, j_n)$ be n -tuples of nonnegative integers. Then we say that the monomial \mathbf{x}^I is smaller in the **dominance order** than \mathbf{x}^J , denoted by $\mathbf{x}^I \leq_{\text{dom}} \mathbf{x}^J$, if the associated partitions satisfy

$$\begin{aligned} \lambda_1(I) &\leq \lambda_1(J), \\ \lambda_1(I) + \lambda_2(I) &\leq \lambda_1(J) + \lambda_2(J), \\ &\vdots \\ \lambda_1(I) + \cdots + \lambda_n(I) &\leq \lambda_1(J) + \cdots + \lambda_n(J). \end{aligned}$$

If $\mathbf{x}^I \leq_{\text{dom}} \mathbf{x}^J$ and $\mathbf{x}^J \leq_{\text{dom}} \mathbf{x}^I$ we say $\mathbf{x}^I =_{\text{dom}} \mathbf{x}^J$. (Note that $\mathbf{x}^I =_{\text{dom}} \mathbf{x}^J$ if and only if I and J are permutations of each other; see Exercise 8 in this chapter.)

We explain next how we rewrite the orbit sum of an arbitrary monomial in terms of orbit sums of special monomials and s_n .

Let \mathbf{x}^I be a monomial. Let $\lambda(I)$ be the associated partition. Assume that \mathbf{x}^I does not satisfy the first condition for being special. Then there is a "gap" in the associated partition. We denote by

$$t_I = \min\{i \mid \lambda_i(I) - \lambda_{i+1}(I) > 1\}$$

the first occurrence of such a gap. We define the **reduced monomial of \mathbf{x}^I** to be $\mathbf{x}^{I_{\text{red}}}$, and denote it by

$$\lambda(I_{\text{red}}) = (\lambda_1(I) - 1, \dots, \lambda_{t_I}(I) - 1, \lambda_{t_I+1}(I), \dots, \lambda_n(I)).$$

²Note that s_n is not special. However, all other elementary symmetric polynomials are (sums of) orbit sums of special monomials.

Thus we obtained the exponent sequence I_{red} from I by lowering the t_I largest exponents by 1. Note that the gap in the new partition is smaller,

$$\lambda_{t_I}(I_{\text{red}}) - \lambda_{t_I+1}(I_{\text{red}}) = \lambda_{t_I}(I) - \lambda_{t_I+1}(I) - 1,$$

thus the reduced monomial $\mathbf{x}^{I_{\text{red}}}$ is “closer” to being special than \mathbf{x}^I was.

Example 4.19. Let us reconsider the example $x_1^1 x_2^4 x_3^1 \in \mathbb{F}[x_1, \dots, x_4]$. We have $t_I = 1$. The reduced monomial is $x_1^1 x_2^3 x_3^1$ with associated partition $(3 \geq 1 \geq 1 \geq 0)$. This is still not special. However, if we reduce this again (again $t_{I_{\text{red}}} = 1$) we obtain the special monomial $x_1^1 x_2^2 x_3^1$ with associated partition $(2 \geq 1 \geq 1 \geq 0)$.

We want to use this reduction method in our algorithm. So that this can work we need to show that this reduction method leads to monomials that are lower with respect to the dominance order than the original one. This is the content of the next proposition.

Proposition 4.20. *Let $I = (i_1, \dots, i_n)$ be an n -tuple of nonnegative integers. Any monomial \mathbf{x}^J occurring in $o(\mathbf{x}^{I_{\text{red}}}) \cdot s_{t_I}$ is lower in the dominance order than \mathbf{x}^I . We have equality if and only if \mathbf{x}^J is a term in $o(\mathbf{x}^I)$.*

Proof. Denote by $t = t_I$ the first occurrence of a gap in the partition of I . Let \mathbf{x}^J be a monomial occurring in $o(\mathbf{x}^{I_{\text{red}}}) \cdot s_t$. Let $I_{\text{red}} = (i_1^{\text{red}}, \dots, i_n^{\text{red}})$. We have

$$\mathbf{x}^J = \mathbf{x}^K x_{l_1} \cdots x_{l_t},$$

where

$$(\star) \quad x_1^{k_1} \cdots x_n^{k_n} = \mathbf{x}^K = g(\mathbf{x}^{I_{\text{red}}}) = x_1^{i_1^{\text{red}} g^{-1}(1)} \cdots x_n^{i_n^{\text{red}} g^{-1}(n)}$$

for some element $g \in G$. Thus K is a permutation of I_{red} , and $\lambda(K) = \lambda(I_{\text{red}})$. Moreover, note that

$$(\ast) \quad j_l = \begin{cases} k_l & \text{if } l \notin \{l_1, \dots, l_t\}, \\ k_l + 1 & \text{if } l \in \{l_1, \dots, l_t\}. \end{cases}$$

We express the associated partition $\lambda(J) = (\lambda_1(J) \geq \cdots \geq \lambda_n(J))$ in terms of the original partition $\lambda(I) = (\lambda_1(I) \geq \cdots \geq \lambda_n(I))$ as follows.

Set $\mathcal{L} = \{l_1, \dots, l_t\} \subseteq \{1, \dots, n\}$. Denote by $\chi_{\mathcal{L}}$ the **characteristic function of the subset \mathcal{L}** , i.e.,

$$\chi_{\mathcal{L}} : \{1, \dots, n\} \longrightarrow \{0, 1\}, \quad \chi_{\mathcal{L}}(i) = \begin{cases} 1 & \text{if } i \in \mathcal{L}, \\ 0 & \text{if } i \in \{1, \dots, n\} \setminus \mathcal{L}. \end{cases}$$

By (\star) we have

$$k_i = i_{g^{-1}(i)}^{\text{red}}, \quad i = 1, \dots, n,$$

for some $g \in G$. Thus (\star) says that

$$\begin{aligned} j_i &= \begin{cases} i_{g^{-1}(i)}^{\text{red}} & \text{if } i \notin \{l_1, \dots, l_t\}, \\ i_{g^{-1}(i)}^{\text{red}} + 1 & \text{if } i \in \{l_1, \dots, l_t\} \end{cases} \\ &= i_{g^{-1}(i)}^{\text{red}} + \chi_{\mathcal{L}}(i). \end{aligned}$$

This means the following for the associated partition $\lambda(J)$:

$$(\dagger) \quad \lambda_i(J) = \begin{cases} \lambda_{g^{-1}(i)}(I) - 1 + \chi_{\mathcal{L}}(i) & \text{if } 1 \leq i \leq t, \\ \lambda_{g^{-1}(i)}(I) + \chi_{\mathcal{L}}(i) & \text{if } t+1 \leq i \leq n. \end{cases}$$

Next, in order to show that $J \leq_{\text{dom}} I$, we need to show that

$$\sum_{i=1}^s \lambda_i(J) \leq \sum_{i=1}^s \lambda_i(I)$$

for all $s = 1, \dots, n$. Assume that $s \leq t$. We have

$$\begin{aligned} \sum_{i=1}^s \lambda_i(J) &\stackrel{(1)}{=} \sum_{i=1}^s (\lambda_{g^{-1}(i)}(I) - 1 + \chi_{\mathcal{L}}(i)) \\ &= \left(\sum_{i=1}^s \lambda_{g^{-1}(i)}(I) \right) - s + \sum_{i=1}^s \chi_{\mathcal{L}}(i) \\ &\stackrel{(2)}{\leq} \sum_{i=1}^s \lambda_{g^{-1}(i)}(I) \stackrel{(3)}{\leq} \sum_{i=1}^s \lambda_i(I), \end{aligned}$$

where (1) is true because of (\dagger) , (2) is true because $\sum_{i=1}^s \chi_{\mathcal{L}}(i) \leq s$, and (3) is true because reordering of $\lambda(I) = (\lambda_1(I) \geq \dots \geq \lambda_n(I))$ can only lead to a smaller sum for the first s terms.

If $s > t$, then with (†) we obtain

$$\begin{aligned}
\sum_{i=1}^s \lambda_i(J) &= \sum_{i=1}^t \lambda_i(J) + \sum_{i=t+1}^s \lambda_i(J) \\
&= \sum_{i=1}^t \left(\lambda_{g^{-1}(i)}(I) - 1 + \chi_{\mathcal{L}}(i) \right) + \sum_{i=t+1}^s \left(\lambda_{g^{-1}(i)}(I) + \chi_{\mathcal{L}}(i) \right) \\
&= \left(\sum_{i=1}^t \lambda_{g^{-1}(i)}(I) \right) - t + \sum_{i=1}^s \chi_{\mathcal{L}}(i) + \left(\sum_{i=t+1}^s \lambda_{g^{-1}(i)}(I) \right) \\
&= \sum_{i=1}^s \lambda_{g^{-1}(i)}(I) - t + \sum_{i=1}^s \chi_{\mathcal{L}}(i).
\end{aligned}$$

Therefore

$$\sum_{i=1}^s \lambda_i(J) = \sum_{i=1}^s \lambda_{g^{-1}(i)}(I) - t + \sum_{i=1}^s \chi_{\mathcal{L}}(i) \stackrel{(1)}{\leq} \sum_{i=1}^s \lambda_{g^{-1}(i)}(I) = \sum_{i=1}^s \lambda_i(I),$$

where (1) follows from $\sum_{i=1}^s \chi_{\mathcal{L}}(i) \leq t$. Finally, we have to show that \mathbf{x}^J is a term in $o(\mathbf{x}^I)$ if and only if $\mathbf{x}^I =_{\text{dom}} \mathbf{x}^J$.

If \mathbf{x}^J is a term in $o(\mathbf{x}^I)$, then I and J are permutations of each other and their associated partitions are equal: $\lambda(I) = \lambda(J)$. Thus of course $\mathbf{x}^I =_{\text{dom}} \mathbf{x}^J$.

To prove the converse, observe that if $\mathbf{x}^I =_{\text{dom}} \mathbf{x}^J$, then $\lambda(I) = \lambda(J)$. Therefore, J is a permutation of I . We need to prove that there is an element $g \in G$ that permutes I into J . To this end, note that by definition the monomial \mathbf{x}^J occurs in the orbit sum $o(\mathbf{x}^{I_{\text{red}}}) \cdot s_t$. Thus, we subtract 1 from the t largest exponents in the exponent sequence I to obtain $\mathbf{x}^{I_{\text{red}}}$. Then we permute the result by an element g of G to get some monomial in the orbit of $\mathbf{x}^{I_{\text{red}}}$. Finally we add 1 to the same t_I exponents in the result as we subtracted in the first step, because we are multiplying with s_t . Since the resulting monomial is \mathbf{x}^J we have that the element $g \in G$ permutes I into J . Thus \mathbf{x}^J is a term of $o(\mathbf{x}^I)$. \square

We illustrate this result with an example.

Example 4.21. Let $\rho : \mathbb{Z}/2 \rightarrow \mathrm{GL}(4, \mathbb{R})$ be the faithful representation afforded by the matrix

$$\begin{bmatrix} 0 & 1 & & 0 \\ 1 & 0 & & 0 \\ & & 0 & 1 \\ 0 & & 1 & 0 \end{bmatrix}.$$

Let $\mathbf{x}^I = x_1x_2^3x_3$; thus $I = (1, 3, 1, 0)$. This is a nonspecial monomial with orbit sum

$$o(\mathbf{x}^I) = x_1x_2^3x_3 + x_1^3x_2x_3.$$

The gap occurs at $t = 1$, and its reduced exponent sequence is

$$I_{\mathrm{red}} = (1, 2, 1, 0).$$

We compute

$$\begin{aligned} o(x_1x_2^2x_3)s_1 &= (x_1x_2^2x_3 + x_1^2x_2x_3)(x_1 + \cdots + x_4) \\ &= x_1^2x_2^2x_3 + x_1^3x_2x_3 + x_1x_2^3x_3 + x_1^2x_2^2x_3 \\ &\quad + x_1x_2^2x_3^2 + x_1^2x_2x_3^2 + x_1x_2^2x_3x_4 + x_1^2x_2x_3x_4. \end{aligned}$$

The exponent sequences occurring in this polynomial have partitions

$$(2, 2, 1, 0), \quad (2, 1, 1, 1), \quad \text{or} \quad (3, 1, 1, 0).$$

The sequences are lower in dominance order than the exponent sequence of our original \mathbf{x}^I . Also note that the two monomials

$$x_1^3x_2x_3 \quad \text{and} \quad x_1x_2^3x_3$$

with the same dominance order occur in $o(\mathbf{x}^I)$. Hence,

$$o(\mathbf{x}^I) = o(x_1x_2^2x_3)s_1 - o(x_1^2x_2^2x_3) - o(x_1x_2^2x_3^2) - o(x_1^2x_2x_3x_4).$$

Furthermore, the first three orbit sums on the right hand side are orbit sums of special monomials. Thus we are left with the problem of breaking up the last orbit sum:

$$o(x_1^2x_2x_3x_4) = o(x_1)s_4.$$

Theorem 4.22 (M. Göbel). *Let $\rho : G \hookrightarrow \mathrm{GL}(n, \mathbb{F})$ be a permutation representation of a finite group. Then the ring of invariants $\mathbb{F}[V]^G$ is generated as an algebra by the top elementary symmetric function $s_n = x_1 \cdots x_n$ and the orbit sums of special monomials.*

Proof. By Proposition 4.12 we need to show that the orbit sum of any monomial can be written as a polynomial in s_n and orbit sums of special monomials. We denote by $A \subseteq \mathbb{F}[V]^G$ the subalgebra generated by s_n and all orbit sums of special monomials.

Let $\mathbf{x}^I \in \mathbb{F}[V]$ be a nonspecial monomial. We use induction on the dominance order to show that $o(\mathbf{x}^I) \in A$.

The monomials with smallest dominance order are those of degree zero. Thus our induction starts.

Assume that the result is proven for all monomials smaller in dominance order than \mathbf{x}^I .

If $\lambda_n(I) \neq 0$, then \mathbf{x}^I is divisible by s_n . We have

$$\mathbf{x}^I = \mathbf{x}^{I'} \cdot s_n^{\lambda_n(I)},$$

where $\mathbf{x}^{I'}$ is no longer divisible by s_n . Since s_n is invariant under G , we have that

$$o(\mathbf{x}^I) = o(\mathbf{x}^{I'}) \cdot s_n^{\lambda_n(I)}.$$

Thus it is enough to show the result for $\mathbf{x}^{I'}$. In other words we assume, without loss of generality, that $\lambda_n(I) = 0$. Since \mathbf{x}^I was assumed to be nonspecial, there is a gap in the associated partition $\lambda(I)$ at t_I . We set

$$o(\mathbf{x}^I) = o(\mathbf{x}^{I_{\text{red}}}) \cdot s_{t_I} - r(\mathbf{x}^I).$$

The t_I th elementary symmetric polynomial s_{t_I} is a sum of orbit sums of special monomials, because $t_I < n$. Thus $s_{t_I} \in A$. Moreover, the orbit sum $o(\mathbf{x}^{I_{\text{red}}})$ belongs to A by induction, because the reduced monomial is strictly lower in dominance order. Therefore, $o(\mathbf{x}^{I_{\text{red}}}) \cdot s_{t_I}$ belongs to A . Finally, we have

$$r(\mathbf{x}^I) = o(\mathbf{x}^{I_{\text{red}}}) \cdot s_{t_I} - o(\mathbf{x}^I).$$

Thus by Proposition 4.20 all the monomials that occur in $r(\mathbf{x}^I)$ are lower in the dominance order than \mathbf{x}^I . Hence $r(\mathbf{x}^I)$ is in A by induction, and therefore so is $o(\mathbf{x}^I)$. \square

What have we gained by this result?

First, special monomials have degree at most $\binom{n}{2}$. Thus we have the following corollary.

Corollary 4.23 (Göbel's Bound). *Let $\rho : G \hookrightarrow \mathrm{GL}(n, \mathbb{F})$ be a permutation representation of a finite group. Then the ring of invariants $\mathbb{F}[V]^G$ is generated as an algebra by homogeneous polynomials of degree at most $\max\{n, \binom{n}{2}\}$.*

Proof. The maximal degree of a special monomial is $\binom{n}{2}$. The degree of s_n is n . \square

Note that Göbel's bound is valid for all permutation representations *independent* of the group or the ground field. However, it depends on n , the size of the representation (or the size of the set being permuted). We will see in Chapter 6 that in characteristic zero there is always an a priori degree bound on the generators. This bound will be called **Noether's bound**.

The next example illustrates that Göbel's bound is sharp, i.e., there exists a permutation representation of a finite group so that any minimal generating set contains an element of highest possible order:

Example 4.24 (Defining Representation of the Alternating Group). Consider the alternating group A_n , $n \geq 4$, in its defining representation. Certainly we have $s_1, \dots, s_n \in \mathbb{F}[x_1, \dots, x_n]^{A_n}$. Furthermore, the **Vandermonde determinant**

$$\nabla_n = \prod_{i < j} (x_i - x_j)$$

is invariant under A_n , because A_n consists of all even permutations. The polynomial ∇_n is not symmetric, because for any transposition $\tau \in \Sigma_n$ we have that

$$\tau \nabla_n = -\nabla_n.$$

Therefore

$$\mathbb{F}[s_1, \dots, s_n] \subsetneq \mathbb{F}[s_1, \dots, s_n, \nabla_n] \subseteq \mathbb{F}[x_1, \dots, x_n]^{A_n}.$$

By what we have proven in this section, we know that in order to find the full ring of invariants we “just” have to calculate the orbit sums of all special monomials and put them together with s_n into an algebra. If you do that (or write a small computer routine that does that for you) you find that $\mathbb{F}[V]^{A_n}$ is equal to $\mathbb{F}[s_1, \dots, s_n, \nabla_n]$. This is of course not very satisfying. It is one of the goals of the next chapters

to find methods to determine the ring of invariants more effectively. For now, note that ∇_n has degree $\binom{n}{2}$. Thus Göbel's bound is sharp. See Exercise 13 in this chapter for the case of $n = 3$.

The preceding results also tell us that a ring of permutation invariants is always finitely generated, because they are only finitely many special monomials (even though their number grows rapidly with the number of variables).

Corollary 4.25. *Let $\rho : G \hookrightarrow \mathrm{GL}(n, \mathbb{F})$ be a permutation representation of a finite group. Then the ring of invariants $\mathbb{F}[V]^G$ is finitely generated as an algebra.*

This statement remains true for all invariant rings of finite groups. We will present two proofs of this fact in this text: an algorithmic proof in Chapter 6 and an abstract proof in Chapter 10.

Finally, note that the results of this chapter yield an *algorithm* to calculate any ring of permutation invariants:

Let $\rho : G \hookrightarrow \mathrm{GL}(n, \mathbb{F})$ be a permutation representation of a finite group G .

first: Calculate the orbit sums of all special monomials.

second: The desired ring is the \mathbb{F} -algebra generated by all orbit sums calculated in the first step and s_n .

Alternatively, you could calculate the orbit sums of *all* monomials of degree at most $\max\{n, \binom{n}{2}\}$.

Note carefully that this algorithm *terminates*, because there are only finitely many special monomials, resp. finitely many monomials of degree at most $\max\{n, \binom{n}{2}\}$.

Example 4.26. Recall the regular representation of $\mathbb{Z}/4$ from Example 4.1:

$$\rho : \mathbb{Z}/4 \longrightarrow \mathrm{GL}(4, \mathbb{F}), \quad 3 \mapsto \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

By Göbel's Theorem the ring of invariants is generated by orbit sums of special monomials and $s_4 = x_1x_2x_3x_4$. Since $n = 4$, special monomials have degree at most 6. The possible partitions are

$$(1 \geq 0 \geq 0 \geq 0), (1 \geq 1 \geq 0 \geq 0), (2 \geq 1 \geq 0 \geq 0), (1 \geq 1 \geq 1 \geq 0), \\ (2 \geq 1 \geq 1 \geq 0), (2 \geq 2 \geq 1 \geq 0), \text{ and } (3 \geq 2 \geq 1 \geq 0).$$

So the following is a complete list of algebra generators (where we noted when two orbit sums are obviously equal): In degree one there is only the first elementary symmetric function

$$o(x_1) = o(x_2) = o(x_3) = o(x_4) = s_1.$$

In degree two there are two orbit sums of special monomials (their sum is s_2 - why is that supposed to be so?)

$$o(x_1x_2) = o(x_2x_3) = o(x_3x_4) = o(x_4x_1) \\ = x_1x_2 + x_2x_3 + x_3x_4 + x_4x_1, \text{ and} \\ o(x_1x_3) = o(x_2x_4) = x_1x_3 + x_2x_4.$$

In degree three we find

$$o(x_1^2x_2) = o(x_2^2x_3) = o(x_3^2x_4) = o(x_4^2x_1) \\ = x_1^2x_2 + x_2^2x_3 + x_3^2x_4 + x_4^2x_1, \\ o(x_1^2x_3) = o(x_2^2x_4) = x_1^2x_3 + x_2^2x_4, \\ o(x_1^2x_4) = o(x_2^2x_1) = o(x_3^2x_2) = o(x_4^2x_3) \\ = x_1^2x_4 + x_2^2x_1 + x_3^2x_2 + x_4^2x_3, \text{ and} \\ o(x_1x_2x_3) = o(x_2x_3x_4) = o(x_3x_4x_1) = o(x_4x_1x_2) \\ = x_1x_2x_3 + x_2x_3x_4 + x_3x_4x_1 + x_4x_1x_2 \\ = s_3.$$

In degree four we have

$$o(x_1^2x_2x_3) = o(x_2^2x_3x_4) = o(x_3^2x_4x_1) = o(x_4^2x_1x_2) \\ = x_1^2x_2x_3 + x_2^2x_3x_4 + x_3^2x_4x_1 + x_4^2x_1x_2, \\ o(x_1x_2^2x_3) = o(x_2x_3^2x_4) = o(x_3x_4^2x_1) = o(x_4x_1^2x_2) \\ = x_1x_2^2x_3 + x_2x_3^2x_4 + x_3x_4^2x_1 + x_4x_1^2x_2, \text{ and} \\ o(x_1x_2x_3^2) = o(x_2x_3x_4^2) = o(x_3x_4x_1^2) = o(x_4x_1x_2^2) \\ = x_1x_2x_3^2 + x_2x_3x_4^2 + x_3x_4x_1^2 + x_4x_1x_2^2.$$

In degree five we have

$$\begin{aligned}
o(x_1^2 x_2^2 x_3) &= o(x_2^2 x_3^2 x_4) = o(x_3^2 x_4^2 x_1) = o(x_4^2 x_1^2 x_2) \\
&= x_1^2 x_2^2 x_3 + x_2^2 x_3^2 x_4 + x_3^2 x_4^2 x_1 + x_4^2 x_1^2 x_2, \\
o(x_1^2 x_2^2 x_4) &= o(x_2^2 x_3^2 x_1) = o(x_3^2 x_4^2 x_2) = o(x_4^2 x_1^2 x_3) \\
&= x_1^2 x_2^2 x_4 + x_2^2 x_3^2 x_1 + x_3^2 x_4^2 x_2 + x_4^2 x_1^2 x_3, \text{ and} \\
o(x_1^2 x_2 x_3^2) &= o(x_2^2 x_3 x_4^2) = o(x_3^2 x_4 x_1^2) = o(x_4^2 x_1 x_2^2) \\
&= x_1^2 x_2 x_3^2 + x_2^2 x_3 x_4^2 + x_3^2 x_4 x_1^2 + x_4^2 x_1 x_2^2.
\end{aligned}$$

Finally, in degree six we have

$$\begin{aligned}
o(x_1^3 x_2^2 x_3) &= o(x_2^3 x_3^2 x_4) = o(x_3^3 x_4^2 x_1) = o(x_4^3 x_1^2 x_2) \\
&= x_1^3 x_2^2 x_3 + x_2^3 x_3^2 x_4 + x_3^3 x_4^2 x_1 + x_4^3 x_1^2 x_2, \\
o(x_1^3 x_2^2 x_4) &= o(x_2^3 x_3^2 x_1) = o(x_3^3 x_4^2 x_2) = o(x_4^3 x_1^2 x_3) \\
&= x_1^3 x_2^2 x_4 + x_2^3 x_3^2 x_1 + x_3^3 x_4^2 x_2 + x_4^3 x_1^2 x_3, \\
o(x_1^3 x_2 x_3^2) &= o(x_2^3 x_1^2 x_3) = o(x_3^3 x_1^2 x_4) = o(x_4^3 x_2^2 x_1) \\
&= x_1^3 x_2 x_3^2 + x_2^3 x_1^2 x_3 + x_3^3 x_1^2 x_4 + x_4^3 x_2^2 x_1, \\
o(x_1^3 x_3^2 x_4) &= o(x_2^3 x_4^2 x_1) = o(x_3^3 x_1^2 x_4) = o(x_4^3 x_2^2 x_1) \\
&= x_1^3 x_3^2 x_4 + x_2^3 x_4^2 x_1 + x_3^3 x_1^2 x_4 + x_4^3 x_2^2 x_1, \\
o(x_1^3 x_4^2 x_2) &= o(x_2^3 x_1^2 x_3) = o(x_3^3 x_2^2 x_4) = o(x_4^3 x_3^2 x_1) \\
&= x_1^3 x_4^2 x_2 + x_2^3 x_1^2 x_3 + x_3^3 x_2^2 x_4 + x_4^3 x_3^2 x_1, \text{ and} \\
o(x_1^3 x_4^2 x_3) &= o(x_2^3 x_1^2 x_4) = o(x_3^3 x_2^2 x_1) = o(x_4^3 x_3^2 x_2) \\
&= x_1^3 x_4^2 x_3 + x_2^3 x_1^2 x_4 + x_3^3 x_2^2 x_1 + x_4^3 x_3^2 x_2.
\end{aligned}$$

Note that there are several orbit sums with the same associated partition. Note also that we do not know whether we really need all of those: there might be an orbit sum that can be expressed in terms of the others, like, e.g.,

$$o(x_1^2 x_2 x_3) = s_3 s_1 - 4s_4 - o(x_1^2 x_3 x_4) - o(x_1^2 x_2 x_4).$$

So Göbel's Theorem gives us a complete set of generators, but some of them might be redundant.

We close this chapter with another example.

Example 4.27. The dihedral group of order 10 has a five-dimensional permutation representation

$$\rho : D_{10} \hookrightarrow \mathrm{GL}(5, \mathbb{F})$$

afforded by the matrices

$$D = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

Following Göbel's bound we calculate the orbit sums of all special monomials up to degree $10 = \binom{5}{2}$. We obtain in degree one only s_1 ; in degree two we have s_2 . The third elementary symmetric function splits into two invariants, namely

$$x_1x_2x_3 + x_2x_3x_4 + x_3x_4x_5 + x_4x_5x_1 + x_5x_1x_2$$

and

$$x_1x_2x_4 + x_2x_3x_5 + x_3x_4x_1 + x_4x_5x_2 + x_5x_1x_3.$$

Furthermore, we have the orbit sums of the monomials with partition $(2 \geq 1 \geq 0 \geq 0 \geq 0)$. In degree four there is the fourth elementary symmetric function s_4 and the orbit sums of the monomials with partition $(2 \geq 1 \geq 1 \geq 0 \geq 0)$. In degree five there is the fifth elementary symmetric function s_5 and the orbit sums of the monomials with partitions $(2 \geq 1 \geq 1 \geq 1 \geq 0)$ and $(2 \geq 2 \geq 1 \geq 0 \geq 0)$. In degree six we need the orbit sums of monomials with partitions $(2 \geq 2 \geq 1 \geq 1 \geq 0)$ and $(3 \geq 2 \geq 1 \geq 0 \geq 0)$; in degree seven the orbit sums of monomials with partitions $(2 \geq 2 \geq 2 \geq 1 \geq 0)$ and $(3 \geq 2 \geq 1 \geq 1 \geq 0)$; in degree eight $(3 \geq 2 \geq 2 \geq 1 \geq 0)$; in degree nine $(3 \geq 3 \geq 2 \geq 1 \geq 0)$ and in degree ten $(4 \geq 3 \geq 2 \geq 1 \geq 0)$; see Exercise 14 in this chapter.

4.4. Exercises

- (1) Find the regular representation of V_4 .
- (2) Find the regular representation of Q_8 .
- (3) Find the regular representation of D_8 .
- (4) Find a permutation representation of $\mathbb{Z}/6$ of degree 6 that is not the regular representation.

- (5) Find a degree 4 representation of $\mathbb{Z}/4$ that is not the regular representation.
- (6) Show that lexicographic order is a partial order.
- (7) Show that dominance order is a partial order.
- (8) Show that $\mathbf{x}^I =_{\text{dom}} \mathbf{x}^J$ if and only if I is a permutation of J .
- (9) Consider the three monomials $x_1^3x_2x_3^5$, $x_1^6x_2x_3$, and $x_1^7x_2x_3$ in $\mathbb{C}[x_1, \dots, x_4]$.
- Which one is the largest in lexicographic order?
 - Which one is the largest in dominance order?
 - Are they special? If not, find the corresponding reduced monomials.

- (10) Rewrite the following polynomials in $\mathbb{C}[x_1, x_2, x_3]^{\Sigma_3}$ in terms of the elementary symmetric polynomials:

$$f_1 = x_1^2x_2 + x_2^2x_3 + x_3^2x_1 + x_1^2x_3 + x_2^2x_1 + x_3^2x_2,$$

$$f_2 = x_1^3x_2x_3 + x_1x_2^3x_3 + x_1x_2x_3^3,$$

$$f_3 = x_1^4 + x_2^4 + x_3^4.$$

(Why do you know that these polynomials are invariant under the Σ_3 -action?)

- (11) Consider the representation of $\mathbb{Z}/2$ of Example 4.21. Rewrite the orbit sum of $x_1x_2^4x_3$ as a polynomial in orbit sums of special monomials. Do the same for the monomial $x_1x_3^4x_4^7$.
- (12) Find the ring of invariants of the $\mathbb{Z}/2$ -representation given in Example 4.8.
- (13) Find the ring of invariants of the regular representation of $\mathbb{Z}/3$. (HINT: This representation coincides with the defining representation of the alternating group in three letters.) Show that there are invariants that can be written in more than one way in terms of the algebra generators.
- (14) Complete Example 4.27:
- Show that the orbit length of a monomial is 1, 2, 5, or 10.
 - Convince yourself that the partitions given in that example are all that we have to consider.
 - Show that there are 30 monomials with the partition $(2 \geq 1 \geq 1 \geq 0 \geq 0)$.
 - Count the number of special monomials of degree at least 5.

- (v) Make a complete list of generators of the ring of invariants.
- (15) Find the ring of invariants of the regular representation of V_4 .
- (16) Let $\mathcal{T} = \{t_1 < t_2 < \cdots < t_k\}$ be an ordered subset of the set $\{1, \dots, n-1\}$. Set $t_{k+1} = n$. We define the **Young subgroup** $Y_{\mathcal{T}}$ of Σ_n associated to the descent set \mathcal{T} by

$$Y_{\mathcal{T}} = \{\sigma \in \Sigma_n \mid \sigma([t_i, t_{i+1}]) = [t_i, t_{i+1}] \forall i = 1, \dots, k\},$$

where $[t_i, t_{i+1}] = \{t_i, t_i + 1, \dots, t_{i+1}\}$ denotes the interval of integers from t_i to t_{i+1} . (In other words, we put n things into k boxes and then permute the things in the boxes independently.)

- (i) Show that $Y_{\mathcal{T}}$ is indeed a group.
- (ii) What is the order of $Y_{\mathcal{T}}$?
- (iii) Take the defining representation of Σ_n and restrict it to $Y_{\mathcal{T}}$. What is the corresponding ring of invariants?
- (17) Prove the following generalization of Newton's formulae (Proposition 4.7): Let $l_1, \dots, l_r \in V^*$ be linear forms. Then

$$i s_i(l_1, \dots, l_r) = \sum_{k=1}^i (-1)^{k-1} p_k(l_1, \dots, l_r) s_{i-k}(l_1, \dots, l_r),$$

where $s_0(l_1, \dots, l_r) = 1$, $s_i(l_1, \dots, l_r)$ is the i th elementary symmetric polynomial in the l_i 's, and $p_i(l_1, \dots, l_r)$ is the i th symmetric power sum in the l_i 's.

- (18) Read about Sir Isaac Newton [1646-1723].
- (19) Read about Edward Waring [1734-1798].
- (20) Read about Carl Friedrich Gauss [1777-1855].
- (21) Who gave the Viète polynomials their name?
- (22) The Vandermonde determinant is named after Alexandre-Théophile Vandermonde [1735-1796]. Why is the Vandermonde determinant a *determinant*?
- (23) Alfred Young [1873-1940] gave Young groups their name. Read about him.

Application: Decay of a Spinless Particle

We want to illustrate the results we achieved so far with an example from particle physics taken from the diploma thesis of Dirk Engelmann.¹

Consider the decay of a spinless boson into four identical spinless bosons. Each of these four final particles has a 4-momentum p_i ($i = 1, \dots, 4$). Mathematically speaking the p_i 's are just 4-dimensional vectors $p_i = (E_i, p_{i1}, \dots, p_{i3})$, where the first component describes the total energy and the last three components describe the 3-momentum. We obtain the six basic Minkowski scalar products

$$p_i p_j = E_i E_j - (p_{i1} p_{j1} + p_{i2} p_{j2} + p_{i3} p_{j3})$$

for $1 \leq i < j \leq 4$, which encode the basic physical properties. Since our four bosons are identical, any permutation of them does not change the physical situation. However, say we interchange boson 1 and boson 2, then the scalar product $p_1 p_3$ is interchanged with $p_2 p_3$. Thus we obtain a permutation of the set of basic scalar products. In other words, setting $v_{ij} = p_i p_j$, we are looking at a Σ_4 -action on the 6-dimensional vector space spanned by v_{ij} for $1 \leq i < j \leq 4$

¹Dirk Engelmann, *Polynominalvarianten einer Darstellung der symmetrischen Gruppe*, University of Göttingen, 1995.

induced by the defining representation of Σ_4 :

$$\sigma(v_{ij}) = \begin{cases} v_{\sigma(i)\sigma(j)} & \text{if } \sigma(i) < \sigma(j), \\ v_{\sigma(j)\sigma(i)} & \text{otherwise.} \end{cases}$$

Now you want to know which expressions in the v_{ij} 's are invariant under this Σ_4 -action.

Let us step back and assume we have three final identical particles. Then we have to study a Σ_3 -action on a 3-dimensional vector space spanned by v_{23}, v_{13}, v_{12} . Which representation is that? Let us check that: We keep the basis vectors in the above order. Then we have

$$\rho : \Sigma_3 \longrightarrow \text{GL}(3, \mathbb{C}),$$

where

$$\rho(12) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \rho(13) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

This is just the defining representation of the symmetric group on three letters. By Theorem 4.13 its ring of invariants $\mathbb{C}[x_{23}, x_{13}, x_{12}]^{\Sigma_3}$ is generated by the three elementary symmetric functions in the basis elements x_{23}, x_{13}, x_{12} .

However, for four final particles the situation is a bit more complicated: Our vector space has dimension 6 with basis elements

$$v_{12}, v_{13}, v_{14}, v_{23}, v_{24}, v_{34}.$$

Our representation is

$$\rho : \Sigma_4 \longrightarrow \text{GL}(6, \mathbb{C})$$

with

$$\rho(12) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \rho(13) = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \quad \text{and}$$

$$\rho(14) = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

So, what is the ring of invariants? According to Göbel's Theorem, Theorem 4.22, we have to check the orbit sums of all special monomials. They have degree at most 15, and there are 332 of them! In his diploma thesis Dirk Engelmann developed methods to treat representations like this for arbitrary symmetric groups: Let V be a $\binom{n}{2}$ -dimensional complex vector space. Denote its basis by v_{ij} for $1 \leq i < j \leq n$. Then consider the representation of the symmetric group on n letters

$$\rho : \Sigma_n \longrightarrow \mathrm{GL}\left(\binom{n}{2}, \mathbb{C}\right)$$

given by

$$\sigma(v_{ij}) = \begin{cases} v_{\sigma(i)\sigma(j)} & \text{if } \sigma(i) < \sigma(j), \\ v_{\sigma(j)\sigma(i)} & \text{otherwise} \end{cases}$$

for all $\sigma \in \Sigma_n$. Now find its ring of invariants! In our case, $n = 4$, he obtained the following set of \mathbb{C} -algebra generators of the ring of invariants $\mathbb{C}[x_{12}, \dots, x_{34}]^{\Sigma_4}$:

$$s_1 = x_{12} + \dots + x_{34},$$

$$p_2 = x_{12}^2 + \dots + x_{34}^2,$$

$$p_3 = x_{12}^3 + \dots + x_{34}^3,$$

$$p_4 = x_{12}^4 + \dots + x_{34}^4,$$

$$o(x_{12}x_{34}) = x_{12}x_{34} + x_{13}x_{24} + x_{14}x_{23},$$

$$o(x_{12}x_{13}x_{14}) = x_{12}x_{13}x_{14} + x_{12}x_{23}x_{24} + x_{13}x_{23}x_{34} + x_{14}x_{24}x_{34}$$

and the orbit sums

$$o(x_{12}^2x_{13}), o(x_{12}^2x_{13}x_{23}), o(x_{12}^4x_{13}), o(x_{12}^4x_{13}^2), \text{ and } o(x_{12}^3x_{13}^3x_{23}^2x_{14}).$$

We note that the first elementary symmetric function s_1 and the power sums p_2, p_3 , and p_4 are invariant under the full symmetric group on six letters. Furthermore, we note that some of his invariants are not orbit sums of *special* monomials, because he had developed

other methods avoiding the calculation of all orbit sums of special monomials and better suited for this type of Σ_n -action.

Application: Counting Weighted Graphs

The representation of the symmetric group Σ_n in the preceding application appears also in graph theory as we show in this section.¹

A **graph** Γ consists of a set of **vertices**

$$V(\Gamma) = \{v_1, v_2, \dots\}$$

and a set of **edges**

$$E(\Gamma) = \{e_{ij} \text{ if } v_i, v_j \text{ are connected}\}$$

between them. Two graphs Γ and Γ' are **isomorphic** if there exists a bijective map

$$\phi : V(\Gamma) \longrightarrow V(\Gamma')$$

between the sets of vertices such that it induces a bijection

$$\tilde{\phi} : E(\Gamma) \longrightarrow E(\Gamma'), e_{ij} \mapsto e'_{\phi(i)\phi(j)}$$

between the sets of edges.

A **weighted graph** Γ is a graph such that every edge e_{ij} has a weight $m_{ij} \in \mathbb{C}$. Two weighted graphs are **weighted isomorphic** if the bijection $\tilde{\phi}$ maps edges onto edges of the same weight.

¹The material of this section is taken out of the PhD thesis of Nicolas M. Thiéry, *Invariants algébriques de graphes et reconstruction. Une étude expérimentale.*, PhD thesis, University of Lyon I, 1999.

We want to count the number of isomorphism classes of weighted graphs with, say, n vertices. This translates into counting invariants of the same type of representations we encountered in the previous application:

Consider the ring of polynomials

$$\mathbb{C}[x_{ij} | 1 \leq i < j \leq n]$$

in $\binom{n}{2}$ variables. Then Σ_n acts on $\mathbb{C}[x_{ij} | i \neq j, i, j = 1, \dots, n]$ by acting on the indices

$$\sigma(x_{ij}) = \begin{cases} x_{\sigma(i)\sigma(j)} & \text{if } \sigma(i) < \sigma(j), \\ x_{\sigma(j)\sigma(i)} & \text{otherwise.} \end{cases}$$

We arrange the basis elements x_{ij} in a symmetric matrix

$$X_{ij} = \begin{cases} x_{ij} & \text{for } i < j, \\ x_{ji} & \text{for } i > j, \\ 0 & \text{otherwise} \end{cases}$$

with zeros on the diagonal. The symmetric group acts by conjugation, $\sigma X_{ij} \sigma^{-1}$, where we have identified $\sigma \in \Sigma_n$ with its image under the defining representation. This is exactly the representation of Σ_n we obtain when considering the decay of a particle. In graph theory this becomes useful by the following observation:

Assume that you have two weighted isomorphic graphs Γ and Γ' with n vertices. The isomorphism $\phi : V(\Gamma) \longrightarrow V(\Gamma')$ is nothing but a permutation of the vertices inducing a permutation on the edges as well. Thus setting $\mathbf{m} = (m_{ij})$ we obtain

$$f(\mathbf{m}) = f(\mathbf{m}')$$

for all invariant polynomials $f \in \mathbb{C}[x_{ij} | 1 \leq i < j \leq n]^{\Sigma_n}$.

The converse is also true: Assume that you have two weighted graphs Γ and Γ' such that $f(\mathbf{m}) = f(\mathbf{m}')$ for all polynomials $f \in \mathbb{C}[x_{ij} | 1 \leq i < j \leq n]^{\Sigma_n}$. Then

$$\mathbf{m} = \sigma(\mathbf{m}')$$

for some $\sigma \in \Sigma_n$. Thus σ provides us with the desired isomorphism between Γ and Γ' .