1

First Steps

Invariant theory seeks to determine whether a (mathematical) object can be obtained from some other object by the action of some group. One way to answer this question is to find some functions that map from the class of objects to some field (or more generally some ring). Invariants are functions which take the same value on any two objects which are related by an element of the group. Thus if we can find any invariant which takes different values on two objects, then these two objects cannot be related by an element of the group. Ideally, we hope to find enough invariants to separate all objects which are not related by any group element. This means we want to find a (finite) set of invariants $f_1, f_2, \ldots, f_r$ with the property that if two objects are not related by the group action then at least one of these r invariants takes different values on the two objects in question.

For example, suppose we wish to determine whether two triangles are congruent, that is, whether one can be obtained from the other by translation, rotation, reflection or a combination of these operations. One useful invariant is the area function: two triangles with different areas cannot be congruent. On the other hand, the (unordered) set of three functions which give the lengths of the three sides are sufficient: two different triangles having sides of the same lengths must be congruent.

For us, the mathematical objects are elements of some vector space with a group action and the invariants will be those regular functions on the vector space that are constant on each of the group orbits.

We begin with some basic material on the action of groups on vector spaces and their coordinate rings, followed by a simple illustrative example. There are excellent references available: Benson [6], Derksen and Kemper [26], Neusel [85], Neusel and Smith [86] and Smith [103]. We also note that many advances in modular invariant theory have been made due to the programming language MAGMA [10], especially the invariant theory packages developed by Gregor Kemper.

H.E.A. Eddy Campbell, D. Wehlau, *Modular Invariant Theory*,
Encyclopaedia of Mathematical Sciences 139,
DOI 10.1007/978-3-642-17404-9_1,

1.1 Groups Acting on Vector Spaces and Coordinate Rings

We begin with a finite dimensional representation ρ of a group G over a field $\mathbb{K}$, i.e., a group homomorphism

$$\rho : G \to GL(V)$$

where V is a finite dimensional vector space over $\mathbb{K}$. In this book we will always denote the characteristic of the field (which may be 0) by p.

For us, the group G is always assumed to be finite of order denoted $|G|$. The representation of G is said to be a *modular* representation if p divides $|\rho(G)|$ and a *non-modular* one if not. Many questions which are well understood in the non-modular case are much less well understood in the modular case. One of the main reasons for this is that Maschke's Theorem fails to hold for modular representations. That is, modular representations may not be completely reducible and usually are not. Researchers have substituted techniques from algebraic geometry, commutative algebra, and group cohomology (including Steenrod operations) in an effort to make up for this deficiency. The main technique used in this book is the fact from representation theory that the cyclic group of order p has only finitely many indecomposable inequivalent representations in characteristic p.

The representation defines a left action of the group G on V. Given $\sigma \in G$ and $\mathbf{v} \in V$ we write $\sigma(\mathbf{v})$ for the vector $\rho(\sigma)(\mathbf{v})$, the result of applying $\rho(\sigma)$ to $\mathbf{v}$. Very often, the representation is fixed throughout an example and we make little reference to it.

We denote the set of vectors fixed (pointwise) by the group G by

$$V^G = \{\mathbf{v} \in V \mid \sigma(\mathbf{v}) = \mathbf{v}, \text{ for all } \sigma \in G\},$$

and for a subset X of V, we denote by

$$G_X = \{\sigma \in G \mid \sigma(v) = v, \text{ for all } v \in X\}$$

the isotropy or stabilizer subgroup of X. Usually, if $X = \{v\}$ is a singleton set we will write G_v to denote G_X.

Now consider V^*, the vector space dual to V. This is the set, $\mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$, of linear functionals from V to $\mathbb{K}$. Recall that $x : V \to \mathbb{K}$ is said to be a linear functional if $x(a\mathbf{v} + b\mathbf{w}) = ax(\mathbf{v}) + bx(\mathbf{w})$ for all $\mathbf{v}, \mathbf{w} \in V$, and all $a, b \in \mathbb{K}$. Of course, we have $\dim_{\mathbb{K}}(V^*) = \dim_{\mathbb{K}}(V)$.

The action of G on V determined by ρ naturally induces a left action of G on V^* as follows. Let $x \in V^*$ be any linear functional on V and let $\sigma \in G$. Then $\sigma(x)$ should be another linear functional on V. This new linear functional is defined by $(\sigma(x))(\mathbf{v}) := x(\sigma^{-1}(\mathbf{v}))$. In this definition we use σ^{-1} instead of σ in order to obtain a left action (and not a right action) of G on V^*. This new representation of G is often referred to as the dual representation.

Lemma 1.1.1. *Suppose we have a fixed representation $\rho : G \to \mathrm{GL}(V)$ and consider also $\rho^* : G \to \mathrm{GL}(V^*)$. In general, for $\sigma \in G$ the matrix representing $\rho(\sigma) \in \mathrm{GL}(V)$ with respect to a fixed basis is the transpose inverse of the matrix representing $\rho^*(\sigma)$ with respect to the dual basis.* □

Associated to the vector space V is its *coordinate ring* also called its *ring of regular functions*. This ring, denoted $\mathbb{K}[V]$, is a major object of study in algebraic geometry. We may define $\mathbb{K}[V]$ in a number of equivalent ways.

Here is a very concrete definition of the coordinate ring of V. Let

$$\{x_1, x_2, \ldots, x_n\}$$

be a fixed basis of V^*. Then $\mathbb{K}[V]$ is the polynomial ring in n variables:

$$\mathbb{K}[V] = \mathbb{K}[x_1, x_2, \ldots, x_n].$$

This is a useful description of $\mathbb{K}[V]$. For an *exponent* sequence $I = (i_1, \ldots i_n)$ consisting of non-negative integers, we define the monomial

$$x^I = x_1^{i_1} \cdots x_n^{i_n},$$

We say that x^I has degree $i_1 + \cdots + i_n$ and we denote the degree of x^I by $\deg(x^I)$ or even $\deg(I)$. As usual, we say that a polynomial $f = \sum a_j x^{I_j}$ for $a_i \in \mathbb{K}$ is homogeneous of degree d if each of its monomials, x^{I_j}, is of degree d. We observe that $\mathbb{K}[V]$ is naturally graded by degree: we may write

$$\mathbb{K}[V] = \oplus_{d \geq 0} \mathbb{K}[V]_d$$

where $\mathbb{K}[V]_d$ denotes the subspace of homogeneous polynomials of degree d (including the zero polynomial). We also observe that $\mathbb{K}[V]$ is a graded algebra. This just means that each $\mathbb{K}[V]_d$ is a subspace and that if $f \in \mathbb{K}[V]_d$ and $f' \in \mathbb{K}[V]_{d'}$ then $ff' \in \mathbb{K}[V]_{d+d'}$.

If V is a direct sum, $V = U \oplus W$ then we have a finer grading indexed by $\mathbb{N}^2$ on $\mathbb{K}[V]$ induced by the isomorphism $\mathbb{K}[V] \cong \mathbb{K}[U] \otimes_{\mathbb{K}} \mathbb{K}[W]$ given by

$$\mathbb{K}[V]_{(d,d')} = \mathbb{K}[U]_d \otimes_{\mathbb{K}} \mathbb{K}[W]_{d'}.$$

More generally, if $V = W_1 \oplus W_2 \oplus \cdots \oplus W_m$ then $\mathbb{K}[V]$ has a $\mathbb{N}^m$-grading given by

$$\mathbb{K}[V]_{(d_1,d_2,\ldots,d_m)} = \mathbb{K}[W_1]_{d_1} \otimes \mathbb{K}[W_2]_{d_2} \otimes \cdots \otimes \mathbb{K}[W_m]_{d_m} .$$

Thus if $f \in \mathbb{K}[V]_{(d_1,d_2,\ldots,d_m)}$, then

$$f(t_1 v_1, t_2 v_2, \ldots, t_m v_m) = t_1^{d_1} t_2^{d_2} \cdots t_m^{d_m} f(v_1, v_2, \ldots, v_m)$$

for all $t_1, t_2, \ldots, t_m \in \mathbb{K}$. We say that elements of $\mathbb{K}[V]_{(d_1,d_2,\ldots,d_m)}$ are multi-homogeneous. If each W_i is G-stable, then the G-action will stabilize each $\mathbb{N}^m$-graded summand $\mathbb{K}[V]_{(d_1,d_2,\ldots,d_m)}$ of $\mathbb{K}[V]$.

From an abstract point of view, if $\mathbb{K}$ is infinite we may define $\mathbb{K}[V]$ as a ring of functions:

$$\mathbb{K}[V] := \{f : V \to \mathbb{K} \mid f \text{ is a regular function on } V\}.$$

A function f is *regular* on V if f may be written as a polynomial in some (and hence every) basis of linear functionals on V.

We note that in order to view $\mathbb{K}[V]$ as a ring of functions on V we require that $\mathbb{K}$ be infinite. If $\mathbb{K}$ is finite, for example, if $\mathbb{K} = \mathbb{F}_p$, the field with p elements, then the two different polynomials x_1 and x_1^p in $\mathbb{K}[V]$ determine the same function on V. Let $\overline{\mathbb{K}}$ denote an algebraic closure of $\mathbb{K}$ and let $\overline{V} = \overline{\mathbb{K}} \otimes V$. The inclusion $\mathbb{K} \subset \overline{\mathbb{K}}$ induces an inclusion $V \subset \overline{V}$. Thus $\mathbb{K}[V] \subseteq \overline{\mathbb{K}}[\overline{V}]$ and two elements of $\mathbb{K}[V]$ are equal if and only if they determine the same function on $\overline{V}$.

We may also define $\mathbb{K}[V]$ as the symmetric algebra on V^*:

$$\mathbb{K}[V] = S^\bullet(V^*).$$

The action of G on V given by ρ also naturally induces an action of G on $\mathbb{K}[V]$. We may describe this action in two ways according to the description we use for $\mathbb{K}[V]$. In terms of polynomials we merely extend the action of G on V^* additively and multiplicatively. That is, let $\sigma \in G$ and $f, f' \in \mathbb{K}[V]$. Thus $\sigma(f + f') = \sigma(f) + \sigma(f')$ and $\sigma(ff') = (\sigma(f))(\sigma(f'))$.

Equivalently, if we regard the elements of $\mathbb{K}[V]$ as functions on V we may define this action of G on $\mathbb{K}[V]$ via $(\sigma(f))(\mathbf{v}) := f(\sigma^{-1}(\mathbf{v}))$.

The main object of study in invariant theory is the collection of polynomial functions on V left fixed by all of G. This collection of functions forms a ring, denoted $\mathbb{K}[V]^G$:

$$\mathbb{K}[V]^G := \{f \in \mathbb{K}[V] \mid \sigma(f) = f \text{ for all } \sigma \in G\}.$$

We observe that if a polynomial f is fixed by both σ and $\tau \in G$, then f is also fixed by $\sigma\tau$. We may conclude, therefore, that if f is invariant with respect to every element of some set of generators for G, then $f \in \mathbb{K}[V]^G$.

1.1.1 V Versus V^*

A common question that arises is why we insist upon considering the action of G upon V^* and $\mathbb{K}[V]$ rather than on the symmetric algebra on V, $S^\bullet(V)$. In order to answer this question, consider the following example.

Example 1.1.2. Let $G = C_p \times C_p$, the elementary Abelian p-group of order p^2. We consider a three dimensional representation of G given by

$$G = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & 0 & 1 \end{pmatrix} \;\middle|\; a, b \in \mathbb{F}_p \right\} \subset \mathrm{GL}(V).$$

Here G is generated by the two elements given by taking $(a, b) = (1, 0)$ and $(a, b) = (0, 1)$. These two elements are easily seen to be of order p and to commute. Thus G is indeed isomorphic to $C_p \times C_p$. We examine the geometry of the action of G on V by considering the orbits under this action. Let $\mathbf{v} = (v_1, v_2, v_3) \in V \cong \mathbb{F}_p^3$. If $v_1 \neq 0$ then we see that $G \cdot \mathbf{v} = \{(v_1, v_2 + av_1, v_3 + bv_1) \mid a, b \in \mathbb{F}_p\}$ consists of p^2 points. Conversely, if $v_1 = 0$ then the orbit of $\mathbf{v}$ consists of the single point $\mathbf{v}$.

If $\sigma \in G$ is represented by a matrix A in $\mathrm{GL}(V)$ with respect to the standard basis, then the matrix of σ in $\mathrm{GL}(V^*)$ with respect to the dual basis is given by $(A^T)^{-1}$. Thus working with the basis of V^* dual to the standard basis of V, we see that

$$G = \left\{ \begin{pmatrix} 1 & -a & -b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \;\middle|\; a, b \in \mathbb{F}_p \right\} \subset \mathrm{GL}(V^*).$$

Let $\{e_1, e_2, e_3\}$ be the standard basis of V and let $\{x_1, x_2, x_3\}$ denote the dual basis of V^*. The geometry of G acting on V is reflected in the invariant functions in $\mathbb{F}_p[V]^G = S^\bullet(V^*)^G = \mathbb{F}[x_1, n_2, n_3]$ where $n_2 = x_2^p - x_1^{p-1}x_2$ and $n_3 = x_3^p - x_1^{p-1}x_3$. If we consider a point $\mathbf{v}$ with $0 \neq v_1 \in \mathbb{F}_p$ then the two functions n_2 and n_3 are both constant on these orbits. Moreover, it is not too difficult to see that, if $\mathbf{v}' \in \overline{V} \cong \overline{\mathbb{F}_p}^3$ with $x_1(\mathbf{v}) = x_1(\mathbf{v}')$, $n_2(\mathbf{v}) = n_2(\mathbf{v}')$ and $n_3(\mathbf{v}) = n_3(\mathbf{v}')$, then $\mathbf{v}' \in G\mathbf{v}$.

Using $S^\bullet(V)^G$ instead, we would have found $S^\bullet(V)^G = \mathbb{F}_p[f_1, e_2, e_3]$ where f_1 is a cubic expression beginning $f_1 = e_1^3 + \ldots$. In particular, these do not correspond to functions which are constant on the orbits of G.

This example shows why we are interested in both the matrix representation of G on V and also on V^*. Examining the former allows us to see the geometry of the group action. Examining the latter allows us to understand which polynomials are invariants. Rather than writing out both matrices for a group element σ, we will often compromise by writing out the matrix A^{-1} of σ^{-1} in $\mathrm{GL}(V)$. This shows us directly how σ^{-1} is acting on V and allows us to study the orbits in V. The transpose of this matrix shows how σ acts on V^* and thus we may understand the action of σ on V^* by considering the rows of A^{-1} and the action of A^{-1} on row vectors by right multiplication.

A dramatic illustration of the difference between the group actions on V and V^* is provided by the following subgroup of $\mathrm{GL}(V)$ where V is a seven dimensional vector space over $\mathbb{F}_p$, the field of order p. We define

$$\sigma(a, b, c, d) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ a & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & b & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & c & 0 & 0 & 1 & 0 \\ d & d & d & 0 & 0 & 0 & 1 \end{pmatrix}$$

and we take $G = \{\sigma(a,b,c,d) \mid a,b,c,d \in \mathbb{F}_p\} \subset \mathrm{GL}(V)$. We will show in Example 8.0.8 that $\mathbb{F}[V]^G$ is a polynomial ring.

On the other hand, consider the group $H \subset \mathrm{GL}(V)$ consisting of the transposes of the elements of G acting on V, that is, the group of matrices

$$\tau(a,b,c,d) = \begin{pmatrix} 1&0&0&a&0&0&d \\ 0&1&0&0&b&0&d \\ 0&0&1&0&0&c&d \\ 0&0&0&1&0&0&0 \\ 0&0&0&0&1&0&0 \\ 0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&1 \end{pmatrix}.$$

We will show in Example 8.0.9 that $\mathbb{F}[V]^H$ is not Cohen-Macaulay.

Note that both groups are generated by reflections (reflections are defined below in §1.5).The definitions of polynomial rings and Cohen-Macaulay rings may be found in §2.3 and §2.8 respectively.

1.2 Constructing Invariants

One general method to construct invariants of finite groups is as follows. Let $f \in \mathbb{K}[V]$. Then the *transfer* or *trace* of f is defined as

$$\mathrm{Tr}(f) = \mathrm{Tr}^G(f) := \sum_{g \in G} \sigma(f).$$

Similarly, the *norm* of f is defined by

$$\mathbf{N}(f) = \mathbf{N}^G(f) := \prod_{g \in G} \sigma(f).$$

We also have occasion to use relative versions of these constructions. Suppose H is a subgroup of G and we have a polynomial f which is H-invariant. Then we choose a fixed set of left coset representatives $G/H := \{\sigma_1, \sigma_2, \ldots, \sigma_r\}$ and define

$$\mathrm{Tr}_H^G(f) := \sum_{\ell=1}^{r} \sigma_\ell(f)$$

and

$$\mathbf{N}_H^G(f) := \prod_{\ell=1}^{r} \sigma_\ell(f).$$

It is easy to see that for $f \in \mathbb{K}[V]^H$ the elements $\mathrm{Tr}_H^G(f)$ and $\mathbf{N}_H^G(f)$ are independent of the choice of $\sigma_1, \sigma_2, \ldots, \sigma_r$. However, for general f this is not true. For this reason, it is often useful to take H to be the isotropy subgroup

G_f. Of course, for any finite group G, subgroup H, and $f \in \mathbb{K}[V]^H$ we have that $\mathrm{Tr}_H^G(f)$ and $\mathbf{N}_H^G(f)$ are both G-invariants. Note that $\mathrm{Tr}^G(f) = \mathrm{Tr}_{\{e\}}^G(f)$ and $\mathbf{N}^G(f) = \mathbf{N}_{\{e\}}^G(f)$.

Still more generally, consider an element $f \in \mathbb{K}[V]$. We define the *G-orbit* of f to be $\{\sigma(f) \mid \sigma \in G\}$, denoted Gf. A slightly different way to describe the orbit of f is to use the isotropy subgroup G_f of f. We have $Gf = \{\sigma(f) \mid \sigma \in G/G_f\}$ where G/G_f denotes a set of (left) coset representatives of G_f in G.

Suppose, then, that $|Gf| = m$. From here, we can form the polynomial

$$\mathcal{S}_f(\lambda) = \prod_{h \in Gf} (\lambda - h) = \sum_{i=0}^{m} (-1)^i s_i \lambda^{m-i},$$

where $s_i \in \mathbb{K}[V]^G$. The coefficients are elementary symmetric functions in the elements of Gf. That is, if we write $Gf = \{f_1, \ldots, f_m\}$, then

$$\begin{aligned} s_1(f) &= f_1 + f_2 + \cdots f_m \ , \\ s_2(f) &= f_1 f_2 + f_1 f_3 + \cdots + f_{m-1} f_m \ , \\ &\vdots \\ s_m(f) &= f_1 f_2 \cdots f_m \ . \end{aligned}$$

For any finite group G, and $f \in \mathbb{K}[V]$ we have that

$$\begin{aligned} \mathrm{Tr}^G(f) &= |G_f| s_1(f), \\ \mathbf{N}^G(f) &= s_m(f)^{|G_f|}. \end{aligned}$$

1.3 On Structures and Fundamental Questions

The problems we will consider fall roughly into two classes:

1. Find generators for $\mathbb{K}[V]^G$. Failing that, find an upper bound for the largest degree of an element of a homogenous minimal generating set.
2. Determine the structure of $\mathbb{K}[V]^G$. For example, determine those groups G for which the ring of invariants $\mathbb{K}[V]^G$ is a polynomial algebra, a hypersurface, or a Cohen-Macaulay ring.

Both questions are interesting for either specific groups, or for classes of groups. In general, much more is known in the non-modular case than in the modular case.

1.4 Bounds for Generating Sets

Emmy Noether proved (see Theorem 3.1.2) that the ring of invariants of a representation V of a finite group acting is always generated as an algebra by a *finite* collection of homogeneous invariants $f_1, f_2, \ldots f_t$. Using the

graded Nakayama lemma (Lemma 2.10.1) we see that the number $\beta(V,G) := \max\{\deg(f_i) \mid 1 \le i \le t\}$ is independent of the choice of generators provided t is minimal. This number $\beta(V,G)$ is called the *Noether number* for V.

Noether showed that generators of degree at most $|G|$ are required when $p = 0$. For non-modular groups with $p > |G|$, this theorem is still true. Richman and others have shown Noether's original bound, $\beta(V,G) \le |G|$, applies if G is solvable. Smith [103][pg 175], Fleischmann [39], and others have shown that for non-modular groups, $\mathbb{K}[V]^G$ is generated in degrees at most $\dim_{\mathbb{K}}(V)(|G|-1)$. For an overview of this topic, see Wehlau's paper [111]. Here we need $\dim_{\mathbb{K}}(V) > 1$ and $|G| > 1$.

There was a conjecture that non-modular groups have rings of invariants that are generated in degrees less than or equal $|G|$. The difference between the known bound and this conjectural bound was known as the problem of Noether's gap: is there a non-modular group in the gap or not? In 1999, Fleischmann gave a beautiful and clever variation of Noether's original argument that showed the conjecture was true (see [39]). Independently, Fogarty [42] proved the same result. Below we give a simplified version of Fogarty's proof, due to Benson, see Theorem 3.5.1.

It is proved by Campbell, Geramita, Hughes, Shank and Wehlau in [17] that if $\mathbb{K}[V]^G$ is a hypersurface, then this ring is generated in degrees less than or equal to $|G|$. More generally, Broer [12] has shown that if $\mathbb{K}[V]^G$ is Cohen-Macaulay, then this ring of invariants is generated by elements of degree at most $\dim_{\mathbb{K}}(V)(|G|-1)$. G. Kemper has made the conjecture that Noether's degree bound, $|G|$, applies whenever $\mathbb{K}[V]^G$ is Cohen-Macaulay.

Symonds [106], using work of Karagueuzian and Symonds [62] has proved that

Theorem 1.4.1. *If $\mathbb{K}$ is finite and G is a non-trivial finite group acting on V with $\dim_{\mathbb{K}}(V) > 1$, then $\mathbb{K}[V]^G$ is generated in degrees less than or equal to $\dim_{\mathbb{K}}(V)(|G|-1)$.*

A synopsis of this work is given in §3.6.

1.5 On the Structure of $\mathbb{K}[V]^G$: The Non-modular Case

The invariant theory of finite groups is much better understood in the non-modular case. For example, in this situation, a complete characterization of those representations for which $\mathbb{K}[V]^G$ is polynomial is known. To state this characterization we need the following definition.

Definition 1.5.1. *Let V be a representation of G defined over a field $\mathbb{K}$. Then $\sigma \in G$ is a* reflection *if $\dim V^\sigma = \dim V - 1$. Over a field of characteristic p, a reflection of order p is called a* transvection.

Classically, a (real) reflection was defined as an element σ with single non-trivial eigenvalue -1, and a (complex) reflection as an element with single

non-trivial eigenvalue a (complex) root of unity. What we have defined as a "reflection" was originally called a "pseudo-reflection". This older terminology is still used by some authors.

The following famous and beautiful theorem follows from the work of Coxeter [24], Shephard and Todd [101], Chevalley [22], and Serre [95]. To prove one direction, that groups generated by reflections over $\mathbb{C}$ have polynomial invariant rings, Shephard and Todd classified all such representations and showed that in each case the ring of invariants is polynomial. Unaware of their work, Chevalley [22] proved in 1955 that for representations over $\mathbb{R}$ generated by reflections of order 2 the ring of invariants is always polynomial. Chevalley's proof is truly beautiful, short and does not rely on any classification. Serre who was familiar with the work of Shephard and Todd observed that Chevalley's proof works for all groups generated by reflections over $\mathbb{C}$ not just reflections of order 2. He also proved a partial converse valid over any field, see below. We describe a new proof by Dufresne of this result, see Section 12.2.

Theorem 1.5.2. *Let G be a finite group with $|G|$ invertible in the field $\mathbb{K}$. Then $\mathbb{K}[V]^G$ is a polynomial algebra if and only if the action of G on V is generated by reflections.*

Theorem 1.5.3. *Let G be a finite group represented over $\mathbb{F}$. If $\mathbb{K}[V]^G$ is a polynomial algebra then the action of G on V is generated by reflections.*

Aside from examples and special cases (see for example Nakajima's Theorem 8.0.7), the characterization of representations of finite groups with polynomial rings of invariants remains one of the most important open problems in modular invariant theory.

There are other wonderful theorems concerning characterizations of hypersurfaces (Nakajima), Gorenstein rings (Watanabe), or Cohen-Macaulay rings (Hochster and Eagon) in the non-modular case.

In the modular case, we note the theorem of Kemper [65]: a *bi-reflection* is an element $\sigma \in G$ with $\mathrm{Im}(\sigma - 1 : V \to V)$ of dimension less than or equal to 2.

Theorem 1.5.4. *Let G be a finite group with $|G|$ represented over the field $\mathbb{F}$ of characteristic $p > 0$ with $p \mid |G|$. If $\mathbb{K}[V]^G$ is Cohen-Macaulay then if the action of G on V is generated by bi-reflections.*

This topic is explored in more depth in §9.2. It remains an open problem to characterize those modular bi-reflection groups whose rings of invariants are Cohen-Macaulay.

1.6 Structure of $\mathbb{K}[V]^G$: Modular Case

J.P. Serre proved one direction of Theorem 1.5.2 holds in the modular case. He showed that whenever $\mathbb{K}[V]^G$ is a polynomial ring, the action of G on V must

be generated by reflections. Examples of reflection groups whose invariant rings are not polynomial are known. See for example, §8.2.

Nakajima has characterized those p-groups with polynomial rings of invariants when $\mathbb{K} = \mathbb{F}_p$ is the prime field of order p. Roughly speaking, he shows that such groups resemble the ring of invariants of the full Upper Triangular group. He gave examples of elementary Abelian reflection p-groups with non-Cohen-Macaulay invariant rings, a somewhat simpler example is the example mentioned above at the end of §1.1. Nakajima's characterization fails over larger fields, as shown by an example due to Stong, see §8.1.

Kemper and Malle have examined the class of irreducible representations of modular reflection groups and determined which have polynomial rings of invariants. Unfortunately, irreducible representations are few and far between. We summarize their work in §8.3.

Much work remains to be done on characterizing groups with polynomial rings of invariants.

1.7 Invariant Fraction Fields

It will be useful on occasion to study the fraction fields denoted $\mathrm{Quot}(\mathbb{K}[V])$ or $\mathbb{K}(V)$ and $\mathbb{K}(V)^G$ of the domains $\mathbb{K}[V]$ and $\mathbb{K}[V]^G$, respectively; in some situations we encounter, it is useful to recall the results of Galois Theory. It is not difficult to see that $(\mathbb{K}(V))^G = \mathrm{Quot}(\mathbb{K}[V]^G)$. For, given an invariant fraction $\frac{f}{f'} \in \mathbb{K}(V)^G$, we may write

$$\frac{f}{f'} = \frac{f \prod_{\sigma \neq 1} \sigma(f')}{f' \prod_{\sigma \neq 1} \sigma(f')} = \frac{f \prod_{\sigma \neq 1} \sigma(f')}{\mathbf{N}^G(f')}$$

and note that the denominator of the right hand side is invariant. Since the fraction itself is also invariant, the numerator of the right hand side is invariant as claimed.

Then we have the diagram

$$\begin{array}{ccc} \mathbb{K}[V]^G & \hookrightarrow & \mathbb{K}[V] \\ \downarrow & & \downarrow \\ \mathbb{K}(V)^G & \hookrightarrow & \mathbb{K}(V) \end{array}$$

and we see that the bottom row of this diagram tells us that $\mathbb{K}(V)$ is a Galois extension of $\mathbb{K}(V)^G$, that is, there exist $q = |G|$-many rational functions $a_i = \frac{f_i}{f'_i}$ such that $\{a_1, \ldots a_q\}$ is a basis for $\mathbb{K}(V)$ as a vector space over $\mathbb{K}(V)^G$. Furthermore, the induced G-action on

$$\mathbb{K}(V) = \oplus_{i=1}^{q} \mathbb{K}(V)^G a_i$$

is the regular representation of G.

It is a famous question of Noether's whether or not $\mathbb{K}(V)^G$ is purely transcendental; this is the question of whether or not there are $n = \dim(V)$ elements $a_i \in \mathbb{K}(V)^G$ such that $\mathbb{K}(V)^G = \mathbb{K}(a_1, \ldots, a_n)$. The answer to this question is negative in general. However, if $p > 0$ and G is a p-group, then $\mathbb{K}(V)^G$ is purely transcendental (see [81]). We will revisit this question in section §7.6.

1.8 Vector Invariants

Consider the coordinate ring of $m\,V = \oplus^m V$ with the diagonal action of G. The ring $\mathbb{K}[m\,V]^G$ is called a *ring of vector invariants* of G. Rings of vector invariants provide an important class of examples and counterexamples.

In [19], Campbell and Hughes give generators, as conjectured by Richman [92], for $\mathbb{F}_p[m\,V_2]^{C_p}$ where C_p denotes the cyclic group of order p, and V_2 denotes its 2 dimensional indecomposable representation. An easy corollary is the fact, first observed by Richman, that this invariant ring requires a generator of degree $m(p-1)$. Therefore, Noether's degree bound, $|G|$, does not hold for p-groups.

Kemper has proved that, if G is any modular group, then $\mathbb{F}[m\,V]^G$ is not Cohen-Macaulay for all sufficiently large m. In every example known, taking $m = 3$ is sufficiently large to obtain a non-Cohen-Macaulay ring of invariants.

If G is a p-group and $m \geq 3$, then $\mathbb{F}[mV]^G$ is not Cohen-Macaulay, see 9.2.3. This is an important corollary of the result (see 9.2.2) that if $\mathbb{K}[V]^G$ is Cohen-Macaulay, then G is generated by bi-reflections. Here an element $\sigma \in G$ is called a *bi-reflection* if $\dim V^\sigma \geq \dim V - 2$. This theorem shows us how rarely we may expect to encounter Cohen-Macaulay rings as the invariants of p-groups.

1.9 Polarization and Restitution

Consider the maps $\Delta : V \to m\,V = \underbrace{V \oplus V \oplus \cdots \oplus V}_{m \text{ copies}}$ and $\phi : m\,V \to V$ given by $\Delta(v) = (v, v, \ldots, v)$ and $\phi(v_1, v_2, \ldots, v_m) = v_1 + v_2 + \cdots + v_m$. Both of these maps are $\mathrm{GL}(V)$-equivariant where $\mathrm{GL}(V)$ acts diagonally on $m\,V$.

These two maps naturally induce ring maps $\Delta^* : \mathbb{F}[m\,V] \to \mathbb{F}[V]$ and $\phi^* : \mathbb{F}[V] \to \mathbb{F}[m\,V]$ given by $(\Delta^*(F))(v) = F(\Delta(v)) = F(v, v, \ldots, v)$ and $(\phi^*(f))(v_1, v_2, \ldots, v_m) = f(\phi(v_1, v_2, \ldots, v_m)) = f(v_1 + v_2 + \cdots + v_m)$.

Let $f \in \mathbb{F}[V]_d$. Using the $\mathbb{N}^m$-grading on $\mathbb{F}[m\,V]$ we have

$$\phi^*(f) = \sum_{i_1+i_2+\cdots+i_m=d} f_{(i_1,i_2,\ldots,i_m)}$$

where each $f_{(i_1,i_2,\ldots,i_m)} \in \mathbb{F}[m\,V]_{(i_1,i_2,\ldots,i_m)}$. These polynomials $f_{(i_1,i_2,\ldots,i_m)}$ are called *partial polarizations* of f and we write

$$\mathcal{P}\mathrm{ol}^m(f) = \left\{ f_{(i_1,i_2,\ldots,i_m)} \mid i_1 + i_2 + \cdots + i_m = d \right\}$$

to denote the set of all such partial polarizations.

In order to compute individual polarizations, we take m indeterminates $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m)$, and consider $\mathbf{v} = (v_1, v_2, \ldots, v_m)$ where each v_i represents a generic element of V. We write $\lambda\mathbf{v} = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_m v_m$, and then we have

$$\begin{aligned} f(\lambda\mathbf{v}) &= \phi^*(f)(\lambda_1 v_1, \lambda_2 v_2, \ldots, \lambda_m v_m) \\ &= \sum_{i_1+i_2+\ldots i_m=d} \lambda_1^{i_1}\lambda_2^{i_2}\cdots\lambda_m^{i_m} f_{(i_1,i_2,\ldots,i_m)}(v_1, v_2, \ldots v_m) \\ &= \sum_{|I|} \lambda^I f_I(\mathbf{v}) \end{aligned}$$

with $|I| = i_1 + i_2 + \cdots + i_m = d$ where

$$f_I \in \mathbb{K}[m\,V]_I = \mathbb{K}[V]_{i_1} \otimes \mathbb{K}[V]_{i_2} \otimes \cdots \otimes \mathbb{K}[V]_{i_m} \subset \mathbb{K}[m\,V]_d\ .$$

As a special case, we may take $m = d = \deg(f)$ and $(i_1, i_2, \ldots, i_m) = (1, 1, \ldots, 1)$ to get the *full polarization* of f denoted

$$\mathcal{P}(f) = f_{(1,1,\ldots,1)} = f_{\text{multi-linear}} \in \mathbb{K}[d\,V].$$

Lemma 1.9.1. *The mapping $f \mapsto f_{(i_1,i_2,\ldots,i_m)}$ is $\mathrm{GL}(V)$-equivariant. In particular, if G is any subgroup of $\mathrm{GL}(V)$ and $f \in \mathbb{K}[V]^G$, then $\mathcal{P}\mathrm{ol}^m(f) \subset \mathbb{K}[m\,V]^G$.*

Proof. Let $\sigma \in \mathrm{GL}(V)$. We need to show $(\sigma f)_I = \sigma(f_I)$. The former is defined by the equation

$$(\sigma f)(\lambda\mathbf{v}) = \sum_I \lambda^I (\sigma f)_I(\mathbf{v}).$$

But

$$(\sigma f)(\lambda\mathbf{v}) = f(\lambda\sigma^{-1}\mathbf{v}) = \sum_I \lambda^I f_I(\sigma^{-1}\mathbf{v}).$$

Therefore,

$$(\sigma f)_I(\mathbf{v}) = f_I(\sigma^{-1}(\mathbf{v})) = (\sigma f_I)(\mathbf{v}).$$

□

Lemma 1.9.2. *The full polarization $\mathcal{P}(f)$ of f is a symmetric function, i.e.,*

$$\mathcal{P}(f)(\tau(\mathbf{v})) = \mathcal{P}(f)(\mathbf{v})$$

where $\tau(\mathbf{v}) = (v_{\tau(1)}, \ldots, v_{\tau(m)})$ for all $\tau \in \Sigma_m$.

Proof. Since $\tau(\lambda\mathbf{v}) = (\lambda_{\tau(1)}v_{\tau(1)} + \lambda_{\tau(2)}v_{\tau(2)} + \cdots + \lambda_{\tau(m)}v_{\tau(m)})$ we have

$$f(\lambda\mathbf{v}) = f(\tau(\lambda\mathbf{v}))$$

for all $\tau \in \Sigma_d$. □

The map induced by Δ is called *restitution* and denoted by $\mathcal{R}$ or by $\mathcal{R}_m$. It is defined by $\mathcal{R} : \mathbb{K}[m\,V] \to \mathbb{K}[V]$ and $\mathcal{R}(F)(v) = F(\underbrace{v, v, \ldots, v}_{m})$.

The following lemma is expressed in terms of the multinomial coefficient $\binom{d}{I} := \frac{d!}{i_1!i_2!\cdots i_m!}$ where $I = (i_1, i_2, \ldots, i_m)$.

Lemma 1.9.3. *Let $f \in \mathbb{K}[V]_d$ be homogeneous of degree d, and consider the sequence of positive integers $i_1, i_2, \ldots, i_m$ with $i_1 + i_2 + \cdots + i_m = d$. Then*

$$\mathcal{R}(f_{(i_1,i_2,\ldots,i_m)}) = \binom{d}{i_1\ i_2\ \ldots\ i_m} f\ .$$

In particular,

$$\mathcal{R}\mathcal{P}(f) = d!f.$$

Proof. Setting $\mathbf{v} = (\mathbf{w}, \mathbf{w}, \ldots, \mathbf{w})$ we have $\lambda\mathbf{v} = |\lambda|\mathbf{w}$ where $|\lambda| = \lambda_1 + \lambda_2 + \cdots + \lambda_d$. Therefore,

$$f(\lambda\mathbf{v}) = f(|\lambda|\mathbf{w}) = |\lambda|^d f(\mathbf{w}) = \sum_{|I|=d} \binom{d}{I} \lambda^I f(\mathbf{w})$$

Conversely.

$$f(\lambda\mathbf{v}) = \sum_{|I|=d} \lambda^I f_I(\mathbf{w}, \mathbf{w}, \ldots \mathbf{w}) = \sum_{|I|=d} \lambda^I R f_I(\mathbf{w})\ .$$

Comparing the coefficients we see $\binom{d}{I} f = \mathcal{R} f_I$. □

Remark 1.9.4. If $d!$ is invertible in $\mathbb{K}$, then $f = \mathcal{R}\mathcal{P}(f/d!)$ lies in the image of $\mathcal{R}$. In particular, if $f \in \mathbb{K}[V]_d^G$ and d is invertible in $\mathbb{K}$, then $f \in \mathcal{R}(\mathbb{F}[d\,V]^G)$.

The following example illustrates polarization and restitution.

Example 1.9.5. Let $\mathbb{K}$ be a field of any characteristic. Consider the usual three dimensional permutation representation V of Σ_3, the symmetric group on three letters. Let $\{x, y, z\}$ be a permutation basis for V^*. It is well known that if $\mathbb{K}$ has characteristic zero, then $\mathbb{K}[V]^{\Sigma_3}$ is the polynomial ring $\mathbb{K}[s_1, s_2, s_3]$ where $s_1 = x + y + z$, $s_2 = xy + xz + yz$ and $s_3 = xyz$. This result is also true when $\mathbb{K}$ has positive characteristic, even for characteristics 2 and 3. We will outline one proof of this result in §3.2 and give another proof in §5.1.1. Here we consider the ring of vector invariants $\mathbb{K}[2\,V]^{\Sigma_3}$. Weyl [112] proved that the polarizations of the elementary symmetric functions $f = s_1, g = s_2, h = s_3$

suffice to generate $\mathbb{K}[2\,V]^{\Sigma_3}$ if $6 = |\Sigma_3|$ is invertible in $\mathbb{K}$. In fact, he proved that if V is the usual permutation representation of Σ_n, then for any n and any m the polarizations of the elementary symmetric polynomials $s_1, s_2, \ldots, s_n$ generate the ring $\mathbb{K}[m\,V]^{\Sigma_n}$ provided only that $n!$ is invertible in $\mathbb{K}$. Here we have

$$\begin{aligned} f(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2) &= f(\lambda_1 x_1 + \lambda_2 x_2, \lambda_1 y_1 + \lambda_2 y_2, \lambda_1 z_1 + \lambda_2 z_2) \\ &= \lambda_1(x_1 + y_1 + z_1) + \lambda_2(x_2 + y_2 + z_2). \end{aligned}$$

Thus $\mathcal{P}\mathrm{ol}^2(f) = \{f_{10}, f_{01}\}$ where

$$\begin{aligned} f_{10} &= x_1 + y_1 + z_1 \\ f_{01} &= x_2 + y_2 + z_2. \end{aligned}$$

Similarly,

$$\begin{aligned} g(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2) &= g(\lambda_1 x_1 + \lambda_2 x_2, \lambda_1 y_1 + \lambda_2 y_2, \lambda_1 z_1 + \lambda_2 z_2) \\ &= \lambda_1^2(x_1y_1 + x_1z_1 + y_1z_1) \\ &\quad + \lambda_1\lambda_2(x_1y_2 + x_1z_2 + y_1x_2 + y_1z_2 + z_1x_2 + z_1y_2) \\ &\quad + \lambda_2^2(x_2y_2 + x_2z_2 + y_2z_2). \end{aligned}$$

Thus $\mathcal{P}\mathrm{ol}^2(g) = \{g_{20}, g_{11}, g_{02}\}$ where

$$\begin{aligned} g_{20} &= x_1y_1 + x_1z_1 + y_1z_1, \\ g_{11} &= x_1y_2 + x_1z_2 + y_1x_2 + y_1z_2 + z_1x_2 + z_1y_2, \\ g_{02} &= x_2y_2 + x_2z_2 + y_2z_2. \end{aligned}$$

Here g_{11} is the full polarization $\mathcal{P}(g)$. Finally

$$\begin{aligned} &h(\lambda_1 x_1 + \lambda_2 x_2, \lambda_1 y_1 + \lambda_2 y_2, \lambda_1 z_1 + \lambda_2 z_2) \\ &\quad = \lambda_1^3(x_1y_1z_1) + \lambda_1^2\lambda_2(x_1y_1z_2 + x_1y_2z_1 + x_2y_1z_1) \\ &\qquad + \lambda_1\lambda_2^2(x_1y_2z_2 + x_2y_1z_2 + x_2y_2z_1) + \lambda_2^3x_2y_2z_2. \end{aligned}$$

Hence $\mathcal{P}\mathrm{ol}^2(h) = \{h_{30}, h_{21}, h_{12}, h_{03}\}$ where

$$\begin{aligned} h_{30} &= x_1y_1z_1, \\ h_{21} &= x_1y_1z_2 + x_1y_2z_1 + x_2y_1z_1, \\ h_{12} &= x_1y_2z_2 + x_2y_1z_2 + x_2y_2z_1, \\ h_{03} &= x_2y_2z_2. \end{aligned}$$

Weyl's result tells us that if the characteristic of $\mathbb{K}$ is neither 2 nor 3, then $\mathbb{K}[2\,V]^{\Sigma_3}$ is generated by the nine invariants

$$f_{10}, f_{01}, g_{20}, g_{11}, g_{02}, h_{30}, h_{21}, h_{12}, h_{03}.$$

It turns out that these nine invariants also generate $\mathbb{K}[2\,V]^{\Sigma_3}$ if $\mathbb{K}$ has characteristic 2. The identity

$$3(x_1y_1z_2^2 + y_1z_1x_2^2 + x_1z_1y_2^2) = f_{10}^2g_{02} - f_{10}f_{01}g_{11} + f_{10}h_{12} + g_{11}^2 \\ - 2f_{10}h_{12} + f_{01}^2g_{11} - 4g_{20}g_{02} + 2f_{01}h_{21}$$

shows how to express the invariant $k := x_1y_1z_2^2 + y_1z_1x_2^2 + x_1z_1y_2^2$ in terms of the polarized elementary symmetric functions when 3 is invertible. However, over a field of characteristic 3, this identify expresses an algebraic relation among the polarized elementary symmetric functions. In fact, over a field of characteristic 3, it is not possible to express k as a polynomial in the nine polarized elementary symmetric functions. In fact, it turns out that the nine polarized elementary symmetric functions together with the invariant k form a minimal generating set for $\mathbb{K}[2\,V]^{\Sigma_3}$ when $\mathbb{K}$ has characteristic 3.

Polarization and restitution may be defined more generally, as follows. Given a multi-homogeneous function

$$f \in \mathbb{K}[W_1 \oplus W_2 \oplus \cdots \oplus W_t]_{(i_1,i_2,\ldots,i_t)}$$

and given positive integers $m_1, m_2, \ldots, m_t$, we define maps

$$\Delta : W_1 \oplus W_2 \oplus \cdots \oplus W_t \to m_1\,W_1 \oplus m_2\,W_2 \oplus \cdots \oplus m_t\,W_t$$

and

$$\phi : m_1\,W_1 \oplus m_2\,W_2 \oplus \cdots \oplus m_t\,W_t \to W_1 \oplus W_2 \oplus \cdots \oplus W_t$$

given by

$$\Delta(v_1, v_2, \ldots, v_t) = (\underbrace{v_1, v_1, \ldots, v_1}_{m_1}, \underbrace{v_2, v_2, \ldots, v_2}_{m_2}, \ldots, \underbrace{v_t, v_t, \ldots, v_t}_{m_t})$$

and

$$\phi(v_{11}, v_{12}, \ldots, v_{1m_1}, \ldots, v_{t1}, v_{t2}, \ldots, v_{tm_t}) = (\sum_{j=1}^{m_1} v_{1j}, \sum_{j=1}^{m_2} v_{2j}, \ldots, \sum_{j=1}^{m_t} v_{tj}).$$

As above, these induce $\mathrm{GL}(W_1) \times \mathrm{GL}(W_2) \times \cdots \times \mathrm{GL}(W_t)$-equivariant maps $\phi^* : \mathbb{K}[W_1 \oplus W_2 \oplus \cdots \oplus W_t] \to \mathbb{K}[m_1\,W_1 \oplus m_2\,W_2 \oplus \cdots \oplus m_t\,W_t]$ and $\Delta^* : \mathbb{K}[m_1\,W_1 \oplus m_2\,W_2 \oplus \cdots \oplus m_t\,W_t] \to \mathbb{K}[W_1 \oplus W_2 \oplus \cdots \oplus W_t]$.

Given $f \in \mathbb{K}[W_1 \oplus W_2 \oplus \cdots \oplus W_t]$, the multi-homogeneous components of $\phi^*(f)$ are the partial polarizations of f. We denote the full set of these partial polarizations by $\mathcal{P}\mathrm{ol}^{m_1,m_2,\ldots,m_t}(f)$.

As above, we distinguish as a special case, the full polarization of f. This is the unique multi-linear partial polarization and we again denote it by $\mathcal{P}(f)$. The full polarization may also be described as follows. For each $k = 1, 2, \ldots, t$, we let $\mathcal{P}_k : \mathbb{K}[W_k]_{d_k} \to \mathbb{K}[d_kW_k]_{(1,1,\ldots,1)}$ denote the full polarization operator

from the k^{th} copy of V as defined earlier. Then we put $\mathcal{P} = \mathcal{P}_t\mathcal{P}_{t-1}\cdots\mathcal{P}_2\mathcal{P}_1$ which is given by

$$\mathcal{P} : \mathbb{K}[W_1 \oplus W_2 \oplus \cdots \oplus W_t]_{(d_1,d_2,\ldots,d_t)} \to \mathbb{K}[d_1W_1 \oplus d_2W_2 \oplus \cdots \oplus d_tW_t]_{(1,1,\ldots,1)}.$$

It is easy to see that these more general partial polarization operators are $\mathrm{GL}(W_1) \times \mathrm{GL}(W_2) \times \cdots \times \mathrm{GL}(W_t)-$equivariant and that $\mathcal{P}(f)$ is symmetric (invariant) under the action of $\Sigma_{d_1} \times \Sigma_{d_2} \times \cdots \times \Sigma_{d_t}$. We define a generalized restitution operator $\mathcal{R} = \mathcal{R}_{(r_1,r_2,\ldots,r_t)} = \mathcal{R}_t \circ \mathcal{R}_{t-1} \circ \cdots \circ \mathcal{R}_t \circ \mathcal{R}_1$ similarly:

$$\mathcal{R} : \mathbb{K}[r_1W_1 \oplus r_2W_2 \oplus \cdots \oplus r_tW_t] \to \mathbb{K}[W_1 \oplus W_2 \oplus \cdots \oplus W_t];$$

so that

$$\mathcal{R}(F)(\mathbf{v}_1,\ldots,\mathbf{v}_t) = F(\underbrace{\mathbf{v}_1,\ldots,\mathbf{v}_1}_{r_1},\underbrace{\mathbf{v}_2,\ldots,\mathbf{v}_2}_{r_2},\ldots,\underbrace{\mathbf{v}_t,\ldots,\mathbf{v}_t}_{r_t})$$

for $F \in \mathbb{K}[r_1W_1 \oplus r_2W_2 \oplus \cdots \oplus r_tW_t]$. Then $\mathcal{R}(\mathcal{P}(f)) = d_1!d_2!\cdots d_t!f$ if $f \in \mathbb{K}[W_1 \oplus W_2 \oplus \cdots \oplus W_t]_{(d_1,d_2,\ldots,d_t)}$. Note that unlike polarization, restitution is an algebra homomorphism.

1.10 The Role of the Cyclic Group C_p in Characteristic p

In many respects, the characteristic p invariant theory of the cyclic group C_p of order p plays a central role in modular invariant theory. In this book, we will spend considerable effort developing our understanding of C_p-invariants in characteristic p. To partially explain the importance of C_p, we begin with the following two very useful lemmas.

Lemma 1.10.1. *Suppose H is a normal subgroup of G with quotient group G/H. Let V be a representation of G. Then G/H acts naturally on V^H and $V^G = (V^H)^{G/H}$.* □

We will use Lemma 1.10.1 in the proof of the next lemma. However, its main use will be when we apply it to a normal subgroup H of a group G acting on a coordinate ring $\mathbb{K}[V]$. Then we have $\mathbb{K}[V]^G = (\mathbb{K}[V]^H)^{G/H}$. This is the topic of Chapter 14.

Lemma 1.10.2. *Let G be any p-group for p a prime and let H be any maximal proper subgroup. Then H is normal in G necessarily of index p. Hence if G is generated by H and σ, we have $G/H = C_p$ generated by $\bar{\sigma}$, the image of σ in G/H.* □

The preceding lemma shows that for any p-group G, we may construct a composition series, that is, construct a tower of groups G_i, each normal in the next such that $G_{i+1}/G_i \cong C_p$ with $G_0 = \{e\}$ and $G_m = G$.

We record this result as the following lemma.

Lemma 1.10.3. *Suppose G is a p-group. Then G is solvable with all composition factors isomorphic to C_p. That is,*

$$\{e\} = G_0 \lhd G_1 \lhd G_2 \lhd \ldots \lhd G_m = G$$

with $G_i/G_{i-1} \cong C_p$ for all $i = 1, 2, \ldots, m$. □

A consequence of this lemma is that we may compute the invariants of a p-group G by repeatedly computing invariants under an action of the cyclic group C_p. Given $\{e\} = G_0 \lhd G_1 \lhd G_2 \lhd \ldots \lhd G_m = G$, we proceed as follows. First, we compute $R_1 := \mathbb{K}[V]^{G_1}$ where $G_1 \cong C_p$. Then we compute $R_2 = \mathbb{K}[V]^{G_2} = (\mathbb{K}[V]^{G_1})^{G_2/G_1} \cong R_1^{C_p}$. Continuing in this manner we compute $R_{j+1} = \mathbb{K}[V]^{G_{j+1}} = (\mathbb{K}[V]^{G_j})^{G_{j+1}/G_j} \cong R_j^{C_p}$ for $j = 0, 1, \ldots, m$. This yields $R_{m+1} = \mathbb{K}[V]^G$. Thus, in theory at least, any composition series of G provides an inductive method of computing the G-invariants. Of course, this so-called "ladder method" is applicable to any solvable group.

In practice, this method runs into difficulties particularly for representations of G over a field $\mathbb{K}$ of characteristic p, see §14. Heuristically, the problems occur because C_p is acting on $\mathbb{K}[V]^H$ which is most often not polynomial and, in particular, is not of the form $\mathbb{K}[W]$. For modular representations, this presents special difficulties. Chapter 14 discusses this technique in detail and shows how we may use group cohomology to handle the extra difficulties that arise in the modular case.

1.11 C_p Represented on a 2 Dimensional Vector Space in Characteristic p

As a simple example of a modular group action, consider the vector space V_2 of dimension 2 over a field $\mathbb{F}$ of characteristic $p > 0$ with basis $\{e_1, e_2\}$. We start with a lengthy but elementary proof which illustrates part of the attraction of modular invariant theory. Namely, that it is possible to prove some theorems using only basic techniques.

Let C_p denote the cyclic group of order p generated by σ. Consider the matrix

$$\tau = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

inside $\mathrm{GL}(2, \mathbb{F})$ where $\mathbb{F}$ is a field of characteristic p. It is easy to show, using induction, that

$$\tau^i = \begin{pmatrix} 1 & 0 \\ i & 1 \end{pmatrix}.$$

Therefore, we obtain a representation $\rho : C_p \to \mathrm{GL}(V_2)$ given by the rule $\rho(\sigma^i) = \tau^i$. We have $\sigma(e_1) = \tau(e_1) = e_1 + e_2$ and $\sigma(e_2) = \tau(e_2) = e_2$.

Let $\{x, y\}$ be the basis for V_2^* dual to $\{e_1, e_2\}$. Then $\sigma(x) = x$ and $\sigma(y) = -x + y$.

We see immediately that the polynomial x is an invariant. Moreover, since $(y+x)^p = y^p + x^p$, the polynomial $N = y^p - x^{p-1}y$ is another example of an invariant. We will see below, in Theorem 1.11.2, that for this representation, these two invariants are the two most important invariants.

Lemma 1.11.1. $\mathbf{N}^{C_p}(y) = y^p - x^{p-1}y$. □

Our goal is to show that the ring of C_p-invariants is the algebra generated by the two invariants N and x:

Theorem 1.11.2. $\mathbb{F}[V_2]^{C_p} = \mathbb{F}[x, N]$.

Before proving Theorem 1.11.2, we need some preliminary results.

Lemma 1.11.3. *Let* $f \in \mathbb{F}[V_2]$. *Then* $\deg_y(\sigma(f)) = \deg_y(f)$.

Proof. Let m denote $\deg_y(f)$ and write $f = a_m y^m + a_{m-1}y^{m-1} + \cdots + a_0$ where $a_i \in \mathbb{F}[x]$ and $a_m \neq 0$. Then

$$\begin{aligned}\sigma(f) &= \sigma(a_m)(\sigma(y)^m) + \sigma(a_{m-1})(\sigma(y)^{m-1}) + \cdots + \sigma(a_0)\\ &= a_m(y-x)^m + a_{m-1}(y-x)^{m-1} + \cdots + a_0\\ &= a_m y^m + \text{terms of lower order in } y\end{aligned}$$

Thus $\deg_y(\sigma(f)) = m$. □

Since N is monic when considered as a polynomial in the variable y with coefficients from $\mathbb{F}[x]$, we may divide any polynomial $f \in R$ by N to get $f = qN + r$ where $q, r \in \mathbb{F}[x, y]$ are unique with $\deg_y(r) < p = \deg_y(N)$.

Lemma 1.11.4. *If* $f \in \mathbb{F}[V_2]^G$ *and* $f = qN + r$ *with* $\deg_y r < p$, *then* $q, r \in \mathbb{F}[V_2]^G$.

Proof. First we note that it is enough to show that q and r are σ-invariant since σ generates C_p.

We have $f = \sigma(f) = (\sigma \cdot q)(\sigma(N)) + (\sigma \cdot r) = (\sigma \cdot q)N + (\sigma \cdot r)$. Since $\deg_y(\sigma \cdot r) = \deg_y(r) < p$, by the uniqueness of remainders and quotients we must have $\sigma \cdot r = r$ and $\sigma \cdot q = q$. □

Now we need a result concerning the partial differential operator $\frac{\partial}{\partial y}$.

Lemma 1.11.5. *If* $f \in \mathbb{F}[x, y]^G$, *then* $\frac{\partial}{\partial y}(f) \in \mathbb{F}[x, y]^G$.

Proof. We note that it is sufficient to show that if $f \in \mathbb{F}[V]$, then $\sigma(\frac{\partial}{\partial y}(f)) = \frac{\partial}{\partial y}(\sigma f)$. Further, we note that both σ and $\frac{\partial}{\partial y}$ are $\mathbb{F}$-linear maps. Therefore, to show that they commute we need only show that they commute on monomials:

$$\sigma(\frac{\partial}{\partial y}(x^a y^b)) = \sigma(bx^a y^{b-1}) = bx^a(y-x)^{b-1} \text{ and}$$

$$\frac{\partial}{\partial y}(\sigma \cdot x^a y^b) = \frac{\partial}{\partial y}(x^a(y-x)^b) = bx^a(y-x)^{b-1}$$

□

Remark 1.11.6. The lemma above also follows from the Leibnitz' rule:

$$\frac{\partial}{\partial y}(f_1 f_2) = \frac{\partial}{\partial y}(f_1) f_2 + f_1 \frac{\partial}{\partial y}(f_2)$$

We now give the proof of Theorem 1.11.2.

Proof. Clearly, $\mathbb{F}[V_2]^{C_p} \supseteq \mathbb{F}[N, x]$. Thus it suffices to prove that each invariant, f, is contained in $\mathbb{F}[x, N]$. We prove this by induction on $\deg_y(f)$.

If $\deg_y(f) = 0$, then $f \in \mathbb{F}[x] \subset \mathbb{F}[x, N]$.

Next, suppose $\deg_y(f) = d$ and that every invariant, whose degree in y is less than d, lies in $\mathbb{F}[x, N]$. Write $f = q \cdot N + r$, where $m := \deg_y(r) < p$. We will now show $m = 0$. Assume, by way of contradiction, that $m \geq 1$ and consider the invariant h defined by

$$h := \frac{\partial^{m-1}(r)}{\partial y^{m-1}}.$$

Then $h = ay + b$, where a is a non-zero scalar and $b \in \mathbb{F}[x]$. But $h = \sigma(ay + b) = a(y - x) + b = ay + b - ax$ and this contradiction shows that we must have $m = 0$. Therefore, $f = q \cdot N + r$ where $q \in \mathbb{F}[V_2]^{C_p}$, $\deg_y(q) = d - p$ and $r \in \mathbb{F}[x]$. By the induction hypothesis, $q \in \mathbb{F}[x, N]$ and thus $f \in \mathbb{F}[x, N]$. □

In later chapters, we will discuss some elements of commutative algebra and we will be able to give a simpler proof of this result. An outline of this simpler proof is as follows. Since $\{x, N\}$ is a *homogeneous system of parameters* for $\mathbb{F}[V_2]^{C_p}$ (see §2.6 for details) and the product of their degrees equals the order of the group, then Theorem 1.11.2 follows from Theorem 3.1.6.

We pause here for a discussion of history and philosophy. The invariant theory of polynomials began in characteristic 0.

For example, at the time of Newton and Vandermonde, there was intense interest in generalizing the famous quadratic formula. The problem was to find the roots of high degree polynomials in one variable with integer coefficients by means of radicals. One method of study was to suppose the solution and study which polynomials arise: perhaps, it was thought, all of them. Let $x_1, \ldots, x_n$ denote the roots of a polynomial $f(t) = a_n t^n + \cdots + a_1 t + a_0$ for integers a_i. This means, of course, that

$$f(t) = a_n \prod_{i=0}^{n} (t - x_i)$$

and $a_0, a_1, \ldots, a_{n-1}$ are the elementary symmetric polynomials in $x_1, x_2, \ldots, x_n$. We note that the right hand side is invariant under any permutation of the variables. Therefore, in order to understand solutions of such equations by means of radicals it is possibly useful to understand the invariants of permutations of n-variables. Such considerations lead to the invariant theory of

other groups, and many books on invariant theory begin with this particular situation.

For another approach, we note that Theorem 1.11.2 tells us that the ring of invariants $\mathbb{F}[V_2]^{C_p}$ is a polynomial algebra. Characterizing the modular groups with this property, where the rings of invariants are again polynomial algebras, is still an open question, perhaps the most important open problem in this area of research. By way of contrast, the characterization of such groups in the non-modular case (Theorem 1.5.2) is one of the best-known and beautiful results in classical invariant theory. We are attracted to such results because an important property of the original algebra is preserved under the action of the group.

It is not difficult to show that $\mathbb{F}(V_2)^{C_p}$ and $\mathbb{F}(x, \mathrm{N})$ are equal using Galois Theory. It is often not too difficult to discover sets of invariants with the property that the quotient field they generate is the quotient field of the ring of invariants. Heuristically, as one is led to believe by Noether's question, fewer generators are needed at the quotient field level to generate the field of invariants. In any event, Theorem 1.11.2 is attractive also in this sense — that the generators needed for the field of invariants suffice to generate the ring of invariants as well.

The example, while simple and straightforward, offers a glimpse into the world of invariant theory. We seek to discover which invariants generate the ring of invariants and we are interested in the algebraic structures that are exhibited by the invariant ring.

1.12 A Further Example: C_p Represented on $2\,V_2$ in Characteristic p

Here we compute the ring of invariants of the group C_p on $2\,V_2 := V_2 \oplus V_2$. Here we are considering the action of C_p determined by

$$\sigma = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

We introduce

$$\sigma_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and

$$\sigma_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Thus $\sigma = \sigma_2\sigma_1^{-1}$.

Let $\{x_1, y_1, x_2, y_2\}$ be the basis for the vector space $(2\,V_2)^*$ dual to the standard basis of $2\,V_2$. Thus $\sigma_i(x_j) = x_j$ for $1 \le i, j \le 2$, $\sigma_i(y_j) = y_j$ for $1 \le j \ne i \le 2$, $\sigma_1(y_1) = y_1 + x_1$ and $\sigma_2(y_2) = y_2 - x_2$. Define G to be the group generated by σ_1 and σ_2, so that $G = C_p \times C_p$ and H to be the group generated by $\sigma = \sigma_2\sigma_1^{-1}$ so that $H = C_p$. We want to compute $\mathbb{F}[2\,V_2]^H$.

Given the example just computed, it is easy to see that

$$\mathbb{F}[2\,V_2]^G = \mathbb{F}[V_2]^{C_p} \otimes \mathbb{F}[V_2]^{C_p} = \mathbb{F}[x_1, N(y_1), x_2, N(y_2)].$$

Now we consider the diagram

$$\begin{array}{ccccc} \mathbb{F}(2\,V_2)^G & \hookrightarrow & \mathbb{F}(2\,V_2)^H & \hookrightarrow & \mathbb{F}(2\,V_2) \\ \updownarrow & & \updownarrow & & \updownarrow \\ \mathbb{F}[2\,V_2]^G & \hookrightarrow & \mathbb{F}[2\,V_2]^H & \hookrightarrow & \mathbb{F}[2\,V_2]. \end{array}$$

By Galois Theory, we have that the field $\mathbb{F}(2\,V_2)^H$ is an extension of the field $\mathbb{F}(2\,V_2)^G$ of degree p; that is, $\mathbb{F}(2\,V_2)^H$ as vector space over $\mathbb{F}(2\,V_2)^G$ has dimension p. In order to exploit this property, we need to find an element of $\mathbb{F}(2\,V_2)^H$ that lies outside of $\mathbb{F}(2\,V_2)^G$.

It is easy to see that the element $u = x_1y_2 - x_2y_1$ is an invariant of least degree in $\mathbb{F}[2\,V_2]^H$ outside $\mathbb{F}[2\,V_2]^G$, and therefore $\mathbb{F}(2\,V_2)^H$ has basis

$$\left\{1, u, u^2, \ldots, u^{p-1}\right\}.$$

We see that

$$u^p = x_1^p N(y_2) - x_2^p N(y_1) + x_1^{p-1}x_2^{p-1}u\ .$$

Theorem 1.12.1.

$$\mathbb{F}[2\,V_2]^{C_p} = \mathbb{F}[x_1, x_2, \mathbf{N}(y_1), \mathbf{N}(y_2), u]\ .$$

In fact,

$$\mathbb{F}[2\,V_2]^{C_p} = \oplus_{i=0}^{p-1}\mathbb{F}[x_1, x_2, \mathbf{N}(y_1), \mathbf{N}(y_2)]u^i\ .$$

Proof. Our challenge is to show that $\left\{1, u, u^2, \ldots, u^{p-1}\right\}$ is a basis for $\mathbb{F}[2\,V_2]^H$ as a module over $\mathbb{F}[2\,V_2]^G$.

Since G is Abelian, we have that H is normal in G and hence

$$\mathbb{F}[2\,V_2]^G = (\mathbb{F}[2\,V_2]^H)^{G/H}.$$

We note that the image of either σ_1 or σ_2 generates $G/H = C_p$.

We define

$$\begin{aligned}\Delta_1 &= \sigma_1 - \mathrm{Id}\\ \Delta_2 &= \sigma_2 - \mathrm{Id} \quad \text{and}\\ \Delta \ &= \sigma - \mathrm{Id}.\end{aligned}$$

Consider $f \in \mathbb{F}[2\,V_2]^H$ and note that $\sigma(f) = f$ implies $\sigma_1(f) = \sigma_2(f)$ and thus $\Delta_1(f) = \Delta_2(f)$. In particular, $f \in \mathbb{F}[2\,V_2]^G$ if and only if $\Delta_1(f) = 0$. Also

$$\sigma(\Delta_1(f)) = \sigma_2\sigma_1^{-1}(\sigma_1(f) - f) = \sigma_2(f) - \sigma(f) = \sigma_1(f) - f = \Delta_1(f).$$

Thus $\Delta_1 : \mathbb{F}[2\,V_2]^H \to \mathbb{F}[2\,V_2]^H$.

Lemma 1.12.2. *If $f \in \mathbb{F}[2\,V_2]^H$, then $\Delta_1(f) = x_1x_2f'$ for some $f' \in \mathbb{F}[2\,V_2]^H$.*

Proof. For any $f \in \mathbb{F}[2\,V_2]$ we write $f = \sum_{\ell=0}^{d} f_\ell y_1^\ell$ with $f_\ell \in \mathbb{F}[x_1, x_2, y_2]$ for $0 \le \ell \le d$ in order to see that $\Delta_1(f) = x_1f'$. Similarly, $\Delta_2(f) = x_2f''$. If $f \in \mathbb{F}[2\,V_2]^H$, then $\sigma_1(f) = \sigma_2(f)$, and so $x_1f' = x_2f''$. But x_1 and x_2 are co-prime in $\mathbb{F}[2\,V_2]$ and so $\Delta_1(f) = x_1x_2f'''$ for some $f''' \in \mathbb{F}[2\,V_2]$. Since both $\Delta_1(f)$ and x_1x_2 are H-invariant, we see that $f''' \in \mathbb{F}[2\,V_2]^H$. □

We now finish the proof of Theorem 1.12.1. Since $\Delta_1^p = (\sigma_1 - \mathrm{Id})^p = \sigma_1^p - \mathrm{Id} = 0$ we see that $\Delta_1^p(f) = 0$ for all $f \in \mathbb{F}[2\,V_2]$. Thus given $0 \ne f \in \mathbb{F}[2\,V_2]^H$ there must exist an ℓ, $0 \le \ell < p$ with the property that $0 \ne \Delta_1^\ell(f) \in \mathbb{F}[2\,V_2]^H$ and $\Delta_1^{\ell+1}(f) = 0$. We claim that then $f = \sum_{m=0}^{\ell} f_m u^m$ for $f_m \in \mathbb{F}[2\,V_2]^G$. We proceed by induction on ℓ. If $\ell = 0$ then $\Delta_1(f) = 0$ which implies $f \in \mathbb{F}[2\,V_2]^G$ as we observed above. For the general case, we write $\Delta_1(f) = x_1x_2f'$ with $f' \in \mathbb{F}[2\,V_2]^H$ and observe that $f' = \sum_{m=0}^{\ell-1} f'_m u^m$ with all $f'_m \in \mathbb{F}[2\,V_2]^G$ by induction. Now consider

$$\begin{aligned}\Delta_1^\ell(f + uf'/\ell) &= \Delta_1^\ell\left(f + \frac{1}{\ell}\sum_{m=0}^{\ell-1} f'_m u^{m+1}\right)\\ &= \Delta_1^{\ell-1}\left(\Delta_1(f) + \frac{1}{\ell}\sum_{m=0}^{\ell-1} f'_m \Delta_1(u^{m+1})\right)\\ &= \Delta_1^{\ell-1}\left(x_1x_2f' + \frac{1}{\ell}f'_{\ell-1}\Delta_1(u^\ell) + \frac{1}{\ell}\sum_{m=0}^{\ell-2} f'_m \Delta_1(u^{m+1})\right)\\ &= \Delta_1^{\ell-1}\Bigg(\sum_{m=0}^{\ell-1} x_1x_2f'_m u^m \frac{1}{\ell}f'_{\ell-1}\sum_{i=0}^{\ell-1}\binom{\ell}{i}u^i(-x_1x_2)^{\ell-i}\\ &\qquad + \frac{1}{\ell}\sum_{m=0}^{\ell-2} f'_m \Delta_1(u^{m+1})\Bigg)\end{aligned}$$

$$= \Delta_1^{\ell-1}\Big(\sum_{m=0}^{\ell-2} x_1x_2f'_mu^m + \frac{1}{\ell}f'_{\ell-1}\sum_{i=0}^{\ell-2}\binom{\ell}{i}u^i(-x_1x_2)^{\ell-i} + \frac{1}{\ell}\sum_{m=0}^{\ell-2} f'_m\Delta_1(u^{m+1})\Big)$$

$$= \Delta_1^{\ell-1}\Big(\sum_{m=0}^{\ell-2} h_mu^m\Big)$$

where $h_m \in \mathbb{F}[2\,V_2]^G$ for $m = 1, 2, \ldots, \ell-2$ and this final expression is equal to zero since, as is easily verified, $\Delta_1^t(u^s) = 0$ whenever $t > s$. Therefore, $f + uf'/\ell \in \oplus_{m=0}^{\ell-1}\mathbb{F}[2\,V_2]^Gu^m$ by the induction hypothesis. Thus $f \in \oplus_{m=0}^{\ell}\mathbb{F}[2\,V_2]^Gu^m$ which proves Theorem 1.12.1. □

1.13 The Vector Invariants of V_2

Given a representation V of a group G and an integer $m \geq 2$, a ring of invariants $\mathbb{K}[m\,V]^G$ is called a ring of vector invariants. A theorem providing an explicit description of $\mathbb{K}[m\,V]^G$ for all $m \geq 1$ is called a *first fundamental (or main) theorem* for V. The following first fundamental theorem for V_2 was conjectured by David Richman and proved by Campbell and Hughes, see [19]. Their proof is technical and uses a deep result about the rank of zero-one matrices in characteristic p. We will give a shorter proof which uses ideas from this book and which has the advantage that it yields more than just a generating set; it yields a SAGBI basis as we shall see in §7.4.

Theorem 1.13.1. *Let $G = C_p = \langle\sigma\rangle$ act on $V = m\,V_2$. Let $\{y_i, x_i\}$ denote a basis for the i^{th} copy of V_2^* in V^* where $\sigma(y_i) = y_i + x_i$ and $\sigma(x_i) = x_i$. Thus $\{x_1, y_1, x_2, y_2, \ldots, x_m, x_m\}$ is an upper triangular basis for V^*. Then the ring of invariants $\mathbb{F}[m\,V_2]^{C_p}$ is generated by the following invariants:*

1. *x_i for $i = 1, 2, \ldots, m$.*
2. *$\mathbf{N}^{C_p}(y_i) = y_i^p - x_i^{p-1}y_i$ for $i = 1, 2, \ldots, m$.*
3. *$u_{ij} = x_jy_i - x_iy_j$ for $1 \leq i < j \leq m$.*
4. *$\mathrm{Tr}^{C_p}(y_1^{a_1}y_2^{a_2}\ldots y_m^{a_m})$ where $0 \leq a_i < p$ for $i = 1, 2, \ldots, m$.*

Remark 1.13.2. Shank and Wehlau [99] showed that if $a_1+a_2+\cdots+a_m \leq 2(p-1)$, then $\mathrm{Tr}^{C_p}(y_1^{a_1}y_2^{a_2}\ldots y_m^{a_m})$ lies in the subalgebra generated by $x_1, x_2, \ldots, x_m$ and u_{ij} with $1 \leq i < j \leq m$. Additionally, they also showed that if we exclude invariants of this form, the remaining invariants *minimally* generate $\mathbb{F}[m\,V_2]^{C_p}$.

The following example illustrates Theorem 1.13.1.

Example 1.13.3. If we take $m = 3$ and $\mathbb{F}$ a field of characteristic $p = 3$, then Theorem 1.13.1 tells us that $\mathbb{F}[3\,V_2]^{C_3}$ is generated by x_1, x_2, x_3, $\mathbf{N}(y_1), \mathbf{N}(y_2), \mathbf{N}(y_3)$, u_{12}, u_{13}, u_{23} and some transfers.

It is straightforward to compute

$$\begin{aligned}
\mathrm{Tr}^{C_3}(y_i) &= 0 && \text{for } i = 1, 2, 3;\\
\mathrm{Tr}^{C_3}(y_i y_j) &= -x_i x_j && \text{for } 1 \le i, j \le 3;\\
\mathrm{Tr}^{C_3}(y_i^2 y_j) &= x_i u_{ji} && \text{for } 1 \le i \ne j \le 3;\\
\mathrm{Tr}^{C_3}(y_1 y_2 y_3) &= x_1 u_{23} - x_3 u_{12}; &&\\
\mathrm{Tr}^{C_3}(y_i^2 y_j^2) &= -u_{ij} - x_i^2 x_j^2 && \text{for } 1 \le i < j \le 3;\\
\mathrm{Tr}(y_i y_1 y_2 y_3) &= -u_{ij} u_{ik} - x_i^2 x_j x_k && \text{where } \{i, j, k\} = \{1, 2, 3\}.
\end{aligned}$$

Thus we see, in agreement with Remark 1.13.2, that $\mathbb{F}[3\,V_2]^{C_3}$ is minimally generated by

$$x_1,\ x_2,\ x_3,\ \mathbf{N}(y_1),\ \mathbf{N}(y_2),\ \mathbf{N}(y_3),\ u_{12},\ u_{13},\ u_{23},\ \mathrm{Tr}^{C_3}(y_1^2 y_2^2 y_3),$$
$$\mathrm{Tr}^{C_3}(y_1^2 y_2 y_3^2),\ \mathrm{Tr}^{C_3}(y_1 y_2^2 y_3^2) \text{ and } \mathrm{Tr}^{C_3}(y_1^2 y_2^2 y_3^2).$$

Remark 1.13.4. The proof of Theorem 1.13.1 (see §7.4) shows in particular that the invariant $\mathrm{Tr}^{C_p}(y_1^{p-1} y_2^{p-1} \dots y_m^{p-1})$ cannot be expressed using only invariants of lower degree and thus the Noether $\beta(m\,V_2, C_p) \ge m(p-1)$. (Of course, the theorem also shows that we have equality here.) Similarly, in Corollary 7.7.3, we show that (over a field of characteristic p) $\beta(m\,V_r, C_p) > m(p-1)$ if $r \ge 3$. Thus, unlike the non-modular situation, in the modular setting, there can be no general bound on $\beta(V, G)$ independent of V. In fact, fix any field $\mathbb{K}$ and any linear algebraic group G and consider $\beta(G) := \sup\{\beta(V, G) \mid G \le \mathrm{GL}_{\mathbb{K}}(V)\}$. Bryant and Kemper [15] showed that $\beta(G)$ is finite if and only if G is a finite group whose order is invertible in $\mathbb{K}$.

http://www.springer.com/978-3-642-17403-2

Modular Invariant Theory
Campbell, H.E.A.E.; Wehlau, D.
2011, XIV, 234 p., Hardcover
ISBN: 978-3-642-17403-2