Encyclopaedia of Mathematical Sciences 139

Invariant Theory and Algebraic Transformation Groups VIII

Revaz V. Gamkrelidze · Vladimir I. Popov *Subseries Editors*

H. E. A. Eddy Campbell
David L. Wehlau

Modular Invariant Theory

Springer

Encyclopaedia of Mathematical Sciences
Volume 139

Invariant Theory and Algebraic Transformation Groups VIII

H.E.A. Eddy Campbell • David L. Wehlau

Modular Invariant Theory

Springer

H.E.A. Eddy Campbell
University of New Brunswick
Sir Howard Douglas Hall
Dept. Mathematics
Bailey Drive 3
Fredericton, New Brunswick
Canada E3B 5A3
heac@unb.ca

David L. Wehlau
Royal Military College of Canada
Dept. Mathematics & Computer Science
Kingston, Ontario
Canada K7K 7B4
wehlau@rmc.ca

Founding editor of the Encyclopaedia of Mathematical Sciences: Revaz V. Gamkrelidze

ISSN 0938-0396
ISBN 978-3-642-17403-2 e-ISBN 978-3-642-17404-9
DOI 10.1007/978-3-642-17404-9
Springer Heidelberg Dordrecht London New York

Mathematics Subject Classification (2010): 13A50

Cover design: deblik

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

For Diane Mary Brennan, Ian Alexander Brennan, Colin Cameron Brennan,
Graham Harold James Brennan and Maggie Orion Cameron.

For Charlene Lynn Janeway and Megan Melinda Jane Wehlau.

Preface

At the time we write this book there are several excellent references available which discuss various aspects of modular invariant theory from various points of view: Benson [6]; Derksen and Kemper [26]; Neusel [85]; Neusel and Smith [86]; and Smith [103]. In this book, we concentrate our attention on the modular invariant theory of finite groups. We have included various techniques for determining the structure of and generators for modular rings of invariants, while attempting to avoid too much overlap with the existing literature. An important goal has been to illustrate many topics with detailed examples. We have contrasted the differences between the modular and non-modular cases, and provided instances of our guiding philosophies and analogies. We have included a quick survey of the elements of algebraic geometry and commutative algebra as they apply to invariant theory. Readers who are familiar with these topics may safely skip this chapter.

We wish to thank our principal collaborators over the years with whom we have had so much pleasure exploring this fascinating subject: Ian Hughes, Gregor Kemper, R. James Shank, John Harris as well as our students and friends, Jianjun Chuai, Greg Smith, Mike Roth, Brandon Fodden, Emilie Dufresne, Asia Matthews and Chester Weatherby. In particular we thank John Harris, R. James Shank, Jianjun Chuai, Mike Roth, Emilie Dufresne, Asia Matthews, Chester Weatherby and Tristram Bogart for reading draft chapters and pointing out errors and suggesting improvements. We also thank Marie-José Bertin for clarifying the history of her own work to us.

Finally, our thanks go to the anonymous referees for many helpful and constructive remarks.

Fredericton, New Brunswick
Kingston, Ontario
August 2010

H. E. A. Eddy Campbell
David L. Wehlau

Contents

Index of notations

$\mathbb{N}$		the natural numbers, including 0.
$\mathbb{Z}$		the integers.
$\mathbb{Q}$		the rational numbers.
$\mathbb{R}$		the real numbers.
$\mathbb{C}$		the complex numbers.
$\mathbb{F}$		a field of characteristic $p \neq 0$.
$\mathbb{F}^\times$		$\mathbb{F} \setminus \{0\}$.
$\mathbb{F}_p$		the prime field of order p.
$\mathbb{K}$		a field of arbitrary characteristic.
$\overline{\mathbb{K}}$		the algebraic closure of $\mathbb{K}$.
$\otimes$		$\otimes_{\mathbb{K}}$ or $\otimes_{\mathbb{F}}$.
V		a vector space over $\mathbb{F}$ or $\mathbb{K}$, most often of dimension n.
$\mathrm{GL}(V)$		the general linear group of V.
$\mathrm{GL}_n(\mathbb{K})$		the linear automorphisms of $\mathbb{K}^n$.
$\mathrm{SL}(V)$		the special linear group of V.
$\mathrm{SL}_n(\mathbb{K})$		the linear automorphisms of $\mathbb{K}^n$ of determinant 1.
G		a group, often a subgroup of $\mathrm{GL}(V)$.
σ, τ		group elements.
$\sigma(v), \sigma \cdot v$	2	the image $v \in V$ under the action of σ.
V^G	2	the subspace of V fixed (point-wise) by the action of G.
G_X	2	the stabilizer or isotropy subgroup of $X \subset V$.
G_v	2	the stabilizer or isotropy subgroup of $v \in V$.
V^*		the hom-dual $\hom_{\mathbb{K}}(V, \mathbb{K})$.
x^I	3	the monomial $x_1^{i_1} \cdots x_n^{i_n}$ for $I = (i_1, \dots, i_n\}$.
$\mathbb{F}[V]$	3	the coordinate ring of V.
$\mathbb{F}[V]^G$	4	the invariant ring.
$\mathrm{Tr}^G(f)$	6	the trace of $f \in \mathbb{F}[V]$.
$\mathrm{Tr}^G_H(f)$	6	the trace of $f \in \mathbb{F}[V]$ relative to a subgroup $H \subset G$.
$\mathrm{N}(f)^G$	6	the norm of f.
$\mathrm{N}^G_H(f)$	7	the norm of f relative to a subgroup $H \subset G$.
Gf	7	the orbit of f under G, i.e., $Gf := \{\sigma(f) \mid \sigma \in G\}$.
$\beta(G, V)$	8	the Noether degree bound for the invariant ring.
$\lvert G \rvert$		the order of G.
V^σ	8	the elements of V fixed by σ.
$\mathbb{K}(V)^G$	10	the field of fractions of $\mathbb{K}[V]^G$.
$\mathrm{Quot}(\mathbb{K}[V]^G)$	10	the field of fractions of $\mathbb{K}[V]^G$.
mV	11	the vector space $V^{\oplus m}$.
$\mathcal{P}\mathrm{ol}$	11	polarization.
$\mathcal{R}$	11	restitution.
$\deg_y(f)$	18	the degree of f in the variable y.
Δ_σ	22	the twisted derivation $\sigma - 1$.
$\mathcal{V}(T)$	26	the zero-set of the polynomials in T.
$\mathcal{I}(X)$	26	the ideal of the variety or set X.

$\sqrt{I}$	26	the radical of the ideal I.
$\mathrm{Spec}(S)$	27	the space of all prime ideals of S.
$\mathrm{MaxSpec}(S)$	27	the space of all maximal prime ideals of S.
R		a graded commutative ring with 1.
R_f	29	the localization of R at f.
$\mathcal{H}(R, \lambda)$	34	the Hilbert series of R.
$V/\!\!/G$	42	the quotient variety.
$(V/\!\!/G)_{\mathrm{good}}$	56	the "good" locus.
$\mathrm{LM}(f)$	83	the lead monomial of f (with respect to some ordering).
$\mathrm{LT}(f)$	83	the lead term of f.
$\mathrm{LT}(R)$	85	the lead term algebra of R.
$\mathbb{K}[V]_G$	99	the ring of coinvariants.
$G_a(\mathbb{C})$	125	the additive group of $\mathbb{C}$.
$\mathrm{RL}_n(\mathbb{F})$	150	the subgroup of $\mathrm{GL}_n(\mathbb{F})$ generated by its reflections.
$\mathrm{GU}_n(\mathbb{F})$	150	the unitary group of dimension n over $\mathbb{F}$.
$\mathrm{RU}_n(\mathbb{F})$	150	the subgroup of $\mathrm{GU}_n(\mathbb{F})$ generated by its reflections.
$\mathrm{GO}_n(\mathbb{F})$	150	the general orthogonal group over $\mathbb{F}$.
$\mathrm{SO}_n(\mathbb{F})$	150	the special orthogonal group over $\mathbb{F}$.
$\Omega_n^{\pm}(\mathbb{F})$	150	the commutator subgroup of $\mathrm{GO}_n(\mathbb{F})$.

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, © Springer-Verlag Berlin Heidelberg 2011

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

$$f(t_1v_1, t_2v_2, \ldots, t_mv_m) = t_1^{d_1} t_2^{d_2} \cdots t_m^{d_m} f(v_1, v_2, \ldots, v_m)$$

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 multihomogeneous. 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 $f_1=e_1^3+\dots$. 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

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

and

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

It is easy to see that for $f \in \mathbb{K}[V]^H$ the elements $\mathrm{Tr}^G_H(f)$ and $\mathbf{N}^G_H(f)$ 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 $\mathrm{Tr}^G_H(f)$ and $\mathbf{N}^G_H(f)$ are both G-invariants. Note that $\mathrm{Tr}^G(f) = \mathrm{Tr}^G_{\{e\}}(f)$ and $\mathbf{N}^G(f) = \mathbf{N}^G_{\{e\}}(f)$.

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 $\operatorname{Im}(\sigma - 1 : V \to V)$ 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

$$\begin{tikzcd}
\mathbb{K}[V]^G \arrow[r, hook] \arrow[d, hook] & \mathbb{K}[V] \arrow[d, hook] \\
\mathbb{K}(V)^G \arrow[r, hook] & \mathbb{K}(V)
\end{tikzcd}$$

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 $\{a_1, \dots a_q\}$ 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 $\varDelta : V \to m\,V = \underbrace{V \oplus V \oplus \cdots \oplus V}_{m \text{ copies}}$ and $\phi : m\,V \to V$ given by $\varDelta(v) = (v, v, \ldots, v)$ 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 $\varDelta^* : \mathbb{F}[m\,V] \to \mathbb{F}[V]$ and $\phi^* : \mathbb{F}[V] \to \mathbb{F}[m\,V]$ given by $(\varDelta^*(F))(v) = F(\varDelta(v)) = F(v, v, \ldots, v)$ 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 $f_{(i_1,i_2,\dots,i_m)} \in \mathbb{F}[m\,V]_{(i_1,i_2,\dots,i_m)}$. These polynomials $f_{(i_1,i_2,\dots,i_m)}$ are called *partial polarizations* of f and we write

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

to denote the set of all such partial polarizations.

In order to compute individual polarizations, we take m indeterminates $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)$, and consider $\mathbf{v} = (v_1, v_2, \dots, v_m)$ 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

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

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 $(i_1, i_2, \dots, i_m) = (1, 1, \dots, 1)$ to get the *full polarization* of f denoted

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

Lemma 1.9.1. *The mapping $f \mapsto f_{(i_1,i_2,\dots,i_m)}$ 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 $\tau(\mathbf{v}) = (v_{\tau(1)}, \dots, v_{\tau(m)})$ 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

$$\begin{aligned} f(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2) &= f(\lambda_1x_1 + \lambda_2x_2, \lambda_1y_1 + \lambda_2y_2, \lambda_1z_1 + \lambda_2z_2) \\ &= \lambda_1(x_1 + y_1 + z_1) + \lambda_2(x_2 + y_2 + z_2). \end{aligned}$$

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,

$$\begin{aligned} g(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2) &= g(\lambda_1x_1 + \lambda_2x_2, \lambda_1y_1 + \lambda_2y_2, \lambda_1z_1 + \lambda_2z_2) \\ &= \lambda_1^2(x_1y_1 + x_1z_1 + y_1z_1) \\ &\quad + \lambda_1\lambda_2(x_1y_2 + x_1z_2 + y_1x_2 + y_1z_2 + z_1x_2 + z_1y_2) \\ &\quad + \lambda_2^2(x_2y_2 + x_2z_2 + y_2z_2). \end{aligned}$$

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

$$\begin{aligned} &h(\lambda_1x_1 + \lambda_2x_2, \lambda_1y_1 + \lambda_2y_2, \lambda_1z_1 + \lambda_2z_2) \\ &\quad = \lambda_1^3(x_1y_1z_1) + \lambda_1^2\lambda_2(x_1y_1z_2 + x_1y_2z_1 + x_2y_1z_1) \\ &\qquad + \lambda_1\lambda_2^2(x_1y_2z_2 + x_2y_1z_2 + x_2y_2z_1) + \lambda_2^3x_2y_2z_2. \end{aligned}$$

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

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

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 $\mathcal{P}_k : \mathbb{K}[W_k]_{d_k} \to \mathbb{K}[d_k W_k]_{(1,1,\ldots,1)}$ 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,*

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

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 $\{e\} = G_0 \triangleleft G_1 \triangleleft G_2 \triangleleft \ldots \triangleleft G_m = G$, 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:

$$\begin{aligned} \sigma(\frac{\partial}{\partial y}(x^a y^b)) &= \sigma(b x^a y^{b-1}) = b x^a (y-x)^{b-1} \text{ and} \\ \frac{\partial}{\partial y}(\sigma \cdot x^a y^b) &= \frac{\partial}{\partial y}(x^a (y-x)^b) = b x^a (y-x)^{b-1} \end{aligned}$$

□

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 $1 \leq i, j \leq 2$, $\sigma_i(y_j) = y_j$ for $1 \leq j \neq i \leq 2$, $\sigma_1(y_1) = y_1 + x_1$ 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

$$\begin{tikzcd}
\mathbb{F}(2\,V_2)^G \arrow[r, hook] & \mathbb{F}(2\,V_2)^H \arrow[r, hook] & \mathbb{F}(2\,V_2) \\
\mathbb{F}[2\,V_2]^G \arrow[r, hook] \arrow[u, hook] & \mathbb{F}[2\,V_2]^H \arrow[r, hook] \arrow[u, hook] & \mathbb{F}[2\,V_2]. \arrow[u, hook]
\end{tikzcd}$$

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

$$\begin{aligned} \varDelta_1 &= \sigma_1 - \mathrm{Id} \\ \varDelta_2 &= \sigma_2 - \mathrm{Id} \quad \text{and} \\ \varDelta \ \ &= \sigma - \mathrm{Id}. \end{aligned}$$

Consider $f \in \mathbb{F}[2\,V_2]^H$ and note that $\sigma(f) = f$ implies $\sigma_1(f) = \sigma_2(f)$ and thus $\varDelta_1(f) = \varDelta_2(f)$. In particular, $f \in \mathbb{F}[2\,V_2]^G$ if and only if $\varDelta_1(f) = 0$. Also

$$\sigma(\varDelta_1(f)) = \sigma_2\sigma_1^{-1}(\sigma_1(f) - f) = \sigma_2(f) - \sigma(f) = \sigma_1(f) - f = \varDelta_1(f).$$

Thus $\varDelta_1 : \mathbb{F}[2\,V_2]^H \to \mathbb{F}[2\,V_2]^H$.

Lemma 1.12.2. *If $f \in \mathbb{F}[2\,V_2]^H$, then $\varDelta_1(f) = x_1x_2f'$ 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 $\varDelta_1(f) = x_1f'$. Similarly, $\varDelta_2(f) = x_2f''$. 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 $\varDelta_1(f) = x_1x_2f'''$ for some $f''' \in \mathbb{F}[2\,V_2]$. Since both $\varDelta_1(f)$ 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 $\varDelta_1^p = (\sigma_1 - \mathrm{Id})^p = \sigma_1^p - \mathrm{Id} = 0$ we see that $\varDelta_1^p(f) = 0$ 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 $0 \ne \varDelta_1^\ell(f) \in \mathbb{F}[2\,V_2]^H$ and $\varDelta_1^{\ell+1}(f) = 0$. 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 $\varDelta_1(f) = 0$ which implies $f \in \mathbb{F}[2\,V_2]^G$ as we observed above. For the general case, we write $\varDelta_1(f) = x_1x_2f'$ 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

$$\begin{aligned} \varDelta_1^\ell(f + uf'/\ell) &= \varDelta_1^\ell\left(f + \frac{1}{\ell}\sum_{m=0}^{\ell-1} f'_m u^{m+1}\right) \\ &= \varDelta_1^{\ell-1}\left(\varDelta_1(f) + \frac{1}{\ell}\sum_{m=0}^{\ell-1} f'_m \varDelta_1(u^{m+1})\right) \\ &= \varDelta_1^{\ell-1}\left(x_1x_2f' + \frac{1}{\ell}f'_{\ell-1}\varDelta_1(u^\ell) + \frac{1}{\ell}\sum_{m=0}^{\ell-2} f'_m \varDelta_1(u^{m+1})\right) \\ &= \varDelta_1^{\ell-1}\left(\sum_{m=0}^{\ell-1} x_1x_2f'_m u^m \frac{1}{\ell}f'_{\ell-1}\sum_{i=0}^{\ell-1}\binom{\ell}{i}u^i(-x_1x_2)^{\ell-i}\right. \\ &\qquad \left. + \frac{1}{\ell}\sum_{m=0}^{\ell-2} f'_m \varDelta_1(u^{m+1})\right) \end{aligned}$$

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

$$= \Delta_1^{\ell-1}\Bigg(\sum_{m=0}^{\ell-2} h_m u^m\Bigg)$$

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, $f + uf'/\ell \in \oplus_{m=0}^{\ell-1}\mathbb{F}[2\,V_2]^G u^m$ by the induction hypothesis. Thus $f \in \oplus_{m=0}^{\ell}\mathbb{F}[2\,V_2]^G u^m$ 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 $\mathrm{Tr}^{C_p}(y_1^{p-1} y_2^{p-1} \ldots y_m^{p-1})$ 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 $\mathbb{K}$.

2

Elements of Algebraic Geometry and Commutative Algebra

In this chapter we summarize the basic elements of algebraic geometry and commutative algebra that are useful in the study of (modular) invariant theory. Normally, these techniques are most useful in questions about the structure of rings of invariants, and, as well, they seem to be most useful in proving theorems that hold true for all groups, modular or not. It is worth noting that a large fraction (as much as one third — see the paper of Fisher [37, Page 146]) of the papers in mathematics in the latter stages of the 19th century were studies of invariant theory. It is worth noting as well that commutative algebra was invented, discovered if you prefer, by Hilbert, in order to clarify and understand invariant theory more fully, see Fisher [37]. There are several excellent references, including Atiyah and Macdonald [5], Dummit and Foote [32], Eisenbud [33], Lang [74], Matsumura [79], and a forthcoming book by Kemper, [67].

2.1 The Zariski Topology

Algebraic geometry uses algebra to study geometric objects and vice versa. Given a vector space V defined over a field $\mathbb{K}$ we study the $\mathbb{K}$-valued polynomial functions on V. Let $\{x_1, x_2, \ldots, x_n\}$ be a basis for $V^* = \hom_{\mathbb{K}}(V, \mathbb{K})$, the vector space of linear functionals on V. We consider the polynomial ring $S = \mathbb{K}[x_1, x_2, \ldots, x_n]$. Here each x_i is naturally a (linear) function on V and we consider a polynomial in S to be a function on V by using the usual pointwise definitions for products and sums of functions. Thus each element of $f \in S$ is a $\mathbb{K}$-valued function, $f : V \to \mathbb{K}$ with domain V.

We would like to use the geometry of V and the action of G on V to study S and its structure as a G-module. However, there is a problem for us since we will most often be working over a *finite* field $\mathbb{F}$. If $\mathbb{F}$ is finite then so is V and thus there are only finitely many functions from V to $\mathbb{F}$. However, the polynomial ring S is always infinite. For example, if $\mathbb{F} = \mathbb{F}_p$, the field of order p, then the two polynomials x_1 and x_1^p represent the same function from V to

H.E.A. Eddy Campbell, D. Wehlau, *Modular Invariant Theory*,
Encyclopaedia of Mathematical Sciences 139,
DOI 10.1007/978-3-642-17404-9_2, © Springer-Verlag Berlin Heidelberg 2011

$\mathbb{F}_p$. There are a few ways around this difficulty. We will proceed by extending our field $\mathbb{K}$ to an algebraically closed field $\overline{\mathbb{K}}$. We will work with both $\mathbb{K}$ and $\overline{\mathbb{K}}$ and then interpret what our results say about invariants over $\mathbb{K}$. We write $\overline{V} := V \otimes_{\mathbb{K}} \overline{\mathbb{K}}$.

We define the *coordinate ring* of V by $\mathbb{K}[V] := \mathbb{K}[x_1, x_2, \ldots, x_n]$ and the coordinate ring of $\overline{V}$ by $\overline{\mathbb{K}}[\overline{V}] := \overline{\mathbb{K}}[x_1, x_2, \ldots, x_n]$. In this section, the connection of this coordinate ring $\overline{\mathbb{K}}[\overline{V}]$, to the geometry of $\overline{V}$, is explored.

Given a set T of polynomials in S we define $\mathcal{V}(T) = \mathcal{V}_{\overline{V}}(T)$ to be the subset of $\overline{V}$ on which every element of T vanishes, that is,

$$\mathcal{V}_{\overline{V}}(T) = \mathcal{V}(T) = \{\mathbf{v} \in \overline{V} \mid f(\mathbf{v}) = 0 \text{ for all } f \in T\}.$$

We say that $\mathcal{V}(T)$ is the *variety* that is "cut out" by the set T. A subset X of $\overline{V}$ which can be realized as $X = \mathcal{V}(T)$ for some subset T of S is called an *algebraic subset* of $\overline{V}$.

Given such an algebraic subset X of $\overline{V}$ we may view functions on $\overline{V}$ as functions on X via restriction. By definition, the coordinate ring of the subvariety X is the set of such restricted functions. Thus restriction gives a surjective algebra map, res : $\overline{\mathbb{K}}[\overline{V}] \to \overline{\mathbb{K}}[X]$. Hence $\overline{\mathbb{K}}[X] \cong \overline{\mathbb{K}}[\overline{V}]/\mathcal{I}(X)$ where the ideal $\mathcal{I}(X)$ is the kernel of res. Thus $\mathcal{I}(X) = \mathcal{I}_S(X)$ is the subset of polynomials in S vanishing on X, that is,

$$\mathcal{I}_S(X) = \mathcal{I}(X) = \{f \in S \mid f(\mathbf{v}) = 0 \text{ for all } \mathbf{v} \in X\}.$$

The following celebrated theorem of Hilbert holds, see [5, pg 85].

Theorem 2.1.1. *(Hilbert's Nullstellensatz) Let $S = \mathbb{K}[x_1, x_2, \ldots, x_n]$, where $\mathbb{K}$ is algebraically closed. If I is an ideal of S then $\mathcal{I}(\mathcal{V}(I)) = \sqrt{I}$.*

The following two results will be useful later on.

Lemma 2.1.2. *Let $v_1, v_2, \ldots, v_m$ be distinct elements of the $\mathbb{K}$-vector space V. Then there exists a (non-homogeneous) $h \in \mathbb{K}[V]$ such that*

$$h(v_i) = \begin{cases} 1 & \text{if } i = 1, \\ 0 & \text{if } i \neq 1 \end{cases}$$

Corollary 2.1.3. *Let $v_1, v_2, \ldots, v_m$ be distinct elements of the $\mathbb{K}$-vector space V. Let $c_1, c_2, \ldots, c_m \in \mathbb{K}$ be given. Then there exists $f \in \mathbb{K}[V]$ such that $f(v_i) = c_i$ for $i = 1, 2, \ldots, m$.*

The concept of the coordinate ring of V can be extended to any algebraic subset $X \subseteq V$ by considering the functions on V restricted to X. Of course, any function of $\mathcal{I}(X)$ when restricted to X gives the zero function. Thus two functions $f_1, f_2 \in S = \mathbb{K}[V]$ have the same restriction to X if and only if $f_1 - f_2 \in \mathcal{I}(X)$. Thus studying the restrictions of elements of S to X is equivalent to studying the ring $S/\mathcal{I}(X)$. This ring, $S/\mathcal{I}(X)$, is called the coordinate ring of the subset X and is denoted $\mathbb{K}[X]$.

We note that $\mathbb{K}[X]$ will fail to be a domain when $I(X)$ is not prime, i.e., when X is not irreducible. However, $\mathbb{K}[X]$ does not have nilpotent elements because $\mathcal{I}(X)$ is a radical ideal. In fact, as we shall see in the next section, every Noetherian $\mathbb{K}$ algebra without nilpotent elements can be realized as the coordinate ring of some variety X.

2.2 The Topological Space Spec(S)

Given a coordinate ring $S = \overline{\mathbb{K}}[\overline{V}]$ as above, we consider two sets; MaxSpec(S) denotes the set of all maximal ideals of S, and Spec(S) denotes the set of all prime ideals of S. Minimal closed subsets of $\overline{V}$ correspond to the maximal ideals of S so that there is a one-to-one correspondence between the points of $\overline{V}$ and the elements of MaxSpec(S). This correspondence identifies $\overline{V}$ with MaxSpec(S). Under this identification, the Zariski topology on $\overline{V}$ makes MaxSpec(S) into a topological space.

This topology passes to Spec(S) as well: $M \subset \text{Spec}(S)$ is closed if and only if there exists an ideal J of S such that M consists of all the prime ideals of S which contain J, i.e, if and only if $M = \{I \in \text{Spec}(S) \mid I \supseteq J\}$. Hence a point of the topological space Spec(S), i.e., a prime ideal of S, is closed in this topology if and only if it is a maximal ideal of S. This topology on Spec(S) is also referred to as the *Zariski topology.*

We now return to the question of non-algebraically closed fields. Often, our field $\mathbb{K}$ is not algebraically closed, for example, it is often a finite field. Since our main object of study is $\mathbb{K}[V] = \mathbb{K}[x_1, x_2, \ldots, x_n]$, we prefer not to merely extend the base field and study $\overline{\mathbb{K}}[\overline{V}]$. Instead, we compromise and consider the elements of $\mathbb{K}[x_1, x_2, \ldots, x_n]$ as functions from $\overline{V}$ to $\overline{\mathbb{K}}$, i.e., we view $\mathbb{K}[V]$ as a sub-ring of $\overline{\mathbb{K}}[\overline{V}]$. With this bigger domain, each of the (infinitely many) elements of $\mathbb{K}[V]$ corresponds to a different function on $\overline{V}$. This allows us to treat the elements of $\mathbb{K}[V]$ solely as functions and so to relate the algebraic properties of $\mathbb{K}[V]$ to the geometric properties of $\overline{V}$.

2.3 Noetherian Rings

We will always study rings that are commutative with a 1. There are many good references including Atiyah and Macdonald [5], Bruns and Herzog [14], Dummit and Foote [32], Eisenbud [34], Land [75], Matsumura [80], Zariski and Samuel [116] and [117].

We assume that R is non-negatively graded, that is $R = \oplus_{d=0}^{\infty} R_d$ where each R_d is a $\mathbb{K}$ vector space and $R_i \cdot R_j \subseteq R_{i+j}$ for all $i, j \in \mathbb{N}$. We further suppose that R is connected, i.e., that $R_0 = \mathbb{K}$. Finally, we will also assume that R is finitely generated. This means that there is a finite collection of elements $Z = \{f_1, f_2, \ldots, f_t\} \subset R$ such that every element of R may be written as a $\mathbb{K}$ linear combination of products whose factors are elements of Z: $R = \mathbb{K}[f_1, f_2, \ldots, f_t]$. If R is a graded connected finitely generated $\mathbb{K}$-algebra and all of the generators $f_1, f_2, \ldots, f_t$ may be taken in R_1 then we say that R is a *standard graded* $\mathbb{K}$-algebra. We note that a finitely generated $\mathbb{K}$-algebra is always *Noetherian*, i.e., every ideal I of R is finitely generated: $I = (g_1, g_2, \ldots, g_s) = Rg_1 + Rg_2 + \cdots + Rg_s$.

An element $h \in R$ is homogeneous (of degree d) if there exists d such that $h \in R_d$. An ideal I of R is *graded* (or *homogeneous*) if it generated by homogeneous elements. Equivalently, I is graded if whenever $h \in I$ and we write $h = h_0 + h_1 + \cdots + h_d$ with $h_i \in R_i$, we have all $h_1, h_2, \ldots, h_d \in I$. The ideal $R_+ := \oplus_{d=1}^{\infty} R_d$ is the unique maximal homogeneous ideal in R.

The prototypical example of a finitely generated $\mathbb{K}$-algebra is the polynomial ring $\mathbb{K}[x_1, x_2, \ldots, x_n]$, where the x_i are (algebraically independent) indeterminates. If $I \subset R$ is a graded ideal then the quotient ring R/I is also a graded ring via $(R/I)_d = R_d/I_d$. Furthermore, if R is a standard graded algebra, then R/I is also standard graded. In fact, it can be shown that every standard graded $\mathbb{K}$-algebra arises in this manner.

Suppose R is a finitely generated $\mathbb{K}$-algebra which is a domain. Let $f_1, f_2, \ldots, f_m$ denote the set of generators for R. Consider the polynomial ring $S = \mathbb{K}[x_1, x_2, \ldots, x_m]$. We may view S as the coordinate ring of the $\mathbb{K}$-vector space V of dimension m, i.e., $S = \mathbb{K}[V]$ where we identify the indeterminates $x_1, x_2, \ldots, x_m$ with a basis of V^*. Then we have a surjection

$$\pi : S \to R$$

given by $\pi(x_i) = f_i$. Let I denote the kernel of π. Since we have assumed that R is a domain, I is a prime ideal of S. Thus we may view R as the coordinate ring of an (irreducible) algebraic subset X of V. Here $X = \mathcal{V}_V(I)$. The map of varieties which is dual to the ring π is an embedding of X into $V \cong \mathbb{K}^m$. For this reason, if R is minimally generated by m elements we call m the *embedding dimension* of R, denoted $\mathrm{edim}(R)$.

Let R be a Noetherian ring and suppose $\wp$ is a prime ideal in R. We define the height of $\wp$ to be i if $\wp_0 \subsetneq \wp_1 \subsetneq \cdots \subsetneq \wp_i = \wp$ is a chain of prime ideals of maximal length of R contained in $\wp$. More generally, if I is any ideal of R, then the height of I is defined by $\mathrm{height}(I) = \min\{\mathrm{height}(\wp) \mid \wp \supseteq I, \wp \text{ is prime}\}$. Finally, the Krull dimension of R is the supremum of the heights of all prime ideals in R.

Example 2.3.1. A polynomial algebra $\mathbb{K}[x_1, x_2, \ldots, x_n]$ has Krull dimension n.

Of course, the embedding dimension of a finitely generated $\mathbb{K}$ algebra R is always greater than or equal to its Krull dimension with equality if and only R is a polynomial ring. A Noetherian ring R is called a *hypersurface ring* or just a *hypersurface* if $\mathrm{edim}(R) \leq \mathrm{Krull\ \ dim}(R) + 1$.

2.4 Localization and Fields of Fractions

Let R be a commutative ring and let $T \subset R$ be a subset which is closed under multiplication and which contains 1. We introduce an equivalence relation $\sim$ on $R \times T$ by putting

$$(r,t) \sim (r',t') \iff s(rt' - r't) = 0 \text{ for some } s \in T \ .$$

We write the fraction a/b to denote the equivalence class of (a,b) and $T^{-1}R$ to denote $(R \times T)/\sim$ which is called the localization of R at T. The operations $a/b + c/d := (ad + bc)/bd$ and $(a/b)(c/d) = ac/bd$ endow $T^{-1}R$ with a ring structure.

An important example of this construction is $T = \{f^m \mid m \in \mathbb{N}\}$ where f is some element of R which is not a zero divisor. In this situation we write R_f to denote $T^{-1}R$.

Another important example arises from a prime ideal $\wp$ in R. Since $\wp$ is prime, its complement $T = R \setminus \wp$ is a multiplicatively closed set containing 1. We write $R_\wp$ for $T^{-1}R$. In our setting, the natural map $R \to R_\wp$ carrying a to $\frac{a}{1}$ is injective.

Furthermore, there is a bijection (induced by the map $R \to R_\wp$) between the set of prime ideals of $R_\wp$ and the set of prime ideals of R which are contained in $\wp$. One of the most important features of localization at $\wp$ is that $R_\wp$ is a local ring, i.e., $R_\wp$ has a unique maximal ideal.

If R is a domain, then (0) is a prime ideal. Then $R_{(0)}$ is just the field of fractions or quotient field of R, which we denote by $\mathrm{Quot}(R)$.

An important property of quotient fields we will exploit is the following. Suppose that $R \subseteq S$ are domains which are finitely generated algebras over the same coefficient field, and that R and S share the same Krull dimension. Then $\mathrm{Quot}(S)$ is a finite dimensional $\mathrm{Quot}(R)$ vector space.

2.5 Integral Extensions

Suppose that R is a subring of S. An element y of S is *integral* over R if there is a monic polynomial $f(X) = X^r + f_{r-1}X^{r-1} + \cdots + f_0$ with $f_1, f_2, \ldots, f_r \in R$ such that $f(y) = 0$. We say that S is an *integral extension* of R or simply that S is *integral over* R if every element y of S is integral over R. The *integral closure* of R in S is the set of elements of S which are integral over R. The *normalization* of a domain R is its integral closure in its field of fractions $\mathrm{Quot}(R)$. A domain R is *integrally closed* or *normal* if it equals its normalization.

Proposition 2.5.1. *Let R be a subring of S and suppose $y \in S$. Then y is integral over R if and only if the ring $R[y]$ is a finitely generated R-module. Furthermore, if S is a finitely generated R-algebra, then S is integral over R if and only if S is a finitely generated R-module.*

We will use the following well-known theorem many times.

Theorem 2.5.2. *Suppose that S is an integral extension of R.*

1. Lying Over: *Let $\mathbf{p}$ be a prime ideal of R. Then there exists a prime ideal $\mathbf{q}$ of S with $\mathbf{q} \cap R = \mathbf{p}$. Moreover, $\mathbf{p}$ is a maximal ideal if and only if $\mathbf{q}$ is. We say that* $\mathbf{q}$ lies over $\mathbf{p}$.
2. Going-Up: *Let $\mathbf{p}_1 \subsetneq \mathbf{p}_2 \subsetneq \cdots \subsetneq \mathbf{p}_t$ be an ascending chain of prime ideals in R. Suppose $t > s$ and $\mathbf{q}_1 \subsetneq \mathbf{q}_2 \subsetneq \cdots \subsetneq \mathbf{q}_s$ is an ascending chain of prime ideals in S with $\mathbf{q}_i \cap R = \mathbf{p}_i$ for $i = 1, 2, \ldots, s$. Then the ascending chain of primes in S can be extended, i.e., there exist prime ideals $\mathbf{q}_{s+1}, \ldots, \mathbf{q}_t$ of S such that $\mathbf{q}_1 \subsetneq \mathbf{q}_2 \subsetneq \cdots \subsetneq \mathbf{q}_t$ and $\mathbf{q}_i \cap R = \mathbf{p}_i$ for all $i = 1, 2, \ldots, t$.*
3. Going-Down: *Further suppose that R is integrally closed. Let $\mathbf{p}_1 \supsetneq \mathbf{p}_2 \supsetneq \cdots \supsetneq \mathbf{p}_t$ be a descending chain of prime ideals in R. Suppose $t > s$ and $\mathbf{q}_1 \supsetneq \mathbf{q}_2 \supsetneq \cdots \supsetneq \mathbf{q}_s$ is a descending chain of prime ideals in S with $\mathbf{q}_i \cap R = \mathbf{p}_i$ for $i = 1, 2, \ldots, s$. Then the descending chain of primes in S can be extended, i.e., there exist prime ideals $\mathbf{q}_{s+1}, \ldots, \mathbf{q}_t$ of S such that $\mathbf{q}_1 \supsetneq \mathbf{q}_2 \supsetneq \cdots \supsetneq \mathbf{q}_t$ and $\mathbf{q}_i \cap R = \mathbf{p}_i$ for all $i = 1, 2, \ldots, t$.*

2.6 Homogeneous Systems of Parameters

Let R be a finitely generated $\mathbb{K}$ algebra. A homogeneous system of parameters of positive degree for R is a set of homogeneous elements $\{f_1, f_2, \ldots, f_n\}$ of R where n is the Krull dimension of R and with the property that R is finitely generated as a module over the ring $A = \mathbb{K}[f_1, f_2, \ldots, f_n]$. Equivalently, R is integral over $\mathbb{K}[f_1, \ldots, f_n]$.

The Noether Normalization Lemma, a fundamental theorem of modern algebra, asserts that homogeneous systems of parameters always exist. This result is used repeatedly in the study of the polynomial invariants of finite groups. Noether's Normalization Lemma was originally proved by Hilbert but has acquired its name since Emmy Noether stated and proved it as a lemma in her 1926 paper [88] in order to prove her main result that rings of invariants of finite groups are finitely generated algebras (Theorem 3.1.2).

Theorem 2.6.1 (Noether Normalization Lemma). *Let A be a finitely generated graded connected $\mathbb{K}$-algebra. Then A has a homogeneous system of parameters.*

Remark 2.6.2. If A is a finitely generated graded $\mathbb{K}$-algebra of Krull dimension n, and $\{h_1, h_2, \ldots, h_n\}$ is a (any) homogeneous system of parameters, then we will refer to $B = \mathbb{K}[h_1, h_2, \ldots, h_n]$ as a *Noether Normalization* of A.

A sequence $h_1, h_2, \ldots, h_t$ in R is a *partial homogeneous system of parameters* if it can be extended to a homogenous system of parameters $h_1, h_2 \ldots, h_t, h_{t+1}, \ldots, h_n$. The following lemma is a very useful characterization of homogeneous systems of parameters.

Lemma 2.6.3. *Let $A \subset \mathbb{K}[V]$ be a Noetherian graded subring of Krull dimension $n = \dim V$. Suppose $\{h_1, h_2, \ldots, h_n\}$ is a set of homogeneous elements of A. Then $h_1, h_2, \ldots, h_n$ is a homogeneous system of parameters for A if and only if $\mathcal{V}_{\overline{V}}(h_1, h_2, \ldots, h_n) = \{0\}$.*

Proof. An ideal $(f_1, f_2, \ldots, f_t)$ generated by t (homogeneous) elements has height less than or equal to t with equality if and only if $(f_1, f_2, \ldots, f_t)$ is a partial (homogeneous) system of parameters. Thus $h_1, h_2, \ldots, h_n$ is a homogeneous system of parameters if and only if $\text{height}(h_1, h_2, \ldots, h_n) = n$ if and only if $\dim \mathcal{V}(h_1, h_2, \ldots, h_n) = 0$ if and only if $\mathcal{V}_{\overline{V}}(h_1, h_2, \ldots, h_n)$ is a finite set of points. Now since each h_i is homogeneous, $\mathcal{V}_{\overline{V}}(h_1, h_2, \ldots, h_n)$ is stable under scaling, i.e., if $v \in \mathcal{V}_{\overline{V}}(h_1, h_2, \ldots, h_n)$ then $\lambda v \in \mathcal{V}_{\overline{V}}(h_1, h_2, \ldots, h_n)$ for all $\lambda \in \overline{\mathbb{K}}$. In particular, if $\mathcal{V}_{\overline{V}}(h_1, h_2, \ldots, h_n)$ contains any point other than the origin then $\mathcal{V}_{\overline{V}}(h_1, h_2, \ldots, h_n)$ contains a line and is thus infinite. This shows that $\mathcal{V}_{\overline{V}}(h_1, h_2, \ldots, h_n)$ is finite if and only if $\mathcal{V}_{\overline{V}}(h_1, h_2, \ldots, h_n) = \{0\}$. □

2.7 Regular Sequences

A sequence $\{r_1, \ldots, r_s\}$ in R is called a *regular sequence* if $(r_1, \ldots, r_s)R \neq R$ and r_i is not a zero-divisor in $R/(r_1, \ldots, r_{i-1})R$ for $1 \leq i \leq s$. (For $i = 1$ this means r_1 is not a zero-divisor in R.) We call s the *length* of the sequence.

If R is a Noetherian ring and $r_1, r_2, \ldots$ is any regular sequence in R then the sequence of ideals $R(r_1) \subsetneq R(r_1, r_2) \subsetneq \ldots$ in R is strictly increasing and so must have finite length. A regular sequence is maximal if it cannot be extended to a longer regular sequence.

Regular sequences do not necessarily remain regular when permuted, see Kaplansky [59, pg. 102], but the situation is nicer for homogeneous regular sequences.

Proposition 2.7.1. *Let R be a finitely generated commutative algebra graded over $\mathbb{K}$ of any characteristic $p \geq 0$. Then any permutation of a homogeneous regular sequence $z_1, z_2, \ldots, z_k$ is again a (homogeneous) regular system.*

Proof. Consider the case where w, z is a regular sequence, that is, suppose that w is not a zero-divisor in R, that z is not a zero-divisor in $R/(w)$ and suppose that $rz = 0$ for some r in R. We note that since z is homogeneous, we may assume that r is homogeneous of minimal degree with this property. However, since z is not a zero-divisor modulo w, we have that $r = sw$ for some homogeneous element $s \in R$. But then $rz = swz = 0$, which implies $sz = 0$ contradicting the minimal degree property of r. Therefore, z cannot be a zero-divisor in R. Now suppose the image $\overline{w}$ of w in $R/(z)$ is a zero-divisor. Then there are homogeneous elements $q, r \in R$ with $rw = qz$. But z is not a zero-divisor modulo (w), so it must be that $q = tw$ for some homogeneous $t \in R$, and then $rw = twz$, so $r = tz$, since w is not a zero-divisor in R, and therefore, $\overline{r} = 0$ as required.

In the general case, because the group of permutations is generated by transpositions of the form, $(\ell, \ell+1)$, it is sufficient to show that

$$z_1, \ldots, z_{\ell-1}, z_{\ell+1}, z_\ell, \ldots, z_k$$

is regular. Now $z_1, \ldots, z_{\ell-1}$ is regular, so it suffices to show that $z_{\ell+1}, z_\ell, \ldots, z_n$ is regular in $R' = R/(z_1, \ldots, z_{\ell-1})$. But $z_{\ell+2}, \ldots, z_n$ is regular in $R'' = R/(z_1, z_2, \ldots, z_\ell, z_{\ell+1})$, so the result follows from the previous paragraph. □

2.8 Cohen-Macaulay Rings

The depth(R) of a ring local ring $(R, \mathbf{m})$ is defined as the maximal length of a regular sequence in $\mathbf{m}$. It can be shown that in a graded Noetherian ring R any two homogeneous maximal regular sequences have the same length. If R is a graded Noetherian ring we put depth(R) = depth(R_+). The depth of a ring is always less than or equal to its Krull dimension. A local Noetherian ring $(R, \mathbf{m})$ is said to be Cohen-Macaulay if depth$(\mathbf{m})$ = height$(\mathbf{m})$ and a general Noetherian ring R is said to be Cohen-Macaulay if the localization of R at each of its prime ideals is Cohen-Macaulay.

It can be shown (see for example [6, Section 4.3]) that a Noetherian graded algebra A is Cohen-Macaulay if and only if there is some homogeneous system of parameters for A which is also a regular sequence. In our setting it is often easier to use this condition to show that a ring is Cohen-Macaulay.

For example, it is easy to show that $x_1, x_2, \ldots, x_n$ is a regular sequence in $\mathbb{K}[x_1, x_2, \ldots, x_n]$ which shows that $\mathbb{K}[x_1, x_2, \ldots, x_n]$ is a Cohen-Macaulay ring.

Cohen-Macaulay rings enjoy a number of useful properties. For non-modular representations of finite groups, a theorem due to Hochster and Eagon [53] shows that the ring of invariants is always Cohen-Macaulay. We will show in Chapter 9, Theorem 9.2.2, that very few modular representations have a Cohen-Macaulay ring of invariants. This is one of the main reasons why modular rings of invariants are more difficult to handle.

We now give a fundamental theorem of Noetherian algebra. The usual proofs of this theorem use homological algebra, e.g. Smith [102, Corollary 6.7.7]). We give a (previously unpublished) proof of Ian Hughes which does not use any homological algebra. In our opinion, Hughes proof is more straightforward than any of the other proofs of which we are aware.

Theorem 2.8.1. *Let A be a finitely generated connected graded $\mathbb{K}$-algebra which is Cohen-Macaulay. Then every homogeneous system of parameters for A is a regular sequence for A.*

Proof. The proof is by induction on $n = \dim A$. If $n = 0$ there is nothing to prove. Suppose $n = 1$. This implies that there exists a homogeneous element $u \in A$ with u not a zero divisor (and $u \notin A_0$). We let y be a homogeneous system of parameters for A. We must show that y is not a zero divisor. Now u satisfies some monic polynomial of degree r (say) with coefficients in $\mathbb{K}[y]$. Since both u and y are homogeneous, this implies that if $a \in A$ with $ay = 0$, then $au^r = 0$. But u is not a zero divisor and so $a = 0$. This shows that y is not a zero divisor as required.

We now suppose that $n > 1$. Let $y_1, y_2, \ldots., y_n$ be a homogeneous system of parameters for A. We first show that A has a homogeneous system of parameters which is a regular sequence for A which is contained in $\mathbb{K}[y_1, y_2, \ldots, y_n]$. Since A is Cohen-Macaulay, it has some homogeneous system of parameters which is a regular sequence. Let u denote the first element in this regular sequence and for $x \in A$ write $\overline{x}$ to denote the image of x in $A/(u)$. Now $A/(u)$ is Cohen-Macaulay with $\dim A/(u) = n - 1$. Since A is integral over $\mathbb{K}[y_1, y_2, \ldots, y_n]$, we know $A/(u)$ is integral over $\mathbb{K}[\overline{y_1}, \overline{y_2}, \ldots, \overline{y_n}]$. Therefore, by the Noether Normalization Lemma, there are homogeneous elements $u_2, u_3, \ldots, u_n \in \mathbb{K}[y_1, y_2, \ldots, y_n]$ such that $\overline{u_2}, \overline{u_3}, \ldots, \overline{u_n}$ is a homogeneous system of parameters for $A/(u)$. By induction, this is a regular sequence for $A/(u)$ and so $u, u_2, u_3, \ldots, u_n$ is a homogeneous system of parameters for A which is a regular sequence for A. But then $u_2, u, u_3, \ldots, u_n$ is a regular sequence for A by Proposition 2.7.1. Note that $u_2 \in \mathbb{K}[y_1, y_2, \ldots, y_n]$. We repeat the above argument using u_2 in the role of u and conclude that A has homogeneous system of parameters $z_1 = u_2, z_2, \ldots, z_n$ contained in $\mathbb{K}[y_1, y_2, \ldots, y_n]$ which is a regular sequence for A.

Now y_1 is integral over $\mathbb{K}[z_1, z_2, \ldots, z_n]$ and so y_1 satisfies a monic polynomial of degree r (say) with coefficients in $\mathbb{K}[z_1, z_2, \ldots, z_n] \subseteq \mathbb{K}[y_1, y_2, \ldots, y_n]$. Thus $y_1^r = -\sum_{i=0}^{r-1} a_i y_1^i$ with $a_i \in \mathbb{K}[z_1, z_2, \ldots, z_n]$. Expressing each a_i as a polynomial in $y_1, y_2, \ldots, y_n$ we see that one of the terms a_i includes the monomial y_1^{r-i}. From this, using the homogeneity of the y_i and of the z_i, it follows that there exists some k, and some non-zero scalar λ for which we have $\lambda z_k = y_1^s + \sum_{i=0}^{s-1} b_i y_1^i$ with homogeneous $b_i \in \mathbb{K}[y_2, \ldots, y_n]$ for $i = 0, 1, \ldots, s-1$. For $x \in A$, we now denote by $\overline{x}$ the image of x in $A/(z_k)$. Then the above implies that $\overline{y_1}$ is integral over $\mathbb{K}[\overline{y_2}, \overline{y_3}, \ldots, \overline{y_n}]$. Thus $A/(z_k)$ is integral over $\mathbb{K}[\overline{y_2}, \overline{y_3}, \ldots, \overline{y_n}]$ and so, since $\dim A/(z_k) = n - 1$, we have that $\overline{y_2}, \overline{y_3}, \ldots, \overline{y_n}$ is a homogeneous system of parameters for $A/(z_k)$. So, by induction, it is a regular sequence for $A/(z_k)$. This implies that $z_k, y_2, \ldots, y_n$ is a regular sequence for A. By Proposition 2.7.1, $y_2, y_3, \ldots, y_n, z_k$ is also a regular sequence for A. But $z_k = \lambda^{-1} y_1^s + b$ with b in the ideal $A(y_2, y_3, \ldots y_n)$ of A. Thus $y_2, y_3, \ldots, y_n, y_1^s$ is a regular sequence for A which implies that $y_1, y_2, \ldots, y_n$ is also a regular sequence for A. □

The following theorem (see [6, Theorem 4.3.5] for a proof) gives two more important characterizations of Cohen-Macaulay algebras.

Theorem 2.8.2. *Let A be a graded connected Noetherian $\mathbb{K}$-algebra. The following are all equivalent.*

1. *A is a Cohen-Macaulay ring of Krull dimension n.*

2. *There exists a homogeneous system of parameters* $y_1, y_2, \ldots, y_n$ *in* A *such that* A *is a free* $\mathbb{K}[y_1, y_2, \ldots, y_n]$*-module.*
3. A *is a free* $\mathbb{K}[y_1, y_2, \ldots, y_n]$*-module for every homogeneous system of parameters* $y_1, y_2, \ldots, y_n$ *in* A*.*

2.9 The Hilbert Series

Let R be a finitely generated graded $\mathbb{K}$-algebra. The fact that R is finitely generated implies that each of its homogenous components R_d is a finite dimensional $\mathbb{K}$ vector space. We define the *Hilbert series* of R to be the power series

$$\mathcal{H}(R, \lambda) = \sum_{i \geq 0} \dim_{\mathbb{K}}(R_i)\lambda^i.$$

This series is sometimes called the Poincaré series of R as well.

For example, suppose $R = \mathbb{K}[h_1, \ldots, h_n]$ is a polynomial algebra on generators of degrees d_i. Then

$$\mathcal{H}(R, \lambda) = \prod_{i=1}^{n} \frac{1}{(1 - \lambda^{d_i})}.$$

This is apparent when we consider that

$$\frac{1}{1 - \lambda^d} = 1 + \lambda^d + \lambda^{2d} + \cdots + \lambda^{md} + \cdots .$$

Suppose R is a graded Cohen-Macaulay ring and $h_1, h_2, \ldots, h_n$ is a homogeneous system of parameters for R. Putting $H = \mathbb{K}[h_1, h_2, \ldots, h_n]$ we have a *Hironaka decomposition*

$$R = \oplus_{i=1}^{r} H f_i$$

on generators f_i, of degrees m_i for $i = 1, \ldots, r$. Then

$$\mathcal{H}(R, \lambda) = \frac{\lambda^{m_1} + \cdots + \lambda^{m_r}}{\prod_{i=1}^{n}(1 - \lambda^{d_i})}.$$

We are now in a position to define what it means for a Noetherian graded domain R over a field $\mathbb{K}$ to be *Gorenstein*. Namely we require that R is Cohen-Macaulay and that there exists an integer m such that

$$\mathcal{H}(R, \lambda^{-1}) = \lambda^m (-1)^{\dim R} \mathcal{H}(R, \lambda).$$

Lemma 2.9.1. 1. *A polynomial algebra* R *has a symmetric Hilbert series in the sense above, hence is Gorenstein.*
2. *A hypersurface has a symmetric Hilbert series in the sense above, hence is Gorenstein.*

□

2.10 Graded Nakayama Lemma

Let R be a finitely generated graded connected $\mathbb{K}$-algebra. Let M be a non-negatively graded R-module. An element $f \in M$ is clearly not required as a generator of M if f lies in the submodule R_+M where $R_+ = \oplus_{d=1}^{\infty} R_d$. We let $Q(M)$ denote the (graded) vector space $Q(M) := M/R_+M$. Here by a *graded vector space* we mean a vector space W which is a direct sum of vector spaces $W = \oplus_{d=0}^{\infty} W_d$ where elements of W_d are said to be homogenous of degree d.

Proposition 2.10.1 (Graded Nakayama Lemma). *Let R be a finitely generated graded connected $\mathbb{K}$-algebra. Let M be a finitely generated non-negatively graded R-module. Then the homogeneous elements $f_1, f_2, \ldots, f_r$ generate M as an R-module if and only if their natural images $\bar{f}_1, \bar{f}_2, \ldots, \bar{f}_r$ span the $\mathbb{K}$ vector space $Q(M) = M/R_+M$. Moreover, the elements $f_1, f_2, \ldots, f_r$ minimally generate M if and only if $\{\bar{f}_1, \bar{f}_2, \ldots, \bar{f}_r\}$ is a vector space basis of $Q(M)$.*

Proof. Suppose that $\{\bar{f}_1, \bar{f}_2, \ldots, \bar{f}_r\}$ spans $Q(M)$ and let $N := \sum_{i=1}^r Rf_i$ be the submodule generated by $f_1, f_2, \ldots, f_r$. Note that since the f_i are homogeneous, N is a graded submodule of M. Assume by way of contradiction that N is a proper submodule of M and let d denote the least degree in which $N_d \subsetneq M_d$. Choose $m \in M_d \setminus N_d$ and write $\bar{m} = \sum_{i=1}^r c_i \bar{f}_i$ with $c_i \in \mathbb{K}$. Then $m = \sum_{i=1}^r c_i f_i + \sum_{j=1}^t g_j m'_j$ for some homogeneous $g_j \in R_+$ and $m'_j \in M$. By projecting this equation onto degree d we may assume $\deg(g_j m'_j) = d$ for $i = 1, 2, \ldots, t$. Thus $\deg(m'_j) < d$ for all j since $\deg g_j \geq 1$. Therefore $m'_j \in N$ for all $j = 1, 2, \ldots, t$ by the definition of d. This shows that $m \in N$, contradicting our assumption.

Conversely it is clear that if $f_1, f_2, \ldots, f_r$ generate M then their natural images span $Q(M)$.

The final assertion of the Proposition follows immediately. □

Note that any (homogeneous) lift of any homogeneous vector space basis for M/R_+M determines a minimal homogeneous generating set for M as an R-module. In particular, consider $M = R_+$. Let $\{f_1, f_2, \ldots, f_r\}$ be homogeneous elements of R_+ whose natural images $\{\bar{f}_1, \bar{f}_2, \ldots, \bar{f}_r\}$ form a vector space basis of $Q(R_+)$. Then $f_1, f_2, \ldots, f_r$ minimally generate R as a $\mathbb{K}$ algebra. To see this consider that $R_+ = \sum_{i=1}^r Rf_i$. Thus if $f \in R_d$ then $f = \sum_{i=1}^r g_i f_i$ for some homogeneous elements $g_i \in R_+$. By induction on degree, we see that $g_i \in \mathbb{K}[f_1, f_2, \ldots, f_r]$ and thus $f \in \mathbb{K}[f_1, f_2, \ldots, f_r]$. Conversely, it is clear that an element of R not lying in R_+^2 cannot be written as a polynomial in lower degree elements of R and thus any homogenous algebra generating set must surject onto to a spanning set for $Q(R_+)$.

In summary, we see that a homogeneous minimal generating set for R as an $\mathbb{K}$-algebra is obtained by lifting a homogeneous vector space basis of R_+/R_+^2. From this we see that such homogeneous minimal generating sets for R are not unique but that the number of generators of any given degree is fixed. We also note that $\dim_{\mathbb{F}}(R_+/R_+^2)$ is the embedding dimension of R.

We apply the graded Nakayama lemma to a ring of invariants $R = \mathbb{K}[V]^G$ to introduce terminology as follows. An invariant lying in R_+^2 may be written as a sum of products of other invariants of lower degree. For this reason, an invariant lying in R_+^2 is called *decomposable*. Conversely, an invariant (with zero constant term) not lying in R_+^2 is *indecomposable* and lies in some set of algebra generators for $\mathbb{K}[V]^G$.

2.11 Hilbert Syzygy Theorem

Theorem 2.11.1 (Hilbert Syzygy Theorem). *Let M be any finitely generated graded module over a polynomial ring $A = \mathbb{K}[x_1, x_2, \ldots, x_n]$. Then there is a finite graded resolution of M by free graded A-modules*

$$0 \longrightarrow M_k \xrightarrow{\rho_k} M_{k-1} \xrightarrow{\rho_{k-1}} \cdots \xrightarrow{\rho_2} M_1 \xrightarrow{\rho_1} M_0 \xrightarrow{\rho_0} M \longrightarrow 0$$

Moreover, we may take $k \leq n$.

Note that here we do *not* assume that $\deg(x_\ell) = 1$. We denote by I_ℓ the kernel of the map $\rho_\ell : M_\ell \to M_{\ell-1}$. The theorem proceeds by choosing homogeneous generators for $I_{\ell-1}$ and a free module M_ℓ mapping onto it. Therefore the maps at each stage depend upon this choice. If at each stage we choose a homogeneous *minimal* generating set for the kernel I_ℓ of ρ_ℓ then we obtain a homogeneous minimal free resolution.

Although the maps depend upon our choice of ideal generators, the number k for which I_k is first zero does not, provided we choose a homogeneous minimal generating set for each ideal I_ℓ, i.e., provided we work with a minimal free resolution. This number measures, in some sense, the complexity of M. It is called the projective dimension of M as an A-module and denoted $\mathrm{pd}(M)$.

We define the *Betti* numbers, $\beta_\ell(M)$, of M to be the rank of the ℓ^{th} syzygy module: $\beta_\ell(M) := \mathrm{rank}_A(M_\ell)$ when the free resolution of M is minimal. For example, $\beta_1(M)$ counts the number of minimal relations among any set of minimal homogeneous generators of M. Like the length of the resolution, k, the Betti numbers $\beta_\ell(M)$ do not depend upon the choices made in constructing a minimal free resolution of M. Using the grading on the M_ℓ we may define the *graded Betti numbers*, $\beta_{\ell j}(M) :=$ the number of minimal generators of M_ℓ which have degree j. Again, the graded Betti numbers do not depend upon the choices made in constructing a minimal free resolution of M. Of course, $\sum_{j\geq 0} \beta_{\ell j} = \beta_\ell$ for all ℓ. Thus $\mathrm{pd}(M) = \max\{i \mid \beta_i(M) \neq 0\}$.

Since the maps in the resolution preserve degree and since the resolution is everywhere exact, the Hilbert syzygy theorem yields

$$\mathcal{H}(M, \lambda) = \sum_{i=0}^{k} (-1)^i \mathcal{H}(M_i, \lambda).$$

Now we have $\mathcal{H}(A,\lambda) = \prod_{\ell=1}^{n} \frac{1}{1-\lambda^{\deg(x_\ell)}}$. Since each M_i is a free A-module, we may write $M_i = \oplus_{j=1}^{r_i} A m_{ij}$ for some $m_{i1}, m_{i2}, \ldots, m_{ir_i} \in M_i$. Therefore, $\mathcal{H}(M_i,\lambda) = (\lambda^{\deg(m_{i1})} + \lambda^{\deg(m_{i2})} + \cdots + \lambda^{\deg(m_{ir_i})}) / \prod_{\ell=1}^{n}(1-\lambda^{\deg(x_\ell)})$. In terms of the graded Betti numbers this is $\mathcal{H}(M_i,\lambda) = (\sum_j \beta_{ij}\lambda^j)/\prod_{\ell=1}^{n}(1 - \lambda^{\deg(x_\ell)})$ and

$$\mathcal{H}(M,\lambda) = \frac{\sum_{ij}(-1)^i \beta_{ij}\lambda^j}{\prod_{\ell=1}^{n}(1-\lambda^{\deg(x_\ell)})}.$$

In particular, this shows that when A is a polynomial ring, the Hilbert series of any finitely generated graded A-module may be expressed as a rational function, i.e., as the quotient of two polynomials. By Noether's Normalization Lemma, every Noetherian graded ring R is finitely generated over some polynomial ring A, and thus $\mathcal{H}(R,\lambda)$ is always a rational function of λ. Furthermore, any finitely generated R-module M is also a finitely generated A-module and thus $\mathcal{H}(M,\lambda)$ is a rational function of λ.

An important example of the above for us will be when $M = R$ is a Noetherian ring. For example, when $R = \mathbb{K}[V]^G$. Let $\{f_1,\ldots,f_s\}$ denote a minimal homogeneous generating set for R. Let $A = \mathbb{K}[x_1, x_2, \ldots, x_s]$ denote the polynomial algebra on generators x_i where we declare that $\deg(x_i) = \deg(f_i)$ for all $i = 1, 2, \ldots, s$. In our notation above, then, we have $s = \mathrm{edim}(R) = \dim(A)$. Let ρ_0 denote the surjective algebra map from $A = M_0$ to R determined by $\rho_0(x_i) = f_i$ for $i = 1, 2, \ldots, s$. The map ρ_0 provides R with the structure of a finitely generated A-module and thus we may consider its resolution by syzygies. This turns out to be an effective way to study R and in particular to compute $\mathcal{H}(R,\lambda)$. In Chapter 13 we will study this technique in detail.

3

Applications of Commutative Algebra to Invariant Theory

In this chapter, we use some of the basic commutative algebra discussed in the previous chapter to develop some basic results about rings of invariants.

Let R be a Noetherian graded $\mathbb{K}$-algebra. Suppose that G is a group of degree-preserving automorphisms of R. Then G acts on R_d and we denote the elements of R_d fixed by G by R_d^G and we define $R^G = \oplus_{d \geq 0} R_i^G$, so that R^G is a graded connected algebra.

If G acts on R as degree-preserving automorphisms, then it is easy to see that G also acts on $\mathrm{Quot}(R)$ as a group of automorphisms of the field. We may therefore form the invariant subfield $\mathrm{Quot}(R)^G$.

Lemma 3.0.1. *For any finite group G, we have $\mathrm{Quot}(R)^G = \mathrm{Quot}(R^G)$. Consequently, the extension $\mathrm{Quot}(R)^G \subset \mathrm{Quot}(R)$ is Galois, with group G and so $\mathrm{Quot}(R)$ has dimension $|G|$ as a $\mathrm{Quot}(R)^G$ vector space.*

Proof. If $r/s \in \mathrm{Quot}(R)^G$ with $r, s \in R$, then

$$\frac{r}{s} := \frac{r \cdot \prod_{e \neq \sigma \in G} \sigma(s)}{\mathbf{N}^G(s)}.$$

Since the fraction itself and the denominator are both invariant, the numerator is also, finishing the proof. □

Proposition 3.0.2. *If R is a unique factorization domain, then R is integrally closed. In particular, $\mathbb{F}[V]$ is integrally closed.*

Proof. Suppose that $r/s \in \mathrm{Quot}(R^G)$ is integral over R. We suppose that r and s do not share any common factor. There exist $a_0, a_1, \ldots, a_{t-1} \in R$ such that $(r/s)^t + a_{t-1}(r/s)^{t-1} + \cdots + a_1(r/s) + a_0 = 0$. Therefore $r^t = -s(a_{t-1}r^{t-1} + a_{t-2}r^{t-2}s + \cdots + a_1 r s^{t-2} + a_0 s^{t-1})$. Thus every irreducible factor of s divides r^t and so must also divide r. But since r and s share no common factors, this implies that s must be a unit. Thus $r/s \in R$. □

Proposition 3.0.3. *Let R be an integrally closed domain on which the finite group G acts. Then R^G is also integrally closed.*

H.E.A. Eddy Campbell, D. Wehlau, *Modular Invariant Theory*,
Encyclopaedia of Mathematical Sciences 139,
DOI 10.1007/978-3-642-17404-9_3, © Springer-Verlag Berlin Heidelberg 2011

Proof. Suppose $r/s \in \mathrm{Quot}(R^G)$ is integral over R^G. Then r/s is integral over R and lies in $\mathrm{Quot}(R)^G \subseteq \mathrm{Quot}(R)$. Since R is integrally closed, this implies that $r/s \in R$ and since r/s is G-invariant, $r/s \in R^G$. □

Proposition 3.0.4. *Let G be any finite group acting on a finitely generated algebra R as a group of automorphisms. Then R is integral over R^G. In particular, R is a finitely generated R^G-module.*

Proof. Suppose R is generated by elements $\{r_1, \ldots, r_s\}$. For t an indeterminate, consider the polynomials $F_i(t) \in R[t]$ defined as

$$F_i(t) = \prod_{\sigma \in G} (t - \sigma(r_i)) = \sum_{j=0}^{|G|} a_{i,j} t^i.$$

We extend the action of G to automorphisms of $R[t]$ by defining $\sigma(t) = t$ for all $\sigma \in G$. With this convention we see from the definition of $F_i(t)$ that the polynomials $F_i(t)$ are invariant. Thus each coefficient $a_{i,j}$ on the right-hand side is invariant. Furthermore, we have shown that each of the generators of R is integral over R^G, and therefore R is integral over R^G as claimed. The final assertion of the proposition follows using Proposition 2.5.1. □

Lemma 3.0.5. *Using the notation of the above proof,*

$$\{r_1^{i_1} \cdots r_s^{i_s} \mid 0 \leq i_j \leq |G| - 1\}$$

spans R as a module over R^G. □

The following important Corollary of Proposition 3.0.4 is immediate.

Corollary 3.0.6. *Let G be a finite group. Let $f_1, f_2, \ldots, f_n \in \mathbb{F}[V]^G$ be a sequence of homogeneous invariants. Then $f_1, f_2, \ldots, f_n$ is a homogenous system of parameters for $\mathbb{F}[V]^G$ if and only if it is a homogenous system of parameters for $\mathbb{F}[V]$.*

Remark 3.0.7. Proposition 3.0.4 implies that if G is a finite subgroup of $\mathrm{GL}(V)$ then $\dim \mathbb{K}[V]^G = \dim \mathbb{K}[V] = \dim_{\mathbb{K}} V$.

It follows that the Krull dimension of R^G is the same as the Krull dimension of R.

3.1 Homogeneous Systems of Parameters

There are some especially useful homogeneous systems of parameters for $\mathbb{F}[V]^G$. If our group is a permutation group, then the elementary symmetric functions form a homogeneous system of parameters, see Section 3.2. If $\mathbb{F}$ is finite then we may always use the Dickson invariants as a homogeneous system of parameters, see Section 3.3. If our representation is lower triangular, then Múi has constructed a homogeneous system of parameters, see Section 3.4. In particular, Múi's result applies for any modular representation of a p-group (by Proposition 4.0.2).

If the field $\mathbb{F}$ is infinite, then we can construct a homogeneous system of parameters by a method due to Dade, see [104]. First, we show that there must exist a basis $\{x_1, \ldots, x_n\}$ for V^* which has the property that, for any n-tuple $\sigma_1, \ldots, \sigma_n$ of elements of G, the set of linear forms $\{\sigma_1(x_1), \ldots, \sigma_n(x_n)\}$ is linearly independent. We begin by choosing any non-zero element $x_1 \in V^*$. We choose the remaining basis elements $x_2, x_3, \ldots, x_n$ inductively as follows. Having chosen $x_1, x_2, \ldots, x_i$ we consider the union of proper subspaces of V^* defined by

$$K_i := \bigcup_{\sigma_1,\sigma_2,\ldots,\sigma_i \in G} \mathrm{span}_{\mathbb{F}}\{\sigma_1(x_1), \sigma_2(x_2), , \ldots, \sigma_i(x_i)\} \ .$$

construction depends upon the observation that, as a finite union of proper subspaces, K_i, cannot be all of V^* when $\mathbb{F}$ is infinite. We note also that K_i is stable under the action of G by definition. While $i < n$ we may choose a linear form $x_{i+1} \in V^* \setminus K_i$. Such an element $x_i + 1$ has the property that $\sigma(x_{i+1}) \in V^* \setminus K_i$ for all $\sigma \in G$. Now form the polynomials

$$f_i = \mathbf{N}^G_{G_{x_i}}(x_i),$$

for $i = 1, 2, \ldots, n$. Each f_i is of degree less than or equal $|G|$. To see that $f_1, f_2, \ldots, f_n$ is a homogeneous system of parameters, suppose $f_i(\mathbf{v}) = 0$ for all $i = 1, 2, \ldots, n$. Since each f_i is a product of linear forms, there must exist, for each i, a group element $\sigma_i \in G$ with $\sigma_i(x_i)(\mathbf{v}) = 0$. But the set $\{\sigma_1(x_1), \sigma_2(x_2), \ldots, \sigma_n(x_n)\}$ is linearly independent in V^* and hence it is linearly independent also in $\overline{V}^* = \overline{\mathbb{F}} \otimes V^*$, and hence $\mathbf{v} = 0$, as required.

Remark 3.1.1. If we are working over a finite field $\mathbb{F}$ we may work in the algebraic closure $\widetilde{\mathbb{F}}$ of $\mathbb{F}$ to use Dade's construction to construct a homogeneous system of parameters, $f_1, f_2, \ldots, f_n$ in $\widetilde{\mathbb{F}}[V]^G$. Note that the homogeneous system of parameters will lie in $\mathbb{F}'[V]^G$ for a finite overfield $\mathbb{F}' \supset \mathbb{F}$ containing the finitely many coefficients of the f_i.

Probably the most famous result of Invariant Theory is David Hilbert's proof [52] that (in modern terminology) the ring of invariants of a (geometrically) reductive group is finitely generated. However, his proof does not apply to modular representations since these fail to be linearly reductive. For a good discussion of these issues, see Derksen and Kemper's book [26, section 2.2]. However, Emmy Noether [88] proved the following result.

Theorem 3.1.2. *The ring of invariants $\mathbb{F}[V]^G$ of a finite group is finitely generated.*

Proof. By the graded Nakayama lemma, Lemma 2.10.1, we need to prove that $Q(\mathbb{F}[V]^G) = \mathbb{F}[V]^G/(\mathbb{F}[V]^G_+)^2$ is a finite dimensional $\mathbb{F}$ vector space. Let $\overline{\mathbb{F}}$ be the algebraic closure of $\mathbb{F}$ and consider $\overline{V} := V \otimes_{\mathbb{F}} \overline{\mathbb{F}}$. It is clear that $Q(\overline{\mathbb{F}}[\overline{V}]^G) \cong Q(\mathbb{F}[V]^G) \otimes_{\mathbb{F}} \overline{\mathbb{F}}$. Therefore, it suffices to show that $Q(\mathbb{F}[\overline{V}]^G)$ is a finite dimensional $\overline{\mathbb{F}}$ vector space. Hence, replacing $\mathbb{F}$ by $\overline{\mathbb{F}}$ if necessary we may assume that $\mathbb{F}$ is infinite. Therefore, by Dade's algorithm, there exists a homogeneous system of parameters $f_1, f_2, \ldots, f_n$ for $\mathbb{F}[V]^G$. By Corollary 3.0.6, these elements $f_1, f_2, \ldots, f_n$ also form a homogeneous system of parameters for $\mathbb{F}[V]$. Let A denote the polynomial ring $A = \mathbb{F}[f_1, f_2, \ldots, f_n]$. Since $\mathbb{F}[V]$ is a Cohen-Macaulay ring, there exists $h_1, h_2, \ldots, h_t \in \mathbb{F}[V]$ such that $\mathbb{F}[V] \cong \oplus_{i=1}^t A h_i$. Since $\mathbb{F}[V]$ is a finitely generated A-module and A is a Noetherian ring, we see that $\mathbb{F}[V]$ is a Noetherian A-module. It is clear that $\mathbb{F}[V]^G$ is an A-module and thus as an A-submodule of a Noetherian A-module, $\mathbb{F}[V]^G$ is also a Noetherian A-module. Thus there exists a finite set $g_1, g_2, \ldots, g_r \in \mathbb{F}[V]^G$ with $\mathbb{F}[V]^G = \sum_{j=1}^r A g_j$. Therefore $\mathbb{F}[V]^G = \mathbb{F}[f_1, f_2, \ldots, f_n, g_1, g_2, \ldots, g_r]$ is a finitely generated $\mathbb{F}$-algebra. □

In the notation of the above proof, the elements $f_1, f_2, \ldots, f_n$ are traditionally called *primary invariants* and the elements $g_1, g_2, \ldots, g_r$ *secondary invariants*. It is important to notice that there are infinitely many choices for both the primary and secondary invariants. These two terms merely describe the role of certain invariants within a particular fixed generating set.

Since $\mathbb{F}[V]^G$ is finitely generated, its spectrum, $\mathrm{Spec}(\mathbb{F}[V]^G)$ is an affine variety which we denote by $V/\!/G$. Dual to the natural inclusion of algebras is the natural surjection of varieties

$$\pi_{V,G} : V \to V/\!/G.$$

The fibres of this surjection consist of G-orbits: $\pi_{V,G}^{-1}(\pi_{V,G}(v)) = Gv$. The geometry and topology of $V/\!/G$ reveals information about the ring of invariants, $\mathbb{F}[V]^G$, and vice versa.

Proposition 3.1.3. *[105, Proposition 2.2.3] Let $\{f_1, f_2, \ldots, f_n\}$ be any homogeneous system of parameters for the polynomial ring $S = \mathbb{F}[x_1, x_2, \ldots, x_n]$ and suppose the degree of f_i is d_i. Then S is a free $R = \mathbb{F}[f_1, f_2, \ldots, f_n]$-module on $\prod_{i=1}^n d_i$ many generators.*

Proof. Since S is Cohen-Macaulay, we have a Hironaka decomposition

$$S = \oplus_{i=1}^m R\eta_i$$

for some homogeneous $\eta_1, \eta_2, \ldots, \eta_m \in S$. Using this decomposition we see that the Hilbert series of S is given by

$$\begin{aligned}\mathcal{H}(S,\lambda) &= \sum_{j=1}^{m} \mathcal{H}(R,\lambda)\lambda^{\deg(\eta_j)} \\ &= \frac{\sum_{j=1}^{m} \lambda^{\deg(\eta_j)}}{\prod_{i=1}^{n}(1-\lambda^{d_i})}\end{aligned}$$

Since $\mathcal{H}(S,\lambda) = \prod_{i=1}^{n}(1-\lambda)^{-1}$ we see that

$$\begin{aligned}\sum_{j=1}^{m} \lambda^{\deg(\eta_j)} &= \frac{\prod_{i=1}^{n}(1-\lambda^{d_i})}{\prod_{i=1}^{n}(1-\lambda)} \\ &= \prod_{i=1}^{n}(1+\lambda+\lambda^2+\ldots{}_+\lambda^{d_i-1})\end{aligned}$$

Setting $\lambda = 1$ here gives $m = \prod_{i=1}^{n} d_i$. □

Corollary 3.1.4. *[26, Theorem 3.7.1] Let $\{f_1, f_2, \ldots, f_n\}$ be any homogeneous system of parameters of degrees $d_1, d_2, \ldots, d_n$ for $\mathbb{F}[V]^G$. Then $\mathbb{F}[V]^G$ is a free $R = \mathbb{F}[f_1, \ldots, f_n]$-module if and only if it has $(\prod_{i=1}^{n} d_i)/|G|$ many generators as an R-module. If $\mathbb{F}[V]^G$ is not free as an R-module, more generators are required.*

Proof. Using Proposition 3.1.3 we have

$$\mathbb{F}[V] = \oplus_{i=1}^{m} R\eta_i$$

where $m = \prod_{i=1}^{n} d_i$. Hence $\{\eta_1, \eta_2, \ldots, \eta_m\}$ is a vector space basis for $\mathbb{F}(V)$ as an Quot(R)-vector space. Thus the degree of the field extension $\mathbb{F}(V)$ over Quot(R) is $\prod_{i=1}^{n} d_i$. By Lemma 3.0.1, the field extension $\mathbb{F}(V)$ over Quot$(\mathbb{F}[V]^G)$ has degree $|G|$. Therefore, the degree of the subfield extension Quot$(\mathbb{F}[V]^G)$ over Quot(R) is $(\prod_{i=1}^{n} d_i)/|G|$. Thus $\mathbb{F}(V)^G$ has dimension $(\prod_{i=1}^{n} d_i)/|G|$ as a $\mathbb{F}(f_1, f_2, \ldots, f_n)$-vector space. Suppose $\mathbb{F}[V]^G = \sum_{j=1}^{s} Rh_j$ where s is minimal. Then $\{h_1, h_2, \ldots, h_s\}$ spans $\mathbb{F}(V)^G$ as a Quot(R)-vector space and thus $s \geq (\prod_{i=1}^{n} d_i)/|G|$.

If $s > (\prod_{i=1}^{n} d_i)/|G|$, then $\{h_1, h_2, \ldots, h_s\} \subset \mathbb{F}(V)^G$ is an Quot(R) linearly dependent set. Therefore, we may write $\sum_{j=1}^{s} \frac{a_j}{b_j} h_j = 0$ for some $a_j, b_j \in R$ with at least one $a_j \neq 0$. Clearing denominators gives $\sum_{j=1}^{s} a'_j h_j = 0$ with $a'_1, a'_2, \ldots, a'_s \in R$ not all zero, showing that $\mathbb{F}[V]^G$ is not a free R-module (and thus that $\mathbb{F}[V]^G$ is not Cohen-Macaulay). □

Remark 3.1.5. In the above proof we showed that the field extension $\mathbb{F}(V)^G$ over Quot(R) has degree $(\prod_{i=1}^{n} d_i)/|G|$. This shows that if $\{f_1, \ldots, f_n\}$ is any homogeneous system of parameters of degrees $d_1, \ldots, d_n$ for $\mathbb{F}[V]^G$, then $|G|$ divides $\prod_{i=1}^{n} d_i$.

Corollary 3.1.6. *Let $\{f_1, f_2, \ldots, f_n\}$ be any homogeneous system of parameters of degrees $d_1, d_2, \ldots, d_n$ for $\mathbb{F}[V]^G$. Then*

$$\prod_{i=1}^{n} d_i = |G| \text{ if and only if } \mathbb{F}[V]^G = \mathbb{F}[f_1, f_2, \ldots, f_n] \ .$$

3.2 Symmetric Functions

Consider $G = \Sigma_n \subset Gl(V)$ acting as all permutations of a basis $\{x_1, \ldots, x_n\}$ for V^*. We note that $Gx_1 = \{x_1, \ldots, x_n\}$ and that $\mathcal{H}_{x_1}(t) = \prod_{i=1}^{n}(t - x_i) = \sum_{j=0}^{n}(-1)^j s_j t^{n-j}$. Here, the s_j is the j-th elementary symmetric function

$$s_j = \sum_{1 \le i_1 < i_2 < \cdots < i_j \le n} x_{i_1} x_{i_2} \cdots x_{i_j} \ .$$

The following theorem is well known. It can be proven easily using Corollary 3.1.6. We will give a more constructive proof in §5.1.1.

Theorem 3.2.1. *Let $\mathbb{K}$ be a field of any characteristic and let the symmetric group on n letters, Σ_n, act on the n dimensional $\mathbb{K}$ vector space V by permuting a basis. Let $\{x_1, x_2, \ldots, x_n\}$ denote the dual basis of V^*. Then $\mathbb{K}[V]^{\Sigma_n} = \mathbb{K}[s_1, s_2, \ldots, s_n]$ where $s_i = s_i(x_1, x_2, \ldots, x_n)$ is the i^{th} elementary symmetric function in $\{x_1, x_2, \ldots, x_n\}$ for $i = 1, 2, \ldots, n$.*

Of course, the elementary symmetric functions enjoy many beautiful properties, and symmetric functions occur in many different situations in mathematics.

The following theorem is well-known. We will prove a stronger version of this result in Theorem 5.1.11.

Theorem 3.2.2.

$$\mathbb{F}[x_1, \ldots, x_n]^{\Sigma_n} = \mathbb{F}[s_1, \ldots s_n].$$

Proof. We begin by showing that the set $\{s_1, \ldots, s_n\}$ forms a homogeneous system of parameters. By Lemma 2.6.3, we may do this by showing that the only common zero (working over the algebraic closure of $\mathbb{F}$) of the functions $\{s_1, \ldots, s_n\}$ is the origin. We proceed by induction on n. When $n = 1$, there is nothing to prove.

Consider the algebra surjection determined by

$$\theta : \mathbb{F}[x_1, \ldots, x_n] \to \mathbb{F}[x_1, \ldots, x_{n-1}]$$
$$x_i \mapsto \begin{cases} x_i & \text{if } i < n \\ 0 & \text{if } i = n \end{cases}$$

We are using $s_1, s_2, \ldots, s_n$ to denote the elementary symmetric functions in $x_1, x_2, \ldots, x_n$. To denote the elementary symmetric functions in $x_1, x_2, \ldots, x_{n-1}$

we will use $s'_1, s'_2, \ldots, s'_{n-1}$. Note that $s'_i = \theta(s_i)$ for $i = 1, 2, \ldots, n-1$ and $\theta(s_n) = 0$. We assume by induction that the only common zero of $\{s'_1, \ldots, s'_{n-1}\}$ is the origin. Now consider a point $(a_1, \ldots, a_n)$ at which $s_1, \ldots, s_n$ all vanish. In particular, $s_n = x_1 \cdots x_n$ vanishes. We may assume without loss of generality that $a_n = 0$. Therefore, $0 = s_i(a_1, \ldots, a_n) = s_i(a_1, \ldots, a_{n-1}, 0) = \theta(s_i)(a_1, \ldots, a_{n-1})$ from which it follows by induction that $0 = a_1 = \cdots = a_{n-1}$.

We note that s_i has degree i and therefore, $\prod_{i=1}^n \deg(s_i) = n! = |\Sigma_n|$. Hence the theorem follows from applying Corollary 3.1.6. □

Of course, there are many interesting questions involving symmetric functions, see for example Macdonald [78] or Sturmfels [105]. For example, given a symmetric function, how can it be written in terms of the elementary symmetric functions? This question is answered by using SAGBI bases, see §5.1.1. As well, there are other generating sets for the symmetric functions, for example, in characteristic zero we may use the power sums

$$h_i = x_1^i + \ldots x_n^i,$$

for $1 \leq i \leq n$. However, the power sums do not necessarily generate the full ring of symmetric functions over a field of positive characteristic. For example, it is easy to see that $x_1 x_2 \notin \mathbb{F}_2[x_1 + x_2, x_1^2 + x_2^2]$. Other important questions concern the Hironaka decomposition of $\mathbb{F}[V]$ as a module over $\mathbb{F}[V]^{\Sigma_n}$. This topic is considered in §6.1.

3.3 The Dickson Invariants

Suppose $\mathbb{F}$ is a finite field of order $q = p^s$ and let V be an n-dimensional vector space over $\mathbb{F}$. We consider $G = \mathrm{GL}(V)$. Then any vector $v \in V^* \setminus \{0\}$ has $\mathcal{O}_G(v) = V^* \setminus \{0\}$. Now we obtain

$$F_n(t) = \prod_{w \in V^*} (t - w) = \sum_{i=0}^{n} (-1)^{n-i} d_{i,n} t^{q^{n-i}}.$$

The $d_{i,n}$ are known as the Dickson invariants, and they enjoy many beautiful properties.

For example, when $p = 2$ and $n = 2$ we have

$$F_2(t) = t(t+x)(t+y)(t+x+y)$$

so that $d_{1,2} = x^2 + xy + y^2$ and $d_{2,2} = xy(x+y) = x^2y + xy^2$. More generally, we have

Lemma 3.3.1. *1.* $F_n(t) = F_{n-1}^q(t) - F_{n-1}^{q-1}(x_n) F_{n-1}(t)$.

2. $d_{i,n} = d_{i,n-1}^q - d_{i-1,n-1} F_{n-1}^{q-1}(x_n)$.

□

We note that $\deg(d_{i,n}) = q^n - q^{n-i}$. Therefore, we have

$$\prod_{i=1}^{n} \deg(d_{i,n}) = |\operatorname{GL}(V)|.$$

This latter calculation is pretty. There are q^n vectors in V, and any one of them may be identified with the first row of a matrix in $\operatorname{GL}(V)$ excepting the zero vector. Hence there are $q^n - 1$ choices for the first row. Similarly, the second row corresponds to vectors in V that are linearly independent of the first, and there are $q^n - q$ choices for these. And so on.

Lemma 3.3.2. *The set $\{d_{1,n}, \ldots, d_{n,n}\}$ is a homogeneous system of parameters for $\mathbb{F}[V]$, hence*

$$\mathbb{F}[d_{1,n}, \ldots, d_{n,n}] = \mathbb{F}[V]^G.$$

□

3.4 Upper Triangular Invariants

Consider $G = U_n(\mathbb{F}_q)$, the group of all upper triangular matrices with 1's along the diagonal and entries from the finite field $\mathbb{F}_q$ of order $q = p^s$. This group was shown to have a polynomial ring of invariants by Múi [82] through his construction of a generating set. Let $\{x_1, x_2, \ldots, x_n\}$ be the basis of V^* dual to the standard basis of $V = \mathbb{F}_q^n$. Note that $\mathcal{O}_G(x_i) = x_i + V_{i-1}$ where V_{i-1} denotes the subspace of V^* with basis $\{x_1, x_2, \ldots, x_{i-1}\}$. We define

$$h_i = \prod_{v \in V_{i-1}} (x_i + v)$$

and observe that h_i is homogeneous of degree q^{i-1}. Note that $h_i = F_{i-1}(x_i)$, where $F_{i-1}(t)$ is defined in §3.3. Therefore, $\prod_{i=1}^n \deg(h_i) = |U_n(\mathbb{F}_q)|$.

When $p = 2$, we get $h_1 = x$, $h_2 = y(y + x) = y^2 + xy$. For arbitrary p, we have $h_2 = \prod_{\alpha \in \mathbb{F}_p} (y + \alpha x) = y^p - x^{p-1}y$.

Lemma 3.4.1. $\mathbb{F}_q[h_1, \ldots, h_n] = \mathbb{F}_q[V]^G$. □

3.5 Noether's Bound

In this section we will prove that the ring of invariants for a non-modular group G is always generated by invariants of degree at most $|G|$. Noether originally proved this over fields of characteristic 0. Here we will prove this under the weaker hypothesis that $|G|$ is invertible in $\mathbb{F}$. This was proved independently by P. Fleischmann [39] and J. Fogarty [42] in 2000. D. Benson (see [26, pg. 109]) simplified Fogarty's proof and here we present this simplified version.

Theorem 3.5.1. *Let V be a representation of a finite group G. If $|G|$ is invertible in $\mathbb{K}$, then $\mathbb{K}[V]^G$ is generated by invariants of degree at most $|G|$.*

Proof. Put $m := |G|$ and $[m] = \{1, 2, \ldots, m\}$. We begin by considering $\mathbb{K}[V]_+ := \sum_{d=1}^{\infty} \mathbb{K}[V]_d$, the unique homogeneous maximal ideal of $\mathbb{K}[V]$. We first show that its m^{th} power $(\mathbb{K}[V]_+)^m$ is a subset of the Hilbert ideal $I := (\mathbb{K}[V]_+^G)\mathbb{K}[V]$, the ideal of $\mathbb{K}[V]$ generated by all the homogeneous invariants of positive degree.

To see this, take any $f_1, f_2, \ldots, f_m \in \mathbb{K}[V]_+$. Write $G = \{\sigma_1, \sigma_2, \ldots, \sigma_m\}$, let $\sigma \in G$ and consider the product $\prod_{i=1}^m (f_i - (\sigma\sigma_i)(f_i)) = 0$. Expanding this expression and summing over all $\sigma \in G$ gives:

$$\sum_{A \subseteq [m]} (-1)^{m-|A|} h_A \prod_{i \in A} f_i = 0$$

where $h_A := \sum_{\sigma \in G} \prod_{i \in [m] \setminus A} \sigma(\sigma_i f_i) \in \mathbb{K}[V]^G$.

The summand corresponding to $A = [m]$ in the above is $|G|\, f_1 f_2 \cdots f_m$ with $h_{[m]} = |G|$. For all other subsets A, we have $h_A \in \mathbb{K}[V]_+^G$ and thus $|G|\, f_1 f_2 \cdots f_m$ lies in the Hilbert ideal I. Since $|G|$ is invertible, $f_1 f_2 \cdots f_m \in I$ and thus $(\mathbb{K}[V]_+)^m \subseteq I$.

Since $\mathbb{K}[V]$ is a Noetherian ring, there exist finitely many homogeneous invariants $h_1, h_2, \ldots, h_r \in \mathbb{K}[V]^G$ which generate the Hilbert ideal I. Without loss of generality we may assume that $\{h_1, h_2, \ldots, h_r\}$ is a minimal such set of invariants. Note that if $\deg h_i > m$, then $h_i = \sum_{j=1}^n h_{ij} x_j$ where each h_{ij} is a homogeneous element of $\mathbb{K}[V] = \mathbb{K}[x_1, x_2, \ldots, x_n]$ with $\deg(h_{ij}) \geq m$, i.e., where $h_{ij} \in (\mathbb{K}[V]_+)^m \subset I$. Since $m \leq \deg h_{ij} < \deg h_i$, we see that h_{ij} lies in the ideal of $\mathbb{K}[V]$ generated by $h_1, h_2, \ldots, \hat{h_i}, \ldots, h_r$. Thus if $\deg h_i > m$, then h_i is not required as a generator of I. Thus our assumption that $h_1, h_2, \ldots, h_r$ minimally generate I implies that $\deg h_i \leq m$ for all $i = 1, 2, \ldots, r$.

Consider any invariant $f \in \mathbb{K}[V]^G$ with $\deg(f) > m$. Since $\deg(f) > m$, we see that $f \in (\mathbb{K}[V]_+)^m \subseteq I$ and we may write $f = \sum_{i=1}^r k_i h_i$ where each k_i is a homogeneous element of $\mathbb{K}[V]_+$. Since $|G|$ is invertible, we may average over the orbit to obtain

$$\begin{aligned} f = \frac{1}{|G|} \sum_{j=1}^m \sigma_j(f) &= \frac{1}{|G|} \sum_{j=1}^m \sum_{i=1}^r \sigma_j(k_i h_i) = \frac{1}{|G|} \sum_{j=1}^m \sum_{i=1}^r \sigma_j(k_i) h_i \\ &= \frac{1}{|G|} \sum_{i=1}^r \sum_{j=1}^m \sigma_j(k_i) h_i. \end{aligned}$$

Since $\sum_{j=1}^m \sigma_j(k_i) \in \mathbb{K}[V]_+^G$, this shows that f is a decomposable invariant and so cannot be part of any minimal algebra generating set for $\mathbb{K}[V]^G$. □

The above proof shows that in the non-modular case, the Hilbert ideal is generated by homogeneous elements of degree at most $|G|$. Kemper [26, Conjecture 3.8.6 (b)] has made the following conjecture.

Conjecture 3.5.2 (Kemper). Let V be a representation of a finite group G. The Hilbert ideal $(\mathbb{K}[V]^G_+)\mathbb{K}[V]$ is generated by homogeneous elements of degree at most $|G|$.

3.6 Representations of Modular Groups and Noether's Bound

In this section, we describe work of Karagueuzian and Symonds [62] and Symonds [106].

Theorem 3.6.1. *If $\mathbb{F}$ is finite of characteristic p and G is a finite group, then $\mathbb{F}[V]$ has only finitely many isomorphism types of indecomposable summands.*

The theorem has real force when G is modular, for then G has, in general, a "wild" representation theory: it is not possible to classify the indecomposable representations of G.

Example 3.6.2. We recall the following material from the paper of Elmer and Fleischmann [36], where a much more complete discussion occurs. Let $G = C_2 \times C_2$ be generated by elements σ, τ. It is known that this group has countably many isomorphism classes of indecomposable representations. Consider the 3-dimensional representation, V, of G over a field, $\mathbb{F}$, of characteristic 2 given by

$$\rho(\sigma) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \rho(\tau) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We let U denote the trivial 1-dimensional representation of G and we let W denote the regular 4-dimensional representation of G. Then the symmetric algebra of V over $\mathbb{F}$ contains the indecomposable summands U, V and W and no others.

Recalling the notation and result of section §3.4, we saw that the invariants of the upper triangular group $G = U_n(\mathbb{F}_q)$ are generated by the orbit polynomials

$$h_i = \prod_{v \in V_{i-1}} (x_i + v), \quad \text{where } V_{i-1} \text{ is the span of } \{x_1, \ldots, x_{i-1}\}.$$

That is, $\mathbb{F}[x_1, x_2, \ldots, x_n]^G = \mathbb{F}[h_1, h_2, \ldots, h_n]$. Let I denote the integer sequence $(1, 2, \ldots, n)$ and let $I' = (1, 2, \ldots, n-1)$, and, for any subsequence $J \subset I'$, we define $J^c = I \setminus J$. Finally, we define

$$H(J) = \mathbb{F}[h_\ell \mid \ell \in J^c].$$

We note that G acts trivially on $H(J)$ for any J. Given any p-group P, we may embed P in G (see §4.0.2), and hence $H(J)$ is P-invariant under this embedding for any J. We define finite dimensional graded $\mathbb{F}P$-modules $X(P, J) \subset \mathbb{F}[V]$. Then

Theorem 3.6.3. *There is an isomorphism of graded $\mathbb{F}P$-modules*

$$\mathbb{F}[V] \cong \oplus_{J \subset I'} H(J) \otimes_{\mathbb{F}} X(P,J).$$

This structure theorem states that $\mathbb{F}[V]$ contains one copy of $X(P,J)$ for each monomial of $H(J)$. The general case of a modular group follows.

Remark 3.6.4. We note that there are $2^{n-1} - 1$ such subsets, for any P. In [20], Campbell and Selick constructed 2^n subsets $M(i)$ of $\mathbb{F}[V]$, each of which is an injective module over Steenrod's algebra $\mathcal{A}$. The module $M(0)$ is the ring of invariants of the non-modular group $\mathbb{F}^*$. The relationship between the modules $X(P,J)$ is not known to the authors of this book.

Example 3.6.5. Let G denote the group $U_2(\mathbb{F})$, so that $\mathbb{F}[V]^G = \mathbb{F}[h_1,h_2]$, where $h_1 = x_1$ and $h_2 = x_2^q - x_1^{q-1}x_2$. We set $T = \mathbb{F}[V]/(h_2)$. Then, as modules over $\mathbb{F}G$, we have

$$\mathbb{F}[V] = \mathbb{F}[h_2] \otimes_{\mathbb{F}} T \quad \text{where} \quad T = X(G,\{1\}) \oplus X(G,\{\emptyset\}) \ .$$

Then, as modules over $\mathbb{F}G$, we have

$$\mathbb{F}[V] = \mathbb{F}[h_1,h_2] \otimes X(G,\emptyset) \oplus \mathbb{F}[h_2] \otimes_{\mathbb{F}} T(G,\{1\})$$

where $X(G,\emptyset)$ is the module spanned by $\{1\}$ and $T(G,\{1\})$ is the module spanned by $\left\{x_2, x_2^2, \ldots, x_2^{q-1}\right\}$. The interested reader should compare to §7.2.

Example 3.6.6. Karagueuzian and Symonds [61, §2, extended example]: the case $n = 3$ and $p = 3$. Let G denote the group $U_3(\mathbb{F})$, so that $\mathbb{F}[V]^G = \mathbb{F}[h_1,h_2,h_3]$ where $|h_i| = 3^{i-1}$. We set $T = \mathbb{F}[V]/(h_3)$. Then, as modules over $\mathbb{F}G$ we have

$$\begin{aligned}\mathbb{F}[V] = \mathbb{F}[h_1,h_2,h_3] \otimes T(G,\emptyset) \oplus \mathbb{F}[h_1,h_3] \otimes X(G,\{2\}) \\ \oplus \mathbb{F}[h_2,h_3] \otimes X(G,\{1\}) \oplus \mathbb{F}[h_3] \otimes X(G,\{1,2\}) \ .\end{aligned}$$

Theorem 3.6.7. *Suppose $\{f_1, f_2, \ldots, f_n\}$ is a homogeneous system of parameters for the ring of invariants of the modular group G. Then $\mathbb{F}[V]^G$ is generated by module generators over $\mathbb{F}[f_1, f_2, \ldots, f_n]$ of degrees less than or equal to*

$$\sum_{i=1}^{n} (|f_i| - 1) \ .$$

Moreover, the relations among the module generators have degrees less than or equal to

$$1 + \sum_{i=1}^{n} (|f_i| - 1).$$

Finally, the degree of $\mathcal{P}(G,V,t)$ as a rational function in t is at most $-n$.

Corollary 3.6.8. *[106] If $\mathbb{F}$ is finite and G is a non-trivial finite group acting on V with $\dim_{\mathbb{F}}(V) > 1$, then $\mathbb{F}[V]^G$ is generated in degrees less than or equal to $\dim_{\mathbb{F}}(V)(|G| - 1)$.*

These conclusions follow from the following theorem of Symonds, [106]

Theorem 3.6.9. *The invariant ring $\mathbb{F}[V]^G$ has Castelnuovo-Mumford regularity 0.*

Remark 3.6.10. Symonds proves more: if G acts on $\mathbb{F}[x_1, x_2, \ldots, x_n]$ by homogeneous linear substitutions of the x_i's, then the invariant ring has Castelnuovo-Mumford regularity 0.

3.7 Molien's Theorem

The following theorem of Molien, provides a constructive method to compute the Hilbert series of a ring of invariants of a finite group in characteristic 0.

Theorem 3.7.1. (Molien) *Let $\mathbb{K}$ be a field of characteristic zero.*

$$\mathcal{H}(\mathbb{K}[V]^G, \lambda) = \frac{1}{|G|}(\sum_{\sigma \in G} \frac{1}{\det(Id_V - \lambda\sigma)}) \ .$$

Proof. Since the dimension of a vector space is unchanged under field extensions, we may assume that $\mathbb{K}$ is algebraically closed.

Let W be any $\mathbb{K}$ representation of G. By the orthogonality relations for characters, see for example the book of Dummit and Foote [32][Chapter 18, Theorems 15 and 16] or Lang [75][XVIII §5]), we have $\dim_{\mathbb{K}}(W^G) = \frac{1}{|G|}\sum_{\sigma \in G} \mathrm{trace}(\sigma_{|W})$.

Take $\sigma \in G$ and consider the matrix $A \in \mathrm{GL}(V)$ representing the action of σ on V. Considering the Jordan Normal form of A we see that since A has finite order and $\mathbb{K}$ is characteristic zero, the matrix A must be diagonalizable. Let $\rho_1, \rho_2, \ldots, \rho_n$ denote the n eigenvalues of A. Thus, working with a basis of eigenvectors, we may assume that $A = \mathrm{diag}(\rho_1, \rho_2, \ldots, \rho_n)$. Write $R := \mathbb{K}[V]$. Let $x_1, x_2, \ldots, x_n$ be the dual basis of V^*. Thus the full set of monomials of degree m, $\{x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n} \mid a_1 + a_2 + \cdots + a_n = m\}$ forms an eigenbasis for the action of A on R_m The trace of the action of σ on R_m is thus

$$\mathrm{trace}(\sigma_{|R_m}) = \sum_{a_1+a_2+\cdots+a_n=m} \rho_1^{-a_1} \rho_2^{-a_2} \cdots \rho_n^{-a_n} .$$

Therefore,

$$\begin{aligned}\frac{1}{\det(\mathrm{Id}_V-\lambda\sigma)} &= \frac{1}{(1-\rho_1\lambda)(1-\rho_2\lambda)\cdots(1-\rho_n\lambda)}\\ &= \prod_{i=1}^{n}(1+\rho_i\lambda+\rho_i^2\lambda^2+\dots)\\ &= \sum_{m=0}^{\infty}\left(\sum_{a_1+a_2+\dots+a_n=m}\rho_1^{a_1}\rho_2^{a_2}\cdots\rho_n^{a_n}\right)\lambda^m\\ &= \sum_{m=0}^{\infty}\left(\mathrm{trace}(\sigma^{-1}_{|R_m})\right)\lambda^m\ .\end{aligned}$$

Thus,

$$\begin{aligned}\mathcal{H}(R^G,\lambda) &= \sum_{m=0}^{\infty}(\dim_{\mathbb{K}} R^G_m)\lambda^m\\ &= \sum_{m=0}^{\infty}\left(\frac{1}{|G|}\sum_{\sigma\in G}\mathrm{trace}(\sigma_{|R_m})\right)\lambda^m\\ &= \frac{1}{|G|}\sum_{\sigma\in G}\frac{1}{\det(\mathrm{Id}_V-\lambda\sigma^{-1})}\\ &= \frac{1}{|G|}\sum_{\sigma\in G}\frac{1}{\det(\mathrm{Id}_V-\lambda\sigma)}\ .\end{aligned}$$

□

Remark 3.7.2. Suppose $|G|^{-1}\in\mathbb{K}$, that is, G is a non-modular group. Elements of representation theory give us a complex representation of G, known as the Brauer lift, which shares the same Hilbert series as $\mathbb{K}[V]^G$, see [25][§82]. Thus Molien's Theorem can be used to compute the Hilbert series of the ring of invariants for any non-modular representation of a finite group.

Remark 3.7.3. Note that if σ and τ are two elements of G lying in the same conjugacy class, then σ and τ have the same eigenvalues on V and therefore,

$$\frac{1}{\det(\mathrm{Id}_V-\lambda\sigma)}=\frac{1}{\det(\mathrm{Id}_V-\lambda\tau)}.$$

3.7.1 The Hilbert Series of the Regular Representation of the Klein Group

Consider

$$\sigma=\sigma^{-1}=\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix}\ \text{and}\ \tau=\tau^{-1}=\begin{pmatrix}0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0\end{pmatrix}.$$

The group $G = \{I_4, \sigma, \tau, \sigma\tau\}$ is the Klein 4-group. We consider the ring of invariants $\mathbb{K}[V]^G$ where $\mathbb{K}$ is a field of any characteristic.

Since G acts by permutations on V, the Hilbert series $\mathcal{H}(\mathbb{K}[V]^G, t)$ is the same for all characteristics. Using Molien's Theorem we compute

$$\begin{aligned}\mathcal{H}(\mathbb{K}[V]^G, \lambda) &= \frac{1}{4}\left(\frac{1}{(1-\lambda)^4} + \frac{3}{(1-\lambda^2)^2}\right)\\ &= \frac{1}{4}(1+\lambda+\lambda^2+\lambda^3+\lambda^4+\dots)^4 + \frac{3}{4}(1+\lambda^2+\lambda^4+\lambda^6+\dots)^2\\ &= \frac{1}{4}(1+4\lambda+10\lambda^2+20\lambda^3+35\lambda^4+\dots)\\ &\quad+\frac{3}{4}(1+2\lambda^2+3\lambda^4+\dots)\\ &= 1+\lambda+4\lambda^2+5\lambda^3+11\lambda^4+\dots\end{aligned}$$

Note that

$$\begin{aligned}\mathcal{H}(\mathbb{K}[V]^G, \lambda) &= \frac{1}{4}\left(\frac{1}{(1-\lambda)^4} + \frac{3}{(1-\lambda^2)^2}\right)\\ &= \frac{(1+\lambda)^2 + 3(1-\lambda)^2}{4(1-\lambda)^4(1+\lambda)^2}\\ &= \frac{1-\lambda+\lambda^2}{(1-\lambda)^4(1+\lambda)^2}\\ &= \frac{1+\lambda^3}{(1-\lambda)(1-\lambda^2)^3}\ .\end{aligned}$$

This is the Hilbert series we would expect to find if $\mathbb{K}[V]^G$ possessed a homogeneous system of parameters in degrees 1, 2, 2, 2 over which the invariants are a rank 2 free module with generators in degrees 0 and 3. Suppose there is such a homogeneous system of parameters. Looking in degrees 1 and 2 we see that (essentially) the only possibility for such a system is a, b_1, b_2, b_3 where $a = x_1+x_2+x_3+x_4$, $b_1 = x_1x_2+x_3x_4$, $b_2 = x_1x_3+x_2x_4$ and $b_3 = x_1x_4+x_2x_3$. We use Lemma 2.6.3 to study this. Suppose $v = (v_1, v_2, v_3, v_4)$ is a point in $\overline{V}$ at which a, b_1, b_2, b_3 all vanish, that is,

$$\begin{aligned}v_1+v_2+v_3+v_4 &= 0\\ v_1v_2+v_3v_4 &= 0\\ v_1v_3+v_2v_4 &= 0\\ v_1v_4+v_2v_3 &= 0\end{aligned}$$

If $v_1 = 0$, then we have $v_2v_3 = v_2v_4 = v_3v_4 = 0$ from which we conclude two of these three are also zero. But $a(v) = 0$ forces the fourth to also be zero. So we may assume by symmetry that none of the $v_\ell = 0$. Assume there is a point v where all 4 coordinates are non-zero. Then we have

$$v_2 = -\frac{v_3 v_4}{v_1}$$
$$v_3 = -\frac{v_2 v_4}{v_1}$$
$$v_4 = -\frac{v_2 v_3}{v_1}$$

It follows that $v_1^2 = v_2^2 = v_3^2 = v_4^2$. Hence $0 = a(v)^2 = 4v_1^2$. In characteristic not 2 we get a contradiction and so can conclude that a, b_1, b_2, b_3 is a homogeneous system of parameters. However in characteristic 2 we see that $v = (1, 1, 1, 1)$ is a point where a, b_1, b_2, b_3 all vanish and hence they are not a homogeneous system of parameters in this characteristic.

However, the elementary symmetric functions form a homogeneous system of parameters in all characteristics. The corresponding form of the Hilbert series is

$$\mathcal{H}(\mathbb{K}[V]^G, \lambda) = \frac{1 + 2\lambda^2 + 2\lambda^4 + \lambda^6}{(1-\lambda)(1-\lambda^2)(1-\lambda^3)(1-\lambda^4)}$$

As we will see later in §4.7, there is a set of module generators for $\mathbb{K}[V]^G$ over $\mathbb{K}[V]^{\Sigma_4}$ of degrees 0, 2, 2, 4, 4 and 6 valid in any characteristic.

3.7.2 The Hilbert Series of the Regular Representation of C_4

Consider the regular representation of the cyclic group $G := C_4$ with generator σ. The matrix of σ is given by

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

We want to compute the ring of invariants $\mathbb{K}[V]^G$ where $\mathbb{K}$ is a field of any characteristic.

As in the previous example, the Hilbert series $\mathcal{H}(\mathbb{K}[V]^G, \lambda)$ is the same in all characteristics. Using Molien's Theorem we compute

$$\begin{aligned}\mathcal{H}(\mathbb{K}[V]^G, \lambda) &= \frac{1}{4}\left(\frac{1}{(1-\lambda)^4} + \frac{1}{(1-\lambda^2)^2} + \frac{2}{(1-\lambda^4)}\right) \\ &= 1 + \lambda + 3\lambda^2 + 5\lambda^3 + 10\lambda^4 + 14\lambda^5 + 22\lambda^6 + \dots \\ &= \frac{1 + \lambda^2 + \lambda^3 + 2\lambda^4 + \lambda^5}{(1-\lambda)(1-\lambda^2)(1-\lambda^3)(1-\lambda^4)}\end{aligned}$$

This suggests that the ring of invariants $\mathbb{K}[V]^G$ is a free module over the ring of symmetric functions on generators of degrees 0, 2, 3, 4, 4, and 5. We will see in §4.8 that this is the case in every characteristic except 2. In fact, in characteristic 2, this ring of invariants is not even Cohen-Macaulay. Bertin [9] provided this ring of invariants as the first example of a unique factorization domain that was not Cohen-Macaulay answering a question of Pierre Samuel.

3.8 Rings of Invariants of p-Groups Are Unique Factorization Domains

Theorem 3.8.1. *Let $\mathbb{F}$ be a field of characteristic $p > 0$ and let P be a p-group. Then $\mathbb{F}[V]^P$ is a unique factorization domain.*

Proof. We proceed by induction on degree. If $f \in \mathbb{F}[V]^P$ has degree 1 then clearly, f is prime and there is nothing to prove.

Suppose then, that $f \in \mathbb{F}[V]^P$ with $\deg(f) > 1$. Decompose f as a product of primes in $\mathbb{F}[V]$: $f = f_1 f_2 \cdots f_t$. Take $\sigma \in P$. Since $\sigma(f) = f$ we must have $\sigma(f_1) = \lambda f_j$ for some j with $1 \leq j \leq t$ and some $\lambda \in \mathbb{F}^\times$. Without loss of generality, we may suppose that $\{1, 2, \ldots, s\} = \{j \mid \exists\, \sigma \in P, \lambda \in \mathbb{F}^*$ with $\sigma(f_1) = \lambda f_j\}$. Note that if $\sigma \in P$ and $1 \leq i \leq s$, then $\sigma(f_i) = \lambda f_j$ for some j with $1 \leq j \leq s$ and some $\lambda \in \mathbb{F}^*$.

Define $h_1 := f_1 f_2 \cdots f_s$. Then for each $\sigma \in P$ we must have $\sigma(h_1) = \lambda(\sigma) h_1$ for some $\lambda(\sigma) \in \mathbb{F}^\times$. Now take $\sigma, \tau \in P$. Then

$$\begin{aligned}(\sigma\tau)(h_1) = \sigma(\tau(h_1)) = \sigma(\lambda(\tau)h_1) = \lambda(\tau)(\sigma(h_1)) = \lambda(\tau)(\lambda(\sigma)h_1)\\ = \lambda(\sigma)\lambda(\tau)h_1.\end{aligned}$$

Thus $\lambda(\sigma\tau) = \lambda(\sigma)\lambda(\tau)$. This shows that

$$\lambda : P \to \mathbb{F}^\times$$

is a linear character of P. In particular, $1 = \lambda(e) = \lambda(\sigma^p) = \lambda(\sigma)^p$ and therefore $(1 - \lambda(\sigma))^p = 0$, whence $\lambda(\sigma) = 1$ for all $\sigma \in P$.

This shows that h_1 lies in $\mathbb{F}[V]^P$. Moreover, for each $j = 1, 2, \ldots, s$ there exists $\sigma \in P$ and $\lambda \in \mathbb{F}^\times$ such that $\sigma(f_1) = \lambda f_j$. This shows that h_i is an irreducible element in $\mathbb{F}[V]^P$. Now by induction, the invariant f/h_1 may be uniquely factored into a product of irreducibles in $\mathbb{F}[V]^P$. Thus $f/h_1 = h_2 h_3 \cdots h_q$ where h_j is irreducible in $\mathbb{F}[V]^P$ for $2 \leq j \leq q$. Therefore, we have factored $f = h_1 h_2 \cdots h_q$ into a product of irreducibles.

Suppose now that $f = h'_1 h'_2 \cdots h'_r$ is another factorization of f into irreducibles in $\mathbb{F}[V]^P$. Working in $\mathbb{F}[V]$ we see that since the prime f_1 of $\mathbb{F}[V]$ divides f we must have that f_1 divides h'_k for some k. Without loss of generality, we assume that $k = 1$. But since $h'_1 \in \mathbb{F}[V]^P$, this implies that f_j divides h'_1 for all $j = 1, 2, \ldots, s$. Hence h_1 divides the irreducible element h'_1 in $\mathbb{F}[V]^P$ and thus h_1 and h'_1 are associated irreducible elements of $\mathbb{F}[V]^P$. Write $h'_1 = \mu h_1$ for some $\mu \in \mathbb{F}^*$. Then $f/h_1 = h_2 h_3 \ldots h_q = \mu^{-1} h'_2 h'_3 \cdots h'_r$. Now by induction, the element f/h_1 of $\mathbb{F}[V]^P$ has a unique factorization up to ordering and scalar factors. Thus $q = r$ and, renumbering if necessary, h'_j is an associate of h_j for all $j = 2, 3, \ldots, q$. Therefore, $\mathbb{F}[V]^P$ is a unique factorization domain. □

Example 3.8.2. We take $p = 3$ and work over $\mathbb{F}_3$. Let

$$B = \left\{ \begin{pmatrix} a & 0 \\ b & d \end{pmatrix} \mid a, d \in \mathbb{F}_3 \setminus \{0\}, b \in \mathbb{F}_3 \right\}$$

be the set of all invertible 2×2 lower triangular matrices over $\mathbb{F}_3$. Let V be the natural 2 dimensional representation of B.

The p-Sylow subgroup U of B is the set of all the elements of B with 1's along the main diagonal. Thus U is isomorphic to C_3, the cyclic group of order 3. We work with basis $\{e_1, e_2\}$ of V with respect to which the elements of B are lower triangular and also the dual basis $\{x, y\}$ of V^*. Thus e_2 and x are eigenvectors for all the elements of B and fixed points for the elements of U.

We have already seen in Theorem 1.11.2 that $\mathbb{F}_3[V]^U = \mathbb{F}_3[x, N]$ where $N = \mathrm{N}(y) = y^3 - x^2y$. Thus $\mathbb{F}_3[V]^U$ is polynomial and so is clearly a unique factorization domain, as it must be by the theorem.

In contrast, consider $\mathbb{F}_3[V]^B$. Since U is a normal subgroup of B, we have $\mathbb{F}_3[V]^B = (\mathbb{F}_3[V]^U)^{B/U} = \mathbb{F}_3[x, N]^{\{\pm 1\}}$. The non-trivial element $-1 \in B/U$ acts on $\mathbb{F}_3[V]^U$ via $-1 \cdot x = -x$ and $-1 \cdot N = -N$. Thus $\mathbb{F}_3[V]^B$ contains the irreducible elements x^2, xN and N^2. Indeed, it is not too difficult to show that these 3 invariants generate $\mathbb{F}_3[V]^B$. The two irreconcilable factorizations $(xN) \cdot (xN) = (x^2) \cdot (N^2)$ show that $\mathbb{F}_3[V]^B$ is not a unique factorization domain.

3.9 When the Fixed Point Subspace Is Large

In this section, we consider the situation where the group fixes point-wise a subspace of codimension 1 or 2. In both cases, this guarantees that the ring of invariants is especially well-behaved.

We will need the following lemma of J.P. Serre, his "Normality Criterion". Recall that a Noetherian domain R is said to be integrally closed or *normal*, if R is integrally closed in its fraction field $\mathrm{Quot}(R)$. In other words, R is normal precisely if the following condition holds: if $r, r' \in R$ and there is a monic polynomial $f(t) = \sum_{i=0}^{k} a_i t^i$ with coefficients $a_i \in R$ such that $f(\frac{r}{r'}) = 0$, then $\frac{r}{r'} \in R$.

Theorem 3.9.1 (Serre's Normality Criterion). *[80, Theorem 23.8] A Noetherian domain R is integrally closed if and only if the following two conditions hold:*

(R_1) If $\wp \in \mathrm{Spec}(R)$ satisfies $\mathrm{height}(\wp) \leq 1$, then $R_\wp$ is a regular local ring.
(S_2) If $\wp \in \mathrm{Spec}(R)$ satisfies $\mathrm{height}(\wp) \geq 2$, then $\mathrm{depth}(R_\wp) \geq 2$.

Part 1 of the following theorem was first proved by Landweber and Stong [73] with different techniques.

Theorem 3.9.2. *Let G be any subgroup of $\mathrm{GL}(V)$ over any field $\mathbb{K}$. Put $n := \dim_{\mathbb{K}}(V)$. Then*

1. If $\dim_{\mathbb{K}}(V^G) = n-1$*, then* $\mathbb{K}[V]^G$ *is a polynomial algebra.*
2. If $\dim_{\mathbb{K}}(V^G) = n-2$*, then* $\mathbb{K}[V]^G$ *is Cohen-Macaulay.*

Proof. We will prove both assertions simultaneously using a technique brought to our attention by Abraham Broer. First, we note that the question of whether $\mathbb{K}[V]^G$ is polynomial or is Cohen-Macaulay is unchanged under an extension of the underlying field $\mathbb{K}$. For example, this can be seen from Corollaries 3.1.6 and 3.1.4. Thus we may assume that $\mathbb{K}$ is algebraically closed. For part 1, we define

$$(V/\!\!/G)_{\text{good}} := \{\wp \in \operatorname{Spec}\mathbb{K}[V]^G \mid \mathbb{K}[V]^G_{\wp} \text{ is a regular local ring}\}.$$

For part 2, we define

$$(V/\!\!/G)_{\text{good}} := \{\wp \in \operatorname{Spec}\mathbb{K}[V]^G \mid \mathbb{K}[V]^G_{\wp} \text{ is a Cohen-Macaulay local ring}\}.$$

In both cases, the set $(V/\!\!/G)_{\text{good}}$ is an open subset of $V/\!\!/G$. For part 1, this is shown in [71, VI Cor.1.16] and [80, Chapter 13]. For part 2, this is shown in [80, Chapter 8, Prop. (22.C)]. We also need the following two results.

Let A be a finitely generated graded $\mathbb{K}$-algebra with maximal homogeneous ideal $\mathbf{m} = A_+$.

1. A is a polynomial ring if and only if $A_{\mathbf{m}}$ is a regular local ring.
2. A is Cohen-Macaulay if and only if $A_{\mathbf{m}}$ is Cohen-Macaulay.

These statements are 2.2.25 and 2.1.27(c) of [14]. Let $\mathbf{m}$ denote the irrelevant ideal of $\mathbb{K}[V]^G$, i.e., $\mathbf{m} = \mathbb{K}[V]^G_+$. By the above two statements, we need to show that $\mathbf{m} \in (V/\!\!/G)_{\text{good}}$.

Let I denote the prime ideal $I := \mathcal{I}(V^G) \in \operatorname{Spec}(\mathbb{K}[V]^G)$ and note that $\operatorname{height}(I) = \operatorname{codim}(V^G)$. Now $\mathbb{K}[V]^G$ is integrally closed and hence by Theorem 3.9.1, it satisfies the two conditions (R_1) and (S_2). By (R_1), $\mathbb{K}[V]^G_I$ is a regular local ring if $\operatorname{codim} V^G = 1$. If $\operatorname{codim} V^G = 2$, then $2 = \text{Krull}\ \dim(\mathbb{K}[V]^G_I)$ and since the depth of any ring is at most its Krull dimension, we see that (S_2) implies that $2 = \text{Krull}\ \dim(\mathbb{K}[V]^G_I) \geq \operatorname{depth}(\mathbb{K}[V]^G_I) \geq 2$ and thus $\mathbb{K}[V]^G_I$ is Cohen-Macaulay. Thus in both cases $I \in (V/\!\!/G)_{\text{good}}$. In particular, we see that $(V/\!\!/G)_{\text{good}}$ is a non-empty open set.

We claim that $(V/\!\!/G)_{\text{good}}$ contains some maximal ideal $\mathbf{n}$ which contains I. To see this we define $(V/\!\!/G)_{\text{bad}} := V/\!\!/G \setminus (V/\!\!/G)_{\text{good}}$. Let J denote the ideal $J := \mathcal{I}((V/\!\!/G)_{\text{bad}})$. Assume by way of contradiction that every maximal ideal $\mathbf{n}$ containing I lies in $(V/\!\!/G)_{\text{bad}}$, i.e., that every maximal ideal $\mathbf{n}$ containing I also contains J. Since I is a radical ideal, I is the intersection of the maximal ideals which contain it (see [80][Theorem 5.5], cf. the discussion in §2.2) and thus

$$I = \bigcap_{\substack{\mathbf{n} \text{ maximal} \\ \mathbf{n} \supseteq I}} \mathbf{n} \supseteq J.$$

Thus $I \in \mathcal{I}(J) = (V/\!\!/G)_{\text{bad}}$. But this is a contradiction since we know $I \in (V/\!\!/G)_{\text{good}}$. This proves the claim.

Thus there is a maximal ideal $\mathbf{n}$ containing I and with $\mathbf{n} \in (V/\!\!/G)_{\text{good}}$. Let w denote the point of V^G corresponding to the ideal $\mathbf{n}$, i.e., $\{w\} = \mathcal{V}(\mathbf{n})$, and consider the map $\psi : V \to V$ defined by $\psi(v) = v + w$. Dual to this map we have the algebra homomorphism $\psi^* : \mathbb{K}[V] \to \mathbb{K}[V]$ given by $\psi^*(f) = f \circ \psi$. The map ψ^* is G-equivariant since for any $\sigma \in G$ and any $v \in V$ we have

$$\begin{aligned}(\sigma(\psi^*(f)))(v) &= (\psi^*(f))(\sigma^{-1}(v))\\ &= (f \circ \psi)(\sigma^{-1}(v))\\ &= f(\sigma^{-1}(v) + w)\\ &= f(\sigma^{-1}(v + w))\\ &= (\sigma(f))(v + w)\\ &= (\sigma(f) \circ \psi)(v)\\ &= (\psi^*(\sigma(f)))(v),\end{aligned}$$

i.e., $\sigma(\psi^*(f)) = \psi^*(\sigma(f))$. In particular, if $f \in \mathbb{K}[V]^G$, then $\psi^*(f) \in \mathbb{K}[V]^G$. Thus restricting ψ^* gives an algebra automorphism of $\mathbb{K}[V]^G$. Note that $\psi^*(\mathbf{n}) = \mathbf{m}$. Therefore, ψ^* induces an isomorphism of $\mathbb{K}[V]^G_{\mathbf{n}}$ onto $\mathbb{K}[V]^G_{\mathbf{m}}$. Since $\mathbf{n} \in (V/\!\!/G)_{\text{good}}$, this shows that $\mathbf{m} \in (V/\!\!/G)_{\text{good}}$. □

4

Examples

In this chapter we collect a number of interesting examples and basic results in order to illustrate some of the techniques. As usual, our focus is two-fold: either determine a generating set of invariants, or determine the structure of the ring of invariants, or both. We begin with a useful lemma.

Lemma 4.0.1. *Suppose G is a p-group and $\mathbb{F}$ has characteristic p and let V be any positive dimensional representation of G. Then $V^G \neq \{0\}$.*

Proof. By Lemma 1.10.3, there exists a composition series

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

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

We proceed by induction on the length of this composition series. The result is trivially true when $m = 0$ and $G = \{e\}$.

Suppose by induction that $W_\ell := V^{G_\ell} \neq \{0\}$ and consider the action of $G_{\ell+1}/G_\ell$ on W_ℓ. Let σ be a generator of the group $G_{\ell+1}/G_\ell \cong C_p$. Consider the Jordan Normal Form for the action of σ on W_ℓ. Choose a Jordan block with associated eigenvalue λ. Since σ has order p, we have $\lambda^p = 1$. But then $(\lambda - 1)^p = 0$ and therefore, $\lambda = 1$. Therefore, the eigenvector associated to the block is in fact fixed by σ. Since $G_{\ell+1}/G_\ell$ is cyclic, this eigenvector is fixed by all of $G_{\ell+1}/G_\ell$. Hence $V^{G_{\ell+1}} = (V^{G_\ell})^{(G_{\ell+1}/G_\ell)} = W_\ell^{(G_{\ell+1}/G_\ell)} \neq \{0\}$, finishing the proof. □

Fixing a basis for V, we will denote by $U(V)$ the subgroup of $\mathrm{GL}(V)$ consisting of lower triangular matrices with ones along the diagonal. We will also use the notation $\mathrm{U}_n(\mathbb{F})$ to denote this subgroup when V has dimension n.

Proposition 4.0.2. *Let G be a p-group. Suppose V is an n dimensional representation of G defined over any field $\mathbb{F}$ of characteristic p. Then G is conjugate to a subgroup of $\mathrm{U}(V)$.*

H.E.A. Eddy Campbell, D. Wehlau, *Modular Invariant Theory*,
Encyclopaedia of Mathematical Sciences 139,
DOI 10.1007/978-3-642-17404-9_4, © Springer-Verlag Berlin Heidelberg 2011

Proof. We will construct an ordered basis of V with respect to which every element of G is lower triangular. The proof proceeds by induction on the dimension n of V. If V has dimension 1, there is nothing to prove.

In general, by the previous lemma, we note that $W = V^G$ has dimension at least 1. We consider the action of G on V/W. By induction, there is a basis for V/W with respect to which every element of G is lower triangular. Appending a basis for W to the chosen basis for V/W gives a lower triangular basis for V. □

The following result, see Wilkerson [113], provides a very useful homogeneous system of parameters when working with a p-group.

Proposition 4.0.3. *Suppose that the action of G on V is lower triangular with respect to the basis $\{e_1, e_2, \ldots, e_n\}$ of V. Let $\{x_1, x_2, \ldots, x_n\}$ be the dual basis of V^*. Then $\mathbf{N}^G_{G_{x_1}}(x_1), \mathbf{N}^G_{G_{x_2}}(x_2), \ldots, \mathbf{N}^G_{G_{x_n}}(x_n)$ is a homogeneous system of parameters for $\mathbb{F}[V]^G$.*

Proof. The action of G on V^* is upper triangular. Thus for each $i = 1, 2, \ldots, n$ and each $\sigma \in G$, we may write $\sigma(x_i) = \alpha_{\sigma,i} x_i + q_{\sigma,i}$ where $\alpha_{\sigma,i} \in \mathbb{F}^\times$ and $q_{\sigma,i} \in \mathbb{F}[x_1, x_2, \ldots, x_{i-1}]$. Therefore $\mathbf{N}^G_{G_{x_i}}(x_i) = \prod_{\sigma \in G/G_{x_i}} \sigma(x_i) = \alpha_i x_i^{d_i} + q_i$ for some $\alpha_i \in \mathbb{F}^\times$ and $\deg_{x_i}(q_i) < d_i$ where $d_i = [G : G_{x_i}]$. Now we will use Lemma 2.6.3. To see that it applies, we must show that the only point where all n of the norms vanish is the point $0 \in \overline{V}$. To see this, consider $\mathbf{v} = (v_1, v_2, \ldots, v_n) \in \mathcal{V}(I)$ where $I = (\mathbf{N}^G_{G_{x_1}}(x_1), \mathbf{N}^G_{G_{x_2}}(x_2), \ldots, \mathbf{N}^G_{G_{x_n}}(x_n))\mathbb{F}[V]$. Since q_1 must be 0 we have $\mathbf{N}^G_{G_{x_1}}(x_1) = \alpha_1 x_1^{d_1}$ and thus $v_1 = 0$. Thus $q_2(\mathbf{v}) = 0$ and thus $\mathbf{N}^G_{G_{x_2}}(x_2)(\mathbf{v}) = 0$ forces $v_2 = 0$. But then $q_3(\mathbf{v}) = 0$ and so $\mathrm{N}^G(x_3)(\mathbf{v}) = 0$ implies $v_3 = 0$. Continuing in this manner we see that $\mathbf{v} = (0, 0, \ldots, 0)$. Thus by Lemma 2.6.3, we see that

$$\mathbf{N}^G_{G_{x_1}}(x_1), \mathbf{N}^G_{G_{x_2}}(x_2), \ldots, \mathbf{N}^G_{G_{x_n}}(x_n)$$

is homogeneous system of parameters for $\mathbb{F}[V]^G$. □

We are able to prove a more general result in Lemma 6.2.1 using term orders.

The following example shows that $\mathbb{F}[3\,V_2]^{C_p}$ is not Cohen-Macaulay.

Example 4.0.4. Consider the representation $V = 3\,V_2$ of C_p over a field $\mathbb{F}$ of characteristic p. Fix a generator σ of C_p and choose an upper triangular basis $\{x_1, y_1, x_2, y_2, x_3, y_3\}$ for $(3\,V_2)^*$ with $\sigma(y_i) = y_i + x_i$ and $\sigma(x_i) = x_i$ for $i = 1, 2, 3$. By Proposition 4.0.3, $x_1, x_2, x_3, \mathbf{N}(y_1), \mathbf{N}(y_2), \mathbf{N}(y_3)$ is a homogeneous system of parameters for $\mathbb{F}[3\,V_2]^{C_p}$. However, x_1, x_2, x_3 is not a regular sequence in $\mathbb{F}[3\,V_2]^{C_p}$. To see this, we use the elements $u_{ij} = x_i y_j - x_j y_i$ for $1 \le i < j \le 3$. These 3 elements are easily verified to be invariants. Furthermore, the relation $x_1 u_{23} - x_2 u_{13} + x_3 u_{12} = 0$, which is easily verified, shows that $\overline{x_3} \cdot \overline{u_{12}} = 0$ in $\mathbb{F}[V]^{C_p}/(x_1, x_2)$. Using $\deg(u_{12}) = 2$, it is easy to see that $u_{12} \notin \mathbb{F}[V]^{C_p}(x_1, x_2)$. Thus x_3 is a zero divisor in $\mathbb{F}[V]^{C_p}/(x_1, x_2)$. This shows that x_1, x_2, x_3 is a partial homogeneous system of parameters which is not a regular sequence in $\mathbb{F}[V]^{C_p}$ and hence that $\mathbb{F}[V]^{C_p}$ is not Cohen-Macaulay. A similar argument in a different setting is given in Example 9.1.6.

4.1 The Cyclic Group of Order 2, the Regular Representation

Suppose $G = C_2$ and V is a 2 dimensional $\mathbb{K}$ vector space. Let σ be a generator of G and let the action of G on V be given by

$$\sigma = \sigma^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

with respect to a basis $\{e_1, e_2\}$. Let $\{x, y\}$ denote the dual basis of V^*. Thus $\sigma(x^i y^j) = x^j y^i$. In particular, $(xy)^i$ is invariant. If $i \neq j$ then $x^i y^j + x^j y^i$ is invariant. We claim that $\mathbb{K}[V]^G = \mathbb{K}[x + y, xy]$. Note that this representation is the special case $n = 2$ of the standard n dimensional representation of Σ_n. It is clear that $\mathbb{K}[V]^G \supseteq \mathbb{K}[x + y, xy]$. Assume, by way of contradiction, that $\mathbb{K}[x + y, xy]$ is a proper subset of $\mathbb{K}[V]^G$. Let d be minimal such that $\mathbb{K}[x + y, xy]_d \subsetneq \mathbb{K}[V]^G_d$.

Every element $f \in \mathbb{K}[V]^G_d \setminus \mathbb{K}[x + y, xy]_d$ is of the form $f = \sum_{i=t}^{d} c_i x^i y^{d-i}$ where $c_t, c_{t+1}, \ldots, c_d \in \mathbb{K}$ and $c_t \neq 0$. Among all such f, fix one with t maximal. Define $f' := f - c_t (xy)^t (x + y)^{d-2t}$. Then f' is invariant and $f' = \sum_{i=t+1}^{d} c'_i x^i y^{d-i}$ for some scalars c'_i. By the maximality of t, we must have $f' \in \mathbb{K}[x + y, xy]_d$. But then $f = c_t (xy)^t (x + y)^{d-2t} + f' \in \mathbb{K}[x + y, xy]$. This contradiction shows that $\mathbb{K}[V]^G = \mathbb{K}[x + y, xy]$.

Of course, since $C_2 = \Sigma_2$, the symmetric group on 2 letters, we recognize these two invariants as the elementary symmetric functions in x and y. This is the best possible situation in invariant theory, in which the ring of invariant polynomials is again polynomial algebra. We see from this calculation, or from Molien's Theorem, that

$$\mathcal{H}(\mathbb{F}[V]^G, \lambda) = \frac{1}{(1 - \lambda)(1 - \lambda^2)}.$$

It is also worth noting that this calculation is independent of the field.

Note that we could have calculated the Hilbert series in advance, and its form suggests that the ring of invariants is a polynomial ring on two invariants, one of degree 1 and the other of degree 2. The reader should be warned, however, that there are rings R for which the form of the Hilbert series resembles that of a polynomial ring even though R is no such thing, see Stanley's paper [104, page 481].

4.2 A Diagonal Representation of C_2

Suppose $\mathbb{K}$ is a field of characteristic different from 2. Then $\mathbb{K}$ contains a square root of 1, different from 1, namely -1. Suppose the generator, σ, of the cyclic group C_2 acts on V^* via

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

with respect to the basis $\{x, y\}$. Then we note that $\sigma(x^i) = (-1)^i x^i$ so that x^i is invariant if and only if $i = 2j$. Similarly, y^i is invariant if and only if $i = 2j$. Observe that $x^{2j} = (x^2)^j$ and that $\sigma(x^i y^i) = x^i y^i$ is invariant. It isn't hard to prove from here that $\mathbb{K}[V]^G = \mathbb{K}[x^2, y^2, xy]$.

There are a variety of ways to consider this example; we demonstrate two particular viewpoints. First we have the isomorphism

$$\mathbb{K}[V]^G \cong \mathbb{K}[a, b, c]/(c^2 - ab)$$

where $\deg(a) = \deg(b) = \deg(c) = 2$. In another approach, we observe that $\{x^2, y^2\}$ forms a homogeneous system of parameters for $\mathbb{K}[V]^G$, and that $\mathbb{K}[V]^G$ is a free module over $\mathbb{K}[x^2, y^2]$ on the basis $\{1, xy\}$. In particular, we see that $\mathbb{K}[V]^G$ is a Cohen-Macaulay ring, for example by applying Corollary 3.1.4.

The Hilbert series of this ring of invariants is

$$\mathcal{H}(\mathbb{K}[V]^G, \lambda) = \frac{1 - t^4}{(1 - t^2)^2}.$$

4.3 Fraction Fields of Invariants of p-Groups

We recall here material from the paper of Campbell and Chuai [16]. Consider a representation V of a p-group G over a field $\mathbb{F}$ of characteristic $p > 0$. By Proposition 4.0.2, we can choose a basis $\{x_1, x_2, \ldots, x_n\}$ for V^* such that $(\sigma - 1)x_i$ is in the span of $\{x_1, x_2, \ldots, x_{i-1}\}$ for all $\sigma \in G$ and for all $i = 1, 2, \ldots, n$. In particular, we note that x_1 is invariant. We define $R[m] := \mathbb{F}[x_1, x_2, \ldots, x_m]$ for $0 \leq m \leq n$ subject to the convention that $R[0] = \mathbb{F}$. Then G acts on $R[m]$. For any non-zero $f \in R[m]$, we may express f as a polynomial in x_m and write

$$f = f_0 + f_1 x_m + \cdots + f_d x_m^d$$

with $f_i \in R[m-1]$ for all $i = 0, 1, \ldots, d$ and where $f_d \neq 0$. The leading coefficient $f_d \in R[m-1]$ of f plays a prominent role in our analysis and we therefore denote it by $c(f)$. Writing $\sigma(x_m) = x_m + \alpha_{m-1}x_{m-1} + \cdots + \alpha_1 x_1$, we have $\sigma(f) = \sum_{i=0}^{d} \sigma(f_i)(x_m + \alpha_{m-1}x_{m-1} + \cdots + \alpha_1 x_1)^i$. Therefore, $\sigma(c(f)) = c(\sigma f)$ for all $\sigma \in G$. In particular, if f is an invariant, so is $c(f)$.

For each m with $1 \leq m \leq n$, let $\phi_m \in R[m]^G$ denote a fixed homogeneous invariant with the smallest positive degree in x_m among all invariants in $R[m]^G$. The existence of ϕ_m follows from the fact that the set $R[m]^G \backslash R[m-1]$ is non-empty since $N(x_m) := \prod_{\sigma \in G} \sigma(x_m) = x_m^{|G|} + \{\text{terms of lower degree in } x_m\}$ lies in it. We take $\phi_1 = x_1$. The invariants $c_m = c(\phi_m) \in R[m-1]$ will play a special role.

Finally, note that we can make no claim as to the total degree of ϕ_m, in particular, we cannot claim that the total degree of ϕ_m is less than or equal to $|G|$ for all m. If that were true, we would be able to prove that $\mathbb{F}(V)^G$ is generated in degrees less than $|G|$, that is, that the Noether bound holds for the invariant fields of p-groups in characteristic p. We note, however, that Fleischmann, Kemper and Woodcock have proved that the Noether bound does holds for invariant fields for arbitrary representations of any finite group in any characteristic, see [40].

We first prove two lemmas.

Lemma 4.3.1. *For any $f \in R[m]^G$, there exists an integer $k \geq 0$ such that $c_m^k f \in R[m-1]^G[\phi_m]$.*

Proof. We use induction on $\deg_{x_m}(f)$. When $\deg_{x_m}(f) = 0$, there is nothing to prove. So we may assume $\deg_{x_m}(f) = d > 0$.

In the ring $R[m]_{c_m}$, the element ϕ_m/c_m is monic as a polynomial in x_m. Hence we may divide f by ϕ_m/c_m in order to obtain $f = q'(\phi_m/c_m) + r'$ where $q', r' \in R[m]_{c_m}$ with $\deg_{x_m}(r') < \deg_{x_m}(\phi_m)$. Thus

$$f = \sigma(f) = \sigma(q')(\phi_m/c_m) + \sigma(r')$$

for all $\sigma \in G$. Since

$$\deg_{x_m}(\sigma(r')) = \deg_{x_m}(r') < \deg_{x_m}(\phi_m),$$

we see by the uniqueness of remainders that $r' = \sigma(r')$ and hence $q' = \sigma(q')$ for all $\sigma \in G$. Therefore $q', r' \in R[m]_{c_m}^G$. Multiplying by a suitable power of c_m, we see that there exist an integer $s \geq 0$ and polynomials $q = c_m^{s+1}q', r = c_m^s r' \in R[m]^G$ such that $c_m^s f = q\phi_m + r$ where $\deg_{x_m}(r) = \deg_{x_m}(r') < \deg_{x_m}(\phi_m)$. Therefore, $r \in R[m-1]^G$ because ϕ_m has the least positive degree in x_m inside $R[m]^G$. Since $\deg_{x_m}(q) = \deg_{x_m}(f) - \deg_{x_m}(\phi_m)$, we see by induction that $c_m^t q \in R[m-1]^G[\phi_m]$ for some $t \geq 0$. Therefore, for $k = s + t$ we have $c_m^k f \in R[m-1]^G[\phi_m]$, as required. □

We note that it follows immediately from Lemma 4.3.1 that if $c_m = 1$ for all $m = 1, 2, \ldots, n$, then any $f \in R[m]^G$ lies in $\mathbb{F}[\phi_1, \ldots, \phi_m]$ as easily seen by induction on m. We record this observation as

Corollary 4.3.2. *If $c_m = 1$ for all $m = 1, 2, \ldots, n$, then*

$$\mathbb{F}[V]^G = \mathbb{F}[\phi_1, \ldots, \phi_n]$$

is a polynomial ring.

Lemma 4.3.3. *For any finite number of invariants $h_1, \ldots, h_t \in R[m]^G$, there exists a monomial $c = c_1^{k_1} \cdots c_m^{k_m}$ in $c_1, \ldots, c_m$, such that $ch_i \in \mathbb{F}[\phi_1, \ldots, \phi_m]$ for all $i = 1, 2, \ldots, t$.*

Proof. We use induction on m. First let $m = 1$. Since $\phi_1 = x_1$ and $c_1 = 1$, the lemma follows from Corollary 4.3.2. Now assume $m > 1$. By Lemma 4.3.1, there exists an integer $s \geq 0$ such that $c_m^s h_i \in R[m-1]^G[\phi_m]$ for all $i = 1, 2, \ldots, t$. For $i = 1, 2, \ldots, t$, write $c_m^s h_i = \sum_{j=1_d^i} a_{ij}\phi_m^j$, where $a_{ij} \in R[m-1]^G$ for all i and j. Now, since the finite set $\{a_{ij} \mid 1 \leq i \leq t, 1 \leq j \leq i_d\}$ is contained in $R[m-1]^G$, we conclude by induction that there exist $k_1, \ldots, k_{m-1} \geq 0$ such that $c_1^{k_1} \cdots c_{m-1}^{k_{m-1}} a_{ji} \in \mathbb{F}[\phi_1, \ldots, \phi_{m-1}]$ for all i and j. Hence we have, for $c = c_1^{k_1} \cdots c_{m-1}^{k_{m-1}} c_m^s$, that $ch_i \in \mathbb{F}[\phi_1, \ldots, \phi_m]$ for all $i = 1, 2, \ldots, t$, as required. □

The following theorem shows that for a p-group, the invariant field is purely transcendental, a result originally due to Miyata [81]. This formulation, however, is constructive.

Theorem 4.3.4. *Let $G \subseteq \mathrm{U}(V) \subset GL(V)$ be a p-group. Choose any set of homogeneous invariants $\phi_1, \ldots, \phi_n$ with the property that $\phi_m \in R[m]^G$ is of smallest positive degree in x_m for $1 \leq m \leq n$. Then $\mathbb{F}(V)^G = \mathbb{F}(\phi_1, \ldots, \phi_n)$. Furthermore, there exists a non-zero $f \in \mathbb{F}[\phi_1, \ldots, \phi_n]$ such that $\mathbb{F}[V]_f^G = \mathbb{F}[\phi_1, \ldots, \phi_n]_f$.*

Proof. We use the notation from above. For the first part of the theorem, we need only to show that any $h \in \mathbb{F}[V]^G$ lies in $\mathbb{F}(\phi_1, \ldots, \phi_n)$. Assume $h \in R[m]^G \backslash R[m-1]$. By Lemma 4.3.1, there exists an integer $s \geq 0$ such that $c_m^s h \in R[m-1]^G[\phi_m]$. We write $c_m^s h = \sum_{k=1}^d a_k \phi_m^k$, where $a_k \in R[m-1]^G$ for $k = 1, 2, \ldots, d$. By Lemma 4.3.3, there exists some monomial $c^K = c_1^{k_1} \cdots c_{m-1}^{k_{m-1}}$ with $c^K \cdot c_m^s \in \mathbb{F}[\phi_1, \ldots, \phi_{m-1}]$ and $c^K \cdot a_k \in \mathbb{F}[\phi_1, \ldots, \phi_{m-1}]$ for all $k = 1, 2, \ldots, d$. Thus $h \in \mathbb{F}(\phi_1, \ldots, \phi_m) \subseteq \mathbb{F}(\phi_1, \ldots, \phi_n)$.

For the proof of the second part, let $\mathbb{F}[V]^G = \mathbb{F}[g_1, \ldots, g_\ell]$. Then we can write $g_i = h_i/f$ where $h_i \in \mathbb{F}[\phi_1, \ldots, \phi_n]$ for $i = 1, 2, \ldots, \ell$ and $f \in \mathbb{F}[\phi_1, \ldots, \phi_n]$ is non-zero. Then

$$\mathbb{F}[V]^G = \mathbb{F}[h_1/f, h_2/f, \ldots, h_\ell/f] \subseteq \mathbb{F}[\phi_1, \ldots, \phi_n]_f,$$

as required. □

4.4 The Alternating Group

We study A_n, the subgroup of Σ_n consisting of all even permutations acting via its usual n dimensional representation V. An alternating function is a polynomial f with $\sigma(f) = -f$ for all odd permutations σ. It is well known that a function invariant under A_n may be uniquely written as the sum of a symmetric function together with a symmetric function times the discriminant.

Here the discriminant may be described as

$$\Delta_n = \prod_{1 \le i < j \le n} (x_j - x_i),$$

if $p = 0$ or if $p > 2$. When $p = 2$, we use the orbit sum

$$\Delta_n := \mathcal{O}_{A_n}(x_1^{n-1} x_2^{n-2} \cdots x_{n-1}).$$

We have the following Hironaka decomposition

$$\mathbb{K}[V]^{A_n} = \mathbb{K}[V]^{\Sigma_n} \oplus \mathbb{K}[V]^{\Sigma_n} \Delta_n.$$

Hence, $\mathbb{K}[V]^{A_n}$ is generated by $n+1$ elements as an algebra, and so $\mathbb{K}[V]^{A_n}$ is a hypersurface. It is easy to see that Δ_n^2 is invariant under Σ_n when $p = 0$ or $p > 2$.

To see this result in these latter cases, suppose $p = 0$ or $p > 2$, and let σ be any odd permutation so that Σ_n is generated by A_n together with σ. Suppose that $f \in \mathbb{K}[V]^{A_n}$ and define $f_+ := \frac{1}{2}(f + \sigma(f))$ and $f_- := \frac{1}{2}(f - \sigma(f))$. We note that for $\tau \in A_n$, we have $\tau(f_\pm) = \frac{1}{2}(\tau(f) \pm \tau\sigma(f)) = \frac{1}{2}(f \pm \sigma\tau'(f)) = \frac{1}{2}(f \pm \sigma(f)) = f_\pm$ for some $\tau' \in A_n$ because A_n is normal in Σ_n. Thus both f_+ and f_- are invariant under A_n. Clearly f_+ is fixed by σ and hence is a symmetric function, while $\sigma(f_-) = -f_-$. To finish the proof, we need to show that $f_- = f'\Delta_n$ for some symmetric function f'. We observe that $f_-(x_1, x_2, \ldots, x_n) = -f(x_2, x_1, \ldots, x_n)$ since the transposition interchanging x_1 and x_2 is odd. Therefore, f_- vanishes on the hyperplane determined by $x_2 - x_1 = 0$ and therefore, $(x_2 - x_1)$ divides f_-. Similarly, $x_j - x_i$ divides f_- for all $1 \le i < j \le n$. Consequently, $f_- = f'\Delta_n$ for some polynomial f'. Finally, for an odd permutation σ, we have $\sigma(f_-) = \sigma(f')\sigma(\Delta_n) = \sigma(f')(-\Delta_n)$ and we conclude $\sigma(f') = f'$ so that f' is, in fact, symmetric.

If $\mathbb{F} = \mathbb{K}$ has characteristic 2, then $\prod_{1 \le i < j \le n}(x_j - x_i) = \prod_{1 \le i < j \le n}(x_j + x_i)$ is a symmetric function. However, as we noted above, we may define

$$\Delta_n := \mathcal{O}_{A_n}(x_1^{n-1} x_2^{n-2} \cdots x_{n-1})$$

which is not symmetric but is invariant under A_n. Lemma 4.5.3 states that with this definition for Δ_n, the usual Hironaka decomposition of $\mathbb{F}[V]^{A_n}$ as a module over $\mathbb{F}[V]^{\Sigma_n}$ is still valid in characteristic 2.

4.5 Invariants of Permutation Groups

We say that V is a *permutation representation* of G if there exists a basis of V which is permuted by the action of G. We call this basis a permutation basis of V. In other words, V is a permutation representation of G

if $G \subset \Sigma_n \subset \mathrm{GL}(V)$. A key observation is that when we work in the basis of V^* dual to the permutation basis, the elements of G permute the set of all monomials. Therefore, given a monomial x^I, we form the orbit sum $\mathcal{O}_G(x^I) = \sum_{\sigma \in G/H} \sigma(x^I)$ where $H = G_{x^I}$ is the isotropy group of x^I. It is easy to see that $\mathcal{O}_G(x^I)$ is a G-invariant polynomial.

Lemma 4.5.1. *The orbit sums $\mathcal{O}_G(x^I)$ of degree d form a basis for $\mathbb{K}[V]^G_d$.*

Proof. Any $f \in \mathbb{K}[V]_d$ may be written as sum of monomials

$$f = \sum_{\deg(I)=d} a_i x^I .$$

But for any $\sigma \in G$ we have $\sigma(f) = \sum_{\deg(I)=d} a_I \sigma(x^J)$. It follows that if f is G-invariant and $x^J \in Gx^I$ then $a_J = a_I$. The result is immediate. □

Corollary 4.5.2. *Suppose that V is a permutation representation of G. The Hilbert series of $\mathbb{K}[V]^G$ depends only on $G \subset \Sigma_n$ and not on the field $\mathbb{K}$.* □

Lemma 4.5.3. *The Hironaka decomposition*

$$\mathbb{F}[V]^{A_n} = \mathbb{F}[V]^{\Sigma_n} \oplus \mathbb{F}[V]^{\Sigma_n} \Delta_n$$

is valid over a field $\mathbb{F}$ of characteristic 2 with $\Delta_n = \mathcal{O}_{A_n}(x_1^{n-1} x_2^{n-2} \cdots x_{n-1})$. □

4.6 Göbel's Theorem

In this section we give a theorem of M. Göbel [44] which provides a good generating set of orbit sums for the invariants of any permutation representation.

In order to prove Göbel's Theorem we introduce the concept of a gap. Consider a permutation representation V for a group G. Let $\{x_1, x_2, \ldots, x_n\}$ be a permutation basis for V^*. Consider a monomial $m = x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n} = x^A$. We let $\mathrm{set}(A)$ denote the set of exponents $\mathrm{set}(A) := \{a_1, a_2, \ldots, a_n\}$. We define the *height* of x^A, denoted $\mathrm{ht}(A)$ and $\mathrm{ht}(x^A)$, to be the largest exponent, $\mathrm{ht}(A) := \max\{a_i \mid 1 \le i \le n\}$. We also define $\deg(A) := \deg(x^A)$. We say the monomial x^A has a *gap* or a gap at r if there exists a non-negative integer r such that $\{r+1, r+2, \ldots, \mathrm{ht}(A)\} \subseteq \mathrm{set}(A)$ but $r \notin \mathrm{set}(A)$. Note that x^A does not have a gap means that $\mathrm{set}(A) = \{0, 1, 2, \ldots, \mathrm{ht}(A)\}$.

Theorem 4.6.1 (M. Göbel). *Let V be a permutation representation of G. Then*

$$\{\mathcal{O}_G(x^A) \mid x^A \textit{ does not have a gap}\} \cup \{x_1 x_2 \cdots x_n\}$$

is a generating set for $\mathbb{K}[V]^G$.

Proof. Given an exponent sequence $A = (a_1, a_2, \ldots, a_n)$ we write $\Lambda_s(A) := \{j \mid a_j = s\}$ and $\lambda_s(A) := |\Lambda_s(A)|$. We define a partial order on exponent sequences (and the corresponding monomials) as follows. If $\deg(A) > \deg(B)$ then we declare that $A > B$. When $\deg(A) = \deg(B)$, we declare that $A > B$ if there exists an integer t with $\lambda_t(A) > \lambda_t(B)$ and $\lambda_s(A) = \lambda_s(B)$ for all $s > t$. Note that if A and B are two sequences with $A \not> B$ and $B \not> A$, then A and B must lie in the same Σ_n-orbit.

If $0 \notin \operatorname{set}(A)$ then x^A and $\mathcal{O}_G(x^A)$ are both divisible by the invariant $s_n = x_1 x_2 \cdots x_n$. In particular, if $0 \notin \operatorname{set}(A)$ and $A \neq (1, 1, \ldots, 1)$, then $\mathcal{O}_G(x^A)$ is decomposable and so not required in any generating set. Thus we can and will assume from now on that $0 \in \operatorname{set}(A)$.

Suppose $m = x^A$ has gap at r. From m we define a new monomial $\overline{m}$ as follows:

$$\overline{m} = x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n} \text{ where } b_i = \begin{cases} a_i, & \text{if } a_i < r; \\ a_i - 1 & \text{if } a_i > r. \end{cases}$$

Consider the product of orbit sums $\mathcal{O}_G(\overline{m})\mathcal{O}_{\Sigma_n}(m/\overline{m})$. Note that $\mathcal{O}_{\Sigma_n}(m/\overline{m})$ is the d^{th} elementary symmetric function, s_d, where $d = \deg(m/\overline{m})$. Then we can write

$$\mathcal{O}_G(\overline{m})\mathcal{O}_{\Sigma_n}(m/\overline{m}) = \sum_{\alpha \in X} c_\alpha \mathcal{O}_G(m_\alpha)$$

where X is some index set, each m_α is a monomial, each c_α is a positive integer and where if $\alpha, \beta \in X$ are distinct, then $\mathcal{O}_G(m_\alpha) \neq \mathcal{O}_G(m_\beta)$. Note that we can and will assume that each m_α is of the form $m_\alpha = \overline{m} \cdot \rho(m/\overline{m})$ for some $\rho \in \Sigma_n$. To see this, note that $m_\alpha = \sigma(\overline{m}) \cdot \tau(m/\overline{m})$ for some $\sigma \in G$ and $\tau \in \Sigma_n$. Since $\mathcal{O}_G(m_\alpha) = \mathcal{O}_G(\sigma^{-1}(m_\alpha))$, we may replace m_α by $\sigma^{-1}(m_\alpha) = \overline{m} \cdot \sigma^{-1}\tau(m/\overline{m})$. In particular, we will assume that each m_α is divisible by $\overline{m}$.

Consider one of the orbit sums $\mathcal{O}_G(m_\alpha)$ where $\alpha \in X$. We claim that $m_\alpha \leq m$ in the partial order defined above. To see this, write $m = x^A$ and $\overline{m} = x^B$ and $m_\alpha = x^C$. Put $t = \operatorname{ht}(A)$. It is clear that $\operatorname{ht}(B) = t - 1$ and that

$$\Lambda_s(B) = \Lambda_{s+1}(A)$$

for $s = r, r+1, \ldots, t-1$. Also,

$$m_\alpha = \overline{m} \cdot \tau(m/\overline{m}) = (x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n}) \cdot (x_{i_1} x_{i_2} \cdots x_{i_d})$$

for some $\tau \in \Sigma_n$ and some $1 \leq i_1 < i_2 \cdots < i_d \leq n$. Taking $\Delta = \{i_1, i_2, \ldots, i_d\}$ we have

$$\begin{aligned} \Lambda_s(C) &= (\Lambda_s(B) \setminus \Delta) \sqcup (\Lambda_{s-1}(B) \cap \Delta) \\ &= (\Lambda_{s+1}(A) \setminus \Delta) \sqcup (\Lambda_s(A) \cap \Delta) \ , \end{aligned}$$

for each $s = r, r+1, \ldots, t$.

Suppose that $C \geq A$. Then $\lambda_t(C) \geq \lambda_t(A)$. Since $\Lambda_t(C) = \Lambda_t(A) \cap \Delta$, from the equation above, we must have $\Lambda_t(C) = \Lambda_t(A)$ which implies that $\Lambda_t(A) \subseteq \Delta$ and $\lambda_t(A) = \lambda_t(C)$. Hence $\lambda_{t-1}(C) \geq \lambda_{t-1}(A)$ because $C \geq A$. But $\Lambda_{t-1}(C) = (\Lambda_t(A) \setminus \Delta) \sqcup (\Lambda_{t-1}(A) \cap \Delta) = \Lambda_{t-1}(A) \cap \Delta$. Therefore, $\Lambda_{t-1}(C) = \Lambda_{t-1}(A)$, $\Lambda_{t-1}(C) \subset \Delta$ and $\lambda_{t-1}(C) = \lambda_{t-1}(A)$.

Continuing in this manner we find that $\Lambda_s(C) = \Lambda_s(A)$ and $\Lambda_s(A) \subset \Delta$ for all $s = r+1, r+2, \ldots, t$. But $|\Delta| = d = \deg(m/\overline{m}) = \sum_{s=r+1}^{t} \lambda_s(A)$ which implies that $\Delta = \sqcup_{s=r+1}^{t} \Lambda_s(A)$. Therefore,

$$\tau(m/\overline{m}) = \prod_{\substack{i \in \Lambda_s \\ r+1 \leq s \leq t}} x_i.$$

Thus our assumption that $C \geq A$ implies that $C = A$ and that $c_\alpha = 1$. Therefore

$$\mathcal{O}_G(m) = \mathcal{O}_G(\overline{m}) \cdot \mathcal{O}_{\Sigma_n}(m/\overline{m}) - \sum_{\alpha \in Y} c_\alpha \mathcal{O}_G(m_\alpha)$$

where $|Y| = |X| - 1$ and $m_\alpha < m$ for all $\alpha \in Y$. The fact that m has a gap implies that $\deg(\overline{m}) < \deg(m)$. Furthermore, $\deg(\overline{m}) \neq 0$ since $m \neq s_n$ since we have assumed that $0 \in \operatorname{set}(A)$. Thus $1 \leq \deg(\overline{m}) < \deg(m)$.

In summary, the assumption that m has a gap and $m \neq s_n$ implies that $\mathcal{O}_G(m)$ may be expressed as a polynomial in orbit sums of monomials which are smaller than m in our partial order.

Using induction with respect to the partial order this shows that if m has a gap and $m \neq s_n$, then we do not require $\mathcal{O}_G(m)$ as part of a minimal generating set of orbit sums for $\mathbb{K}[V]^G$. Thus a generating set is formed by taking s_n together with the orbit sums of all monomials without a gap. □

Remark 4.6.2. Note that, in fact, the above proof proves the stronger statement that $\mathbb{K}[V]^G$ is generated as a module over $\mathbb{K}[V]^{\Sigma_n}$ by the set $\{\mathcal{O}_G(x^I) \mid I \text{ has no gaps}\}$.

Definition 4.6.3. *An orbit sum $\mathcal{O}_G(x^A)$ is called* special *if either A has no gaps or $A = (1, 1, \ldots, 1)$. Thus Göbel's Theorem asserts that the ring of invariants of a permutation representation is always generated by its special orbit sums.*

Remark 4.6.4. If the permutation action of G on V acts transitively on the permutation basis of V, then the monomial s_n is an indecomposable orbit sum. Conversely, if the permutation basis of V consists of r orbits with $r \geq 2$, then s_n is the product of r lower degree invariant monomials, each without a gap. Thus if G does not act transitively on the permutation basis of V, then the ring of invariants is generated by the orbit sums of the monomials without gaps.

Corollary 4.6.5. *Let V be a permutation representation of G with* $\dim(V) \geq 3$. *Then*

$$\beta(V,G) \leq \binom{\dim V}{2} .$$

Proof. The largest degree monomials without a gap are those monomials m with $\mathrm{set}(m) = \{0, 1, 2, \ldots, n-1\}$. Such monomials have degree $\binom{n}{2}$. Since $n \geq 3$ we have $\binom{n}{2} \geq n = \deg(s_n)$. □

4.7 The Ring of Invariants of the Regular Representation of the Klein Group

In this section and the next we examine in detail two examples in order to illustrate the use of Göbel's Theorem and Hilbert series.

The first example is a continuation of the example we considered in §3.7.1. We are considering the regular representation of the Klein group $G := C_2 \times C_2$ over a field $\mathbb{K}$ of any characteristic. We choose a basis $\{x_1, x_2, x_3, x_4\}$ for V^* with G-action given by

$$\begin{array}{cc} x_1 \overset{\sigma_1}{\leftrightarrow} x_2 & x_1 \overset{\sigma_2}{\leftrightarrow} x_4 \\ x_3 \overset{\sigma_1}{\leftrightarrow} x_4 & x_2 \overset{\sigma_2}{\leftrightarrow} x_3 . \end{array}$$

We begin by analyzing the structure of $\mathbb{K}[V]^G$ as a module over the ring $\mathbb{K}[V]^{\Sigma_4}$ since the elementary symmetric functions s_1, s_2, s_3, s_4 form a homogeneous system of parameters for $\mathbb{K}[V]^G$.

By the proof of Göbel's Theorem (see Remark 4.6.2), we know that the ring of invariants is generated as a module over $\mathbb{K}[s_1, s_2, s_3, s_4]$ by the orbit sums without gaps. In addition to the elementary symmetric functions, the special orbit sums are associated to the decreasing exponent sequences

$$(3,2,1,0),\ (2,2,1,0),\ (2,1,1,0),\ (2,1,0,0),$$

For each such exponent sequence, as well as for the elementary symmetric functions, we need to write the Σ_4 orbit sum as G orbit sums.

In §3.7.1, we showed that

$$\mathcal{H}(\mathbb{K}[V]^G, \lambda) = \frac{1 + 2\lambda^2 + 2\lambda^4 + \lambda^6}{(1-\lambda)(1-\lambda^2)(1-\lambda^3)(1-\lambda^4)}$$

The form of this Hilbert series means that if $\mathbb{K}[V]^G$ is Cohen-Macaulay, it must be generated as a module over $\mathbb{K}[s_1, s_2, s_3, s_4]$ by 6 invariants of degrees 0, 2, 2, 4, 4 and 6. We note that these invariants may be chosen to be orbit sums by Lemma 4.5.1. Of course, if $\mathbb{K}[V]^G$ is not Cohen-Macaulay, then more generators will be needed, see Corollary 3.1.4.

Writing out the Σ_4 special orbit sums as G-orbit sums we get:

$$\begin{aligned}
\mathcal{O}_{\Sigma_4}(x_1) &= s_1 = \mathcal{O}_G(x_1)\\
\mathcal{O}_{\Sigma_4}(x_1x_2) &= s_2 = \mathcal{O}_G(x_1x_2) + \mathcal{O}_G(x_1x_3) + \mathcal{O}_G(x_1x_4)\\
\mathcal{O}_{\Sigma_4}(x_1x_2x_3) &= s_3 = \mathcal{O}_G(x_1x_2x_3)\\
\mathcal{O}_{\Sigma_4}(x_1x_2x_3x_4) &= s_4 = \mathcal{O}_G(x_1x_2x_3x_4)\\
\mathcal{O}_{\Sigma_4}(x_1^2x_2) &= s_1s_2 - 3s_3 = \mathcal{O}_G(x_1^2x_2) + \mathcal{O}_G(x_1^2x_3) + \mathcal{O}_G(x_1^2x_4)\\
\mathcal{O}_{\Sigma_4}(x_1^2x_2x_3) &= s_1s_3 - 4s_4 = \mathcal{O}_G(x_1^2x_2x_3) + \mathcal{O}_G(x_1^2x_2x_4) + \mathcal{O}_G(x_1^2x_3x_4)\\
\mathcal{O}_{\Sigma_4}(x_1^2x_2^2x_3) &= s_2s_3 - 3s_1s_4 = \mathcal{O}_G(x_1^2x_2^2x_3) + \mathcal{O}_G(x_1^2x_2x_3^2) + \mathcal{O}_G(x_1^2x_2x_4^2)\\
\mathcal{O}_{\Sigma_4}(x_1^3x_2^2x_3) &= s_1s_2s_3 - 3s_1^2s_4 - 3s_3^2 + 4s_2s_4\\
&= \mathcal{O}_G(x_1^3x_2^2x_3) + \mathcal{O}_G(x_1^3x_2x_3^2) + \mathcal{O}_G(x_1^3x_2x_4^2) + \mathcal{O}_G(x_1^3x_3^2x_4)\\
&\quad + \mathcal{O}_G(x_1^3x_2^2x_4)
\end{aligned}$$

By Göbel's Theorem, the 20 G-orbit sums which occur on the right hand side of the above equations, together with 1, generate $\mathbb{K}[V]^G$ as a module over $\mathbb{K}[s_1, s_2, s_3, s_4]$. We will now show that the 6 orbit sums 1, $\mathcal{O}_G(x_1x_3)$, $\mathcal{O}_G(x_1x_4)$, $\mathcal{O}_G(x_1^2x_2x_4)$, $\mathcal{O}_G(x_1^2x_3x_4)$, $\mathcal{O}_G(x_1^3x_2^2x_3)$ suffice to generate $\mathbb{K}[V]^G$ as a module over $\mathbb{K}[s_1, s_2, s_3, s_4]$.

Clearly, $\mathcal{O}_G(x_1) = s_1$, $\mathcal{O}_G(x_1x_2x_3) = s_3$ and $\mathcal{O}_G(x_1x_2x_3x_4) = s_4$ are not required as module generators.

The following identities, although tedious, are all easily verified. They show how the remaining 12 candidate orbit sums lie in the $\mathbb{K}[s_1, s_2, s_3, s_4]$-module generated by the 6 orbit sums listed above.

$$\begin{aligned}
\mathcal{O}_G(x_1x_2) &= s_2 - \mathcal{O}_G(x_1x_3) - \mathcal{O}_G(x_1x_4)\\
\mathcal{O}_G(x_1^2x_2) &= -s_3 + s_1(s_2 - \mathcal{O}_G(x_1x_3) - \mathcal{O}_G(x_1x_4))\\
\mathcal{O}_G(x_1^2x_3) &= -s_3 + s_1\mathcal{O}_G(x_1x_3)\\
\mathcal{O}_G(x_1^2x_4) &= -s_3 + s_1\mathcal{O}_G(x_1x_4)\\
\mathcal{O}_G(x_1^2x_2x_3) &= (s_1s_3 - 4s_4) - \mathcal{O}_G(x_1^2x_2x_4) - \mathcal{O}_G(x_1^2x_3x_4)\\
\mathcal{O}_G(x_1^2x_2^2x_3) &= -(s_1s_4 - s_2s_3) + s_3\mathcal{O}_G(x_1x_4) + s_3\mathcal{O}_G(x_1x_3)\\
\mathcal{O}_G(x_1^2x_2x_3^2) &= -(s_1s_4) + s_3\mathcal{O}_G(x_1x_3)\\
\mathcal{O}_G(x_1^2x_2x_4^2) &= -(s_1s_4) + s_3\mathcal{O}_G(x_1x_4)\\
\mathcal{O}_G(x_1^3x_2^2x_4) &= -(s_1^2s_4 - s_1s_2s_3 + s_3^2) - (s_1s_3 - 2s_4)\mathcal{O}_G(x_1x_3)\\
&\quad - (s_1s_3 - 2s_4)\mathcal{O}_G(x_1x_4) - \mathcal{O}_G(x_1^3x_2^2x_3)\\
\mathcal{O}_G(x_1^3x_3x_4^2) &= -(s_1^2s_4 - s_1s_2s_3 + s_3^2) - s_1s_3\mathcal{O}_G(x_1x_3) - s_2\mathcal{O}_G(x_1^2x_2x_4)\\
&\quad - \mathcal{O}_G(x_1^3x_2^2x_3)\\
\mathcal{O}_G(x_1^3x_2x_3^2) &= -(s_1^2s_4 - s_1s_2s_3 + 2s_2s_4 + s_3^2) - 2s_4\mathcal{O}_G(x_1x_4)\\
&\quad - s_2\mathcal{O}_G(x_1^2x_2x_4) - s_2\mathcal{O}_G(x_1^2x_3x_4) - \mathcal{O}_G(x_1^3x_2^2x_3)
\end{aligned}$$

$$\begin{aligned}\mathcal{O}_G(x_1^3x_3^2x_4) &= -(s_1s_2s_3 - 4s_2s_4) + (s_1s_3 - 2s_4)\mathcal{O}_G(x_1x_3) - 2s_4\mathcal{O}_G(x_1x_4)\\ &\quad + s_2\mathcal{O}_G(x_1^2x_2x_4) + s_2\mathcal{O}_G(x_1^2x_3x_4) + \mathcal{O}_G(x_1^3x_2^2x_3)\\ \mathcal{O}_G(x_1^3x_2x_4^2) &= -(s_1s_2s_3 - 2s_2s_4) - s_1s_3\mathcal{O}_G(x_1x_3) + (s_1s_3 - 2s_4)\mathcal{O}_G(x_1x_4)\\ &\quad + s_2\mathcal{O}_G(x_1^2x_2x_4) + \mathcal{O}_G(x_1^3x_2^2x_3)\end{aligned}$$

This shows that $\mathbb{K}[V]^G$ is the $\mathbb{K}[s_1, s_2, s_3, s_4]$-module generated by 1, $\mathcal{O}_G(x_1x_3)$, $\mathcal{O}_G(x_1x_4)$, $\mathcal{O}_G(x_1^2x_2x_4)$, $\mathcal{O}_G(x_1^2x_3x_4)$, $\mathcal{O}_G(x_1^3x_2^2x_3)$. Furthermore, applying Proposition 3.1.4, we see that $\mathbb{K}[V]^G$ is in fact the free $\mathbb{K}[s_1, s_2, s_3, s_4]$-module generated by these 6 orbit sums and thus that $\mathbb{K}[V]^G$ is Cohen-Macaulay.

Next, we turn to the question of a minimal algebra generating set for this example. The above shows that $\mathbb{K}[V]^G$ is generated by the 4 primary invariants s_1, s_2, s_3, s_4 and the 6 secondary invariants listed above. Of course, 1 is not required in an algebra generating set. Also, since $\mathcal{O}_G(x_1^2x_2x_4) = \mathcal{O}_G(x_1x_2)\mathcal{O}_G(x_1x_4)$ and $\mathcal{O}_G(x_1^2x_3x_4) = \mathcal{O}_G(x_1x_3)\mathcal{O}_G(x_1x_4)$, we see that $\mathcal{O}_G(x_1^2x_2x_4)$ and $\mathcal{O}_G(x_1^2x_3x_4)$ are decomposable invariants and so not part of a minimal algebra generating set. Similarly, the identity

$$\begin{aligned}\mathcal{O}_G(x_1^3x_2^2x_3) &= s_3s_2s_1 - 4s_4s_2 + \mathcal{O}_G(x_1x_4)\mathcal{O}_G(x_1x_3)^2 + s_2\mathcal{O}_G(x_1x_4)^2\\ &\quad - s_2^2\mathcal{O}_G(x_1x_4) - s_3s_1\mathcal{O}_G(x_1x_3) + 2s_4\mathcal{O}_G(x_1x_3)\end{aligned}$$

shows that $\mathcal{O}_G(x_1^3x_2^2x_3)$ is decomposable. Therefore, $\mathbb{K}[V]^G$ is generated by $s_1, s_2, \mathcal{O}_G(x_1x_3), \mathcal{O}_G(x_1x_4), s_3, s_4$.

Continuing we may write

$$4s_4 = s_1s_3 - \mathcal{O}_G(x_1x_2)\mathcal{O}_G(x_3x_4) - \mathcal{O}_G(x_1x_3)\mathcal{O}_G(x_2x_4) - \mathcal{O}_G(x_1x_4)\mathcal{O}_G(x_2x_3).$$

First suppose that the characteristic of $\mathbb{K}$ is different from 2. Then $1/4 \in \mathbb{K}$ and so we may use the above identity to write s_4 as a polynomial in the other 5 invariants. Therefore, $\mathbb{K}[V]^G$ is generated by $s_1, s_2, \mathcal{O}_G(x_1x_3), \mathcal{O}_G(x_1x_4), s_3$.

This is the best we can do. If we could get by with only 4 generators, then $\mathbb{K}[V]^G$ would be a polynomial ring. The form of the Hilbert series shows this is not the case. More directly, recall that the Hilbert series begins $\mathcal{H}(\mathbb{K}[V]^G, \lambda) = 1 + \lambda + 4\lambda^2 + 5\lambda^3 + 11\lambda^4 + \dots$. Of course, we need the linear generator s_1. We also need three quadratic generators in addition to s_1^2 to span the degree 2 invariants. These 4 invariants will generate only a 4 dimensional subspace of $\mathbb{K}[V]_3$ and thus a fourth generator of degree 3 is required.

Now suppose that $\mathbb{K}$ has characteristic 2. In this case, 6 algebra generators are required. To see this, consider the point $v = (1, 1, 1, 1)$. The first five generators $s_1, s_2, \mathcal{O}_G(x_1x_3), \mathcal{O}_G(x_1x_4), s_3$ all vanish at v. However $s_4(v) = 1$. Hence s_4 cannot be expressed in terms of the other 5 generators. Using degree arguments it is easy to see that none of the 5 remaining generators can be expressed in the other 4.

4.8 The Ring of Invariants of the Regular Representation of C_4

We consider the regular representation of the cyclic group $G := C_4$ of order 4 over a field $\mathbb{K}$. We choose a permutation basis for V and we denote the corresponding permutation basis for V^* by $\{x_1, x_2, x_3, x_4\}$. So a generator $\sigma \in G$ maps x_1 to x_2, x_2 to x_3, x_3 to x_4, and x_4 to x_1.

As we saw in §3.7.2, the Hilbert series of the ring of invariants $\mathbb{K}[V]^G$ in any characteristic is given by

$$\begin{aligned}\mathcal{H}(\mathbb{K}[V]^G, \lambda) &= \frac{1}{4}\left(\frac{1}{(1-\lambda)^4} + \frac{1}{(1-\lambda^2)^2} + \frac{2}{(1-\lambda^4)}\right)\\ &= 1 + \lambda + 3\lambda^2 + 5\lambda^3 + 10\lambda^4 + 14\lambda^5 + 22\lambda^6 + \dots\\ &= \frac{1+\lambda^2+\lambda^3+2\lambda^4+\lambda^5}{(1-\lambda)(1-\lambda^2)(1-\lambda^3)(1-\lambda^4)}\end{aligned}$$

By Göbel's theorem, we know that $\mathbb{K}[V]^G$ is generated by elements of degrees at most 6. As in the case of the regular representation of the Klein group, we analyze the ring of invariants as a module over

$$\mathbb{K}[x_1, x_2, x_3, x_4]^{\Sigma_4} = \mathbb{K}[s_1, s_2, s_3, s_4],$$

where s_ℓ denotes the elementary symmetric functions. Thus if $\mathbb{K}[V]^G$ is Cohen-Macaulay, the generators for $\mathbb{K}[V]^G$ as a $\mathbb{K}[V]^{\Sigma_4}$-module will have degrees 0, 2, 3, 4, 4, and 5.

By the proof of Göbel's Theorem (see Remark 4.6.2), we know that the ring of invariants is generated as a module over $\mathbb{K}[s_1, s_2, s_3, s_4]$ by the orbit sums without gaps. In addition to the elementary symmetric functions, the special orbit sums — just as in the case of the regular representation of the Klein group — are associated to the decreasing exponent sequences.

$$(3,2,1,0),\ (2,2,1,0),\ (2,1,1,0),\ (2,1,0,0).$$

Proceeding just as before to write the Σ_4 orbit sums of these and the elementary symmetric functions as G orbit sums, we find that there are 19 orbit sums from which to determine module generators:

$$\begin{aligned}\mathcal{O}_{\Sigma_4}(x_1) &= s_1 = \mathcal{O}_G(x_1)\\ \mathcal{O}_{\Sigma_4}(x_1x_2) &= s_2 = \mathcal{O}_G(x_1x_2) + \mathcal{O}_G(x_1x_3)\\ \mathcal{O}_{\Sigma_4}(x_1x_2x_3) &= s_3 = \mathcal{O}_G(x_1x_2x_3)\\ \mathcal{O}_{\Sigma_4}(x_1x_2x_3x_4) &= s_4 = \mathcal{O}_G(x_1x_2x_3x_4)\\ \mathcal{O}_{\Sigma_4}(x_1^2x_2) &= \mathcal{O}_G(x_1^2x_2) + \mathcal{O}_G(x_1^2x_3)\\ \mathcal{O}_{\Sigma_4}(x_1^2x_2x_3) &= \mathcal{O}_G(x_1^2x_2x_3) + \mathcal{O}_G(x_1^2x_2x_4) + \mathcal{O}_G(x_1^2x_3x_4)\\ \mathcal{O}_{\Sigma_4}(x_1^2x_2^2x_3) &= \mathcal{O}_G(x_1^2x_2^2x_3) + \mathcal{O}_G(x_1^2x_2x_3^2) + \mathcal{O}_G(x_1^2x_2^2x_4)\end{aligned}$$

$$\begin{aligned}\mathcal{O}_{\Sigma_4}(x_1^3x_2^2x_3) = {} & \mathcal{O}_G(x_1^3x_2^2x_3) + \mathcal{O}_G(x_1^3x_2^2x_4) + \mathcal{O}_G(x_1^3x_2x_3^2) + \mathcal{O}_G(x_1^3x_3^2x_4)\\ & + \mathcal{O}_G(x_1^3x_2x_4^2) + \mathcal{O}_G(x_1^3x_3x_4^2)\end{aligned}$$

In the Cohen-Macaulay case, we will see that

$$1, \mathcal{O}_G(x_1x_2), \mathcal{O}_G(x_1^2x_2), \mathcal{O}_G(x_1^2x_2x_3), \mathcal{O}_G(x_1^2x_2x_4), \mathcal{O}_G(x_1^2x_2^2x_3)$$

generate $\mathbb{K}[V]^G$ as a $\mathbb{K}[s_1, s_2, s_3, s_4]$-module.

Here are the identities needed to verify this:

$$\begin{aligned}
\mathcal{O}_G(x_1x_3) &= s_2 - \mathcal{O}_G(x_1x_2)\\
\mathcal{O}_G(x_1^2x_3) &= (s_1s_2 - s_3) - s_1\mathcal{O}_G(x_1x_2)\\
\mathcal{O}_G(x_1^2x_3x_4) &= (s_1s_3 - 4s_4) - \mathcal{O}_G(x_1^2x_2x_3) - \mathcal{O}_G(x_1^2x_2x_4)\\
\mathcal{O}_G(x_1^2x_2x_3^2) &= (-s_1s_4 + s_2s_3) - s_3\mathcal{O}_G(x_1x_2))\\
\mathcal{O}_G(x_1^2x_2^2x_4) &= -2s_1s_4 + s_3\mathcal{O}_G(x_1x_2) - \mathcal{O}_G(x_1^2x_2^2x_3)\\
\mathcal{O}_G(x_1^3x_2^2x_4) &= (-s_1^2s_4 + 2s_2s_4) - s_4\mathcal{O}_G(x_1x_2) + s_2\mathcal{O}_G(x_1^2x_2x_4)\\
&\quad - s_1\mathcal{O}_G(x_1^2x_2^2x_3) + \mathcal{O}_G(x_1^3x_2^2x_3)\\
\mathcal{O}_G(x_1^3x_3x_4^2) &= (-s_1^2s_4 - s_3^2) - s_1s_3\mathcal{O}_G(x_1x_2) - s_2\mathcal{O}_G(x_1^2x_2x_4)\\
&\quad - \mathcal{O}_G(x_1^3x_2^2x_3)\\
\mathcal{O}_G(x_1^3x_2x_3^2) &= (-s_1^2s_4 + 2s_2s_4 - s_3^2) + s_4\mathcal{O}_G(x_1x_2) + s_2\mathcal{O}_G(x_1^2x_2x_3)\\
&\quad - \mathcal{O}_G(x_1^3x_2^2x_3)\\
\mathcal{O}_G(x_1^3x_3^2x_4) &= (s_1s_2s_3 - 2s_2s_4) + (-s_1s_3 + s_4)\mathcal{O}_G(x_1x_2) - s_2\mathcal{O}_G(x_1^2x_2x_3)\\
&\quad + \mathcal{O}_G(x_1^3x_2^2x_3)\\
\mathcal{O}_G(x_1^3x_2x_4^2) &= (2s_2s_4 - s_3^2) - s_4\mathcal{O}_G(x_1x_2) + s_1\mathcal{O}_G(x_1^2x_2x_3) - \mathcal{O}_G(x_1^3x_2^2x_3)\\
2\mathcal{O}_G(x_1^3x_2^2x_3) &= s_3\mathcal{O}_G(x_1^2x_2) - s_2\mathcal{O}_G(x_1^2x_2x_4) + s_1\mathcal{O}_G(x_1^2x_2x_3)
\end{aligned}$$

Thus if the characteristic of $\mathbb{K}$ is not 2, then the 6 module generators listed above suffice. However, in characteristic 2, the last equation yields a $\mathbb{K}[s_1, s_2, s_3, s_4]$ linear relation

$$s_3\mathcal{O}_G(x_1^2x_2) - s_2\mathcal{O}_G(x_1^2x_2x_4) + s_1\mathcal{O}_G(x_1^2x_2x_3) = 0$$

among the module generators. By considering the Hilbert series we see that in light of this relation, we need an extra module generator which we may take to be $\mathcal{O}_G(x_1^3x_2^2x_3)$. Thus, in characteristic 2, $\mathbb{K}[V]^G$ is not Cohen-Macaulay and requires 7 secondary generators.

Now we consider algebra generating sets. The identity

$$s_1s_3 - s_2\mathcal{O}_G(x_1x_2) + \mathcal{O}_G(x_1x_2)^2 - \mathcal{O}_G(x_1^2x_2x_4) = 4s_4$$

rewrites as

$$\mathcal{O}_G(x_1^2x_2x_4) = s_1s_3 - s_2\mathcal{O}_G(x_1x_2) + \mathcal{O}_G(x_1x_2)^2 - 4s_4.$$

Hence $\mathcal{O}_G(x_1^2x_2x_4)$ is not required as a generator (in the presence of s_4).

Also,

$$\begin{aligned}2\mathcal{O}_G(x_1^2x_2^2x_3) = &-2s_1s_4 + s_1\mathcal{O}_G(x_1^2x_2x_3) - s_2s_3 - s_2\mathcal{O}_G(x_1^2x_2)\\ &+ 2s_3\mathcal{O}_G(x_1x_2) + \mathcal{O}_G(x_1x_2)\mathcal{O}_G(x_1^2x_2)\end{aligned}$$

which shows that $\mathcal{O}_G(x_1^2x_2^2x_3)$ is decomposable if the characteristic of the field is not 2.

Together with the above relations these two identities show that $\mathbb{F}[V]^G$ is generated as an algebra by

$$s_1, s_2, s_3, s_4, \mathcal{O}_G(x_1x_2), \mathcal{O}_G(x_1^2x_2), \mathcal{O}_G(x_1^2x_2x_3)$$

and (if the characteristic of $\mathbb{K}$ is 2)

$$\mathcal{O}_G(x_1^2x_2^2x_3), \mathcal{O}_G(x_1^3x_2^2x_3)\ .$$

Thus in characteristics different from 2, we see that

$$\mathbb{K}[V]^G = \mathbb{K}[s_1, s_2, \mathcal{O}_G(x_1x_2), s_3, \mathcal{O}_G(x_1^2x_2), \mathcal{O}_G(x_1^2x_2x_3), s_4].$$

Note that the largest required generator has degree 4 which is the order of G.

In characteristic 2, we have

$$s_1^2s_4 + s_3^2 + \mathcal{O}_G(x_1x_2)\mathcal{O}_G(x_1^2x_2x_3) = \mathcal{O}_G(x_1^3x_2^2x_3).$$

Thus $\mathcal{O}_G(x_1^3x_2^2x_3)$ is not required as an algebra generator even in characteristic 2. Thus, in characteristic 2, $\mathbb{K}[V]^G$ is generated by

$$\left\{s_1, s_2, \mathcal{O}_G(x_1x_2), s_3, \mathcal{O}_G(x_1^2x_2), s_4, \mathcal{O}_G(x_1^2x_2x_3), \mathcal{O}_G(x_1^2x_2x_4), \mathcal{O}_G(x_1^2x_2^2x_3)\right\}.$$

From the Hilbert series, we know that $\dim_{\mathbb{K}}[V]_5^G = 14$. However, the extra relation in characteristic 2 given by

$$s_1\mathcal{O}_G(x_1^2x_2x_3) - s_2s_3 - s_2\mathcal{O}_G(x_1^2x_2) + \mathcal{O}_G(x_1x_2)\mathcal{O}_G(x_1^2x_2) = 0$$

(which follows from the above identities) shows that the 8 generators other than $\mathcal{O}_G(x_1^2x_2^2x_3)$ only generate a 13 dimensional subspace of $\mathbb{K}[V]_5$. Thus an algebra generator of degree 5 is required in characteristic 2.

Remark 4.8.1. The example of the regular representation of C_4 over the field $\mathbb{F}_2$ of order 2 is important historically. Pierre Samuel asked the question whether every unique factorization domain was Cohen-Macaulay. His student Marie-José Bertin [9] answered this question in the negative using the example we have just considered. She showed that this ring of invariants is not Cohen-Macaulay and since it is the ring of invariants of a p-group, it is a unique factorization domain. For a good description of this, see [43].

4.9 A 2 Dimensional Representation of C_3, $p = 2$

We consider a representation V of $G = C_3$ with generator σ on V of dimension 2 over $\mathbb{F}_2$. We suppose that the action of σ on V^* with basis $\{x, y\}$ given

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

Then σ acting on $\mathbb{F}_2[V] = \mathbb{F}_2[x, y]$ sends x to y and sends y to $x + y$.

It is straightforward to calculate the ring of invariants for G. First we observe that the Dickson invariants (see §3.3) $r = x^2 + xy + y^2, s = x^2y + xy^2$ form, as always, a homogeneous system for $\mathbb{F}_2[V]^G$. Second we observe that G has index 2 in $\mathrm{GL}_2(\mathbb{F}_2)$, and that $t = x^3 + x^2y + y^3$ is invariant. It isn't hard to see using Galois theory that

$$\mathbb{F}_2[x, y]^G = \mathbb{F}_2[x^2 + xy + y^2, x^2y + xy^2, x^3 + x^2y + y^3],$$

and, therefore, this ring is a hypersurface. As part of this calculation, we note $t^2 = r^3 + s^2 + rs$.

Therefore, we obtain a resolution over the ring $A = \mathbb{F}_2[a, b, c]$ with $\deg(a) = 2$, $\deg(b) = \deg(c) = 3$, and $\rho(a) = x^2 + xy + y^2$, $\rho(b) = x^2y + xy^2$, $\rho(c) = x^3 + x^2y + y^3$. We obtain

$$0 \to A(c^2 + a^3 + b^2 + bc) \to A \to \mathbb{F}_2[x, y]^G \to 0.$$

It follows that

$$\mathcal{H}(\mathbb{F}_2[V]^G, t) = \frac{1 + t^2 + t^4}{(1 - t^3)^2}.$$

4.10 The Three Dimensional Modular Representation of C_p

Suppose $p > 2$ and let $\mathbb{F}$ be a field of characteristic p. Consider the action of $C_p = \langle\sigma\rangle$ on a three dimensional $\mathbb{F}$ vector space V determined by the matrix

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

Note that this matrix has order p and thus does indeed afford a three dimensional representation of $G = C_p$. This representation turns out to be the unique indecomposable representation of C_p of this dimension. We will discuss this question and related matters later in §7.1.

We will compute the ring of invariants of this representation and see that it is a hypersurface. We will give two different computations of this ring of invariants. The first takes advantage of prior knowledge of the Hilbert series which reduces the amount of work considerably. The second is longer but does not rely on any prior knowledge.

We let $\{x, y, z\}$ denote the basis of V^* dual to the standard basis of V. We begin by observing that some invariants are easily constructed: x is invariant, as are $\mathbf{N}^G(y)$ and $\mathbf{N}^G(z)$. Furthermore, it is not hard to see that $d = y^2 - 2xz - xy$ is also invariant. You may well wonder where d came from. There are two considerations, one coming from the form of the Hilbert series, see immediately below, and the other from a consideration of what we refer to as "integral invariants", see §7.5. In the meantime, we note simply that while the polynomial expressions for $\mathbf{N}^G(y)$ and $\mathbf{N}^G(z)$ depend upon p, the expressions for x and d do not: the "same" polynomial x, or d, is invariant for all primes.

Theorem 4.10.1.

$$\mathbb{F}[V]^G = \mathbb{F}[x, \mathbf{N}^{C_p}(y), \mathbf{N}^{C_p}(z), d]$$

The next two sections are devoted to the two proofs of this theorem.

4.10.1 Prior Knowledge of the Hilbert Series

Our first approach to this example uses the work of Almkvist and Fossum [3]. They have shown that there exists a three dimensional representation of C_p in characteristic 0 which has the same Hilbert series as the current ring of invariants $\mathbb{F}[x, y, z]^{C_p}$. Let ξ be a complex (primitive) p-th root of unity and consider the matrix

$$\tau = \begin{pmatrix} \xi^{-2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \xi^2 \end{pmatrix}.$$

For $p > 2$, the matrix τ has order p and the Hilbert series of the associated representation of C_p can be shown to be the same as the Hilbert series of the characteristic p representation. Almkvist and Fossum's work has been extended and explained by Hughes and Kemper in references [54] and [55].

Lemma 4.10.2. *We have*

$$\mathcal{H}(\mathbb{F}[V]^G, \lambda) = \frac{\sum_{i=0}^{p-1} \lambda^{2i}}{(1-\lambda)(1-\lambda^p)^2} = \frac{1+\lambda^p}{(1-\lambda)(1-\lambda^2)(1-\lambda^p)}.$$

□

The first form of the Hilbert series suggests that there is a homogeneous system of parameters for $\mathbb{F}[V]^G$ consisting of one invariant of degree 1 and two of degree p, and that there is a fourth invariant of degree 2, whose powers give module generators for the ring of invariants over the subalgebra generated by the homogeneous system.

The second form of the Hilbert series suggests that there is a homogeneous system of parameters for $\mathbb{F}[V]^G$ consisting of one invariant of degree 1, one invariant of degree 2 and one invariant of degree p, and that there is a fourth invariant of degree p which generates the ring of invariants as a module over the subalgebra generated by the homogeneous system.

Both of these suggestions turn out to be correct, but the reader should be warned that the form of the Hilbert series cannot always be "realized".

It is easy to find three invariants: first of all x is invariant, and we may form $\mathbf{N}^G(y) = y^p - x^{p-1}y$ and $\mathbf{N}^G(z) = z^p - xz^{p-1} - \ldots$. We note that $\mathbf{N}^G(y)$ is $\mathbb{F}$-linear in y, but that $\mathbf{N}^G(z)$ is not. Consideration of this phenomenon eventually leads to the following.

Lemma 4.10.3. *1. Let $x \in V^*$ and suppose G is a p-group acting on V^* with the property that*

$$(G-1)x = \{(\sigma - 1)(x) \mid \sigma \in G\}$$

is a vector space. Then $\mathbf{N}^G(x)$ is a p-polynomial, that is, is the only powers of x which occur are of the form x^{p^t} and hence the polynomial is $\mathbb{F}_p$-linear in x.

2. The coefficients of the various powers of x in the polynomial $\mathbf{N}^G(x)$ are the Dickson invariants in the elements of the vector space $(G-1)(x)$.

3. $\mathbf{N}^G(x)$ is $\mathbb{F}$-linear in x if and only if $(G-1)(x)$ is a vector space.

□

By Proposition 4.0.3, the sequence $x, \mathbf{N}(y), \mathbf{N}(z)$ is a homogeneous system of parameters for $\mathbb{F}[V]^{C_p}$. Therefore, $\mathbb{F}[V]^G$ is finitely generated as a module over the subalgebra $A = \mathbb{F}[x, \mathbf{N}(y), \mathbf{N}(z)]$. As noted above, the form of the Hilbert series suggests the existence of an invariant of degree 2. We note that $\mathbb{F}[V]_2$ has dimension 6, and it is fairly easy to compute the action of σ on $\mathbb{F}[V]_2$ and determine that $d = y^2 - 2xz - xy$ and x^2 span the homogeneous invariants of degree 2. We also note that the proof of Proposition 4.0.3 shows that $\{x, d, \mathbf{N}(z)\}$ is also a homogeneous system of parameters for $\mathbb{F}[V]^{C_p}$.

What relation holds between the generators above? We observe that d^p and $\mathbf{N}(y)^2$ both have terms of the form y^{2p} and this allows the one to be written as a polynomial expression in x, $\mathbf{N}(z)$ and the other. Further, no lower power of d^i can be written as a polynomial in $x, \mathbf{N}(y)$ and $\mathbf{N}(z)$ given the term y^{2i} in d^i. Thus

$$\mathbb{F}[x, d, \mathbf{N}(y), \mathbf{N}(z)] = \oplus_{i=0}^{p-1} \mathbb{F}[x, \mathbf{N}(y), \mathbf{N}(z)] d^i.$$

Similarly, $\mathbf{N}(y)^2 \in \mathbb{F}[x, d, \mathbf{N}(z)$ but $\mathbf{N}(y) \notin \mathbb{F}[x, d, \mathbf{N}(z)$. Hence

$$\mathbb{F}[x, d, \mathbf{N}(y), \mathbf{N}(z)] = \mathbb{F}[x, d, \mathbf{N}(z)] \oplus \mathbb{F}[x, d, \mathbf{N}(z)]\mathbf{N}(y)$$

Using either of these decompositions we can find $\mathcal{H}(\mathbb{F}[x, d, \mathbf{N}(y), \mathbf{N}(z)], \lambda)$. For example, from the second decomposition, we see

$$\begin{aligned}\mathcal{H}(\mathbb{F}[x, d, \mathbf{N}(y), \mathbf{N}(z)], \lambda) &= \mathcal{H}(\mathbb{F}[x, d, \mathbf{N}(z)], \lambda)(1 + \lambda^p) \\ &= \frac{1 + \lambda^p}{(1 - \lambda)(1 - \lambda^2)(1 - \lambda^p)}.\end{aligned}$$

Given the computation of Almkvist and Fossum above which computes the Hilbert series of $\mathbb{F}[V]^G$, we see

$$\mathcal{H}(\mathbb{F}[x, d, \mathbf{N}(y), \mathbf{N}(z)], \lambda) = \mathcal{H}(\mathbb{F}[V]^G, \lambda).$$

Since $\mathbb{F}[x, d, \mathbf{N}(y), \mathbf{N}(z)] \subseteq \mathbb{F}[V]^G$, these two rings must be equal.

We contrast the preceding with the computation below in which none of the individual steps are hard, but there are many steps to take.

4.10.2 Without the Use of the Hilbert Series

Without the information encoded in the Hilbert series we must work harder to compute the ring of invariants. As above, it is easy to construct invariants, but it is much more difficult to know when to stop. That is, it is hard to determine which sets of invariants are generating sets in the absence of other information.

The next theorem tells us that the only invariants in x, y and z which have degree less than p in z must be polynomials in x, d and $\mathbf{N}^G(y) = \mathbf{N}(y)$.

Theorem 4.10.4. *If $f \in \mathbb{F}[V_3]^{C_p}$ and $f = \sum_{i=0}^{j} f_i z^i$ where $f_i \in \mathbb{F}[x, y, d]$ and $j < p$, then $f \in \mathbb{F}[x, d, \mathbf{N}(y)]$.*

Proof. The proof of the theorem is by induction on j.

For $j = 0$, $f = f_0 \in \mathbb{F}[x, y, d]$. Write

$$f = \sum_{i=0}^{n} a_i d^i$$

where $a_i \in \mathbb{F}[x, y]$ and apply delta:

$$0 = \Delta(f) = \sum_{i=0}^{n} \Delta(a_i) d^i.$$

Since $\{x, y, d\}$ is algebraically independent, $\Delta(a_i) = 0$. Thus

$$a_i \in \mathbb{F}[V_2]^{C_p} = \mathbb{F}[x, \mathbf{N}(y)].$$

For $j > 0$, we apply delta to f to get

$$0 = \Delta(f) = \Delta(f_j) z^j + \sigma(f_j)\Delta(z^j) + \dots$$

Since $\deg_z(\Delta(z^j)) < j$, we have $\Delta(f_j) = 0$. Rewriting $\Delta(f)$ gives

$$0 = \Delta(f_j)\sigma(z^j) + f_j\Delta(z^j) + \Delta(f_{j-1})z^{j-1} + \dots$$

or

$$0 = jf_jyz^{j-1} + \Delta(f_{j-1})z^{j-1} + \text{terms with smaller z-degree.}$$

Thus $f_j \in \Delta(\mathbb{F}[x, y, d]) \subset (x)\mathbb{F}[x, y, d]$.

Write $f_j = xh$ with $h \in \mathbb{F}[x, y, d]$. Let $d' = d - zx \in \mathbb{F}[x, y]$ so that $zx = d - d'$.

Then

$$f_jz^j = h(xz)z^{j-1} = hdz^{j-1} - hd'z^{j-1}$$

and

$$f = (hd - hd' + f_{j-1})z^{j-1} + \sum_{i=0}^{j-2} f_iz^i.$$

Thus, by induction, $f \in \mathbb{F}[x, d, \mathbf{N}(y)]$.

□

Since $\mathbf{N}(z)$ is monic when considered as a polynomial in the variable z, we may divide any polynomial f by $\mathbf{N}(z)$ to get $f = q\mathbf{N}(z) + r$ where q and r are unique with $\deg_z r < p$.

Lemma 4.10.5. *If $f \in R^G$ and $f = q\mathbf{N}(z) + r$ with $\deg_z r < p$, then $q, r \in R^G$.*

Proof. We have $f = \sigma \cdot f = (\sigma \cdot q)(\sigma \cdot \mathbf{N}(z)) + (\sigma \cdot r) = (\sigma \cdot q)\mathbf{N}(z) + (\sigma \cdot r)$. Since $\sigma \cdot z = z + y$, $\sigma \cdot y = y + x$ and $\sigma \cdot x = x$, it follows that $\deg_z(\sigma \cdot h) = \deg_z(h)$ for all $h \in R$. In particular, $\deg_z(\sigma \cdot r) = \deg_z(r) < p$. Thus by the uniqueness of remainders and quotients, we must have $\sigma \cdot r = r$ and $\sigma \cdot q = q$. Hence $\sigma^i \cdot r = r$ and $\sigma^i \cdot q = q$ and thus $q, r \in R^G$. □

Consider $f \in \mathbb{F}[x, y, z]^G$ and write $f = q \cdot \mathbf{N}(z) + r$ where $\deg_z(r) < p$. We may also write

$$r = \sum_{i=0}^{p-1} r_iz^i$$

where $r_i \in \mathbb{F}[x, y]$.

Before proceeding to the proof of Theorem 4.10.1, we have the following lemma.

Lemma 4.10.6. *x^i divides r_i.*

Proof.

$$\begin{aligned} r = \sigma(r) &= \sum_{i=0}^{p-1} \sigma(r_i)\sigma(z^i) \\ &= \sum_{i=0}^{p-1} \sigma(r_i) \sum_{j=0}^{i} \binom{i}{j} y^{i-j}z^j \end{aligned}$$

Fix j with $0 \leq j \leq p-1$. Equating the coefficients of z^j in the above equations we get

$$r_j = \sum_{i=j}^{p-1} \binom{i}{j} y^{i-j} \sigma(r_i).$$

Equivalently, we have

$$(j+1)y\sigma(r_{j+1}) = -(\sigma - 1)(r_j) + \sum_{i=j+2}^{p-1} \binom{i}{j} y^{i-j} \sigma(r_i)$$

Applying σ^{-1} to this equation we obtain:

$$(j+1)\sigma^{-1}(y)r_{j+1} = -(\sigma^{-1} - 1)(r_j) + \sum_{i=j+2}^{p-1} \binom{i}{j} \sigma^{-1}(y^{i-j})(r_i) \qquad (4.10.1)$$

Note that $\sigma^{-1}(y) = y - x$ and $\sigma^{-1}(x) = x$. Thus if for some k we have x^k divides r_j, then we also have x^{k+1} divides $(\sigma^{-1} - 1)(r_j)$. Therefore, if x^{k+1} divides r_i for all $i = j+2, \ldots, p-1$ and also x^k divides r_j, then Equation 4.10.1 implies that x^{k+1} divides r_{j+1}.

We proceed by induction. We will prove that x^t divides r_t and that x^{t+1} divides r_i for all $i = t+2, \ldots, p-1$ by induction.

For the initial step, $t = 0$, we consider Equation 4.10.1 with $j = p-2$. From that equation, we see that x divides r_{p-1}. Then the above remark with $k = 0$ and $j = p-3$ shows that x divides r_{p-2}. Continuing to apply this remark we obtain that x divides r_i for all $i = 1, \ldots, p-1$ which completes the initial step of the induction.

For the general step, we assume that x^{t-1} divides r_{t-1} and that x^t divides r_i for all $i = t+1, \ldots, p-1$. But then using the above remark again with $k = t-1$ and $j = t-1$, we obtain that x^t divides r_t. Now applying the remark with $k = t$ and $j = p-2$ shows that $x^t + 1$ divides r_{p-2}. Next applying the remark with $k = t$ and $j = p-3$ shows that $x^t + 1$ divides r_{p-3}.

Continuing in this fashion we obtain that x^{t+1} divides r_i for all $i = t+1, \ldots, p-1$ which completes the proof of the lemma. □

We now give the proof of Theorem 4.10.1.

Proof. As above, we take $f \in \mathbb{F}[x,y,z]^G$ and write $f = q\mathbf{N}(z) + r$. We consider elements of $\mathbb{F}[x,y,z]$ as polynomials in z. In particular,

$$d = xz + (\frac{p-1}{2}y^2 + \frac{p+1}{2}xy).$$

By the lemma, we may divide d into r to obtain $r = d \cdot g + h$ where $\deg_z(h) = 0$. Then $r = \sigma(r) = d \cdot \sigma(g) + \sigma(h)$ with $\deg_z(\sigma(h)) = 0$. Thus by the uniqueness of the remainder, we see that $g = \sigma(g)$ and $h = \sigma(h)$, i.e., $g, h \in \mathbb{F}[x,y,z]^G$.

In summary, we have shown that if $f \in \mathbb{F}[x, y, z]^G$, then we have $f = q \cdot \mathbf{N}(z) + r = q \cdot \mathbf{N}(z) + d \cdot g + h$ where $h \in \mathbb{F}[x, y]^G = \mathbb{F}[x, \mathbf{N}(y)]$. By induction on degree we may show that

$$g, q \in \mathbb{F}[x, d, \mathbf{N}(y), \mathbf{N}(z)]$$

and thus

$$h \in \mathbb{F}[x, d, \mathbf{N}(y), \mathbf{N}(z)].$$

□

5

Monomial Orderings and SAGBI Bases

Let S denote the polynomial ring $\mathbb{K}[x_1, \dots, x_n]$. The set of *monomials* in S is $\mathcal{M} := \{x_1^{a_1} x_2^{a_2} \cdot \dots \cdot x_n^{a_n} \mid a_1, a_2, \dots, a_n \in \mathbb{N}\}$. A *term* in S is an element of the form cm where $c \in \mathbb{K}$ and $m \in \mathcal{M}$.

Definition 5.0.1. *A monomial order is a total order on $\mathcal{M}$ such that*

1. *$m > 1$ for all $m \in \mathcal{M} \setminus \{1\}$, and*
2. *if $m_1 > m_2$, then $mm_1 > mm_2$ for all monomials m, m_1, m_2.*

The following lemma shows that every monomial ordering is a well-ordering, i.e., every non-empty set of monomials contains a least element. This is the key fact we need in order to use monomial orders as the basis of induction proofs.

Lemma 5.0.2. *A monomial ordering is a well-ordering.*

Proof. Consider a non-empty set of monomials $A \subseteq \mathcal{M}$ and let I be the ideal of S generated by A. Then the monomials in I are precisely those monomials in S which are divisible by some monomial of A. Since S is a Noetherian ring, the ideal I is minimally generated by some finite subset $\{m_1, \dots, m_r\}$ of A.

Without loss of generality we assume that m_1 is the smallest monomial in the finite set $\{m_1, \dots, m_r\}$. Since $\{m_1, \dots, m_r\}$ generates I, the set of monomials in I is also characterized as those monomials of S which are divisible by some m_i with $1 \leq i \leq r$. Take any monomial m in A. Then m lies in I and so there exists i such that m_i divides m. Write $m = m'm_i$. Then $1 \leq m'$. This together with $m_1 \leq m_i$ implies $m_1 \leq m'm_1 \leq m'm_i = m$. Therefore, m_1 is less than or equal to every monomial of A. □

Now let f be any non-zero element of S and write f (uniquely) as a linear combination of distinct monomials: $f = c_1 m_1 + c_2 m_2 + \dots + c_r m_r$ where $c_i \in \mathbb{K}$ and $m_i \in \mathcal{M}$ for all $i = 1, 2, \dots, r$. Without loss of generality, assume $m_1 > m_i$ for all $i = 2, 3, \dots, r$. We say that m_1 is the *lead monomial* of f and write $\mathrm{LM}(f) = m_1$. Similarly, we say that $c_1 m_1$ is the *lead term* of f and write

H.E.A. Eddy Campbell, D. Wehlau, *Modular Invariant Theory*,
Encyclopaedia of Mathematical Sciences 139,
DOI 10.1007/978-3-642-17404-9_5, © Springer-Verlag Berlin Heidelberg 2011

$\mathrm{LT}(f) = c_1 m_1$, that c_1 is the *lead coefficient* of f and write $\mathrm{LC}(f) = c_1$. We make the conventions that $\mathrm{LM}(0) = 0$, $\mathrm{LT}(0) = 0$ and $\mathrm{LC}(0) = 0$ although we do not consider 0 a term nor a monomial.

We now give some examples of monomial orderings. In these examples we will use two different monomials $m = x_1^{a_1} x_2^{a_2} \cdot \ldots \cdot x_n^{a_n}$ and $m' = x_1^{a'_1} x_2^{a'_2} \cdot \ldots \cdot x_n^{a'_n}$.

Example 5.0.3 (Lexicographic Ordering). Define $i \geq 1$ by $a_1 = a'_1, a_2 = a'_2, \ldots, a_{i-1} = a'_{i-1}$ but $a_i \neq a'_i$. Then in the lexicographic ordering, $m <_{\mathrm{Lex}} m'$ if and only if $a_i < a'_i$.

Example 5.0.4 (Graded Lexicographic Ordering). Suppose $a_1 + a_2 \ldots + a_n \neq a'_1 + a'_2 \ldots + a'_n$, then $m <_{\mathrm{GrLex}} m'$ if and only if $a_1 + a_2 \ldots + a_n < a'_1 + a'_2 \ldots + a'_n$. On the other hand, if $a_1 + a_2 \ldots + a_n = a'_1 + a'_2 \ldots + a'_n$, then $m <_{\mathrm{grLex}} m'$ if and only if $m <_{\mathrm{Lex}} m'$. Thus if two monomials have different degrees, the one of lower degree is smaller in the graded lexicographic ordering. For two monomials of the same degree, the graded lexicographic ordering yields the same order as the lexicographic ordering.

Example 5.0.5 (Graded Reverse Lexicographic Ordering). There is a j, $1 \leq j \leq n$, such that $a_n = a'_n, a_{n-1} = a'_{n-1}, \ldots, a_{j+1} = a'_{j+1}$ but $a_j \neq a'_j$. Suppose $a_1 + a_2 \ldots + a_n \neq a'_1 + a'_2 \ldots + a'_n$, then in the graded reverse lexicographic ordering, $m <_{\mathrm{GRevLex}} m'$ if and only if $a_1 + a_2 \ldots + a_n < a'_1 + a'_2 \ldots + a'_n$. Conversely, if $a_1 + a_2 \ldots + a_n = a'_1 + a'_2 \ldots + a'_n$, then $m <_{\mathrm{GrevLex}} m'$ if and only if $a_j > a'_j$.

We might be tempted to define a reverse lexicographic ordering by putting $m <_{RevLex} m'$ if and only if $a_j > a'_j$ where j is the greatest value of i such that $a_i \neq a'_i$. Unfortunately, this ordering would not be a monomial ordering since $1 = x_1^0 x_2^0 \cdot \ldots \cdot x_n^0 > m$ for all monomials $m \neq 1$.

Example 5.0.6. Let $f = 3x_1x_3 - 5x_2^2x_4 + 2x_3^3 \in S = \mathbb{K}[x_1, x_2, x_3, x_4]$. In the lexicographic order $\mathrm{LM}(f) = x_1x_3$. In graded lexicographic order $\mathrm{LM}(f) = x_2^2x_4$ and with respect to the graded reverse lexicographic ordering $\mathrm{LM}(f) = x_3^3$.

Example 5.0.7 (Weight Orderings). Let $\mathbf{U} = U_1, U_2, \ldots, U_n$ be a sequence of n elements of $\mathbb{N}^n \subset \mathbb{R}^n$ which are linearly independent over $\mathbb{R}$. From the two monomials, we form two integer sequences $E = (e_1, e_2, \ldots, e_n)$ and $E' = (e'_1, e'_2, \ldots, e'_n)$ in $\mathbb{N}^n$ where e_j is the dot product of U_j with $(a_1, a_2, \ldots, a_n)$ and e'_j is the dot product of U_j with $(a'_1, a'_2, \ldots, a'_n)$. Since $U_1, U_2, \ldots, U_n$ span $\mathbb{R}^n$ and $m \neq m'$, we must have $E \neq E'$. Define $i \geq 1$ by $e_1 = e'_1, e_2 = e'_2, \ldots, e_{i-1} = e'_{i-1}$ but $e_i \neq e'_i$. Then we define $m <_{\mathbf{U}} m'$ if and only if $e_i < e'_i$.

The first three examples are special cases of a weight ordering. For example, if we take $\mathbf{U} = (1, 0, \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, 0, 0, \ldots, 1)$ then the two orderings $<_{\mathbf{U}}$ and $<_{\mathrm{L}ex}$ coincide. We note that the graded lexicographic order and the graded reverse lexicographic order can both be realized as weight orders.

The experience of many people has shown that computer computations done using graded reverse lexicographic order are very often faster than the same computations done in another monomial order. There are also theoretical reasons for preferring the graded reverse lexicographic order when studying the invariants of p-groups. For example, many of the proofs given in Chapter 7 rely upon properties of the graded reverse lexicographic order.

We will assume the reader to be familiar with the elements of the Gröbner basis theory for ideals.

5.1 SAGBI Bases

In invariant theory, we are interested in subalgebras of $S = \mathbb{K}[x_1, \ldots, x_n]$. This leads us to consider SAGBI bases. The word SAGBI is an acronym for Subalgebra Analogue of Gröbner Bases for Ideals. The concept of a SAGBI basis was introduced separately by Robbiano and Sweedler [93] and by Kapur and Madlener [60].

As the name suggests, the theory of SAGBI bases is very much analogous to the theory of Gröbner basis. There is one central difference between the two theories. While we know that every ideal of $S = \mathbb{K}[x_1, x_2, \ldots, x_n]$ has a finite Gröbner basis. But as we shall see there exist finitely generated subalgebras R of S which have no finite SAGBI basis.

Let R be any graded subspace of $S = \mathbb{K}[x_1, x_2, \ldots, x_n]$. As we did for ideals, we denote by $\mathrm{LT}(R)$ the vector space

$$\mathrm{LT}(R) := \mathrm{span}_{\mathbb{K}}\{\mathrm{LM}(f) : f \in R\}\ .$$

One of the most important properties of this subspace is the fact that $\mathcal{H}(R, \lambda) = \mathcal{H}(\mathrm{LT}(R), \lambda)$. Moreover, if R is a subalgebra of S, then $\mathrm{LT}(R)$ is also an algebra which we call the *lead term algebra* of R.

Lemma 5.1.1. *1. $\mathcal{H}(R, \lambda) = \mathcal{H}(\mathrm{LT}(R), \lambda)$.*
2. If R is a subalgebra of S, then $\mathrm{LT}(R)$ is an algebra.

□

Definition 5.1.2. *Let R be a subalgebra of $S = \mathbb{K}[x_1, x_2, \ldots, x_n]$. A subset $\mathcal{B}$ of R is a SAGBI basis for R if the algebra generated by $\{LT(b) \mid b \in \mathcal{B}\}$ is $\mathrm{LT}(R)$.*

The following example shows how a generating set may fail to be a SAGBI basis.

Example 5.1.3. Consider the subalgebra R of $S = \mathbb{K}[x, y, z]$ generated by the three functions $f_1 := xy + y^2$, $f_2 := xy^3 + yz^3$ and $f_3 := 3x^3y$. We use the graded reverse lexicographic monomial ordering with $x > y > z$. Thus $\mathrm{LT}(f_1) = xy$, $\mathrm{LT}(f_2) = xy^3$ and $\mathrm{LT}(f_3) = 3x^3y$.

Notice that $3\,\mathrm{LT}(f_1)^4 = 3(xy)^4 = (xy^3)(3x^3y) = \mathrm{LT}(f_2)\,\mathrm{LT}(f_3)$. This identity among lead terms gives rise to a corresponding difference, $3f_1^4 - f_2f_3$. Evaluating this difference we find

$$3f_1^4 - f_2f_3 = 12x^3y^5 + 18x^2y^6 + 12xy^7 + 3y^8 - 3x^3y^2z^3.$$

Now the lead monomial here, x^3y^5, can be obtained as $\mathrm{LM}(f_1)^2\,\mathrm{LM}(f_2)$ and so we consider

$$\begin{aligned}3f_1^4 - f_2f_3 - 12f_1^2f_2 = {}&-6x^2y^6 + 3y^8 - 3x^3y^2z^3 - 12x^2y^3z^3 - 24xy^4z^3\\ &- 12y^5z^3.\end{aligned}$$

Here the lead monomial $x^2y^6 = \mathrm{LM}(f_2)^2$ and we consider

$$\begin{aligned}3f_1^4 - f_2f_3 - 12f_1^2f_2 + 6f_2^2 = {}&3y^8 - 3x^3y^2z^3 - 12x^2y^3z^3 - 12xy^4z^3\\ &- 12y^5z^3 + 6y^2z^6.\end{aligned}$$

Thus y^8 lies in $\mathrm{LT}(R)$. However, y^8 is not in the algebra generated by the three monomials $\mathrm{LT}(f_i)$, for $i = 1, 2, 3$. Thus the original monomial relation $\mathrm{LM}(f_1)^4 = \mathrm{LM}(f_2)\,\mathrm{LM}(f_3)$ has led us to an element

$$f_4 := 3f_1^4 - f_2f_3 - 12f_1^2f_2 + 6f_2^2$$

of R whose lead monomial is not a consequence of the lead monomials of f_1, f_2, f_3.

Definition 5.1.4. *Let B be a subset of $S = \mathbb{K}[x_1, x_2, \ldots, x_n]$. A* tête-a-tête *(over B) consists of two different factorizations of a monomial with all factors taken from the set $\{\mathrm{LM}(f) \mid f \in B\}$:*

$$\prod_{i=1}^{s} \mathrm{LM}(f_i)^{a_i} = \prod_{i=1}^{s} \mathrm{LM}(f_i)^{b_i}$$

where $a_i \geq 0$, $b_i \geq 0$ and $f_i \in B$ for all $i = 1, 2, \ldots, s$. The tête-a-tête is trivial *if the two factorizations share a common factor greater than 1.*

Given a tête-a-tête, we may choose non-zero constants $c_1, c_2 \in \mathbb{K}$ such that $c_1 \prod_{i=1}^{s} \mathrm{LT}(f_i)^{a_i} = c_2 \prod_{i=1}^{s} \mathrm{LT}(f_i)^{b_i}$. We will call the corresponding difference $c_1 \prod_{i=1}^{s} f_i^{a_i} - c_2 \prod_{i=1}^{s} f_i^{b_i}$ a *tête-a-tête difference.*

Example 5.1.5. We continue with Example 5.1.3. To the three generators f_1, f_2, f_3, we must add f_4 to our SAGBI basis since its lead term, y^8 is not accounted for by the lead terms of f_1, f_2, f_3.

Now we consider tête-a-têtes among the four polynomials f_1, f_2, f_3, f_4. There is only 1 non-trivial new tête-a-tête given by $3f_2^3 - f_3f_4$. Evaluating this difference we get

$$f_5 := 3f_2^3 - f_3f_4 = 3x^6y^3z^3 + 12x^5y^4z^3 + 12x^4y^5z^3 + 12x^3y^6z^3 + 9x^2y^7z^3 - 6x^3y^3z^6 + 9xy^5z^6 + 3y^3z^9.$$

Since $\mathrm{LM}(f_5) = x^6y^3z^3$ cannot be written as a product of the lead monomials of the f_1, f_2, f_3, f_4, we must further extend our SAGBI basis to $\{f_1, f_2, f_3, f_4, f_5\}$. At this stage, we do not find any new non-trivial tête-a-têtes among the lead terms of these five elements. As we shall see, this guarantees that these five elements form a SAGBI basis for R, the subalgebra they generate. That is to say that the subalgebra $\mathrm{LT}(\mathbb{K}[f_1, f_2, f_3])$ is generated by the 5 monomials $\mathrm{LM}(f_i)$ with $i = 1, 2, \ldots, 5$.

With the above examples in mind, we will now describe algorithms associated with finding a SAGBI basis. The first, the subduction algorithm, is the analogue of the reduction algorithm used for Gröbner bases. Given a finite set $\mathcal{B} = \{f_1, f_2, \ldots, f_s\}$ and a non-zero element f, it produces a new element $F = f - P$ where F is such that its lead term cannot be factored over the lead monomials of elements of $\mathcal{B}$ and P is in the subalgebra generated by $\mathcal{B}$.

Algorithm 5.1.6 (Subduction Algorithm)

```
1. Initialize: $F := f$;
2. While $F \neq 0$ do
     If there exists a factorization $\mathrm{LT}(F) = c\prod_{i=1}^{s}\mathrm{LT}(f_i)^{a_i}$
       Then $P := P + c\prod_{i=1}^{s} f_i^{a_i}$; $F := F - c\prod_{i=1}^{s} f_i^{a_i}$;
       Else Return $F, P$;
3. End While
```

The "If" statement in the above algorithm can be implemented by a recursive search. In general, this statement can be rather time consuming to perform.

The result F, produced by the subduction algorithm is called the *subduction* of f (over $\mathcal{B}$). Note that in general, it is not unique since it depends upon the choice of the factorizations chosen each time through the loop.

The next algorithm is not guaranteed to halt on all possible inputs. As we shall see there are (finitely generated) subalgebras of $S = \mathbb{K}[x_1, x_2, \ldots, x_n]$ whose lead term algebras are not finitely generated. Such a subalgebra cannot have a finite SAGBI basis. If however, the subalgebra, R, generated by $B = \{f_1, f_2, \ldots, f_s\}$ does have a finitely generated lead term algebra, then the following algorithm will produce a SAGBI basis $\mathcal{B}$ for R.

Algorithm 5.1.7

```
1. Initialize: $\mathcal{B} := B$;
```

```
2. Repeat
     $\mathcal{B}' := \emptyset$;
     For each non-trivial tête-a-tête difference $d$ over $\mathcal{B}$ do
       $F :=$ the subduction of $d$ over $\mathcal{B}$;
       If $F \neq 0$ then $\mathcal{B}' := \mathcal{B}' \cup \{F\}$;
     End For
     $\mathcal{B} := \mathcal{B} \cup \mathcal{B}'$;
     For each $F$ in $B'$ do
       If the subduction of $F$ over $(B \setminus \{F\})$ is $0$
          then $B := B \setminus \{F\}$;
     End For
3. Until $\mathcal{B}' = \emptyset$
4. Return $\mathcal{B}$
```

Example 5.1.8. Let $\mathbb{F}$ be a field of characteristic $p > 0$ and let V_2 denote the two dimensional representation of the cyclic group C_p of order p. If σ is a generator of C_p, then there is a basis, $\{x, y\}$, of V^*, such that $\sigma(x) = x$ and $\sigma(y) = x + y$. We consider $2\,V_2 = V_2 \oplus V_2$ and its coordinate ring $\mathbb{F}[2\,V_2] = \mathbb{F}[x_1, y_1, x_2, y_2]$. Let R denote the ring $\mathbb{F}[x_1, x_2, N_1, N_2, u_{12}]$ where $N_i = \mathbf{N}(y_i) = y_i^p - x_i^{p-1} y_i$ and $u_{12} = x_2 y_1 - x_1 y_2$. We showed in §1.12 that $\mathbb{F}[2\,V_2]^{C_p} = R$ but we will not need this fact for this example.

First we consider the ring R equipped with the graded reverse lexicographic ordering with $y_1 < x_1 < y_2 < x_2$. Then $\mathrm{LT}(N_1) = -y_i x_i^{p-1}$ for $i = 1, 2$ and $\mathrm{LT}(u_{12}) = -y_2 x_1$.

Define $f_0 := u_{12}$ and inductively define $f_i := x_2^{p^i - p^{i-1}} f_{i-1} - x_1 N_2^{p^{i-1}} \in R$ for $i \geq 1$. It is easy to see using induction that $f_i = -y_2^{p^i} x_1 + x_2^{p^i} y_1$:

$$\begin{aligned} f_i &= x_2^{p^i - p^{i-1}} f_{i-1} - x_1 N_2^{p^{i-1}} \\ &= x_2^{p^i - p^{i-1}} (-y_2^{p^{i-1}} x_1 + x_2^{p^{i-1}} y_1) - x_1 (y_2^p - x_2^{p-1} y_2)^{p^{i-1}} \\ &= -y_2^{p^i} x_1 + x_2^{p^i} y_1. \end{aligned}$$

Thus $y_2^{p^i} x_1 = \mathrm{LM}(f_i) \in \mathrm{LT}(R)$ for all $i \geq 0$.

Conversely, we claim that $y_2^j \notin \mathrm{LT}(R)$ for all $j \geq 1$. To see this, assume, by way of contradiction, that there is an element $h \in R$ with $\mathrm{LT}(h) = y_2^j$ where $j \geq 1$. Since $h \in R$, we know that h may be written as a linear combination of products of x_1, x_2, N_1, N_2 and u_{12}. Because the monomial y_2^j occurs in h, this expression for h must include N_2^m (and we must have $j = pm$). But that would imply that $\mathrm{LM}(h) \geq \mathrm{LM}(N_2^m) = x_2^{(p-1)m} y_2^m$ since there is no other way to factor $x_2^{(p-1)m} y_2^m$ over the lead terms of $x_1, x_2, N_1, N_2, u_{12}$. This contradiction shows that $y_2^j \notin \mathrm{LT}(R)$ for all $j \geq 1$. Therefore, any SAGBI basis for R must be infinite since it must contain $y_2^{p^i} x_1 = \mathrm{LM}(f_i) \in \mathrm{LT}(R)$ for all $i \geq 0$.

We remark that the elements f_i arise as tête-a-têtes when the SAGBI algorithm is run to try to produce a SAGBI basis for this example.

Example 5.1.9. Here we consider the same ring $R = \mathbb{F}[x_1, x_2, N_1, N_2, u_{12}]$ as in the previous example. Again, we will use a graded reverse lexicographic order; however, this time we will put $x_1 < y_1 < x_2 < y_2$. Thus we have $\mathrm{LT}(N_i) = y_i^p$ for $i = 1, 2$ and $\mathrm{LT}(u_{12}) = x_2 y_1$. Here (as in the previous example) there is only one non-trivial tête-a-tête among the original five generators. This is given by $(x_2 y_1)^p = (x_2)^p (y_1)^p$. The corresponding tête-a-tête difference is

$$u_{12}^p - x_2^p N_1 = x_2^p y_1 x_1^{p-1} - y_2^p x_1^p$$

which has the same lead term, $x_2^p y_1 x_1^{p-1}$, as does $x_2^{p-1} x_1^{p-1} u_{12}$. Subtracting this off we get

$$u_{12}^p - x_2^p N_1 - x_2^{p-1} x_1^{p-1} u_{12} = -y_2^p x_1^p + y_2 x_2^{p-1} x_1^p.$$

Finally,

$$u_{12}^p - x_2^p N_1 - x_2^{p-1} x_1^{p-1} u_{12} + x_1^p N_2 = 0$$

and thus the lone tête-a-tête subducts to zero. This implies that the minimal generating set $x_1, x_2, N_1, N_2, u_{12}$ is itself a finite SAGBI basis.

Notice also that we have found the generating relation that the five generators of this four dimensional ring must satisfy, i.e., $R \cong \mathbb{F}[a, b, c, d, e]/I$ where I is the principal ideal generated by

$$e^p - b^p c - a^{p-1} b^{p-1} e + a^p d$$

and $\deg(a) = \deg(b) = 1$, $\deg(c) = \deg(d) = p$ and $\deg(e) = 2$.

5.1.1 Symmetric Polynomials

Here we consider the classical example of symmetric polynomials. Unlike the usual analysis, we will work over an arbitrary field, $\mathbb{K}$.

Recall from §3.2 the *elementary symmetric polynomials* $s_1, s_2, \ldots, s_n$ in n variables. The s_t are implicitly defined by the equation

$$\prod_{j=1}^{n} (\lambda - z_i) = \sum_{t=0}^{n} (-1)^t s_t(z_1, z_2, \ldots, z_n) \lambda^{n-t}$$

in $\mathbb{K}[z_1, z_2, \ldots, z_n][\lambda]$.

Thus explicitly we have

$$\begin{aligned}
s_1(z_1, z_2, \ldots, z_n) &= z_1 + z_2 + \cdots + z_n, \\
s_2(z_1, z_2, \ldots, z_n) &= z_1 z_2 + z_1 z_3 + \cdots + z_{n-1} z_n, \\
&\vdots \\
s_t(z_1, z_2, \ldots, z_n) &= \sum_{i_1 < i_2 < \cdots < i_t} z_{i_1} z_{i_2} \cdots z_{i_t}, \\
&\vdots \\
s_n(z_1, z_2, \ldots, z_n) &= z_1 z_2 \cdots z_n.
\end{aligned}$$

Also $s_0(z_1, z_2, \ldots, z_n) = 1$.

Now we consider the symmetric group on n letters: $G = \Sigma_n$ and its usual n dimensional representation, V. It is well known (at least for characteristic zero) that the elementary symmetric functions in the n elements of a permutation basis for V^* generate the polynomial ring $\mathbb{K}[V]^G$. Indeed, we gave a characteristic free proof of this result in §3.2. Here we will give another proof of this result and show furthermore that the elementary symmetric functions are in fact a SAGBI basis for $\mathbb{K}[V]^G$ with respect to any monomial order.

We begin by choosing a permutation basis for V and we consider the dual basis, B, for V^*. Fix a monomial order, $<$. We label the n elements of the B by $x_1, x_2, \ldots, x_n$ such that $x_i > x_{i+1}$ in the fixed monomial order for $i = 1, 2, \ldots, n-1$. The symmetric group acts on $\mathbb{K}[V]$ by permuting the elements of B via $\sigma(x_i) = x_{\sigma(i)}$ for $\sigma \in \Sigma_n$.

Lemma 5.1.10. $\mathrm{LT}(s_t) = x_1 x_2 \cdots x_t$.

Proof. We proceed by induction on t. It is clear that $\mathrm{LT}(s_1) = x_1$ since $x_1 > x_i$ for all $i \geq 2$. Suppose now that $t \geq 2$ and assume by induction that $\mathrm{LT}(s_{t-1}) = x_1 x_2 \cdots x_{t-1}$. Given any t distinct subscripts, $i_1, i_2, \ldots, i_t$, we need to show that $x_{i_1} x_{i_2} \cdots x_{i_t} \leq x_1 x_2 \cdots x_t$. Without loss of generality, we may assume that $i_t > i_j$ for all $j = 1, 2, \ldots, t-1$. In particular, $i_t \geq t$ and thus $x_{i_t} \leq x_t$. Since $\mathrm{LT}(s_{t-1}) = x_1 x_2 \cdots x_{t-1}$, we know $x_{i_1} x_{i_2} \cdots x_{i_{t-1}} \leq x_1 x_2 \cdots x_{t-1}$. Therefore, $x_{i_1} x_{i_2} \cdots x_{i_{t-1}} x_{i_t} \leq x_1 x_2 \cdots x_{t-1} x_{i_t}$. But from $x_{i_t} \leq x_t$, we see that $x_1 x_2 \cdots x_{t-1} x_{i_t} \leq x_1 x_2 \cdots x_{t-1} x_t$. Therefore, $x_{i_1} x_{i_2} \cdots x_{i_t} \leq x_1 x_2 \cdots x_t$. □

For ease of notation, we define $m_t := \mathrm{LT}(s_t)$ for $t = 1, 2, \ldots, n$.

We now proceed to characterize the lead monomials in $\mathbb{K}[V]^G$. Given any monomial $m' = x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$, there is some permutation $\sigma \in \Sigma_n$ such that $a_{\sigma(1)} \geq a_{\sigma(2)} \geq \cdots \geq a_{\sigma(n)}$. Then $\sigma(m') = x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n}$ where $b_i = \sigma(i)$ satisfies $b_1 \geq b_2 \geq \cdots \geq b_n$. We call a monomial whose exponents are ordered in this way a *descending monomial.* Clearly, the Σ_n-orbit of any monomial contains a unique descending monomial. We claim that this descending monomial is always the largest monomial in the orbit.

To see this, consider a descending monomial $m := x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n}$. Notice that $m = \prod_{t=1}^n m_t^{b_t - b_{t+1}}$ where for ease of notation we have $b_{n+1} = 0$. Thus

$$\begin{aligned} m &= \prod_{t=1}^{n} \mathrm{LM}(s_t)^{b_t - b_{t+1}} \\ &= \mathrm{LM}(\prod_{t=1}^{n} s_t^{b_t - b_{t+1}}) \\ &= \mathrm{LM}(f) \end{aligned}$$

where $f := \prod_{t=1}^n s_t^{b_t - b_{t+1}} \in \mathbb{K}[V]^G$. If $\tau(m)$ is any element of the orbit of m, then $\tau(m)$ is a monomial occurring with non-zero coefficient in $f = \tau(f)$ and thus $\tau(m) \leq m$. This shows that the descending monomial m is the greatest monomial in its orbit.

Since V is a permutation representation, the set consisting of the orbit sums of all monomials forms a vector space basis for $\mathbb{K}[V]^G$. By the above, the lead term in the orbit sum of a monomial, m, is the unique descending monomial in the orbit of m. Therefore, the set of all descending monomials is precisely the set of lead monomials for $\mathbb{K}[V]^G$. Furthermore, the above factorization of the descending monomial m into a product of powers of $m_1, m_2, \ldots, m_n$ shows that $\{s_1, s_2, \ldots, s_n\}$ is a SAGBI basis for $\mathbb{K}[V]^G$. Also, note that subduction provides an algorithm for writing any invariant as a polynomial in the elementary symmetric functions. We record the above results as the following theorem.

Theorem 5.1.11. *Let $\mathbb{K}$ be any field. Let V be the usual n dimensional permutation representation of the symmetric group on n-letters, Σ_n. Then $\mathbb{K}[V]^{\Sigma_n}$ is the polynomial ring on the n elementary symmetric functions: $\mathbb{K}[V]^{\Sigma_n} = \mathbb{K}[s_1, s_2, \ldots, s_n]$. Furthermore, for any monomial ordering on the permutation basis, the n elementary symmetric functions form a SAGBI basis for $\mathbb{K}[V]^{\Sigma_n}$.* □

Remark 5.1.12. This theorem is in fact valid over any commutative ring of coefficients, $\mathbb{K}$, see Bourbaki [11, Ch. 4, Thm. 1, p. 58].

5.2 Finite SAGBI Bases

Unlike the situation for Gröbner bases we cannot always guarantee that a finite SAGBI basis will exist, even for finitely generated subalgebras. We next give an example which illustrates this phenomenon by exhibiting a ring of invariants that fails to have a finite SAGBI basis.

Example 5.2.1. Consider the usual three dimensional representation of the alternating group A_3 over a field $\mathbb{K}$ of any characteristic. We will work with a permutation basis and lexicographic order. Write $A_3 = \langle\sigma\rangle$ and let $\{e_1, e_2, e_3\}$ be a basis of V such that $\sigma^{-1}(e_1) = e_3$, $\sigma^{-1}(e_2) = e_1$ and $\sigma^{-1}(e_3) = e_2$. Let $\{x, y, z\}$ be the dual basis of V^*. Then $\sigma(x) = y$, $\sigma(y) = z$ and $\sigma(z) = x$.

Since this is a permutation representation, the orbit sums form a vector space basis for $\mathbb{K}[V]^{A_3}$. These orbit sums are all of one of the following two forms $x^a y^a z^a$ or $x^a y^b z^c + x^b y^c z^a + x^c y^a z^b$ where $a \geq b$ and $a > c$. In the latter case $x^a y^b z^c$ is the lead term of the orbit sum. In particular $x^{r+1} z^r$ is the leading monomial of an invariant for all $r = 0, 1, 2, \ldots$. We will show that this monomial cannot be written as a product of other lead monomials. Suppose by way of contradiction that $x^{r+1} z^r = (x^{a_1} z^{c_1})(x^{a_2} z^{c_2})$ where $x^{a_1} z^{c_1} = \mathrm{LT}(f_1)$ and $x^{a_2} z^{c_2} = \mathrm{LT}(f_2)$ with $f_1, f_2 \in \mathbb{K}[V]^{A_3}$. Since $f_i = \sigma(f_i)$ we see that $x^{c_i} y^{a_i}$ is a monomial occurring in f_i and thus $x^{a_i} z^{c_i} > x^{c_i} y^{a_i}$. This implies that $a_i \geq c_i + 1$ for $i = 1, 2$. Therefore $r+1 = a_1 + a_2 \geq (c_1+1) + (c_2+1) = r+2$.

This contradiction shows that $x^{r+1}z^r$ cannot be written as a product of lower degree lead monomials. This means that every monomial generating set for the lead term algebra of $\mathbb{K}[V]^{A_3}$ must contain $\{x^{r+1}z^r \mid r = 0, 1, 2, \dots\}$ and thus $\mathbb{K}[V]^{A_3}$ has no finite SAGBI basis (with respect to the basis $\{x, y, z\}$ and lexicographic order).

We will see below that changing the monomial order cannot help. However, first we will show that in a certain setting, there always exists a finite SAGBI basis. In particular, this will mean that if the characteristic of $\mathbb{K}$ is 3, then by changing the basis of V^* in the above, we can get a finite SAGBI basis for $\mathbb{K}[V]^{A_3}$.

The results in the following sequence were proved by Shank and Wehlau, see [99].

Proposition 5.2.2. *Let $A \subseteq \mathbb{F}[x_1, x_2, \dots, x_n]$ be a subalgebra and suppose there exist $h_1, h_2, \dots, h_n \in A$ and positive integers $d_1, d_2, \dots, d_n$ such that $\mathrm{LM}(h_i) = x_i^{d_i}$. Then A has a finite SAGBI basis.*

Proof. Let $H := \mathbb{F}[x_1^{d_1}, x_2^{d_2}, \dots, x_n^{d_n}]$. Since $x_i^{d_i} = \mathrm{LM}(h_i)$ for $i = 1, 2, \dots, n$, we see that $H \subseteq \mathrm{LT}(A)$. Furthermore, since $x_1^{d_1}, x_2^{d_2}, \dots, x_n^{d_n}$ is a homogeneous system of parameters for $\mathbb{F}[x_1, x_2, \dots, x_n]$, we have that $\mathbb{F}[x_1, x_2, \dots, x_n]$ is a finitely generated H-module containing A as a submodule. Since H is a Noetherian algebra, this implies that $\mathrm{LT}(A)$ is a finitely generated H-module. Since $\mathrm{LT}(A)$ is generated by monomials, this implies that there is a finite subset of monomials $m_1, m_2, \dots, m_r$ which generate $\mathrm{LT}(A)$ as an H-module. For each $i = 1, 2, \dots, r$, write $m_i = \mathrm{LT}(f_i)$ with $f_i \in A$. Then the set $\{h_1, h_2, \dots, h_n\} \cup \{f_1, f_2, \dots, f_r\}$ is a SAGBI basis for A. □

Fix any monomial order on $\mathbb{F}[V] = \mathbb{F}[x_1, x_2, \dots, x_n]$ such that $x_1 < x_2 < \cdots < x_n$. Recall that the representation of G on V is called triangular (with respect to the chosen basis of V^*) if $\mathrm{LM}(\sigma(x_i)) = x_i$ for all $i = 1, 2, \dots, n$ and for all $\sigma \in G$.

Theorem 5.2.3. *If the representation of G on V is triangular, then $\mathbb{F}[V]^G$ has a finite SAGBI basis.*

Proof. Let $\{x_1, x_2, \dots, x_n\}$ be the basis of V^* with respect to which the representation of G is triangular. Since $\mathrm{LM}(\sigma(x_i)) = x_i$ for all $\sigma \in G$, we see that the norms of the x_i satisfy $\mathrm{LM}(\mathbf{N}(x_i)) = \mathrm{LM}(\prod_{\sigma \in G} \sigma(x_i)) = \prod_{\sigma \in G} \mathrm{LM}(\sigma(x_i)) = x_i^{|G|}$. Therefore, by the preceding proposition, $\mathbb{F}[V]^G$ has a finite SAGBI basis. □

Combining Proposition 4.0.2 with Theorem 5.2.3 we get the following result.

Theorem 5.2.4. *Suppose that G is a p-group and V is any representation of G defined over a field $\mathbb{F}$ of characteristic p. Then there is a choice of basis and monomial order such that $\mathbb{F}[V]^G$ has a finite SAGBI basis.*

5.3 SAGBI Bases for Permutation Representations

M. Göbel [45] showed that if G acts on $\mathbb{K}[V]$ by permuting the variables then, using a lexicographic order, the ring of invariants, $\mathbb{K}[V]^G$, has a finite SAGBI basis if and only if G is a product of symmetric groups, or what is the same thing, that the permutation group G is generated by reflections. Later [46] he conjectured that this result should extend to any term order. He was able to prove the extension for the case of the alternating group in the paper [47]. The conjecture was proved in (various forms in) [91], [72] and [107]. We give a proof here that uses the geometry of the group representation.

Theorem 5.1.11 showed that the usual permutation representation of Σ_n has a finite SAGBI basis. The goal of this section is to prove that this is essentially the only permutation representation with a finite SAGBI basis.

Definition 5.3.1. *We take $\sqcup_{i=1}^{k} A_i = \{1, 2, \ldots, n\}$ to be any partition of the set $\{1, 2, \ldots, n\}$ into disjoint subsets. The corresponding* Young subgroup *of Σ_n is the group $\Sigma(A_1) \times \Sigma(A_2) \times \ldots \Sigma(A_k)$ where $\Sigma(A_j) = \{\tau \in \Sigma_n \mid \tau(i) = i$ for all $i \notin A_j\}$.*

We begin with a result from group theory.

Lemma 5.3.2. *Let G be a subgroup of Σ_n. Suppose that G is generated by transpositions. Then G is a Young subgroup of Σ_n. Furthermore, if G acts transitively on $\{1, 2, \ldots, n\}$, then $G = \Sigma_n$.*

Proof. We prove the second statement first. Suppose that G is a transitive subgroup of Σ_n generated by transpositions. Let R denote the set of transpositions in G. Without loss of generality, $(1, 2) \in R$. If $n = 2$, then we are done. Thus we suppose that $n \geq 3$. Since G acts transitively, there must be a transposition $(i_1, j_1) \in R$ with $i_1 \leq 2$ and $j_1 \geq 3$. Without loss of generality, $i_1 = 2$ and $j_1 = 3$ and thus we may assume that $(2, 3) \in R$. Thus R contains all permutations on $1, 2, 3$. If $n = 3$, we are done, and thus we may suppose that $n \geq 4$. Again, since G acts transitively, there exists $(i_2, j_2) \in R$ with $i_2 \leq 3$ and $j_2 \geq 4$. Without loss of generality, $i_2 = 3$ and $j_2 = 4$ and thus we may assume that $(3, 4) \in R$. Continuing in this manner we may assume that $(1, 2)$, $(2, 3), \ldots, (n-1, n) \in R$. Thus $G = \Sigma_n$ which proves the second assertion of the theorem.

The first assertion follows easily from the second. □

Suppose that V is a permutation representation of G with permutation basis $\{v_1, v_2, \ldots, v_n\}$. Let $\{x_1, x_2, \ldots, x_n\}$ be the dual basis of V^*. The permutation action of G on V induces a permutation action of G on the monomials in $\mathbb{F}[V]$. This in turn induces a permutation action of G on the set of exponents as follows. Let $(a_1, a_2, \ldots, a_n) \in \mathbb{N}^n$ be a sequence of exponents. Then $\sigma((a_1, a_2, \ldots, a_n)) = (b_1, b_2, \ldots, b_n)$ where $\sigma(x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}) = x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n}$ (and $b_i = a_{\sigma(i)}$).

In this section we will prove the following theorem which was proved independently by Kuroda [72], Reichstein [91], and Thiéry and Thomassè [107].

Theorem 5.3.3. *Suppose that V is a permutation representation of G. Choose a permutation basis of V and let $\{x_1, x_2, \ldots, x_n\}$ be the dual basis of V^*. Fix a monomial ordering on $\mathbb{K}[x_1, x_2, \ldots, x_n]$. Then $\mathbb{K}[V]^G$ has a finite SAGBI basis with respect to this fixed order if and only if the action of G on V is generated by reflections.*

Since the only permutations that are reflections are transpositions, applying Lemma 5.3.2 we obtain the following immediate corollary.

Theorem 5.3.4. *Suppose that V is a permutation representation of G. Choose a permutation basis of V and let $\{x_1, x_2, \ldots, x_n\}$ be the dual basis of V^*. Fix a monomial ordering on $\mathbb{K}[x_1, x_2, \ldots, x_n]$. Then $\mathbb{K}[V]^G$ has a finite SAGBI basis with respect to this fixed order if and only if G acts on V as a Young subgroup of Σ_n.*

Now we begin the proof of Theorem 5.3.3. First we suppose that $\mathbb{K}[V]^G$ has a finite SAGBI basis with respect to the permutation basis and the fixed monomial order. We will show that this implies that the action of G on V is generated by reflections. Throughout this proof we will work with the usual Euclidean topology on $\mathbb{R}^n$. We write $\mathbb{R}_{\geq 0} := \{x \in \mathbb{R} \mid x \geq 0\}$ and $\mathbb{Q}_{\geq 0} := \{x \in \mathbb{Q} \mid x \geq 0\}$.

Definition 5.3.5. *A cone in $\mathbb{R}^n$ is a subset C of $\mathbb{R}^n$ such that $\lambda x \in C$ for all $x \in C$ and for all $\lambda \in \mathbb{R}$ with $\lambda \geq 0$.*

Let $\operatorname{cone}(S) := \{\lambda_1 s_1 + \lambda_2 s_2 + \cdots + \lambda_t s_t \mid t \in \mathbb{N}, s_i \in S, \lambda_i \in \mathbb{R}, \lambda_i \geq 0\}$ denote the convex cone spanned by S.

Let C denote the convex cone $C := \operatorname{cone}(\{(a_1, a_2, \ldots, a_n) \mid \prod_{i=1}^n x_i^{a_i} \in \operatorname{LT}(\mathbb{K}[V]^G)\})$ spanned by the exponents of lead terms of elements of $\mathbb{K}[V]^G$.

Let $\mathcal{B} = \{f_1, f_2, \ldots, f_r\}$ be a finite SAGBI basis for $\mathbb{K}[V]^G$ and let $S_0 := \{(a_1, a_2, \ldots, a_n) \mid \prod_{i=1}^n x_i^{a_i} = \operatorname{LT}(f) \text{ for } f \in \mathcal{B}\}$. Then $C = \operatorname{cone}(S_0)$ which is a closed subset of $\mathbb{R}^n$ since S_0 is finite.

Lemma 5.3.6. *The union of all the G-translates of C is the entire positive orthant:*

$$\mathbb{R}^n_{\geq 0} = \bigcup_{\sigma \in G} \sigma(C)$$

Proof. Clearly, $\cup_{\sigma \in G} \sigma(C) \subseteq \mathbb{R}^n_{\geq 0}$. Since C is closed, $\cup_{\sigma \in G} \sigma(C)$ is also closed. Thus it suffices to prove that $\mathbb{Q}^n_{\geq 0} \subset \cup_{\sigma \in G} \sigma(C)$. Let $a = (a_1, a_2, \ldots, a_n)$ be any element of $\mathbb{Q}^n_{\geq 0}$. Take $t \in \mathbb{N}$ such that $b := t(a_1, a_2, \ldots, a_n) = (b_1, b_2, \ldots, b_n) \in \mathbb{N}^n$. Then the orbit sum $\mathcal{O}_G(\prod_{i=1}^n x_i^{b_i}) \in \mathbb{K}[V]^G$. Let $\prod_{i=1}^n x_i^{c_i}$ denote the corresponding lead term, $\operatorname{LT}(\mathcal{O}_G(\prod_{i=1}^n x_i^{a_i}))$. Then $(c_1, c_2, \ldots, c_n) = \sigma((b_1, b_2, \ldots, b_n))$ for some $\sigma \in G$.

We also know that $(c_1, c_2, \ldots, c_n)$ and $t^{-1}(c_1, c_2, \ldots, c_n)$ both lie in C. Thus

$$(b_1, b_2, \ldots, b_n) = \sigma^{-1}((c_1, c_2, \ldots, c_n))$$

and

$$(a_1, a_2, \ldots, a_n) = \sigma^{-1}((c_1/t, c_2/t, \ldots, c_n/t)) \in \sigma^{-1}(C).$$

Since $(a_1, a_2, \ldots, a_n)$ was an arbitrary element of $\mathbb{Q}^n_{\geq 0}$, this shows that $\mathbb{Q}^n_{\geq 0} \subset \cup_{\sigma \in G} \sigma(C)$. □

Lemma 5.3.7. *Let $a \in \mathbb{Q}^n_{\geq 0} \cap C$. Suppose $\sigma \in G$ is such that $\sigma(a) \in C$. Then $\sigma(a) = a$.*

Proof. Since $\sigma(a) \in C$, we may write $\sigma(a) = \sum_{j=1}^{r} \lambda_j s_j$ where $S_0 = \{s_1, s_2, \ldots, s_r\}$ and $\lambda_j \in \mathbb{R}$ with $\lambda_j \geq 0$ for all $j = 1, 2, \ldots, r$. Put $\lambda := (\lambda_1, \lambda_2, \ldots, \lambda_r) \in \mathbb{R}^r_{\geq 0}$. Consider the set of all solutions $A := \{z = (z_1, z_2, \ldots, z_r) \in \mathbb{R}^r \mid z_1 s_1 + z_2 s_2 + \cdots + z_r s_r = \sigma(a)\}$. Then A is a non-empty affine subspace of $\mathbb{R}^n$. If A is zero dimensional, then A consists of the unique solution λ to a system of rational equations. Thus this unique solution must have $\lambda_j \in \mathbb{Q}$ for all $j = 1, 2, \ldots, r$. On the other hand, if the dimension of A is positive, then there exist points of $A \cap \mathbb{Q}^r$ arbitrarily close to λ. Since $\lambda \in \mathbb{R}^r_{\geq 0}$, we may find a point $\lambda' = (\lambda'_1, \lambda'_2, \ldots, \lambda'_r) \in A \cap \mathbb{Q}^r_{\geq 0}$. Thus we have shown that we may always write $\sigma(a) = \sum_{j=1}^{r} \gamma_j s_j$ where $S_0 = \{s_1, s_2, \ldots, s_r\}$ and $\gamma_j \in \mathbb{Q}$ with $\gamma_j \geq 0$ for all $j = 1, 2, \ldots, r$.

Clearing denominators we may further write $b := t\sigma(a) = \sum_{j=1}^{r} (\gamma'_j) s_j$ where $b := t\sigma(a)$ and $\gamma'_j := t\gamma_j \in \mathbb{N}$ for all $j = 1, 2, \ldots, r$. Therefore $b = \prod_{j=1}^{r} \mathrm{LT}(s_j^{\gamma'_j})$ is a lead monomial of an invariant. Therefore $b \geq \tau(b)$ for all $\sigma \in G$, i.e., $t\sigma(a) \geq \tau(t\sigma(a))$ for all $\sigma \in G$. This implies that $\sigma(a) \geq \tau(\sigma(a))$ for all $\tau \in G$. Thus $\sigma(a) \geq a$.

Similarly $a \geq \sigma(a)$ and this proves that $a = \sigma(a)$. □

Corollary 5.3.8. *Let a lie in the topological interior of C. If $\sigma(a) = a$ for some $\sigma \in G$, then $\sigma = e$.*

Proof. Choose a small (half-) ball U' in C with $\sigma(a) \in U'$. Choose a small ball $U \in C$ with $a \in U$. Since σ^{-1} is continuous, we may choose U small enough to ensure that $U \subseteq \sigma^{-1}(U')$. Then $\sigma(U) \subseteq U'$. There are infinitely many rational points in U and U'. By the preceding lemma, the group element σ must fix each of these rational points, including a basis of $\mathbb{Q}^n$ (and of $\mathbb{R}^n$). Therefore $\sigma = e$. □

Corollary 5.3.9.

$$\bigcup_{\tau \in G \setminus \{\sigma\}} \tau(C) \neq \mathbb{R}^n_{\geq 0}$$

for all $\sigma \in G$.

Proof. Take $a \in \mathbb{Q}^n$ in the topological interior of C. By Corollary 5.3.8, the isotropy group of a is trivial and thus the G-orbit of a consists of $|G|$ many distinct elements. By Lemma 5.3.7, no two of these elements can lie in the same translate of C. Thus all $|G|$ translates of C are required to cover $\mathbb{R}^n_{\geq 0}$. □

Proposition 5.3.10. *Let a lie in the topological boundary of C. Then there exists $\sigma \in G$ with $\sigma \neq e$ such that $\sigma(a) \in C$.*

Proof. Choose an infinite sequence of rational points $a_1, a_2, a_3, \ldots$ outside of C which converge to a. For every i, there exists some $\sigma_i \in G$ such that $\sigma_i(a_i) \in C$ and $\sigma_i \neq 1$. Choose a subsequence $a_{i_1}, a_{i_2}, a_{i_3}, \ldots$ such that $\sigma_{i_j} = \sigma$ for all j. Then $\sigma_{i_1}(a_{i_1}), \sigma_{i_2}(a_{i_2}), \sigma_{i_3}(a_{i_3}), \ldots$ is a sequence in C which converges to $\sigma(a)$. Since C is closed, $\sigma(a) \in C$. □

Corollary 5.3.11. *Let F be a codimension 1 face of the cone C. Let H be the hyperplane spanned by F. Let σ_H denote the automorphism of $\mathbb{R}^n$ which is reflection in H. Then $\sigma_H \in G$.*

Proof. The hyperplane H contains infinitely many rational points. Combining the previous proposition with Lemma 5.3.7 we see that there is a non-trivial $\sigma \in G$ which fixes H pointwise. Since σ is a permutation, it preserves distance and thus we must have $\sigma = \sigma_H$. □

Corollary 5.3.12. *G is generated by reflections.*

Proof. Let G' denote the subgroup of G generated by all the reflections in G. Take a point $a \in \mathbb{Q}^n$ in the topological interior of C and consider its orbit $G \cdot a$. By Corollary 5.3.8, this orbit has size $|G|$. We will show that G' acts transitively on this orbit. Choose a point in the orbit, say $a' \in G \cdot a$. We will show that a' lies in the G'-orbit of a.

Let $\{H_1, H_2, \ldots, H_s\}$ denote the set of reflection hyperplanes corresponding to the reflections in G. This set of reflection hyperplanes, subdivides $\mathbb{R}^n$ into a finite number of connected convex subsets. We consider a path, $\Gamma : [0, 1] \to \mathbb{R}^n_{\geq 0}$ from a to a'. Thus $\Gamma(0) = a$ and $\Gamma(1) = a'$. We choose this path Γ such that for each reflecting hyperplane H_j, the path meets H_j at most once, i.e., $|\Gamma^{-1}(H_j)| \leq 1$. Furthermore, we choose Γ to avoid the intersections of every pair of reflecting hyperplanes H_j, H_k, i.e., we have $\Gamma^{-1}(H_j \cap H_k) = \emptyset$ for all $1 \leq j < k \leq s$. Write $\Gamma^{-1}(\cup_{j=1}^s H_j) = \{t_1, t_2, \ldots, t_r\}$. We choose Γ such that r is minimal. We will prove that a' lies in the G'-orbit of a by induction on the number r.

If $r = 0$, then the path Γ does not cross any of the reflecting hyperplanes. Thus a and a' both lie in C. But since $a' \in G \cdot a$, we must have $a = a'$ by Lemma 5.3.7. Thus $a' \in G' \cdot a$.

Now we handle the general case $r \geq 1$. The point $x := \Gamma(\frac{t_{r-1}+t_r}{2})$ lies in some translate C_1 of C. There is a unique point $a'' \in G \cdot a \cap C_1$. Clearly, we may define a path $\Gamma' : [0, 1] \to \mathbb{R}^n_{\geq 0}$ from a to a'' such that $\Gamma'(t) = \Gamma(t)$ for all

$t \in [0, t_{r-1}]$ and with $\Gamma'(t)$ lying in the interior of C_1 for all $t \in (t_{r-1}, 1]$. Then ${\Gamma'}^{-1}(\cup_{j=1}^{s} H_j) = \{t_1, t_2, \dots, t_{r-1}\}$. Therefore, by induction, we see that a'' lies in the G'-orbit of a. Let H denote the reflecting hyperplane containing $\Gamma(t_r)$. Then a'' and a' lie on opposite sides of H. Furthermore, by Corollary 5.3.11, we know that the reflection σ_H is an element of G'. Since $\sigma_H(a'')$ must lie in C_1, we see that $\sigma_H(a'') = a'$ by Lemma 5.3.7. Thus $a' \in G' \cdot a$ as required.

Since G' acts transitively on the $G \cdot a$, we must have $|G'| \geq |G \cdot a| = |G|$. This proves $G = G'$ is generated by reflections. □

This completes the proof of one direction of Theorem 5.3.3, namely that if V is a permutation representation of G such that $\mathbb{K}[V]^G$ has a finite SAGBI basis, then the action of G on V is generated by reflections.

Now we consider the converse. Let W be an n dimensional permutation representation of G generated by reflections. Lemma 5.3.2 implies that $W \cong W_1 \oplus W_2 \oplus \cdots \oplus W_r$ and $G \cong \Sigma_{d_1} \times \Sigma_{d_2} \times \cdots \times \Sigma_{d_r}$ is a Young subgroup of Σ_n where $d_i = \dim W_i$ for $i = 1, 2, \dots, r$. Therefore, $\mathbb{K}[W]^G = \mathbb{K}[W_1]^{\Sigma_{d_1}} \otimes \mathbb{K}[W_2]^{\Sigma_{d_2}} \otimes \cdots \otimes \mathbb{K}[W_r]^{\Sigma_{d_r}}$. By Theorem 5.1.11, $\mathbb{K}[W_i]^{\Sigma_{d_i}}$ is a polynomial ring with a SAGBI basis $\mathcal{B}_i$ consisting of d_i invariants. Thus $\mathcal{B} = \mathcal{B}_1 \sqcup \mathcal{B}_2 \sqcup \cdots \sqcup \mathcal{B}_r$ is a finite SAGBI basis for $\mathbb{K}[W]^G$. This completes the proof of Theorem 5.3.3.

6

Block Bases

The ideal I of $\mathbb{K}[V]$ generated by the homogeneous invariants of positive degree, $I = \mathbb{K}[V]^G_+ \cdot \mathbb{K}[V]$, is called the *Hilbert ideal.* We have already seen one use of the Hilbert ideal in the proof of Theorem 3.5.1. Not surprisingly, the quotient algebra formed with respect to this ideal is also very useful to consider.

Definition 6.0.1. *The* ring of coinvariants *is the quotient ring*

$$\mathbb{K}[V]_G := \mathbb{K}[V]/(\mathbb{K}[V]^G_+ \cdot \mathbb{K}[V]) \ .$$

Since $\mathbb{K}[V]^G_+ \cdot \mathbb{K}[V]$ is a graded ideal, $\mathbb{K}[V]_G$ is also a graded algebra, and is easily seen to be Artinian, i.e., $\dim_{\mathbb{K}}(\mathbb{K}[V]_G)$ is finite. The ring of coinvariants is a classical object of study. One of the first important questions to ask is which representation of G is given by each of its graded pieces? Another problem is to find a vector space basis for $\mathbb{K}[V]_G$ and then to describe the multiplication with respect to this basis.

We observe that

$$\mathbb{K}[V]_G \cong \mathbb{K} \otimes_{\mathbb{K}[V]^G} \mathbb{K}[V]$$

so that

$$\mathbb{K}[V]^G \otimes_{\mathbb{K}} \mathbb{K}[V]_G \cong \mathbb{K}[V]^G \otimes_{\mathbb{K}} \mathbb{K} \otimes_{\mathbb{K}[V]^G} \mathbb{K}[V] \cong \mathbb{K}[V]$$

and therefore

$$\mathcal{H}(\mathbb{K}[V]^G, \lambda) \cdot \mathcal{H}(\mathbb{K}[V]_G, \lambda) = \mathcal{H}(\mathbb{K}[V], \lambda) \ .$$

Example 6.0.2. Let $G = \Sigma_3$ be the symmetric group on 3 letters, acting in the usual way on a three dimensional vector space V by permuting the basis $\{e_1, e_2, e_3\}$. Let $\{x_1, x_2, x_3\}$ be the dual basis of V^*. Then by Theorem 5.1.11, $\mathbb{K}[V]^G = \mathbb{K}[s_1, s_2, s_3]$ where $s_1 = x_1 + x_2 + x_3$, $s_2 = x_1x_2 + x_1x_3 + x_2x_3$ and $s_3 = x_1x_2x_3$. The Hilbert series of $\mathbb{K}[V]^G$ is given by $\mathcal{H}(\mathbb{K}[V]^G, \lambda) = (1-\lambda)^{-1}(1-\lambda^2)^{-1}(1-\lambda^3)^{-1}$. Of course, $\mathcal{H}(\mathbb{K}[V], \lambda) = (1-\lambda)^{-3}$. Since

H.E.A. Eddy Campbell, D. Wehlau, *Modular Invariant Theory*,
Encyclopaedia of Mathematical Sciences 139,
DOI 10.1007/978-3-642-17404-9_6, © Springer-Verlag Berlin Heidelberg 2011

s_1, s_2, s_3 is a homogeneous system of parameters, $\mathbb{K}[V]$ is a free $\mathbb{K}[V]^G$-module with

$$\begin{aligned}\mathcal{H}(\mathbb{K}[V]_G, \lambda) = \frac{\mathcal{H}(\mathbb{K}[V], \lambda)}{\mathcal{H}(\mathbb{K}[V]^G, \lambda)} &= \frac{(1-\lambda)(1-\lambda^2)(1-\lambda^3)}{(1-\lambda)(1-\lambda)(1-\lambda)} \\ &= (1+\lambda)(1+\lambda+\lambda^2) = 1 + 2\lambda + 2\lambda^2 + \lambda^3\end{aligned}$$

The elements $1, x_1, x_2, x_1x_2, x_1^2, x_1^2x_2$ (or more properly their natural images) form a basis for $\mathbb{K}[V]_G$. Note that these six elements are all the monomial factors of $x_1^2x_2$.

Working in $\mathbb{K}[V]_G$, many of the products of these basis elements are obvious: $(x_1)^2 = x_1^2$, $(x_1)(x_2) = x_1x_2$, $(x_1)(x_1x_2) = x_1^2x_2$, and $(x_2)(x_1^2) = x_1^2x_2$.

Since $x_2^2 = (x_1 + x_2)s_1 - s_2 - x_1^2 - x_1x_2$, we have $(x_2)^2 = -x_1^2 - x_1x_2$ in $\mathbb{K}[V]_G$. Since $x_1^3 = s_3 - x_1s_2 + x_1^2s_1$, we have $(x_1)^3 = 0$ in $\mathbb{K}[V]_G$. Finally, since $x_1x_2^2 \equiv (x_1)(-x_1^2 - x_1x_2) = -x_1^3 - x_1^2x_2$, we have $x_1x_2^2 = -x_1^2x_2$ in $\mathbb{K}[V]_G$.

Definition 6.0.3. *Let $\alpha \in S = \mathbb{K}[x_1, x_2, \ldots, x_n]$ be a monomial and let $R \subset S$ be a graded subalgebra (of Krull dimension n). We say that α* generates a block basis *for S over R if the set of all monomial factors of α is a vector space basis of S/R_+S. Such a basis consisting of all the monomial factors of a single monomial is called a* block basis.

Let $f_1, f_2, \ldots, f_n$ be a homogeneous system of parameters for a Cohen-Macaulay ring S. Let B denote the algebra $B = \mathbb{K}[f_1, f_2, \ldots, f_n]$. Since S is Cohen-Macaulay, S is a free B-module:

$$S = \oplus_{j=1}^t Bh_j$$

for some $h_1, h_2, \ldots, h_t \in S$. Therefore,

$$\mathcal{H}(S, \lambda) = \sum_{j=1}^t \mathcal{H}(Bh_j, \lambda) = \sum_{j=1}^t \mathcal{H}(B, \lambda)\lambda^{\deg(h_j)}.$$

Thus $\mathcal{H}(S, \lambda)/\mathcal{H}(B, \lambda) = \sum_{j=1}^t \lambda^{\deg(h_j)}$. From this we see that the polynomial $\mathcal{H}(S, \lambda)/\mathcal{H}(B, \lambda)$ encodes the degrees of the module generators for S over B. In particular, if $\alpha = x_1^{a_1}x_2^{a_2}\cdots x_n^{a_n}$ generates a block basis for S over B, then

$$\begin{aligned}\mathcal{H}(S, \lambda)/\mathcal{H}(B, \lambda) &= \sum_{i_1=0}^{a_1}\sum_{i_2=0}^{a_2}\cdots\sum_{i_n=0}^{a_n} \lambda^{i_1+i_2+\cdots+i_n} \\ &= \sum_{i_1=0}^{a_1}\sum_{i_2=0}^{a_2}\cdots\sum_{i_n=0}^{a_n} \lambda^{i_1}\lambda^{i_2}\cdots\lambda^{i_n} \\ &= \left(\sum_{i_1=0}^{a_1}\lambda^{i_1}\right)\left(\sum_{i_2=0}^{a_2}\lambda^{i_2}\right)\cdots\left(\sum_{i_n=0}^{a_n}\lambda^{i_n}\right) \\ &= \prod_{j=1}^n \frac{1-\lambda^{a_j+1}}{1-\lambda}\end{aligned}$$

Also, notice that substituting $\lambda = 1$ into $\mathcal{H}(S,\lambda)/\mathcal{H}(B,\lambda) = \sum_{j=1}^{t} \lambda^{\deg(h_j)}$ yields t, the rank of S as a free B-module.

When $\mathbb{K}[V]^G = \mathbb{K}[f_1, f_2, \ldots, f_n]$ is a polynomial ring, its ring of coinvariant $\mathbb{K}[V]_G$ may be studied as follows. We consider the Hironaka decomposition of $\mathbb{K}[V]$ as an $\mathbb{K}[V]^G$-module: $\mathbb{K}[V] = \oplus_{j=1}^{t} \mathbb{K}[V]^G h_j$. In this case, the Hilbert ideal I is easily described: $I = \mathbb{K}[V]_+^G \mathbb{K}[V] = \oplus_{j=1}^{t} \mathbb{K}[V]_+^G h_j$ and thus

$$\mathbb{K}[V]_G = \mathbb{K}[V]/I = \oplus_{j=1}^{t} \mathbb{K} h_j \ .$$

Thus we see that

$$\begin{aligned}
\mathcal{H}(\mathbb{K}[V]_G, \lambda) &= \mathcal{H}(\mathbb{K}[V], \lambda)/\mathcal{H}(\mathbb{K}[V]^G, \lambda) \\
&= \prod_{i=1}^{n} \frac{1}{1-\lambda} \Big/ \prod_{i=1}^{n} \frac{1}{1-\lambda^{\deg(f_i)}} \\
&= \prod_{i=1}^{n} \frac{1-\lambda^{\deg(f_i)}}{1-\lambda} \\
&= \prod_{i=1}^{n} (1 + \lambda + \lambda^2 + \cdots + \lambda^{\deg(f_i)-1})
\end{aligned}$$

when $\mathbb{K}[V]^G$ is a polynomial ring. In particular, evaluating at $\lambda = 1$, we see that, when $\mathbb{K}[V]^G$ is polynomial, the rank of $\mathbb{K}[V]$ as a $\mathbb{K}[V]^G$-module is given by $\prod_{i=1}^{n} \deg(f_i)$. Of course, we have already seen this result in Proposition 3.1.4.

6.1 A Block Basis for the Symmetric Group

Let $S_n = \mathbb{K}[x_1, x_2, \ldots, x_n]$ and consider the natural action of the symmetric group Σ_n on S_n. The Hilbert series of S_n is given by $\mathcal{H}(S_n, \lambda) = \prod_{i=1}^{n} \frac{1}{1-\lambda}$. As we have seen in Section 5.1.1, $S_n^{\Sigma_n} = \mathbb{K}[s_1, s_2, \ldots, s_n]$ where $\deg(s_i) = i$ for $i = 1, 2, \ldots, n$. Thus $\mathcal{H}(S_n^{\Sigma_n}, \lambda) = \prod_{i=1}^{n} \frac{1}{1-\lambda^i}$. Therefore, $\mathcal{H}((S_n)_{\Sigma_n}, \lambda) = \prod_{i=1}^{n} \frac{1-\lambda^i}{1-\lambda}$.

This shows that S_n is a rank $n!$ free $S_n^{\Sigma_n}$-module as also follows from Proposition 3.1.3. It also suggests the possibility that the monomial $\alpha_n = x_2 x_3^2 \cdots x_n^{n-1}$ might generate a block basis for S_n over $S_n^{\Sigma_n}$. We will now show that this is indeed the case. This result is classical, see Artin's book [4, p. 41].

Proposition 6.1.1. *Let* $\alpha_n = x_2 x_3^2 \cdots x_n^{n-1}$ *and let* $\mathcal{B} := \{\beta \mid \beta \text{ divides } \alpha_n\}$. *Then*

$$\mathbb{K}[x_1, x_2, \ldots, x_n] = \bigoplus_{\beta \in \mathcal{B}} \mathbb{K}[x_1, x_2, \ldots, x_n]^{\Sigma_n} \beta \ .$$

Proof. Since we know that S_n has rank $n!$ as an $S_n^{\Sigma_n}$-module, and since there are $n!$ monomials which divide α_n, it suffices to prove that the $S_n^{\Sigma_n}$-module generated by the monomial divisors of α_n contains all of S_n. The proof is by induction on n. For $n = 1$ we have $\Sigma_1 = \{e\}$, $\alpha_1 = 1$ and $S_1 = S_1^{\Sigma_1}$ and hence the result is trivially true. For the inductive step, we assume that the divisors of α_{n-1} form an $S_{n-1}^{\Sigma_{n-1}}$-module basis for S_{n-1}. We need to show that every $f \in S_n$ lies in the $S_n^{\Sigma_n}$-module generated by the factors of α_n. Recall the algebra surjection first discussed in §3.2:

$$\begin{aligned} \theta : S_n &\to S_{n-1} \\ x_i &\mapsto \begin{cases} x_i & \text{if } i < n \\ 0 & \text{if } i = n \end{cases} \end{aligned}$$

Recall that $s_1, s_2, \ldots, s_n$ denote the elementary symmetric functions in $x_1, x_2, \ldots, x_n$. We denote by $s'_1, s'_2, \ldots, s'_{n-1}$ the elementary symmetric functions in $x_1, x_2, \ldots, x_{n-1}$. Recall that $s'_i = \theta(s_i)$ for $i = 1, 2, \ldots, n-1$ and $\theta(s_n) = 0$. In particular, $\theta(S_n^{\Sigma_n}) = S_{n-1}^{\Sigma_{n-1}}$.

Clearly, we may assume that f is homogeneous. We proceed by induction on the degree of f. If $\deg(f) = 0$ then f lies in the $S_n^{\Sigma_n}$-module generated by 1.

Suppose then that $\deg(f) = m > 0$. Write $f = a' + bx_n$ where $a' \in S_{n-1}$ and $b \in S_n$. Then $\theta(f) = a'$. By induction on n, there exist (unique)

$$f'_\beta = f'_\beta(s'_1, s'_2, \ldots, s'_{n-1}) \in S_{n-1}^{\Sigma_{n-1}} = \mathbb{K}[s'_1, s'_2, \ldots, s'_{n-1}]$$

such that

$$a' = \sum_{\beta \text{ divides } \alpha_{n-1}} f'_\beta \beta.$$

Define

$$f_\beta := f'_\beta(s_1, s_2, \ldots, s_{n-1}) \in S_n^{\Sigma_n}$$

and

$$a = \sum_{\beta \text{ divides } \alpha_{n-1}} f_\beta \beta \in S_n.$$

Then

$$\begin{aligned} \theta(a) &= \theta\Big(\sum_{\beta \text{ divides } \alpha_{n-1}} f_\beta \beta\Big) \\ &= \sum_{\beta \text{ divides } \alpha_{n-1}} \theta(f_\beta)\beta \\ &= \sum_{\beta \text{ divides } \alpha_{n-1}} f'_\beta \beta \\ &= a' \\ &= \theta(f) \ . \end{aligned}$$

Therefore, $f - a \in \ker\theta = (x_n)S_n$. Write $f - a = gx_n$ for some $g \in S_n$ with $\deg(g) = m - 1$. By induction on degree, we may write $g = \sum_{\beta \text{ divides } \alpha_n} g_\beta \beta$ for some $g_\beta \in S_n^{\Sigma_n}$. Hence we have

$$\begin{aligned} f &= a + gx_n \\ &= \sum_{\beta \text{ divides } \alpha_{n-1}} f_\beta \beta \;+ \sum_{\beta \text{ divides } \alpha_n} g_\beta x_n \beta \end{aligned}$$

We would be done at this stage were it not for the possibility that one of the monomials $x_n\beta$ occurring in the above expansion of f might not divide α_n. This can only happen if the exponent of x_n in $x_n\beta$ is exactly n. We handle this as follows. Recall that $\prod_{i=1}^{n}(t - x_i) = \sum_{j=0}^{n}(-1)^j s_j t^{n-j}$. Substituting $t = x_n$ into this identity we find $0 = \sum_{j=0}^{n}(-1)^j s_j x_n^{n-j}$ and thus $x_n^n = \sum_{j=1}^{n}(-1)^{j+1} s_j x_n^{n-j}$. For each monomial $x_n\beta$ not dividing α_n, we make the substitution $x_n^n = \sum_{j=1}^{n}(-1)^{j+1} s_j x_n^{n-j}$. This shows that f lies in the $S_n^{\Sigma_n}$-module generated by the monomial factors of α_n and so completes the proof. □

It is well-known that the elementary symmetric polynomials generate the ring of symmetric polynomials. Proposition 6.1.1 gives the deeper result showing how to express any polynomial what so ever in terms of symmetric polynomials.

6.2 Block Bases for p-Groups

The lemma below generalizes Proposition 4.0.3.

Lemma 6.2.1. *Let $f_1, f_2, \ldots, f_n$ be a sequence of homogeneous elements in $S := \mathbb{K}[x_1, x_2, \ldots, x_n]$ with $\deg(f_i) = d_i$ for $i = 1, 2, \ldots, n$. Suppose that with respect to some monomial ordering, $\mathrm{LT}(f_i) = x_i^{d_i}$ for $i = 1, 2, \ldots, n$. Then $f_1, f_2, \ldots, f_n$ is a homogeneous system of parameters in $\mathbb{K}[V]$ and $\alpha := \prod_{i=1}^{n} x_i^{d_i - 1}$ generates a block basis for S over $R := \mathbb{K}[f_1, f_2, \ldots, f_n]$*

Proof. First we show that $f_1, f_2, \ldots, f_n$ is a homogeneous system of parameters in $\mathbb{K}[V]$. We will show this by proving that the only common zero of all these n functions evaluated on $\overline{V} := V \otimes_{\mathbb{K}} \overline{\mathbb{K}}$ is the origin, i.e., by applying Lemma 2.6.3. Let $v = (v_1, v_2, \ldots, v_n) \in \overline{V}$ be such that $f_i(v) = 0$ for all $i = 1, 2, \ldots, n$.

Without loss of generality, $x_1 < x_2 < \cdots < x_n$ in the given monomial ordering. By the multiplicative property of monomial orderings, every monomial of degree d_i which is less than $x_i^{d_i}$ must be divisible by some x_j where $j < i$. In particular, $f_i \equiv x_i^{d_i}$ modulo $(x_1, x_2, \ldots, x_{i-1})$. Thus $f_1 = x_1^{d_1}$ and hence $v_1 = 0$. Therefore $f_2(v) = v_2^{d_2}$ which implies that $v_2 = 0$. Continuing in this manner we see that the only common zero is $v = (0, 0, \ldots, 0)$.

As we observed earlier, we know that

$$\mathcal{H}(R, \lambda) = (1 - \lambda^{d_1})^{-1}(1 - \lambda^{d_2})^{-1} \cdot \cdots \cdot (1 - \lambda^{d_n})^{-1}$$

and therefore,

$$\mathcal{H}(S/R_+S, \lambda) = \prod_{i=1}^{n} \frac{(1-\lambda)^{d_i}}{(1-\lambda)} = \prod_{i=1}^{n}(1 + \lambda + \lambda^2 + \cdots + \lambda^{d_i - 1}).$$

From this we see that the rank of S as an R-module is $\prod_{i=1}^{n} d_i$. Since this rank is equal to the number of monomial factors of α, it suffices to show that these factors span S/R_+S or equivalently that they generate S as an R-module.

Let S' denote the R-module generated by the monomial factors of α and suppose, by way of contradiction, that $S' \subsetneq S$. Choose $h \in S \setminus S'$ such that $\mathrm{LM}(h)$ is the smallest possible. Write $\mathrm{LT}(h) = cx_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$ and write $a_i = q_i d_i + r_i$ where $0 \leq r_i < d_i$ for all $i = 1, 2, \ldots, n$. Then $\prod_{i=1}^{n} x_i^{r_i}$ is a monomial factor of α and thus

$$h' := h - c(\prod_{i=1}^{n} f_i^{q_i})(\prod_{i=1}^{n} x_i^{r_i})$$

lies in $S \setminus S'$ and $\mathrm{LT}(h') < \mathrm{LT}(h)$. This contradiction proves the result. □

Corollary 6.2.2. *Suppose that P is a p-group and that V is a representation of P. Choose a basis $\{x_1, x_2, \ldots, x_n\}$ of V^* with respect to which G is upper triangular. Then $\mathbf{N}^G_{G_{x_1}}(x_1), \mathbf{N}^G_{G_{x_2}}(x_2), \ldots, \mathbf{N}^G_{G_{x_n}}(x_n)$ is a homogeneous system of parameters in $\mathbb{F}[V]^P$ and $\alpha = \prod_{i=1}^{n} x_i^{d_i - 1}$ generates a block basis for $\mathbb{F}[V]$ over $\mathbb{F}[\mathbf{N}^G_{G_{x_1}}(x_1), \mathbf{N}^G_{G_{x_2}}(x_2), \ldots, \mathbf{N}^G_{G_{x_n}}(x_n)$ where $d_i = |G|/|G_{x_i}|$.*

Proof. This follows immediately since (using the graded reverse lexicographic ordering with $x_1 < x_2 < \cdots < x_n$) we have $\mathrm{LM}(\mathbf{N}^G_{G_{x_i}}(x_i)) = x_i^{d_i}$ for all $i = 1, 2, \ldots, n$. □

Example 6.2.3. Consider the group of lower triangular matrices,

$$U_n = U_n(V) \subset \mathrm{GL}_n(\mathbb{F}).$$

Choose the dual (upper triangular) basis $\{x_1, x_2, \ldots, x_n\}$ for V^* and use the graded reverse lexicographic order with $x_1 < x_2 < \cdots < x_n$. We define

$$V(i) := \mathrm{span}_{\mathbb{F}}\{x_1, x_2, \ldots, x_i\} \subset V^*$$

for $1 \leq i \leq n$ and $V(0) := \{0\}$. Then $U_n(V) \cdot x_i = \{x_i + y \mid y \in V(i-1)\}$. Let h_i denote the norm

$$h_i := \mathbf{N}^G_{G_{x_i}} = \prod_{y \in V(i-1)} (x_i + y).$$

Clearly, $\mathrm{LT}(h_i) = x_i^{p^{i-1}}$. Then Corollary 6.2.2 asserts that the monomial $x_2^{p-1} x_3^{p^2-1} \cdots x_n^{p^{n-1}-1}$ generates a block basis for $\mathbb{F}_p[V]$ over $\mathbb{F}_p[V]^{U_n}$.

7

The Cyclic Group C_p

7.1 Representations of C_p in Characteristic p

Let $G = C_p$ denote the cyclic group of order p. In this section, we wish to study the representation theory of C_p over a field $\mathbb{F}$ of characteristic p.

We begin by considering a finite dimensional indecomposable C_p-module, V. In representation terminology, this means we have a finite dimensional indecomposable representation $\rho : C_p \to GL(V)$. We fix a generator σ of C_p. Now since $\sigma^p = 1$, every eigenvalue λ of σ must be a p^{th} root of unity. Thus $0 = \lambda^p - 1 = (\lambda - 1)^p$ and so $\lambda = 1$ is the only p^{th} root of unity in $\mathbb{F}$. Since the only eigenvalue of σ lies in $\mathbb{F}$, we may choose a basis $\{e_1, e_2, \ldots, e_n\}$ of V such that $\rho(\sigma)$ is in (lower) Jordan normal form with respect to this basis. If there are two or more Jordan blocks in this Jordan normal form these blocks yield a decomposition of V into a direct sum of smaller C_p-modules contradicting the indecomposability of V. Thus

$$\rho(\sigma) = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ 1 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & 0 & \ldots & 1 & 1 \end{pmatrix}.$$

Lemma 7.1.1. *The matrix above has order p^ℓ if and only if $p^{\ell-1} < n \le p^\ell$. In particular, this matrix has order p if and only if $1 < n \le p$.* □

We note that $\sigma(e_n) = e_n$ and $\sigma(e_i) = e_i + e_{i+1}$ for $1 \le i \le n-1$. We call such a basis a *triangular* basis of V_n and we say that e_1 is *distinguished*. Notice that the C_p-module generated by e_1 is all of V_n.

Definition 7.1.2. *For each n with $1 \le n \le p$ we denote the indecomposable C_p-module of dimension n by V_n.*

H.E.A. Eddy Campbell, D. Wehlau, *Modular Invariant Theory*,
Encyclopaedia of Mathematical Sciences 139,
DOI 10.1007/978-3-642-17404-9_7, © Springer-Verlag Berlin Heidelberg 2011

The preceding discussion, combined with lemma 7.1.1, proves the

Lemma 7.1.3. *Over any field of characteristic p, there are exactly p^r inequivalent indecomposable representations of C_{p^r}, one of dimension n for each n less than or equal to p^r. Furthermore, we have the following chain of inclusions:*

$$V_1 \subset V_2 \subset \cdots \subset V_{p^r}.$$

□

The indecomposable representations V_n also occur in the following way. Consider the action of σ on V_2, the indecomposable representation of dimension 2:

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

We saw in Theorem 1.11.2 that $\mathbb{F}[V_2]^{C_p} = \mathbb{F}[x, y^p - x^{p-1}y]$. In particular, for $0 \leq d \leq p-1$, we see that $\mathbb{F}[V_2]_d^{C_p}$ is spanned by x^d. Thus $\dim \mathbb{F}[V_2]_d^{C_p} = 1$ for $0 \leq d \leq p-1$ which implies that $\dim \mathbb{F}[V_2]_d$ is an indecomposable C_p-representation for $d < p$. Since $\dim \mathbb{F}[V_2]_d = d+1$ we must have $\mathbb{F}[V_2]_d \cong V_{d+1}$ for $d = 0, 1, \ldots, p-1$.

In general, any p-group has only one projective indecomposable module up to isomorphism in characteristic p, see Benson [7, §3.14]. Here we prove

Lemma 7.1.4. *If the indecomposable C_p-module V_r is projective, then $r = p$. Thus every projective C_p-module is a free C_p-module.*

Remark 7.1.5. We observe that $V_p \cong \mathbb{F}C_p$ and that, therefore, V_p is (isomorphic to) the regular representation of C_p.

Proof. Let $\{e_1, e_2, \ldots, e_p\}$ be a triangular basis for V_p where e_1 is distinguished. Consider the submodule $V_r = \operatorname{span}_{\mathbb{F}}\{e_{p-r+1}, e_{p-r+2}, \ldots, e_p\}$ where $r < p$. Assume, by way of contradiction, that V_r is projective and consider the short exact sequence of C_p-modules:

$$0 \to W \to V_p \xrightarrow{\nu} V_r \to 0$$

given by $\nu(e_i) = \begin{cases} e_{p-r+i}, & \text{if } i \leq r; \\ 0, & \text{if } i > r; \end{cases}$ where $W = \ker(\nu)$.

If V_r is projective, then this sequence splits and thus V_r must have a complement M in V_p and we have $V_p = V_r \oplus M$. But then $V_p^{C_p} = V_r^{C_p} \oplus M^{C_p}$ has dimension at least 2. This contradiction proves that V_r is not projective.

□

Remark 7.1.6. Since V_p is projective, its dual, V_p^*, is an injective module. But since $V_p \cong V_p^*$, we see that V_p is injective.

Remark 7.1.7. We note here that if G is a finite group with non-cyclic p-Sylow subgroups, then there are infinitely many inequivalent indecomposable representations of G over a field of characteristic p. If $p = 2$ and the 2-Sylow subgroups of G are isomorphic to either the Klein 4-group, a dihedral group, a semi-dihedral group or a generalized quaternion group, then the representation theory of G over an infinite field $\mathbb{F}$ of characteristic 2 is *tame*. This means that all but finitely many indecomposable representations of each fixed dimension can be parameterized by essentially one parameter (running over $\mathbb{F}$). Of course, over a finite field, there are always only finitely many representations of a given dimension since $\mathrm{GL}(n, \mathbb{F})$ has only finitely many subgroups for each n and so the notion of tame does not make sense over finite fields. If the p-Sylow subgroups of G are non-cyclic and not the Klein 4-group, nor a dihedral group, nor a semi-dihedral group nor a generalized quaternion group, then the modular representation theory of G is *wild*. This means that the problem of classifying the inequivalent indecomposable representations contains the unsolved problem of simultaneously reducing to Jordan canonical form two linear operators on a finite-dimensional space. Solving this problem is considered hopeless. For details about these results, see, for example the book of Benson [6, § 4.4]. But recall from §3.6 that Symonds has shown that only finitely isomorphism classes of representations of such a group G occur in any decomposition of $\mathbb{F}[V]^G$ into indecomposable G-modules.

An immediate consequence of Lemma 7.1.3 is that the equivalence class of an indecomposable C_p-module, V, is entirely determined by $\dim(V)$. In particular, the indecomposable module $V^* := \hom(V, \mathbb{F})$ is isomorphic to V since $\dim(V^*) = \dim(V)$.

Since we are primarily interested in invariants we often concentrate on the C_p action on V^* rather than on V itself. Accordingly we will usually choose the dual basis $\{x_1, x_2, \dots, x_n\}$ for V^* to the basis $\{e_1, e_2, \dots, e_n\}$ and replace the previously chosen generator σ by the new generator σ^{-1} (which will we again denote by σ. In other words, we choose an upper triangular basis of V^* with $\sigma(x_1) = x_1$ and $\sigma(x_i) = x_i + x_{i-1}$ for $2 \leq i \leq n$. Again, we call x_n a *distinguished variable* since x_n generates the cyclic C_p-module V^*. Observe that $V_n^{C_p}$ has basis $\{e_n\}$ and that $(V_n^*)^{C_p}$ has basis $\{x_1\}$. Note that if $\lambda \in V^*$, then λ is a distinguished variable for V^* if and only if λ restricted to $V_n^{C_p}$ is not identically zero.

As in Chapter 1, we consider the *twisted derivation* $\Delta : \mathbb{F}[V] \to \mathbb{F}[V]$ defined as $\Delta = \sigma - \mathrm{Id}$. Then $\ker(\Delta) = \mathbb{F}[V]^{C_p}$ and for $f, f' \in \mathbb{F}[V]$, we have $\Delta(ff') = \Delta(f)f' - \sigma(f)\Delta(f')$. As well, we note that $\Delta^{p-1} = \sum_{i=0}^{p-1} \sigma^i = \mathrm{Tr}^{C_p}$. In this notation, we observe that $x_i = \Delta^{n-i}(x_n)$.

We note that for such a triangular basis of V_n^*, that the fixed point x_1 is dual to a distinguished element of V_n. Conversely, the distinguished variable x_n is dual to a fixed vector in V_n.

If $W \subset V$ are two G-modules, we say W is a *summand* of V if there exists another G-submodule U of V such that $V = W \oplus U$. In this case, we say that U and W are *complements* of one another. If V is any finite dimensional C_p-module, then V can be decomposed into a direct sum of indecomposable C_p-modules:

$$V \cong m_1V_1 \oplus m_2V_2 \cdots \oplus \ldots m_pV_p$$

where $m_i \in \mathbb{N}$ for all i and $mW := \underbrace{W \oplus W \oplus \cdots \oplus W}_{m \text{ summands}}$. The invariants $\mathbb{F}[m\,W]^{C_p}$, associated to $m\,W$, are referred to as *vector* invariants.

Lemma 7.1.8. *If $V \cong m_1V_1 \oplus m_2V_2 \cdots \oplus \ldots m_pV_p$ and $V \cong m_1'V_1 \oplus m_2'V_2 \cdots \oplus \ldots m_p'V_p$ are two decompositions of V, then $m_i = m_i'$ for all $1 \le i \le p$.*

Proof. Note that the kernel of $\Delta : V_i \to V_i$ is $V_i^{C_p}$ which is one dimensional for all i. Thus

$$\dim(\Delta^j(V_i)) = \begin{cases} 0 & \text{if } j \ge i, \\ i - j & \text{if } j < i. \end{cases}$$

Therefore, $(p-j)m_p + (p-1-j)m_{p-1} + \cdots + m_{j+1} = \dim(\Delta^j(V))$ for all $0 \le j \le p-1$. Clearly, this system of equations uniquely determines the coefficients $m_1, m_2, \ldots, m_p$. □

Remark 7.1.9. We may also prove the above lemma by observing that for each i, the number of Jordan blocks of size i in the Jordan normal form for the matrix of σ is m_i.

Remark 7.1.10. Each indecomposable C_p-module, V_n, satisfies $\dim_{\mathbb{F}}(V_n)^{C_p} = 1$. Therefore, the number of summands occurring in a decomposition of V is given by $m_1 + m_2 + \cdots + m_p = \dim V^{C_p}$.

It is important to observe that such a decomposition is in general not unique, there are many choices for the individual summands. The following example gives a simple illustration of this.

Example 7.1.11. Let $p \ge 5$ and consider three indecomposable C_p modules: $A = \operatorname{span}_{\mathbb{F}}\{a_5, a_4, a_3, a_2, a_1\} \cong V_5$, $B = \operatorname{span}_{\mathbb{F}}\{b_5, b_4, b_3, b_2, b_1\} \cong V_5$ and $C = \operatorname{span}_{\mathbb{F}}\{c_2, c_1\} \cong V_2$. Here we suppose the given bases are triangular with fixed points a_1, b_1 and c_1. Now define $V := A \oplus B \oplus C$.

The C_p-submodule $A' := \operatorname{span}_{\mathbb{F}}\{a_5+b_5+c_2, a_4+b_4+c_1, a_3+b_3, a_2+b_2, a_1+b_1\}$ is isomorphic to V_5. Similarly, $B' := \operatorname{span}_{\mathbb{F}}\{a_5 + 2b_5 + 3c_1, a_4 + 2b_4, a_3 + 2b_3, a_2+2b_2, a_1+2b_1\}$ is isomorphic to V_5. Finally, if we take $C' := \operatorname{span}_{\mathbb{F}}\{a_2+b_1-c_2, a_1-c_1\} \cong V_2$ then we have another decomposition: $V = A' \oplus B' \oplus C'$.

In the non-modular situation, every indecomposable submodule has a complement. This is not the case in the modular setting as the following example illustrates.

Example 7.1.12. We continue with the notation of the previous example. The submodules A, B, C, A', B', C', $A \oplus B$, $A \oplus C$, etc. are all summands of V. $B \oplus C$ is a complement of A.

Consider $D := \operatorname{span}_{\mathbb{F}}\{a_3, a_2, a_1\}$ and $E := \operatorname{span}_{\mathbb{F}}\{a_3 + c_1, a_2, a_1\}$. These two modules do not have any complements and so are not summands of V. Note that while D lies in a summand, A, the submodule E can not be extended to an indecomposable summand.

In general, the decomposition of a tensor product $V_n \otimes V_m$ into indecomposable summands can be somewhat complicated to derive. It is clear that $V_1 \otimes V_n \cong V_n$ and the following lemma shows how to decompose $V_p \otimes V_n$.

Lemma 7.1.13. *$V_p \otimes V_n \cong n\, V_p$.*

Proof. We proceed by induction on n. For $n = 1$, the isomorphism $V_p \otimes V_1 \cong V_p$ is clear. For the general case, we assume the induction hypothesis $V_p \otimes V_{n-1} \cong (n-1)\, V_p$ and consider the short exact sequence

$$0 \longrightarrow V_{n-1} \longrightarrow V_n \longrightarrow V_1 \longrightarrow 0$$

Since V_p is free, hence flat, tensoring with V_p yields a new short exact sequence:

$$0 \longrightarrow V_p \otimes V_{n-1} \longrightarrow V_p \otimes V_n \longrightarrow V_p \otimes V_1 \cong V_p \longrightarrow 0$$

Since V_p is projective, this sequence splits and thus

$$V_p \otimes V_n \cong V_p \otimes V_{n-1} \oplus V_p \cong (n-1)\, V_p \oplus V_p \cong n\, V_p$$

□

We also have a simple formula for the result of tensoring with a copy of V_2.

Lemma 7.1.14.

$$V_n \otimes V_2 \cong \begin{cases} V_2, & \textit{if } n = 1; \\ V_{n-1} \oplus V_{n+1}, & \textit{if } 1 < n < p; \\ 2\, V_p, & \textit{if } n = p. \end{cases}$$

Proof. The case $n = 1$ is clear and the case $n = p$ is a special case of Lemma 7.1.13. Suppose then that $1 < n < p$. We choose triangular bases $\{e_1, e_2, \ldots, e_n\}$ and $\{e'_1, e'_2\}$ for V_n and V_2 respectively. Here e_1 and e'_1 are distinguished and e_n and e'_2 are C_p-fixed. Define $v_1 := e_1 \otimes e'_1$, $w_1 := (n-1)e_1 \otimes e'_2 - e_2 \otimes e'_1$ and $v_t := \Delta^{t-1}(v_1)$ for $1 \leq t \leq n+1$ and $w_t := \Delta^{t-1}(w_1)$ for $1 \leq t \leq n-1$. It is easy to verify that $w_t = (n-t)e_t \otimes e'_2 - e_{t+1} \otimes e'_1 - (t-1)e_{t+1} \otimes e'_2$ for $1 \leq t \leq n-1$ and $v_t = e_t \otimes e'_1 + (t-1)e_{t-1} \otimes e'_2 + (t-1)e_t \otimes e'_2$ for $1 \leq t \leq n$ and $v_{n+1} = ne_n \otimes e'_2$. Since v_{n+1} and w_{n-1} are 2 linearly independent C_p-fixed vectors, we see that $V_2 \otimes V_n$ contains at least 2 indecomposable summands. On the other hand, $V := \mathrm{span}_{\mathbb{F}}\{v_1, v_2, \ldots, v_{n+1}\} \cong V_{n+1}$ and $W := \mathrm{span}_{\mathbb{F}}\{w_1, w_2, \ldots, w_{n-1}\} \cong V_{n-1}$. Furthermore, $V \cap W = \{0\}$ since their fixed lines are linearly independent. Finally, considering dimensions we see that $V_2 \otimes V_n = V \oplus W \cong V_{n+1} \oplus V_{n-1}$. □

Remark 7.1.15. A shorter non-constructive proof using the Jordan normal form of matrices has been given by Hughes and Kemper, [54, Lemma 2.2].

7.2 The C_p-Module Structure of $\mathbb{F}[V_n]$

We now want to consider the symmetric algebra $\mathbb{F}[V]$ as a graded C_p-module, and as a module over $\mathbb{F}[V]^{C_p}$. In this section, we treat the case where $V = V_n$ is indecomposable. In the next section, we will consider decomposable modules V. Although $\mathbb{F}[V_n]$ is infinite dimensional, we have the degree decomposition $\mathbb{F}[V_n] = \bigoplus_{d=0}^{\infty} \mathbb{F}[V_n]_d$ and each $\mathbb{F}[V_n]_d$ is a finite dimensional C_p-module. Our goal is to decompose each $\mathbb{F}[V_n]_d$ as a direct sum: $\mathbb{F}[V_n]_d = m_{d,1}V_1 \oplus m_{d,2}V_2 \oplus \cdots \oplus m_{d,p}V_p$. We fix notation by choosing a triangular basis $\{x_1, x_2, \ldots, x_n\}$ for V_n^* with distinguished variable x_n.

In order to solve the decomposition problem, we will consider $\mathbb{F}[V_n]$ not just as a C_p-module but also as a $\mathbb{F}[V_n]^{C_p}$-module. Suppose that W is an indecomposable summand of the C_p-module $\mathbb{F}[V_n]_d$ and $f \in \mathbb{F}[V_n]^{C_p}$ is a homogeneous invariant of degree r. Then $f \cdot W$ is a C_p-submodule of $\mathbb{F}[V_n]_{d+r}$ isomorphic to W. However, as the following example shows, $f \cdot W$ need not be a summand of $\mathbb{F}[V_n]_{d+r}$.

Example 7.2.1. Consider $\mathbb{F}[V_2]$ as a C_p-module where $\mathbb{F}$ has positive characteristic $p \neq 2$. As usual, we denote the generator of C_p as σ and write $\Delta = \sigma - 1$. Choose a basis $\{x_1, x_2\}$ of $V_2^* = \mathbb{F}[V_2]_1$ such that $\Delta x_1 = 0$ and $\Delta x_2 = x_1$. Note that $\{v_3 := x_2^2, v_2 := 2x_1x_2 + x_1^2, v_1 := 2x_1^2\}$ is a basis of $\mathbb{F}[V_2]_2$ such that $\Delta(v_3) = v_2$, $\Delta(v_2) = v_1$ and $\Delta(v_1) = 0$. Thus $\mathbb{F}[V_2]_1 \cong V_2$ and $\mathbb{F}[V_2]_2 \cong V_3$. Also, $x_1 \in \mathbb{F}[V_2]_1^{C_p}$. However, $x_1 \cdot \mathbb{F}[V_2]_1 \cong V_2 \subset \mathbb{F}[V_2]_2 \cong V_3$ and $x_1 \cdot \mathbb{F}[V_2]_1$ is not a summand of $\mathbb{F}[V_2]_2$.

Note that when $p = 2$, we have $\Delta(v_2) = v_1 = 0$. Thus in characteristic 2, $\mathbb{F}[V_2]_2 \cong V_2 \oplus V_1$ where $\mathrm{span}_{\mathbb{F}}\{x_2^2, x_1^2\} \cong V_2$ and $\mathrm{span}_{\mathbb{F}}\{x_2^2 + x_1x_2\} \cong V_1$.

7.2.1 Sharps and Flats

Let $\mathbb{F}[V_n]^\sharp$ denote the principal ideal of $\mathbb{F}[V_n]$ generated by $\mathbf{N}^{C_p}(x_n)$. Clearly we may assume $n \geq 2$. Note that $\deg_{x_n}(\mathbf{N}^{C_p}(x_n)) = p$ and that $\mathbf{N}^{C_p}(x_n)$ is monic when considered as a polynomial in x_n; therefore, we may divide any given $f \in \mathbb{F}[V_n]$ by $\mathbf{N}^{C_p}(x_n)$ to obtain $f = q \cdot \mathbf{N}^{C_p}(x_n) + r$ for some $q, r \in \mathbb{F}[V_n]$ with $\deg_{x_n}(r) < p$.

Note that if f is invariant, then so are both q and r. To see this, apply σ to the equation $f = q \cdot \mathbf{N}^{C_p}(x_n) + r$ to get $f = \sigma(q) \cdot \mathbf{N}^{C_p}(x_n) + \sigma(r)$.

Since x_n is a distinguished variable, $\deg_{x_n}(\sigma(r)) = \deg_{x_n}(r) < p$. Thus, by the uniqueness of remainders, $\sigma(r) = r$ and therefore $\sigma(q) = q$.

We define $\mathbb{F}[V]^\flat := \{r \in \mathbb{F}[V] \mid \deg_{x_n}(r) < p\}$. Note that both $\mathbb{F}[V_n]^\flat$ and $\mathbb{F}[V_n]^\sharp$ are vector spaces and that as vector spaces, $\mathbb{F}[V_n] = \mathbb{F}[V_n]^\sharp \oplus \mathbb{F}[V_n]^\flat$. Furthermore, both $\mathbb{F}[V_n]^\flat$ and $\mathbb{F}[V_n]^\sharp$ are C_p-stable. Therefore, we have the C_p-module decomposition

$$\mathbb{F}[V_n] = \mathbb{F}[V_n]^\sharp \oplus \mathbb{F}[V_n]^\flat.$$

In contrast to Example 7.2.1, we have

Proposition 7.2.2. *Let W be an indecomposable summand of $\mathbb{F}[V_n]_d$ and let $f \in \mathbb{F}[V_n]^{C_p}$ be a homogeneous invariant of degree r.*

1. *If $W \cong V_p$, then $f \cdot W$ is a summand of $\mathbb{F}[V_n]_{d+r}$.*
2. *If $f = \mathbf{N}(x_n)$, then $f \cdot W$ is a summand of $\mathbb{F}[V_n]_{d+p}$.*

Proof. For the first we assertion, we note that $f \cdot W$ is a C_p-submodule of $\mathbb{F}[V_n]_{d+r}$. Since $\dim(f \cdot W) = p$, we see that $f \cdot W \cong V_p$ is an injective submodule by Remark 7.1.6. Hence the inclusion $fW \hookrightarrow \mathbb{F}[V_n]_{d+r}$ splits and therefore, $\mathbb{F}[V_n]_{d+r} \cong f \cdot W \oplus W'$ for some C_p-module W'.

For the second assertion, write $\mathbb{F}[V_n]_d = W \oplus W'$ for some C_p-module W'. Then

$$\begin{aligned} \mathbb{F}[V_n]_{d+p} &\cong \mathbb{F}[V_n]^\sharp_{d+p} \oplus \mathbb{F}[V_n]^\flat_{d+p} \\ &\cong \mathbf{N}(x_n) \cdot \mathbb{F}[V_n]_d \oplus \mathbb{F}[V_n]^\flat_{d+p} \\ &\cong \mathbf{N}(x_n) \cdot W \oplus \mathbf{N}(x_n) \cdot W' \oplus \mathbb{F}[V_n]^\flat_{d+p} \end{aligned}$$

□

Lemma 7.2.3. *The C_p-module $\mathbb{F}[V_n]^\flat_d$ is free if $d + n \geq p + 1$.*

Proof. The proof is by (downward) induction on n. For the base case $n = p$, we are considering the regular representation V_p and we may choose a basis $\{w_1, w_2, \ldots, w_p\}$ of V_p^* which is (transitively) permuted by the action of C_p. Then the degree d monomials in the variables $\{w_1, w_2, \ldots, w_p\}$ form a vector space basis for $\mathbb{F}[V_p]_d$ and this basis of monomials is again permuted by C_p. Clearly, the only invariant monomials are those of the form $(w_1 w_2 \cdots w_p)^b$ for some non-negative integer b. All other monomials subdivide naturally into C_p-orbits of size p.

Notice that any of the variables w_i may be chosen as a distinguished variable x_p for V_p since each generates V_p^* as a C_p-module. Thus we may take $\mathbf{N}^{C_p}(x_p) = \mathbf{N}^{C_p}(w_i) = \prod_{\tau \in C_p} \tau(w_i) = w_1 w_2 \cdots w_p$. From this we see that $\mathbb{F}[V_p]^\sharp$ is spanned by the monomials divisible by $w_1 w_2 \cdots w_p$, i.e., by the monomials $w_1^{e_1} w_2^{e_2} \cdots w_p^{e_p}$ where all $e_i \geq 1$. Also $\mathbb{F}[V_p]^\flat$ is spanned by the monomials not divisible by $w_1 w_2 \cdots w_p$, i.e., by the monomials $w_1^{e_1} w_2^{e_2} \cdots w_p^{e_p}$

where $e_i = 0$ for at least one i. In particular, every monomial, other than the monomial 1, in the basis of $\mathbb{F}[V_p]^\flat$ lies in a C_p-orbit of order p. Thus for $a \geq 1$ we see that $\mathbb{F}[V_p]^\flat_d$ is a direct sum of copies of V_p, i.e., is a free C_p-module.

For the general case, we suppose that the result holds for $\mathbb{F}[V_{n+1}]^\flat$ and that $d \geq (p+1) - n$. Let $\{x_1, x_2, \ldots, x_{n+1}\}$ be a triangular basis for V^*_{n+1} where x_{n+1} is distinguished. Consider the short exact sequence of C_p-modules

$$0 \longrightarrow \mathbb{F}[V_{n+1}]^\flat_d \xrightarrow{\mu} \mathbb{F}[V_{n+1}]^\flat_{d+1} \xrightarrow{\theta} \mathbb{F}[V_n]^\flat_{d+1} \longrightarrow 0 \ .$$

Here μ is given by multiplication by the C_p-fixed variable x_1 and θ is induced by $\theta(x_i) = \begin{cases} 0, & \text{if } i = 1; \\ x_{i-1}, & \text{if } i \geq 2. \end{cases}$ Now $\mathbb{F}[V_{n+1}]^\flat_d$ and $\mathbb{F}[V_{n+1}]^\flat_{d+1}$ are both free by the induction hypothesis. Thus by Remark 7.1.6, we see that $\mathbb{F}[V_{n+1}]^\flat_d$ is injective. Thus the above sequence splits and $\mathbb{F}[V_n]^\flat_{d+1}$ is projective. But by Lemma 7.1.4, this means that $\mathbb{F}[V_n]^\flat_{d+1}$ is free. □

Theorem 7.2.4. *Let d be a non-negative integer and write $d = qp + r$ where $0 \leq r \leq p-1$. Then $\mathbb{F}[V_n]_d \cong \mathbb{F}[V_n]_r \oplus k\, V_p$ as C_p-modules for some non-negative integer k.*

Proof. The proof is by induction on q. If $q = 0$ the result is trivially true. Suppose $d \geq p$ (so q is at least 1) and consider the C_p-equivariant homomorphism

$$\mu : \mathbb{F}[V_n]_{d-p} \longrightarrow \mathbb{F}[V_n]^\sharp_d$$

given by multiplication by $\mathbf{N}^{C_p}(z)$ for a distinguished variable z. The map μ is onto by definition of $\mathbb{F}[V_n]^\sharp$ and is injective since $\mathbb{F}[V_n]$ is a domain. Thus μ is an isomorphism. Also, since $d \geq p$, we know, by Lemma 7.2.3, that $\mathbb{F}[V_n]^\flat_d$ is free and so we may write $\mathbb{F}[V_n]^\flat_d \cong s\, V_p$ for some s. Thus

$$\begin{aligned} \mathbb{F}[V_n]_d &\cong \mathbb{F}[V_n]^\sharp_d \oplus \mathbb{F}[V_n]^\flat_d \\ &\cong \mathbb{F}[V_n]_{d-p} \oplus s\, V_p \\ &\cong (\mathbb{F}[V_n]_r \oplus t\, V_p) \oplus s\, V_p \end{aligned}$$

for some t where the last isomorphism follows from the induction hypothesis. Hence $\mathbb{F}[V_n]_d \cong \mathbb{F}[V_n]_r \oplus k\, V_p$. Furthermore, we may compute $k = \left(\binom{n+d-1}{d} - \binom{n+r-1}{r}\right)/p$ by comparing dimensions. □

Example 7.2.5. We consider $p = 11$ and $n = 4$. What are we able to say at this point about the decomposition of $\mathbb{F}[V_4]_d$? According to Theorem 7.2.4, if we know the decomposition of $\mathbb{F}[V_4]_d$ into indecomposable representations for $1 \leq d \leq 10$, then we know the decomposition of $\mathbb{F}[V_4]_d$ for all values of d. Furthermore, by Lemma 7.2.3, we know that $\mathbb{F}[V_4]_d = \mathbb{F}[V_4]^\flat_d$ is free for $d = 8, 9$ and 10. As well, we have $\mathbb{F}[V_4]_0 = \mathbb{F} = V_1$ and $\mathbb{F}[V_4]_1 = V_4$. In order to complete the calculation for $d = 2, 3, 4, 5, 6$ and 7, it suffices to compute the matrix representing the action of σ on $\mathbb{F}[V_4]_d$ for these values of d and rewrite them in Jordan Normal Form. This is a straightforward problem of Gaussian elimination. Using a MAGMA script to do this we obtained the results summarized in the following table.

d	dimension	Non-Free Summands	Free Summands
0	1	V_1	
1	4	V_4	
2	10	$V_3 \oplus V_7$	
3	20	$V_4 \oplus V_6 \oplus V_{10}$	
4	35	$V_1 \oplus V_5 \oplus V_7$	$2\,V_{11}$
5	56	$V_4 \oplus V_8$	$4\,V_{11}$
6	84	V_7	$7\,V_{11}$
7	120	V_{10}	$10\,V_{11}$
8	165		$15\,V_{11}$
9	220		$20\,V_{11}$
10	286		$26\,V_{11}$

Now, suppose we wish to know the decomposition of $\mathbb{F}[V_4]_d$ for some large value of d. For $d = 126$ for example, we observe that $126 \equiv 5 \pmod{11}$ and thus $\mathbb{F}[V_4]_{126} \cong V_4 \oplus V_8 \oplus 4\,V_{11} \oplus k\,V_{11}$, where $k = 31768$ again by Theorem 7.2.4.

Note that one can use these arguments to compute the Hilbert series of $\mathbb{F}[V_4]^{C_p}$. That is, we know that the coefficient of λ^a is equal to the number of summands in $\mathbb{F}[V_4]_d$. For example, the coefficient of λ^{126} is 31774. This example is revisited in detail in Example 13.0.2.

7.3 The C_p-Module Structure of $\mathbb{F}[V]$

We now extend these ideas to decomposable C_p-modules, V. This is not too difficult since most of the ideas extend in a straightforward manner.

In order to simplify things, we first observe that if V_1 is a summand of V, say $V = V_1 \oplus V'$, then

$$\mathbb{F}[V] = \mathbb{F}[V_1 \oplus V'] = \mathbb{F}[V_1] \otimes \mathbb{F}[V'] \ ,$$

and more importantly,

$$\mathbb{F}[V]^{C_p} = \mathbb{F}[V_1] \otimes \mathbb{F}[V']^{C_p}.$$

Thus we may easily describe $\mathbb{F}[V]$ and $\mathbb{F}[V]^{C_p}$ once we have understood $\mathbb{F}[V']$ and $\mathbb{F}[V']^{C_p}$. Therefore, by induction, we may assume that V contains no summand isomorphic to V_1. Such a representation V is said to be *reduced.*

Let V be a reduced finite dimensional representation of C_p and decompose V into a direct sum of indecomposable summands:

$$V = V_{n_1} \oplus V_{n_2} \oplus \cdots \oplus V_{n_m} \ .$$

Choose a distinguished variable $z_i \in V_{n_i}^*$ for each $i = 1, 2, \ldots, m$. Then the set $\{z_{i,j} := \Delta^j(z_i) \mid 0 \leq j \leq n_i - 1, 1 \leq i \leq m\}$ is a vector space basis for V^*. We let $N_i = \mathbf{N}^{C_p}(z_i) = \prod_{\tau \in C_p} \tau \cdot z_i$ denote the norm of z_i for $i = 1, 2, \ldots, m$.

Let $f \in \mathbb{F}[V]^{C_p}$. Since N_1 is monic, when considered as a polynomial in the single variable z_1, we may divide N_1 into f to obtain the unique decomposition $f = f_1 N_1 + r_1$ where the remainder r_1 has degree at most $p - 1$ in the variable z_1. Next, we divide r_1 by N_2 to obtain a unique decomposition: $f = f_1 N_1 + f_2 N_2 + r_2$ where $\deg_{z_1}(f_2) < p$, $\deg_{z_1}(r_2) < p$ and $\deg_{z_2}(r_2) < p$. Continuing in this manner we obtain a unique decomposition

$$f = f_1 N_1 + f_2 N_2 + \ldots + f_m N_m + r$$

where $\deg_{z_i}(f_j) < p$ for all $i < j$ and $\deg_{z_i}(r) < p$ for all i. We note that r is unchanged under a reordering of the norms. Therefore, the set $\{N_1, N_2, \ldots, N_m\}$ is a Gröbner basis for the ideal they generate and r is the normal form of f, see [1, Theorem 1.6.7, p. 34]. We will call this the *norm decomposition* of f. Note that the norm decomposition depends upon the choice and order of the z_i but is otherwise unique.

Proposition 7.3.1. *Suppose $f \in \mathbb{F}[V]^{C_p}$ and consider its norm decomposition: $f = f_1 N_1 + f_2 N_2 + \ldots + f_m N_m + r$. Then $f_1, f_2, \ldots, f_t, r \in \mathbb{F}[V]^{C_p}$.*

Proof. Applying σ we have that $f = \sigma(f_1) \cdot N_1 + \sigma(f_2) \cdot N_2 \ldots + \sigma(f_m) \cdot N_m + \sigma(r)$. Since $\deg_{z_i}(\sigma(r)) = \deg_{z_i}(r)$ and $\deg_{z_i}(\sigma(f_j)) = \deg_{z_i}(f_j)$ for all i and j, the uniqueness of the norm decomposition shows that $\sigma(r) = r$ and $\sigma(f_j) = f_j$ for all j. □

As before, denote by $\mathbb{F}[V]^\sharp$ the ideal of $\mathbb{F}[V]$ generated by the norms $N_1, N_2, \ldots, N_m$, and $\mathbb{F}[V]^\flat := \{r \in \mathbb{F}[V] \mid \deg_{z_i}(r) < p \text{ for all } i = 1, 2, \ldots, m\}$. Thus $\mathbb{F}[V]^\flat$ is the set of functions f having all coefficients $f_i = 0$ in its norm decomposition. Note that again $\mathbb{F}[V]^\flat$ and $\mathbb{F}[V]^\sharp$ are both C_p-stable and we have the sharps and flats C_p-module decomposition

$$\mathbb{F}[V] = \mathbb{F}[V]^\sharp \oplus \mathbb{F}[V]^\flat.$$

The decomposition $V = V_{n_1} \oplus V_{n_2} \oplus \cdots \oplus V_{n_m}$ induces an isomorphism $\mathbb{F}[V] \cong \mathbb{F}[V_{n_1}] \otimes \mathbb{F}[V_{n_2}] \otimes \cdots \otimes \mathbb{F}[V_{n_m}]$. This isomorphism in turn yields an $\mathbb{N}^m$ multi-grading on $\mathbb{F}[V]$ given by the degrees in each V_{n_i}:

$$\mathbb{F}[V]_{(d_1, d_2, \ldots, d_m)} = \mathbb{F}[V_{n_1}]_{d_1} \otimes \mathbb{F}[V_{n_2}]_{d_2} \otimes \cdots \otimes \mathbb{F}[V_{n_m}]_{d_m} \ .$$

The action of C_p preserves this grading and thus $\mathbb{F}[V]^{C_p}$, $\mathbb{F}[V]^\sharp$ and $\mathbb{F}[V]^\flat$ each inherit this grading. The following general version of Theorem 7.2.4 follows.

Theorem 7.3.2 (Periodicity Theorem). *Let $V = V_{n_1} \oplus V_{n_2} \oplus \cdots \oplus V_{n_m}$. Let $d_1, d_2, \ldots, d_m$ be non-negative integers and write $d_i = q_i p + r_i$ where $0 \leq r_i \leq p-1$ for $i = 1, 2, \ldots, m$. Then*

$$\mathbb{F}[V]_{(d_1,d_2,\ldots,d_m)} \cong \mathbb{F}[V]_{(r_1,r_2,\ldots,r_m)} \oplus k\,V_p$$

as C_p-modules for some non-negative integer k.

Proof. We will induct on m. Write $V = V_{n_1} \oplus W$ for $W = V_{n_2} \oplus \cdots \oplus V_{n_m}$. By induction, we have

$$\begin{aligned}\mathbb{F}[V]_{(d_1,d_2,\ldots,d_m)} &\cong \mathbb{F}[V_{n_1}]_{d_1} \otimes \mathbb{F}[W]_{(d_2,\ldots,d_m)} \\ &\cong (\mathbb{F}[V_{n_1}]_{r_1} \oplus sV_p) \otimes (\mathbb{F}[W]_{(r_2,\ldots,r_m)} \oplus tV_p) \\ &\cong \mathbb{F}[V]_{(r_1,\ldots,r_m)} \oplus kV_p\end{aligned}$$

as claimed. □

Finally, we want to generalize Lemma 7.2.3. It is easy to see that

$$\mathbb{F}[V]^\flat_{(d_1,d_2,\ldots,d_t)} \cong \mathbb{F}[V_{n_1}]^\flat_{d_1} \otimes \mathbb{F}[V_{n_2}]^\flat_{d_2} \otimes \cdots \otimes \mathbb{F}[V_{n_m}]^\flat_{d_m}.$$

Using Lemma 7.1.13 we see by Lemma 7.2.3 that $\mathbb{F}[V]^\flat_{(d_1,d_2,\ldots,d_m)}$ is free if $d_i + n_i \geq p+1$ for some i with $1 \leq i \leq m$. In particular, if $d = d_1 + d_2 + \cdots + d_m > mp - (n_1 + n_2 + \cdots + n_m)$, then some d_i must exceed $p - n_i$ and therefore, $\mathbb{F}[V]^\flat_d$ is free. This proves following result.

Proposition 7.3.3. *Suppose V is a reduced C_p-module with $m = \dim V^{C_p}$. Let d be a positive integer with $d > m(p-2)$. Then $\mathbb{F}[V]_d \cong \mathbb{F}[V]^\sharp_d \oplus k\,V_p$ for some k.*

7.4 The First Fundamental Theorem for V_2

This section is devoted to finding generators for the vector invariants of V_2, i.e., for $\mathbb{F}[m\,V_2]^{C_p}$ for all $m \geq 1$.

In 1990, Richman [92] conjectured a set of generators for $\mathbb{F}[m\,V_2]^{C_p}$. Campbell and Hughes [19] proved Richman's conjectured generating set is correct in 1997. Their proof itself is somewhat difficult and relies upon a deep result of Wilson concerning the rank of 0-1 matrices in characteristic p.

Here we give a new proof of the first fundamental theorem due to Campbell, Shank and Wehlau [21]. This new proof enjoys a number of advantages. The new proof yields a description of the C_p-module structure of $\mathbb{F}[m\,V_2]$. This description is quite explicit and easily computable. The new proof also provides a SAGBI basis for the ring of invariants. Thirdly, the new proof is simpler and in particular avoids the use of Wilson's theorem.

We will prove the following.

Theorem 7.4.1. *Let $G = C_p$ 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 $\sigma(x_i) = x_i$. 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. *$\mathbf{N}(y_i) = y_i^p - x_i^{p-1} y_i$ for $i = 1, 2, \ldots, m$.*
3. *$u_{ij} = x_i y_j - x_j y_i$ for $1 \le i < j \le m$.*
4. *$\mathrm{Tr}^{C_p}(y_1^{a_1} y_2^{a_2} \ldots y_m^{a_m})$ where $0 \le a_i < p$ for $i = 1, 2, \ldots, m$.*

Moreover, this set of generators is a SAGBI basis for the ring of invariants with respect to the graded reverse lexicographic order determined by $y_1 > x_1 > y_2 > \cdots > x_m$.

A theorem such as this giving explicit algebra generators for all vector invariants of a fixed representation is called a *first fundamental theorem.* A *second fundamental theorem* would give a generating set for the algebraic relations among such generators.

Remark 7.4.2. It was shown by Shank and Wehlau [99] that the invariants of the form $\mathrm{Tr}^{C_p}(y_1^{a_1} y_2^{a_2} \ldots y_m^{a_m})$ with $a_1 + a_2 + \cdots + a_m \le 2(p-1)$ are not required in a minimal generating set, nor in a SAGBI basis. However, none of the other invariants listed in Theorem 7.4.1 can be omitted.

Proof. We now begin the proof of Theorem 7.4.1. Let A denote the ring generated by the invariants given in the statement,

$$A := \mathbb{F}[x_i, \mathbf{N}(y_i), u_{ij}, \mathrm{Tr}(y^A) \mid 1 \le i < j \le m, 0 \le a_i < p] .$$

We need to show that $\mathbb{F}[m\,V_2]_d^{C_p} \subseteq A_d$ for all $d \ge 0$. We proceed by induction on d.

The case $d = 0$ is clear since $\mathbb{F}[m\,V_2]_0 = \mathbb{F}$. Now consider the general case $d > 0$. Take $f \in \mathbb{F}[m\,V_2]_d^{C_p}$ and write $f = f^\sharp + f^\flat$ where $f^\sharp = \sum_{i=1}^m f_i \mathbf{N}(y_i)$ with $f_i \in \mathbb{F}[m\,V_2]_{d-p}^{C_p}$. By induction, each of the $f_i \in A$ and thus $f^\sharp \in A$. Hence we may assume from now on that $f = f^\flat$. Suppose $f = f^\flat \in \mathbb{F}[m\,V_2]_{(d_1, d_2, \ldots, d_m)}^{C_p}$ where $d_1 + d_2 + \cdots + d_m = d$. If there exists i with $d_i \ge p$ then since $\deg_{y_i}(f) < p$ we see that x_i divides f. Hence $f = x_i f'$ where $f' \in \mathbb{F}[m\,V_2]_{d-1}^{C_p}$. Again, by induction, this implies that f' and hence f lies in A. Therefore, we may assume that $d_i < p$ for all $i = 1, 2, \ldots, m$.

We now recall the polarization and restitution operators defined in §1.9. We will take $f \in \mathbb{F}[m\,V_2]_d$ and consider its full polarization $\mathcal{P}(f) = f_{(1,1,\ldots,1)}$. We will also exploit the restitution map

$$\mathcal{R} = \mathcal{R}_{(d_1, d_2, \ldots, d_m)} : \mathbb{F}[d\,V_2]_{(1,1,\ldots,1)} \to \mathbb{F}[m\,V_2]_{(d_1, d_2, \ldots, d_m)}$$

defined by

$$\mathcal{R}(f) = f(\underbrace{\mathbf{v}_1, \ldots, \mathbf{v}_1}_{d_1}, \underbrace{\mathbf{v}_2, \ldots, \mathbf{v}_2}_{d_2}, \ldots, \underbrace{\mathbf{v}_m, \ldots, \mathbf{v}_m}_{d_m})$$

where $d = d_1 + d_2 + \cdots + d_m$.

We have $\mathcal{R}(\mathcal{P}(f)) = \lambda f$ where $\lambda = d_1!d_2! \cdots d_m!$. Since each $d_i < p$ we see that $f = \mathcal{R}(F)$ where $F = \mathcal{P}_{(1,1,\ldots,1)}(\lambda^{-1}f)) \in \mathbb{F}[d\,V_2]^{C_p}_{(1,1,\ldots,1)}$. We will show that $f \in A_d$ by considering the lead monomials of f and F.

We are using $\{x_j, y_j\}$ to denote the usual co-ordinate functions for the j^{th} copy of V_2 in $m\,V_2$. We have $\mathbb{F}[d\,V_2] = \mathbb{F}[d_1V_2 \oplus d_2V_2 \oplus \cdots \oplus d_mV_2]$. We let

$$\{x_{j1}, y_{j1}, x_{j2}, y_{j2}, \ldots, x_{jd_j}, y_{jd_j}\}$$

denote the co-ordinate functions on $d_j\,V_2$. We will use the graded reverse lexicographic order on $\mathbb{F}[m\,V_2]$ with $y_1 > x_1 > y_2 > x_2 > \cdots > y_m > x_m$. Similarly, we use the graded reverse lexicographic order on $d\,V_2$ determined by

$$y_{11} > x_{11} > y_{12} > x_{12} > \cdots > x_{1d_1} > y_{21} > \cdots > x_{dd_m}$$

Thus

$$x_{jk} < y_{jk}, x_{(j+1)k} < x_{jk}, y_{(j+1)k} < y_{jk}, x_{j(k+1)} < x_{jk} \text{ and } y_{j(k+1)} < y_{jk}.$$

With this term order we have the following lemma.

Lemma 7.4.3. *Let $d_1, d_2, \ldots, d_m$ be non-negative integers with $0 \le d_i < p$ for $i = 1, 2, \ldots, m$. Suppose $F \in \mathbb{F}[d\,V_2]^{C_p}_{(1,1,\ldots,1)}$ is a multi-linear invariant of degree $d = d_1 + d_2 + \cdots + d_m$. Then $\mathrm{LM}(F) \in \{\mathrm{LM}(\mathcal{P}(f) \mid f \in A_{(d_1,d_2,\ldots,d_m)}\}$.*

We defer the proof of Lemma 7.4.3 until §7.4.2. Instead, we show why Lemma 7.4.3 suffices to prove the invariants listed in Theorem 7.4.1 do form a generating set for the ring of invariants. To see this, assume by way of contradiction, that f is a homogeneous element of $\mathbb{F}[m\,V_2]^{C_p} \setminus A$. By the above, we have that $f = f^\flat$ is homogeneous of multi-degree $(d_1, d_2, \ldots, d_m)$ where $d_i < p$ for all $i = 1, 2, \ldots, m$. Suppose f is such that $\mathrm{LM}(\mathcal{P}(f))$ is minimal among those $f = f^\flat \in \mathbb{F}[m\,V_2]^{C_p} \setminus A$. Write $f = \mathcal{R}\mathcal{P}(\lambda^{-1}f)$ where $\lambda = d_1!d_2! \cdots d_m! \in \mathbb{F}$. By Lemma 7.4.3, there exists $h \in A_{(d_1,d_2,\ldots,d_m)}$ with $\mathrm{LT}(\mathcal{P}(h)) = \mathrm{LT}(\mathcal{P}(\lambda^{-1}f))$. Define $F' := \mathcal{P}(\lambda^{-1}f) - \mathcal{P}(h)$. Therefore, we have $\mathrm{LM}(F') = \mathrm{LM}(\mathcal{P}(\lambda^{-1}f - h)) < LM(\mathcal{P}(\lambda^{-1}f))$. By the choice of f, this implies that $\lambda^{-1}f - h \in A$. Thus $f = \lambda(\lambda^{-1}f - h) + \lambda h \in A$. This contradiction shows that the invariants listed in the statement of Theorem 7.4.1 do indeed form a generating set for the ring of invariants. □

We will prove that these invariants are not just a generating set but indeed a SAGBI basis at the end of §7.4.2.

7.4.1 Dyck Paths and Multi-Linear Invariants

Before we can prove Lemma 7.4.3, we need to describe the multi-linear C_p-invariants. This amounts to giving a description of the decomposition of the

C_p-module $\otimes^d V_2$ into indecomposable C_p-modules. Each such summand contains a fixed line and these are the multi-linear invariants we seek.

Define non-negative integers $\mu_p^d(k)$ by the direct sum decomposition of the C_p-module $\otimes^d V_2$ over $\mathbb{F}_p$:

$$\bigotimes^d V_2 \cong \bigoplus_{k=1}^{p} \mu_p^d(k)\, V_k \ .$$

We have the following lemma.

Lemma 7.4.4. *Let* $p \geq 3$. *Then*

$$\mu_p^0(k) = \delta_k^1 \textit{ and } \mu_p^1(k) = \delta_k^2,$$

and

$$\mu_p^{d+1}(k) = \begin{cases} \mu_p^d(2), & \textit{if } k = 1; \\ \mu_p^d(k-1) + \mu_p^d(k+1), & \textit{if } 2 \leq k \leq p-2; \\ \mu_p^d(p-2), & \textit{if } k = p-1; \\ \mu_p^d(p-1) + 2\mu_p^d(p), & \textit{if } k = p; \end{cases}$$

for $d \geq 1$.

Proof. The initial conditions are clear. The recursive conditions follow immediately from the three equations from Lemma 7.1.14:

$$\begin{aligned} V_1 \otimes V_2 &= V_2 \\ V_k \otimes V_2 &= V_{k-1} \oplus V_{k+1} \text{ for all } 2 \leq k \leq p-1 \\ V_p \otimes V_2 &= 2\, V_p. \end{aligned}$$

□

We now introduce certain combinatorial objects that we will use to describe the decomposition of $\otimes^d V_2$ into indecomposable summands. At the end of this section, we provide an example illustrating the correspondence between invariants and these combinatorial objects.

A *Dyck path* of length d is a lattice path in the first quadrant of the xy-plane from (0,0) to $(d, 0)$ comprised of a sequence of steps each of the form (1,1) or (1,-1). Steps of the form (1,1) are called *rises* and denoted by R. Steps of the form $(1, -1)$ are called *falls* and denoted by F.

A Dyck path is encoded by a *Dyck word* which is a sequence of R's and F's with an equal number of each. The condition that the path remain in the first quadrant, i.e., on or above the x-axis corresponds to the restriction on the word that among the first t symbols there are never more F's than R's for all $1 \leq t \leq d$.

If $\Gamma = \gamma_1\gamma_2 \ldots \gamma_d$ is a Dyck word, then for each t with $1 \leq t \leq d$, the initial portion of the word, $\gamma_1\gamma_2 \ldots \gamma_t$, is a *partial Dyck word* and the corresponding t steps are called a *partial Dyck path.*

The *path height* or *height* of a (partial) Dyck path is q if the path touches but does not cross the line $y = q$. We will say that a partial Dyck path from (0,0) to (t, h) has *finishing height* h.

An *initially Dyck path* of escape height q and length d is a partial Dyck path from (0,0) to (t, q) for some $t \leq d$ followed by an entirely arbitrary sequence of $d - t$ steps each step being a rise or a fall. Thus an initially Dyck path may wander after touching the line $y = q$ and in particular is not confined to the first quadrant. We denote the set of initially Dyck paths of escape height q and length d by IDP_q^d

We let $\mathrm{PDP}_{\leq q}^d$ denote the set of all partial Dyck paths of length d and height at most q. We write $\mathrm{PDP}_{\leq q}^d(h)$ to denote the set of partial Dyck paths of length d, height at most q and finishing height h.

We will abuse notation and terminology occasionally by identifying a Dyck path with its corresponding word and vice versa. We will denote the set of all finite words (without any restrictions) that can be formed using the alphabet {R,F} by {R,F}*.

Let $\nu_q^d(h) := |\mathrm{PDP}_{\leq q}^d(h)|$ for $1 \leq h \leq q$. We also define $\bar{\nu}_q^d := |\mathrm{IDP}_q^d|$. With this notation we have the following lemma.

Lemma 7.4.5. *Let $q \geq 2$. Then*

$$\nu_q^0(h) = \delta_h^0 \textit{ and } \nu_q^1(h) = \delta_h^1 \ ,$$

$$\bar{\nu}_q^0(h) = 0 \textit{ and } \bar{\nu}_q^1(h) = 0 \ ,$$

and

$$\nu_q^{d+1}(h) = \begin{cases} \nu_q^d(1), & \textit{if } h = 0; \\ \nu_q^d(h-1) + \nu_q^d(h+1), & \textit{if } 1 \leq h \leq q-1; \\ \nu_q^d(q-1), & \textit{if } h = q; \end{cases}$$

and

$$\bar{\nu}_q^{d+1} = \nu_{q-1}^d(q-1) + 2\bar{\nu}_q^d$$

for all $d \geq 1$.

Proof. All of these equations are easily seen to hold except perhaps the final one. Its left-hand term $\bar{\nu}_q^{d+1} = |\mathrm{IDP}_q^{d+1}|$ is the number of initial Dyck paths of length $d+1$ and escape height q. We divide such paths into two classes: those which first achieve height q on their final step and those which achieve height q sometime during the first d steps. Paths in the first class are partial Dyck paths of length d, height at most $q-1$ and finishing height $q-1$ followed by a rise on the $(d+1)^{\mathrm{st}}$ step. There are $\nu_{q-1}^d(q-1) = |\mathrm{PDP}_{\leq q-1}^d(q-1)|$ such paths. The second class consists of initial Dyck paths of escape height q and length d followed by a final step which may be either a rise or a fall. Clearly, there are $2|\mathrm{IDP}_q^d| = 2\bar{\nu}_q^d$ paths of this kind. □

Corollary 7.4.6. *For all $d \in \mathbb{N}$, all primes $p \geq 2$ and all $k = 1, 2, \ldots, p-1$ we have*

$$\mu_p^d(k) = \nu_{p-2}^d(k-1) \quad \textit{and} \quad \mu_p^d(p) = \bar{\nu}_{p-1}^d \ .$$

Proof. Comparing the recursive expressions and initial conditions for $\mu_p^d(k)$ and $\nu_{p-2}^d(k-1)$ and for $\mu_p^d(p)$ and $\bar{\nu}_{p-1}^d$ given in the previous two lemmas makes the result clear for $p \geq 5$.

For $p = 2$, it is easy to see that $\mu_2^d(1) = \nu_0^d(0) = \delta_d^0$ for $d \geq 0$ and $\mu_2^d(2) = 2^{d-1} = \bar{\nu}_1^d$ for $d \geq 1$.

For $p = 3$ and $k = 1, 2$, we have $\mu_3^d(k) = \nu_1^d(k-1) = \begin{cases} 1, & \text{if } k+d \text{ is odd;} \\ 0, & \text{if } k+d \text{ is even.} \end{cases}$

Hence $\mu_3^d(3) = \lfloor \frac{2^d-1}{3} \rfloor$ for $d \geq 0$. From the recursive relation $\bar{\nu}_2^{d+1} = \nu_1^d(1) + 2\bar{\nu}_2^d$ we see that $\bar{\nu}_2^d = \lfloor \frac{2^d-1}{3} \rfloor = \mu_3^d(3)$. □

Given an arbitrary word $\Gamma = \gamma_1\gamma_2 \cdots \gamma_t \in \{\mathrm{R},\mathrm{F}\}^*$ we define a multi-linear monomial $\Lambda(\Gamma) = z_1 z_2 \cdots z_t \in \mathbb{F}[t\,V_2]$ by taking $z_i = x_i$ if $\gamma_i = \mathrm{R}$ and $z_i = y_i$ if $\gamma_i = \mathrm{F}$.

We will show that for $k = 1, 2, \ldots, p-1$, each element of $\Lambda(\mathrm{PDP}_{\leq p-2}^d(k))$ is the lead term of an invariant which spans the fixed line of a summand isomorphic to V_k of $\otimes^d V_2$. Further, we will show that each element of $\Lambda(\mathrm{IDP}_{p-1}^d)$ is the lead term of an invariant in the image of the transfer.

Consider a partial Dyck path of length d, height at most $p-1$ and finishing height k. Let $\Gamma = \gamma_1\gamma_2 \cdots \gamma_d$ denote the corresponding partial Dyck word. We wish to match each of the falls γ_j with a rise $\gamma_{\pi(j)}$. We do this recursively as follows. Choose the smallest value of j such that γ_j is a fall which we have not yet matched with a rise. Let i be maximal such that $i < j$, $\gamma_i = \mathrm{R}$ and $i \neq \pi(\ell)$ for all $\ell < j$. Then we declare that $\pi(j) = i$. Let $I_1 := \{j \mid \gamma_j = \mathrm{F}\}$, $I_2 := \pi(I_1)$ and $I_3 := \{1, 2, \ldots, d\} \setminus (I_1 \cup I_2)$. Note that $|I_3| = k$ and $\gamma_i = \mathrm{R}$ for all $i \in I_3$. We define $\theta(\Gamma) = \left(\prod_{j \in I_1} u_{\pi(j),j}\right) \prod_{i \in I_3} x_i$. We also define $\theta'(\Gamma) = \left(\prod_{j \in I_1} u_{\pi(j),j}\right) \prod_{i \in I_3} y_i$.

This construction associates to each partial Dyck path of length d, height at most $p-1$ and finishing height k a multi-linear invariant $\theta(\Gamma) \in \mathbb{F}[d\,V_2]^{C_p}$. Also,

$$\begin{aligned} \mathrm{LM}(\theta(\Gamma)) &= \left(\prod_{j \in I_1} \mathrm{LM}(u_{\pi(j),i})\right) \prod_{i \in I_3} x_i \\ &= \left(\prod_{j \in I_1} x_{\pi(j)} y_j\right) \prod_{i \in I_3} x_i \\ &= \Lambda(\Gamma). \end{aligned}$$

Next, we suppose $\Gamma = \gamma_1\gamma_2 \cdots \gamma_d$ is a word corresponding to a initially Dyck path of escape height $p-1$ and length d. Then there exists t with $1 \leq$

$t \le d$ such that $\Gamma' := \gamma_1\gamma_2\cdots\gamma_t$ corresponds to a partial Dyck path of finishing height $p-1$. Let t be minimal with this property and let $\Gamma' := \gamma_1\gamma_2\cdots\gamma_t$. We define $\theta(\Gamma) := \mathrm{Tr}(\theta'(\Gamma')\prod_{i=t+1}^d z_i)$ where $z_i = x_i$ if $\gamma_i = \mathrm{R}$ and $z_i = y_i$ if $\gamma_i = \mathrm{F}$.

Thus we have associated a multi-linear invariant $\theta(\Gamma) \in \mathrm{Tr}^{C_p}(\mathbb{F}[d\,V_2])$ to each initially Dyck path of escape height $p-1$ and length d. We have

$$\mathrm{Tr}^{C_p}(\theta'(\Gamma')) = \left(\prod_{j\in I_1} u_{\pi(j),j}\right)\mathrm{Tr}^{C_p}(\prod_{k\in I_3} y_k) = \left(\prod_{j\in I_1} u_{\pi(j),j}\right)\prod_{k\in I_3} x_k$$

since $|I_3| = p-1$. Thus

$$\mathrm{LM}(\mathrm{Tr}^{C_p}(\theta'(\Gamma'))) = \left(\prod_{j\in I_1}\mathrm{LM}(u_{\pi(j),j})\right)\prod_{k\in I_3} x_k = \left(\prod_{j\in I_1} x_{\pi(j)}y_j\right)\prod_{k\in I_3} x_k\ .$$

Therefore,

$$\mathrm{LM}(\mathrm{Tr}^{C_p}(\theta(\Gamma))) = \left(\prod_{j\in I_1} x_{\pi(j)}y_j\right)\prod_{k\in I_3} x_k\prod_{i=t+1}^{m} z_i = \Lambda(\Gamma)$$

where $z_i = x_i$ if $\gamma_i = \mathrm{R}$ and $z_i = y_i$ if $\gamma_i = \mathrm{F}$.

We have

Proposition 7.4.7. *Let $\gamma \in PDP^d_{\le p-2} \sqcup IDP^d_{p-1}$. Then*

1. $\mathrm{LM}(\theta(\gamma)) = \Lambda(\gamma)$.
2. *If $\gamma \in PDP^d_{\le p-2}(h)$, then the invariant $\theta(\gamma)$ lies in $\mathbb{F}[d\,V_2]^{C_p}_{(1,1,\ldots,1)} \cong (\otimes^d V_2)^{C_p}$ and has length $h+1$ for all $0 \le h \le p-2$.*
3. *If $\gamma \in IDP^d_{p-1}$, then the invariant $\theta(\gamma)$ lies in $\mathbb{F}[d\,V_2]^{C_p}_{(1,1,\ldots,1)} \cong (\otimes^d V_2)^{C_p}$ and has length p.*
4. *The set $\theta\Big(PDP^d_{\le p-2} \sqcup IDP^d_{p-1}\Big)$ is a basis for $\mathbb{F}[d\,V_2]^{C_p}_{(1,1,\ldots,1)} \cong (\otimes^d V_2)^{C_p}$.*

Proof. We have already seen that $\mathrm{LM}(\theta(\gamma)) = \Lambda(\gamma)$ and that $\Lambda(\gamma) \in \mathbb{F}[d\,V_2]^{C_p}_{(1,1,\ldots,1)}$. If $\gamma \in \mathrm{PDP}^d_{\le p-2}(h)$, then

$$\begin{aligned}\Delta^{h-1}(\theta'(\Gamma)) &= \Delta^{h-1}\Big(\prod_{j\in I_1} u_{\pi(j),j}\prod_{i\in I_3} y_i\Big) = \Big(\prod_{j\in I_1} u_{\pi(j),j}\Big)\Delta^{h-1}\Big(\prod_{i\in I_3} y_i\Big)\\ &= \Big(\prod_{j\in I_1} u_{\pi(j),j}\Big)(h-1)!\prod_{i\in I_3} x_i = (h-1)!\,\theta(\Gamma).\end{aligned}$$

Therefore, $\ell(\theta(\Gamma)) \ge h$.

Since the lead terms of the invariants in the image of θ are distinct, we see that the image of θ is a linearly independent set. This shows that the image of θ is a basis for $(\otimes^d V_2)^{C_p}$ since $|\sqcup_{h=0}^{p-2}\mathrm{PDP}^d_{\leq p-2}(h) \;\sqcup\; \mathrm{IDP}^d_{p-1}| = \sum_{h=0}^{p-2}\nu^d_{p-2}(h) + \bar{\nu}^d_{p-1} = \sum_{h=1}^{p-1}\mu^d_p(h) + \mu^d_p(p) = \dim_{\mathbb{F}}(\otimes^d V_2)^{C_p}$. Furthermore, this equality shows that $\theta(\mathrm{IDP}^d_{p-1})$ is a basis for $\mathrm{Tr}^{C_p}(\otimes^d V_2)$ and that each element of $\theta(\mathrm{PDP}^d_{\leq p-2}(h))$ spans the fixed line of one summand of $\otimes^d V_2$ isomorphic to V_h. □

7.4.2 Proof of Lemma 7.4.3

In this section we combine some of the above results in order to prove Lemma 7.4.3. After giving this proof we also prove the final assertion of Theorem 7.4.1.

We maintain the notation of the previous two sections.

We will use the following lemma which is easy to prove.

Lemma 7.4.8. *Suppose γ_1 and γ_2 are two monomials in $\mathbb{F}[m\,V_2]_{(d_1,d_2,\ldots,d_m)}$ with $\gamma_1 > \gamma_2$. Then $\mathrm{LM}(\mathcal{P}(\gamma_1)) > \mathrm{LM}(\mathcal{P}(\gamma_2))$.*

For ease of notation, we will write $\mathrm{Tr}^{C_p}(\mathbf{y}^{\mathbf{a}}) = \mathrm{Tr}^{C_p}(y_{11}^{a_{11}} y_{12}^{a_{12}} \cdots y_{md_m}^{a_{md_m}})$ where $\mathbf{a} = (a_{11}, a_{12}, \ldots, a_{md_m})$.

Proof (of Lemma 7.4.3).

Suppose $F \in \mathbb{F}[d\,V_2]^{C_p}_{(1,1,\ldots,1)}$ is a multi-linear invariant of degree d.

By Proposition 7.4.7, we may write

$$\mathrm{LM}(F) = \prod_{(i,j)\in B_1} x_{ij} \prod_{(i,j,k,\ell)\in B_2} \mathrm{LM}(u_{ij,k\ell}) \prod_{\mathbf{a}\in B_3} \mathrm{LM}(\mathrm{Tr}^{C_p}(\mathbf{y}^{\mathbf{a}}))$$

for some index sets B_1, B_2 and B_3. Define

$$f := \prod_{(i,j)\in B_1} \mathcal{R}(x_{ij}) \prod_{(i,j,k,\ell)\in B_2} \mathcal{R}(u_{ij,k\ell}) \prod_{\mathbf{a}\in B_3} \mathcal{R}(\mathrm{Tr}^{C_p}(\mathbf{y}^{\mathbf{a}}))\ .$$

Note that $\mathcal{R}(x_{ij}) = x_i$, $\mathcal{R}(u_{ij,k\ell}) = u_{i,k}$ (or 0 if $i = k$). Also, since $\mathcal{R}$ is C_p-equivariant,

$$\begin{aligned}\mathcal{R}(\mathrm{Tr}^{C_p}(y_{11}^{a_{11}} \cdots y_{md_m}^{a_{md_m}})) &= \mathcal{R}\left(\sum_{k=0}^{p-1} \sigma^k (y_{11}^{a_{11}} \cdots y_{md_m}^{a_{md_m}})\right)\\ &= \sum_{k=0}^{p-1} \sigma^k \mathcal{R}(y_{11}^{a_{11}} \cdots y_{md_m}^{a_{md_m}})\\ &= \mathrm{Tr}^{C_p}(\mathcal{R}(y_{11}^{a_{11}} \cdots y_{md_m}^{a_{md_m}})\\ &= \mathrm{Tr}^{C_p}(y_1^{\alpha_1} y_2^{\alpha_2} \cdots y_m^{\alpha_m})\end{aligned}$$

where $\alpha_i = \sum_{j=1}^{d_i} a_{ij} \leq d_i < p$ for $1 \leq i \leq m$. Thus $f \in A$.

Let $\Gamma_1 := \mathrm{LM}(F)$. By Lemma 7.4.8, $\Gamma_1 = \mathrm{LM}(\mathcal{P}(\gamma_1)) = \mathrm{LM}(\mathcal{P}(f))$ where

$$\gamma_1 = \mathcal{R}(\varGamma_1) = \prod_{(i,j)\in B_1} x_i \prod_{(i,j,k,\ell)\in B_2} u_{i,k} \prod_{\mathbf{a}\in B_3} \mathrm{Tr}^{C_p}(\mathcal{R}(\mathbf{y}^{\mathbf{a}})) = \mathrm{LM}(f)\ .$$

Hence $f \in A$ with $\mathrm{LM}(F) = \mathrm{LM}(\mathcal{P}(f))$ as required. □

Proof (of the second assertion of Theorem 7.4.1). We need to prove that the invariants listed in Theorem 7.4.1 are in fact a SAGBI basis for the ring of invariants. Let $f \in \mathbb{F}[m\,V_2]^{C_p}$. Dividing f by each $\mathbf{N}(y_i)$ we may reduce to the case where $f = f^\flat$. Furthermore, if f is divisible by any x_i, we may replace f by f/x_i. Thus we may reduce to the case where $f = f^\flat \in \mathbb{F}[m\,V_2]^{C_p}_{(d_1,d_2,\ldots,d_m)}$ with $0 \le d_i < p$ for all $i = 1, 2, \ldots, m$.

Next, we note that the above proof demonstrates something stronger than the statement given in Lemma 7.4.3. The proof shows that for every $f \in \mathbb{F}[m\,V_2]^{C_p}_{(d_1,d_2,\ldots,d_m)}$, the lead monomial $\mathrm{LM}(\mathcal{P}(f))$ lies in the algebra generated by $\{\mathrm{LM}(\mathcal{P}(x_i) \mid 1 \le i \le m\} \cup \{\mathrm{LM}(\mathcal{P}(u_{ij})) \mid 1 \le i < j \le m\} \cup \{\mathrm{LM}(\mathcal{P}(\mathrm{Tr}^{C_p}(y_1^{a_1} y_2^{a_2} \cdots y_m^{a_m}))) \mid 0 \le a_i < p \text{ for } i = 1, 2, \ldots, m\}$. From this it follows immediately that the invariants listed in Theorem 7.4.1 form a SAGBI basis. This completes the proof of Theorem 7.4.1. □

We will now give an example to illustrate the ideas in the above proof of Theorem 7.4.1.

Example 7.4.9. We take $p = 5$ and $m = 3$ and consider invariants of multidegree $(1, 1, 2)$. In symbols, we are considering $\mathbb{F}[3\,V_2]^{C_5}_{(1,1,2)}$. We have the full polarization $\mathcal{P} : \mathbb{F}[3\,V_2]_{(1,1,2)} \to \mathbb{F}[4\,V_2]_{(1,1,1,1)}$. To avoid double subscripts, we will abuse notation by using the bases $\{x_1, y_1, x_2, y_2, x_3, y_3\}$ of $(3\,V_2)^*$ and $\{x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4\}$ of $(4\,V_2)^*$. Here $\mathcal{P}(z) = z$ for $z \in \{x_1, y_1, x_2, y_2\}$ and $\mathcal{P}(x_3) = x_3 + x_4$ and $\mathcal{P}(y_3) = y_3 + y_4$. Accordingly, $\mathcal{R}(z) = z$ for $z \in \{x_1, y_1, x_2, y_2, x_3, y_3\}$ and $\mathcal{R}(x_4) = x_3$ and $\mathcal{R}(y_4) = y_3$.

This full polarization equivariantly embeds $\mathbb{F}[3\,V_2]_{(1,1,2)} \cong V_2 \otimes V_2 \otimes S^2(V_2) \cong V_2 \otimes V_2 \otimes V_3$ into $V_2 \otimes V_2 \otimes V_2 \otimes V_2$. The image of this embedding is spanned by the 12 elements $z_1 z_2 z_2$ where $z_1 \in \{x_1, y_1\}$, $z_2 \in \{x_2, y_2\}$ and $z_3 \in \{x_3 x_4, x_3 y_4 + y_3 x_4, y_3 y_4\}$.

By Lemma 7.4.5, we see that there are 5 partial Dyck paths of length $d = 4$ and height at most $p - 2 = 3$. The partial Dyck words corresponding to these 5 paths are $\varGamma_1 = \mathrm{RFRF}$, $\varGamma_2 = \mathrm{RRFF}$, $\varGamma_3 = \mathrm{RRRF}$, $\varGamma_4 = \mathrm{RRFR}$ and $\varGamma_5 = \mathrm{RFRR}$. Furthermore, there is one word corresponding to an initially Dyck path of length $d = 4$ and escape height $p - 1 = 4$. This is the word $\varGamma_6 = \mathrm{RRRR}$. The paths corresponding to the first 5 words have finishing heights 0,0,2, 2 and 2 respectively. Of course, the final word, $\varGamma_6$, corresponds to a path which achieves height $p-1 = 4$. This means that $\otimes^4 V_2 \cong 2\,V_1 \oplus 3\,V_3 \oplus V_5$. This decomposition can also be found using Lemma 7.4.4.

Applying $\varLambda$, we get $\varLambda(\varGamma_1) = x_1 y_2 x_3 y_4$, $\varLambda(\varGamma_2) = x_1 x_2 y_3 y_4$, $\varLambda(\varGamma_3) = x_1 x_2 x_3 y_4$, $\varLambda(\varGamma_4) = x_1 x_2 y_3 x_4$, $\varLambda(\varGamma_5) = x_1 y_2 x_3 x_4$, and $\varLambda(\varGamma_6) = x_1 x_2 x_3 x_4$. We recognize these 6 monomials as the lead terms of the 6 invariants: $F_1 := \theta(\varGamma_1) = u_{12} u_{34}$, $F_2 := \theta(\varGamma_2) = u_{14} u_{23}$, $F_3 := \theta(\varGamma_3) = x_1 x_2 u_{34}$,

$F_4 := \theta(\Gamma_4) = x_1u_{23}x_4$, $F_5 := \theta(\Gamma_5) = u_{12}x_3x_4$, and $F_6 := \theta(\Gamma_6) = x_1x_2x_3x_4 = -\operatorname{Tr}(y_1y_2y_3y_4)$.

Note that $u_{13}u_{24} = u_{12}u_{34} + u_{14}u_{23}$, $x_2u_{13}x_4 = x_1u_{23}x_4 + x_3u_{12}x_4$ and $x_1x_3u_{24} = x_1x_2u_{34} + x_1x_4u_{23}$ are not required as basis elements.

Restituting these 6 invariants we get $\mathcal{R}(F_1) = 0$, $f_2 := \mathcal{R}(F_2) = u_{13}u_{23}$, $\mathcal{R}(F_3) = 0$, $f_4 := \mathcal{R}(F_4) = x_1u_{23}x_3$, $f_5 := \mathcal{R}(F_5) = u_{12}x_3^2$, and $f_6 := \mathcal{R}(F_6) = x_1x_2x_3^2 = -\operatorname{Tr}(y_1y_2y_3^2)$. The two invariants F_1 and F_3 restitute to 0 since these two invariants do not lie in $\mathcal{P}(\mathbb{F}[3\,V_2]_{(1,1,2)})$.

Note that $f_4 = \Delta^2(\frac{1}{2}y_1u_{23}y_3)$ and $f_5 = \Delta^2(\frac{1}{2}u_{12}y_3^2)$ and thus f_4 and f_5 both lie inside copies of V_3. Also, $f_6 = \Delta^4(-y_1y_2y_3^2)$ spans the fixed line in a copy of V_5. Therefore, we see that $\{f_2, f_4, f_5, f_6\}$ forms a basis for $\mathbb{F}[3\,V_2]^{C_5}_{(1,1,2)} \cong (V_2 \otimes V_2 \otimes S^2(V_2))^{C_5} \cong (V_2 \otimes V_2 \otimes V_3)^{C_5} \cong (V_1 \oplus 2\,V_3 \oplus V_5)^{C_5}$.

7.5 Integral Invariants

In this section, we describe natural invariants in characteristic 0 which map to elements of $\mathbb{F}[V_n]^{C_p}$ under mod p reduction. Much of this discussion follows the paper of Shank [98].

Consider

$$\sigma = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 1 & 1 \end{pmatrix}$$

an $n \times n$ matrix with entries in the field of rational numbers $\mathbb{Q}$. Then σ generates an infinite cyclic subgroup of $SL(n, \mathbb{Q})$. We denote this subgroup by $\mathbb{Z}$. Let W_n denote the n dimensional $\mathbb{Q}$ vector space on which σ and hence $\mathbb{Z}$ acts. Let $\{x_1, x_2, \dots, x_n\}$ denote the basis of W_n^* dual to the standard basis.. We consider the ring of invariants $\mathbb{Q}[W_n] = \mathbb{Q}[x_1, \dots, x_n]^{\mathbb{Z}}$. Note that since $\mathbb{Z}$ is infinite, we have no reason to expect that this ring of invariants is finitely generated. Surprisingly, it turns out to be so.

Remark 7.5.1. By way of explaining this latter remark, consider the group $H \cong \mathbb{Z}$ generated by

$$\sigma = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

so that

$$H = \left\{ \begin{bmatrix} 1 & 0 \\ i & 1 \end{bmatrix} \mid i \in \mathbb{Z} \right\}.$$

This group sits naturally in $\mathrm{SL}_2(\mathbb{C})$. The Zariski closure of H is the group

$$G_a(\mathbb{C}) = \left\{ \begin{bmatrix} 1 & 0 \\ z & 1 \end{bmatrix} \mid z \in \mathbb{C} \right\}.$$

Here we use the notation $G_a(\mathbb{C})$ to emphasize that this group is isomorphic to the group of complex numbers under addition. Let V be any representation of $\mathrm{SL}_2(\mathbb{C})$ and let $\mathbb{C}^2$ denote the standard 2-dimensional representation of $\mathrm{SL}_2(\mathbb{C})$ with basis $\{e_1, e_2\}$. Consider the algebra map

$$\rho : \mathbb{C}[V \oplus \mathbb{C}^2] \to \mathbb{C}[V]$$

taking f to the function $\rho(f)$ given by $\rho(f)(v) = f(v, e_2)$. We note immediately that $G_a(\mathbb{C}) = \{\tau \in \mathrm{SL}_2(\mathbb{C}) \mid \tau(e_2) = e_2\}$. Roberts, [94], proved that the map ρ restricted as follows

$$\mathbb{C}[V \oplus \mathbb{C}^2]^{\mathrm{SL}_2(\mathbb{C})} \to \mathbb{C}[V]^{G_a(\mathbb{C})}$$

is an isomorphism of algebras. Since the former ring of invariants is known to be finitely generated by Hilbert's famous result, so also is the latter. For the case needed here where $G = \mathrm{SL}_2(\mathbb{C})$, the finite generation of the ring of invariants was first proved by Gordan [48] or [49]. It is also true that $\mathbb{Q}[V]^{\mathbb{Z}} = \mathbb{Q}[V]^{G_a(\mathbb{Q})}$.

Famously, there is a unique n-dimensional irreducible representation of $\mathrm{SL}_2(\mathbb{C})$ which can be given by the natural action of $\mathrm{SL}_2(\mathbb{C})$ on

$$\left\{X^{n-1}, X^{n-2}Y, \ldots, Y^{n-1}\right\},$$

where $(\mathbb{C}^2)^*$ has basis $\{X, Y\}$. This identification provides a map from $\mathrm{SL}_2(\mathbb{C})$ to $\mathrm{GL}(W_n \otimes \mathbb{C})$. Under this map and with an appropriate choice of basis, the matrix

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

is carried to

$$\begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ 1 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & 0 & \ldots & 1 & 1 \end{pmatrix}.$$

There are excellent references for this material: the primer of Kraft and Procesi [70], and the book of Procesi [90].

Given an element f of $\mathbb{Q}[W_n]^{\mathbb{Z}}$, we may clear denominators by multiplying by an appropriate integer m to obtain $f' = mf \in \mathbb{Z}[x_1, \ldots, x_n]^{\mathbb{Z}}$. Then reduction modulo p gives a map from $\mathbb{Z}[x_1, \ldots, x_n]^{\mathbb{Z}}$ to $\mathbb{F}_p[x_1, \ldots, x_n]^{C_p} \subseteq \mathbb{F}[x_1, \ldots, x_n]^{C_p} = \mathbb{F}[V_n]^{C_p}$. Any element of $\mathbb{F}[V_n]^{C_p}$ that may be constructed in this manner is called a *integral invariant*. Note the abuse of notation here in which $\{x_1, \ldots, x_n\}$ is used to denote the dual bases over both $\mathbb{Q}$ and $\mathbb{F}_p$ and σ is used to represent a generator of both $\mathbb{Z}$ and C_p.

Classical invariant theorists have long recognized the difficulty of describing $\mathbb{C}[V \oplus \mathbb{C}^2]^{\mathrm{SL}_2(\mathbb{C})}$. Even today, the description is complete for only a handful of representations. The material above indicates that describing the integral invariants of $\mathbb{F}_p[V_n]^{C_p}$ is equivalent.

The invariant $x_1 \in \mathbb{Q}[V_n]^{\mathbb{Z}}$ is associated to the invariant $x_1 \in \mathbb{F}[V_n]^{C_p}$ for every field $\mathbb{F}$ of characteristic p and representation V_n of C_p for $n \geq 1$. Similarly,

$$x_2^2 - 2x_1x_3 - x_1x_2 \in \mathbb{Q}[W_n]^{\mathbb{Z}}$$

is associated to an infinite family of invariants $x_2^2 - 2x_1x_3 - x_1x_2 \in \mathbb{F}[V_n]^{C_p}$ for every field $\mathbb{F}$ of characteristic p and representation V_n of C_p for $n \geq 3$.

Shank, [97] and [98], conjectured that every ring of invariants of a C_p-representation, $\mathbb{F}[V_n]^{C_p}$, is generated by the three classes of invariants: norms, transfers and integral invariants. By norms here we mean only the norm $\mathrm{N}(x_n)$ of the distinguished variables $x_n \in V_n^*$.

In order to discuss integral invariants comprehensively, we introduce the (non-finitely generated) ring $\mathbb{Q}[V_\infty] := \mathbb{Q}[x_1, x_2, x_3, \ldots]$ over $\mathbb{Q}$ having infinitely many generators of degree 1. Of course, for any $n \in \mathbb{N}$, we have the surjective map $\theta : \mathbb{Q}[V_\infty] \to \mathbb{Q}[V_n]$ given by

$$\theta(x_i) = \begin{cases} x_i & \text{if } i \leq n, \\ 0 & \text{if } i > n. \end{cases}$$

The integers, $\mathbb{Z}$, with generator σ, act on $\mathbb{Q}[V_\infty]$ via the left action generated by $\sigma(x_i) = \begin{cases} x_i + x_{i-1} & \text{if } i \geq 2, \\ x_1 & \text{if } i = 1. \end{cases}$ As usual, we will use Δ to denote the operator $\sigma - \mathrm{Id}$.

We will equip $\mathbb{Q}[V_\infty]$ with the graded lexicographic ordering with $x_1 < x_2 < x_3 < \ldots$.

Lemma 7.5.2. *Suppose $n \geq m > 1$. Let $\beta \in \mathbb{K}[V_{m-1}]$ be a monomial and write $\beta := x_1^{i_1} x_2^{i_2} \cdots x_{m-1}^{i_{m-1}}$. Then βx_{m-1} and βx_m are consecutive monomials of $\mathbb{K}[V_n]$ in the monomial ordering, i.e., if γ is a monomial of $\mathbb{K}[V_n]$ with $\beta x_{m-1} < \gamma \leq \beta x_m$, then $\gamma = \beta x_m$.*

Proof. Write $\gamma = x_1^{j_1} x_2^{j_2} \cdots x_n^{j_n}$ and let a denote the least integer such that $i_a + \delta_a^{m-1} \neq j_a$. Since $\beta x_{m-1} < \gamma \leq \beta x_m$, we must have $\deg(\gamma) = \deg(\beta) + 1$ and thus $\beta x_{m-1} < \gamma$ implies that $i_a + \delta_a^{m-1} > j_a$. If $a < m-1$, then $i_a > b_a$ would imply that $\beta x_m < \gamma$, contradicting the hypotheses. If $a > m-1$, then $i_{m-1} + 1 = j_{m-1}$ would imply that $\gamma > \beta x_m$, again contradicting the hypothesis. Therefore, $a = m-1$ and we have $i_1 = j_1, i_2 = j_2, \ldots, i_{m-2} = j_{m-2}$ and $i_{m-1} + 1 \neq j_{m-1}$.

From, $\beta x_{m-1} < \gamma$ we see $j_{m-1} < i_{m-1} + 1$. Conversely, from $\gamma \leq \beta x_m$, we see that $j_{m-1} \geq i_{m-1}$. Therefore, $j_{m-1} = i_{m-1}$. Finally, from $\gamma \leq \beta x_m$, we conclude that $j_m \geq 1$ and since $\deg(\gamma) = \deg(\beta) + 1$, we must have $j_m = 1$ and $\gamma = \beta x_m$. □

Lemma 7.5.3. *Let f be any non-zero polynomial in $\mathbb{K}[x_1, x_2, \ldots, x_n]$. Then $\mathrm{LT}(\Delta(f)) < \mathrm{LT}(f)$.*

Proof. It is easy to see that $\mathrm{LT}(\Delta(\beta)) < \beta$ for any monomial β. Applying this to each term in f we see that $\mathrm{LT}(\Delta(f)) < \mathrm{LT}(f)$. □

Theorem 7.5.4. *Suppose that $n \geq m > 1$ and that β is a monomial in $\mathbb{K}[x_1, x_2, \ldots, x_{m-1}]$. Then $\beta x_m \notin \mathrm{LT}(\mathbb{K}[x_1, x_2, \ldots, x_n]^{\mathbb{Z}})$.*

Proof. Suppose that $f \in \mathbb{K}[x_1, x_2, \ldots, x_n]$ with $\mathrm{LT}(f) = \beta x_m$. Write $f = \beta x_m + f'$. Then $\mathrm{LT}(f') < \beta x_m$, i.e, by Lemma 7.5.2, $\mathrm{LT}(f') \leq \beta x_{m-1}$. Now $\Delta(f) = \beta x_{m-1} + \Delta(\beta) x_m + \Delta(\beta) x_{m-1} + \Delta(f')$. We will show that $\mathrm{LT}(\Delta(f)) = \beta x_{m-1} \neq 0$ and thus $f \notin \mathbb{K}[x_1, x_2, \ldots, x_n]^{\mathbb{Z}}$.

First, we see that $\mathrm{LT}(\Delta(f')) < \mathrm{LT}(f') \leq \beta x_{m-1}$. Secondly, $\Delta(\beta) < \beta$ implies that $\Delta(\beta) x_{m-1} < \beta x_{m-1}$. Finally, $\mathrm{LT}(\Delta(\beta)) < \beta$ implies that $\mathrm{LT}(\Delta(\beta) x_m) < \beta x_m$ and thus by Lemma 7.5.2 we get

$$\mathrm{LT}(\beta) x_m = \mathrm{LT}(\Delta(\beta) x_m) \leq \beta x_{m-1}.$$

Since x_m does not divide βx_{m-1}, we see that $\mathrm{LT}(\Delta(\beta) x_m) \neq \beta x_{m-1}$ and therefore, $\mathrm{LT}(\Delta(\beta) x_m) < \beta x_{m-1}$. Thus $\mathrm{LT}(\Delta(f)) = \beta x_{m-1} \neq 0$ as claimed. □

In contrast to Theorem 7.5.4, we have the following.

Theorem 7.5.5. *Consider a monomial*

$$\beta = x_1^{a_1} x_2^{a_2} \cdots x_{m-1}^{a_m - 1} \in \mathbb{F}[x_1, x_2, \ldots, x_{m-1}].$$

Let $a_m \geq 2$ and $n \geq 1 + \sum_{\ell=1}^{m} a_\ell(\ell - 1)$. Then there exists $f \in \mathbb{K}[x_1, x_2, \ldots, x_n]^{\mathbb{Z}}$ with $\mathrm{LT}(f) = \beta x_m^{a_m}$.

Proof. Suppose first that either $\mathbb{K}$ has characteristic 0 or that $a_m < p :=$ characteristic($\mathbb{K}$). Then

$$\mathrm{LT}(\Delta(\beta x_m^{a_m})) = a_m \beta x_{m-1} x_m^{a_m - 1}.$$

We define

$$f_1 := \beta x_m^{a_m} - a_m \beta x_{m-1} x_m^{a_m - 2} x_{m+1}$$

and note that

$$\mathrm{LT}(\Delta(f_1)) < a_m \beta x_{m-1} x_m^{a_m - 1}.$$

For $j \geq 2$, if $\Delta(f_{j-1}) \neq 0$, then we will define f_j inductively from f_{j-1} as follows. Write

$$\mathrm{LT}(\Delta(f_{j-1})) = \gamma x_r^k$$

where $k \geq 1$ and $\gamma \in \mathbb{K}[x_1, x_2, \ldots, x_{r-1}]$. Then define

$$f_j := f_{j-1} - \gamma x_r^{k-1} x_{r+1}.$$

Thus

$$\mathrm{LT}(\Delta(f_j)) < \gamma x_r^k = \mathrm{LT}(\Delta(f_{j-1}))$$

and therefore, the sequence $\mathrm{LT}(\Delta(f_1))$, $\mathrm{LT}(\Delta(f_2)), \ldots$ is strictly decreasing. However, this may not guarantee that the algorithm terminates since the number of variables involved is increasing at the same time. However, since

$$\mathrm{LT}(f_j) = \beta x_m^{a_m}$$

for all j, all that remains to show is that the algorithm terminates with some $f_j \in \mathbb{K}[x_1, x_2, \ldots, x_n]^{\mathbb{Z}}$.

To see that the algorithm does indeed terminate and that the final invariant uses at most n variables, we introduce the concept of the weight of a monomial. A monomial

$$\lambda = x_1^{b_1} x_2^{b_2} \cdots x_s^{b_s}$$

has weight, $\mathrm{wt}(\lambda) := \sum_{\ell=1}^{s} \ell b_\ell$. It is easy to see that every monomial appearing in $\Delta(\lambda)$ has weight strictly smaller than $\mathrm{wt}(\lambda)$. Also, note that $\mathrm{wt}(\gamma x_r^{k-1} x_{r+1}) = \mathrm{wt}(\gamma x_r^k) + 1$. Thus every monomial occurring in each f_j has weight at most $\mathrm{wt}(\beta x_m^{a_m})$. Furthermore, each of these monomials also has degree $\sum_{\ell=1}^{m} a_m$. Thus if

$$\lambda = x_1^{b_1} x_2^{b_2} \cdots x_s^{b_s}$$

with $b_s \geq 1$ appears in some f_j, then $\sum_{\ell=1}^{s} b_\ell = \sum_{\ell=1}^{m} a_\ell$ and $\sum_{\ell=1}^{s} \ell b_\ell \leq \sum_{\ell=1}^{m} \ell a_\ell$. From this we see that the largest value of s that can occur corresponds to the monomial $x_1^t x_n$ where $t = (\sum_{\ell=1}^{m} a_\ell) - 1$ and $t + n = \sum_{\ell=1}^{m} \ell a_\ell$. Thus

$$n = \sum_{\ell=1}^{m} \ell a_\ell - (\sum_{\ell=1}^{m} a_\ell) + 1 = 1 + \sum_{\ell=1}^{m} a_\ell(\ell - 1).$$

In fact, it can be shown that this monomial always does occur with non-zero coefficient in the invariant f.

Finally, we consider the case where $a_m \geq p :=$ characteristic$(\mathbb{K}) > 0$. Dividing a_m by p we may write $a_m = qp + r$ where $0 \leq r < p$. By the above, there exists $h \in \mathbb{F}[x_1, x_2, \ldots, x_n]^{\mathbb{Z}}$ with $\mathrm{LT}(h) = \beta x_m^r$. Since $\mathrm{LT}(\mathrm{N}(x_m)) = x^p$, we see that $\mathrm{N}(x_m)^q h$ is invariant and has lead term $\beta x_m^{a_m}$, as required. □

Definition 7.5.6. *The algorithm used in the above proof to construct an invariant with given lead term was introduced in [98] and is known as* Shank's algorithm.

When we apply Shank's algorithm to the lead term x_2^{i-1}, we produce invariants $s_{i-1} \in \mathbb{F}[V_n]^{C_p}$ for each i, $3 \leq i \leq n$. This family of invariants plays an important role in our analysis of the fraction fields associated to the rings of invariants, hence we will refer to them as the *Shank* invariants.

The algorithm yields the following invariants:

$$
\begin{aligned}
s_{i-1} &= x_2^{i-2}v_0 + x_1x_2^{i-3}v_1 + \cdots + x_1^{i-3}x_2v_{i-3} + x_1^{i-2}v_{i-2}\\
&= \sum_{j=0}^{i-1} x_i^j x_2^{i-2-j} v_j
\end{aligned}
$$

where $v_0 = x_2$, and $v_1 = -(i-1)x_3$. It is not hard to see that, for $1 \leq j \leq i-3$, we have

$$v_j = \alpha_{j,3}x_3 + \cdots + \alpha_{j,j+2}x_{j+2},$$

while

$$v_{i-2} = \alpha_{i-2,2}x_2 + \alpha_{i-2,3}x_3 + \cdots + \alpha_{i-2,i}x_i,$$

for coefficients $\alpha_{j,k} \in \mathbb{F}$. Here are some examples:

$$
\begin{aligned}
s_2 &= x_2(x_2) - x_1x_2 + x_1(-2x_3)\\
s_3 &= x_2^2(x_2) + x_1x_2(-3x_3) + x_1^2(-x_2 + 3x_4)\\
s_4 &= x_2^3(x_2) + x_1x_2^2(-4x_3) + x_1^2x_2(2x_3 - 6x_4)\\
&\quad +x_1^3(-x_2 - 2x_3 - 6x_4 - 8x_5)
\end{aligned}
$$

Of course, as an alternative to the work of Shank, we could proceed to show that the requirement that $\Delta(s_{i-1}) = 0$ gives us a system of equations in the unknowns $\alpha_{j,k}$ which can be shown to have a (unique monic) solution over the integers.

The following lemma ensures that we may isolate x_i from the expression for s_{i-1} once we have inverted x_1.

Lemma 7.5.7. *The coefficient of the term $x_1^{i-2}x_i$ in the integral invariant s_{i-1} is invertible in $\mathbb{F}_p$.*

Proof. The lemma follows through the calculation of the coefficients of the following monomials occurring in $\Delta(s_{i-1})$:

Monomial	**Coefficient**
$x_1^{i-2}x_{i-1}$	$\alpha_{i-1,i} + \alpha_{i-2,i-1}$
$x_1^{i-3}x_2x_{i-2}$	$\alpha_{i-2,i-1} + 2\alpha_{i-3,i-2}$
$\vdots$	$\vdots$
$x_1^{i-j-1}x_2^{j-1}x_{i-j}$	$\alpha_{i-j,i-1-j} + j\alpha_{i-2-j,i-j}$
$\vdots$	$\vdots$
$x_1^2x_2^{i-4}x_3$	$\alpha_{2,4} + (i-3)\alpha_{1,3},$

and now Shank's algorithm tells us that $\alpha_{1,3} = -(i-1)$. We may conclude that the coefficient of $x_1^{i-2}x_i$ is

$$(-1)^i(i-3)!(i-1),$$

which is invertible for $3 \leq i \leq n \leq p$, as required. □

7.6 Invariant Fraction Fields and Localized Invariants

We have good descriptions of the fraction field of a ring of invariants since we may apply Galois theory. But how much information can we expect to obtain about the ring of invariants from its fraction field? This question was examined by Campbell and Chuai [16] and is recalled here. See also §4.3.

We think of the fraction field as the localization of the ring away from the prime ideal (0). Our approach is to begin by considering a ring localized at a fixed point of the C_p-action.

First, let us consider the case $V = V_n$ over a field $\mathbb{F}$ of characteristic p. We have

$$\mathbb{F}[V_n] = \mathbb{F}[x_1, x_2, \ldots, x_n]$$

such that $\sigma(x_i) = x_i + x_{i-1}$, subject to the convention $x_0 = 0$. We use the Shank polynomials

$$s_{i-1} = x_2^{i-1} + \cdots + \alpha_{i-2,2} x_1^{i-1} x_2 + \cdots + (-1)^i (i-1)(i-3)! x_1^{i-2} x_i \ .$$

Since, by Lemma 7.5.7, the coefficient of $x_1^{i-2} x_i$ is invertible in $\mathbb{F}$, this equation can be solved for x_i in the localized ring $\mathbb{F}[V_n]_{x_1}$. For example, we have $s_2 = x_2^2 - x_1 x_2 - 2x_1 x_3$, which we solve for

$$x_3 = \frac{-1}{2x_1}(s_2 - x_2^2 + x_1 x_2).$$

Inductively, we may assume that $x_3, \ldots, x_n$ can be written as functions of $x_1^{\pm 1}$, x_2 and the invariant rational functions s_{i-1}, $3 \leq i \leq n$. That is, we have proved

Proposition 7.6.1.

$$\mathbb{F}[V_n]_{x_1} = \mathbb{F}[x_1^{\pm 1}, x_2, s_{i-1} \mid 3 \leq i \leq n].$$

Consequently,

$$\mathbb{F}[V_n]_{x_1}^{C_p} = \mathbb{F}[x_1^{\pm 1}, N(x_2) = x_2^p - x_1^{p-1} x_2, s_{i-1} \mid 3 \leq i \leq n]$$

and

$$\mathbb{F}(V_n)^{C_p} = \mathbb{F}(x_1, N(x_2), s_{i-1} \mid 3 \leq i \leq n).$$

The proposition reveals clearly the limitations of this method. We note that the invariant fraction field is generated by $\dim(V)$ elements. Further, the ring of invariants localized at the single invariant x_1 is also generated by $\dim(V)$ many elements. That is, as soon as just one suitable element is inverted, the complex and subtle structure of the ring of invariants collapses to this relatively simple structure: no traces are needed to generate the localized rings and all of the generators needed are of degree p or less.

In the event V has more than one non-trivial summand, say

$$V_n \oplus V_m \subset V$$

with $n, m > 1$, we need a particular kind of integral invariant, the degree 2 determinant invariants. Let $\{x_1, x_2, \ldots, x_n\}$ be a triangular basis for V_n^* with x_n distinguished. Similarly let $\{y_1, y_2, \ldots, y_m\}$ denote a triangular bases for V_m^* with y_m distinguished. Then $u = u_{V_n,V_m} = x_1y_2 - x_2y_1$ is invariant, and there is one such an invariant for every pair of non-trivial summands in the indecomposable decomposition of V.

There is one more family of invariants that are needed, again associated to a representation with more than one non-trivial summand, say $V_n \oplus V_m \subset V$ with bases as above and $n \leq m$. Suppose $f = f(x_1, \ldots, x_n) \in \mathbb{F}[V_n]^{C_p}$ has degree d. One of the polarizations of f, $f_{d-1,1}$, has multi-degree $(d-1, 1)$ in $V_n \oplus V_n$. We consider the second copy of V_n here as being embedded as a C_p-submodule of V_m. Thus we may interpret $f_{d-1,1}$ as a bi-homogeneous polynomial of degree $d-1$ in the variables $x_1, x_2, \ldots, x_n$ and degree 1 in $y_1, y_2, \ldots, y_n$. With this interpretation, we denote $f_{d-1,1}$ by f_1. We are interested only in applying this construction to the Shank invariants. In the three examples above, these are

$$\begin{aligned}
(s_{2,n,m})_1 &= -x_2y_1 + 2x_2y_2 - 2x_3y_1 + x_1(-y_2 - 2y_3)\\
(s_{3,n,m})_1 &= 6x_4x_1y_1 - 3x_3x_2y_1 - 3x_3x_1y_2 + 3x_2^2y_2\\
&\quad -3x_2x_1y_3 - 2x_2x_1y_1 + x_1^2(-y_2 + 3y_4)\\
(s_{4,n,m})_1 &= -24x_5x_1^2y_1 + 16x_4x_2x_1y_1 + 8x_4x_1^2y_2\\
&\quad -18x_4x_1^2y_1 - 4x_3x_2^2y_1 - 8x_3x_2x_1y_2 + 4x_3x_2x_1y_1\\
&\quad +2x_3x_1^2y_2 - 6x_3x_1^2y_1 + 4x_2^3y_2 - 4x_2^2x_1y_3\\
&\quad +8x_2x_1^2y_4 + 2x_2x_1^2y_3 - 3x_2x_1^2y_1 - x_1^3(-y_2 - 2y_3 - 6y_4 + 8y_5),
\end{aligned}$$

Note the abuse of notation here: $s_{2,n,m} = s_{2,V_n,V_m}$. Note as well that $s_{i,n,m}$ is only defined for $i \leq \min(n, m)$.

Again, it is critical to note that the coefficient of $x_1^{i-1}y_i$ in $(s_{i,n,m})_1$ is invertible in $\mathbb{F}$. This follows from Lemma 7.5.7.

Decompose V as $V = k_1V_1 \oplus k_2V_2 \oplus \cdots \oplus k_pV_p$ and choose a triangular basis for V^*

$$\{x_{i,j,n} \mid 1 \leq i \leq n,\ 1 \leq j \leq k_n,\ 1 \leq n \leq p\}$$

with the property that $\sigma(x_{i,j,n}) = x_{i,j,n} + x_{i-1,j,n}$, subject to the convention that $x_{0,j,n} = 0$. Thus

$$\mathbb{F}[V] = \mathbb{F}[x_{i,j,n} \mid 1 \leq i \leq n,\ 1 \leq j \leq k_n,\ 1 \leq n \leq p]$$

Now we choose the largest m for which $k_m > 0$ and fix a particular copy $V_{\ell,m}$ of V_m. Let $\{x_1, x_2, \ldots, x_m\}$ denote the usual basis for $V_{\ell,m}^*$. Set $x_1 = x_{1,\ell,m}$ and $x_2 = x_{2,\ell,m}$. Then $\sigma(x_1) = x_1$ and that $\sigma(x_2) = x_2 + x_1$.

For each other distinct non-trivial summand V_n, then, we have $n \leq m$ and we choose the usual basis $\{y_1, \ldots, y_n\}$, that is, we have $\sigma(y_i) = y_i + y_{i-1}$ with $y_0 = 0$. We consider the polarized Shank invariants $(s_{i-1,n,m})_1$ for $3 \leq i \leq n-1$, each of which has a term $x_1^{i-2} y_i$ which occurs with coefficient $(-1)^i (i-1)(i-3)!$, invertible in $\mathbb{F}$. As above, we may inductively solve for y_i, $3 \leq i \leq n$, in the localized ring $\mathbb{F}[V]_{x_1}$. Finally, we consider the determinant invariant

$$u_{n,m} = u_{V_n, V_m} = x_1 y_2 - x_2 y_1$$

and note that we have

$$y_2 = \frac{1}{x_1}(u_{n,m} + x_2 y_1).$$

Hence we have proved the following

Theorem 7.6.2. *Let $V = k_2 V_2 \oplus \cdots \oplus k_p V_p$ be any reduced representation of C_p. Choose a triangular basis $\{x_{i,j,n} \mid 1 \leq i \leq n\}$ for the jth copy of V_n^* for $j = 1, 2, \ldots, k_n$ for $n = 2, 3, \ldots p$. Fix a choice $x_1 = x_{1,\ell,m}$ and $x_2 = x_{2,\ell,m}$ for some ℓ and m the largest m for which $k_m > 0$. Then*

$$\mathbb{F}[V]_{x_1} = \mathbb{F}[x_1^{\pm 1}, x_2, x_{1,j,n}, u_{n,m}, (s_{i,n,m})_1 \mid 1 \leq j \leq k_n,\ 2 \leq n \leq p]$$

and

$$\mathbb{F}(V)^{C_p} = \mathbb{F}(x_1^{\pm 1}, N(x_2), x_{1,j,n}, u_{n,m}, (s_{i,n,m})_1 \mid 1 \leq j \leq k_n,\ 2 \leq n \leq p)\ .$$

7.7 Noether Number for C_p

In this section, we will present a theorem of Fleischmann, Sezer, Shank and Woodcock which gives exact values for the Noether number of the ring of invariants of any representation of C_p. Here we will follow their proof closely. However, we will change it in order to make it self-contained, The proof given below does not rely on results from [100].

Recall that we say that V is reduced if V does not contain a copy of V_1 as a summand.

Theorem 7.7.1. *Let V be a non-trivial reduced representation of C_p, the cyclic group of order p defined over a field $\mathbb{F}$ of characteristic $p > 0$. Let s be the maximum dimension of an indecomposable summand of V. (Thus $2 \leq s \leq p$.) Then*

$$\beta(V, C_p) = \begin{cases} p, & \text{if } V \cong V_2 \text{ or } V \cong 2\,V_2; \\ (p-1)\dim(V^{C_p}), & \text{if } V \cong m\,V_2 \text{ for } m \geq 3; \\ (p-1)\dim(V^{C_p}) + 1, & \text{if } s = 3; \\ (p-1)\dim(V^{C_p}) + p - 2, & \text{if } s \geq 4. \end{cases}$$

Note that $\dim V^{C_p}$ is the number of indecomposable summands in V.

If $s = 2$, then the result follows from the Theorem 7.4.1 (and Remark 7.4.2). Thus we suppose that $s \geq 3$.

We begin the proof of Theorem 7.7.1 by developing lower bounds for $\beta(V, C_p)$. To do this, we consider the two C_p-representations $U = V_4 \oplus m\,V_2$ and $W = V_3 \oplus m\,V_2$. Choose a triangular basis $\{w_0, z_0, y_0, x_0\}$ for the submodule V_4^* of U^* and triangular bases $\{y_i, x_i\}$ $(i = 1, 2, \ldots, m)$ for the duals of the m summands isomorphic to V_2. Thus $\Delta(w_0) = z_0$, $\Delta(z_0) = y_0$, $\Delta(y_i) = x_i$ and $\Delta(x_i) = 0$ for $i = 0, 1, \ldots, m$. Then $\{w_0, z_0, y_0, x_0\} \sqcup \{y_i, x_i \mid i = 1, 2, \ldots, m\}$ is a basis for U^*. Viewing W as a C_p submodule of U we may take $\{z_0, y_0, x_0\} \sqcup \{y_i, x_i \mid i = 1, 2, \ldots, m\}$ as a basis for W^*.

Lemma 7.7.2. *Take $U = V_4 \oplus m\,V_2$ and $W = V_3 \oplus m\,V_2$ as above. Then $\mathrm{Tr}^{C_p}((w_0 y_1 y_2 \cdots y_m)^{p-1} z_0^{p-2})$ is an indecomposable element of $\mathbb{F}[U]^{C_p}$ and $\mathrm{Tr}^{C_p}((z_0 y_1 y_2 \cdots y_m)^{p-1} y_0)$ is an indecomposable element of $\mathbb{F}[W]^{C_p}$.*

Proof. We consider U first. Put $F_m := \mathrm{Tr}^{C_p}((w_0 y_1 y_2 \cdots y_m)^{p-1} z_0^{p-2})$.

We proceed by induction on m. For $m = 0$, we are concerned with $\mathbb{F}[V_4]^{C_p}$. Shank [98] found a set of generators for $\mathbb{F}[V_4]^{C_p}$. We describe this computation at length in Chapter 13. We will use the graded reverse lexicographic order on $\mathbb{F}[V_4]$ determined by $w_0 > z_0 > y_0 > x_0$. Using Lemma 9.0.2, we can show that $\mathrm{LT}(F_0) = -z_0^{2p-3}$. The (non-minimal) generating set for $\mathbb{F}[V_4]^{C_p}$ given by Shank includes invariants whose lead term is of the form z_0^i precisely for the values $i = p-1, p, \ldots, 2p-3$ (including F_0). Furthermore, the only invariant among the generators whose lead term is divisible by w_0 is $\mathbf{N}(w_0)$ with lead term w_0^p. This shows that the largest (with respect to the graded reverse lexicographic order) monomial in $\mathrm{LT}(\mathbb{F}[V_4]^{C_p})_d$ is strictly smaller than z_0^d for all $d = 1, 2, \ldots, p-2$. In particular, there can be no (sequence of) tête-a-têtes among the generators yielding invariants with lead term z_0^d for $d \leq p-1$. This shows that F_0 is indecomposable since its lead term cannot be expressed as a product of lead terms of lower degree invariants. The reader may find a more detailed version of this argument in [96].

For the general case $m \geq 1$, we assume the induction hypothesis that F_{m-1} is a non-zero indecomposable element of $\mathbb{F}[V_4 \oplus (m-1)\,V_2]^{C_p}$. We consider the C_p submodule $U' = V_4 \oplus (m-1)\,V_2 \oplus V_1$ of U whose dual has basis $\{w_0, z_0, y_0, x_0, y_1, x_1, \ldots, y_{m-1}, x_{m-1}, x_m\}$. The inclusion of U' into U induces a surjective, multi-degree preserving, C_p-equivariant algebra map $\pi : \mathbb{F}[U] \to \mathbb{F}[U']$ determined by $\pi(y_m) = x_m$, $\pi(y_{m-1}) = x_{m-1}$ and fixing all the other variables. Now

$$\begin{aligned}
\pi(F_m) &= \pi(\sum_{\sigma \in C_p} \sigma((z_0 y_1 y_2 \cdots y_{m-1} y_m)^{p-1} y_0^{p-2})) \\
&= \sum_{\sigma \in C_p} \pi(\sigma((z_0 y_1 y_2 \cdots y_{m-1} y_m)^{p-1} y_0^{p-2})) \\
&= \sum_{\sigma \in C_p} \sigma(\pi((z_0 y_1 y_2 \cdots y_{m-1} y_m)^{p-1} y_0^{p-2}))
\end{aligned}$$

$$\begin{aligned}
&= \sum_{\sigma\in C_p} \sigma((z_0y_1y_2\cdots y_{m-1}x_m)^{p-1}y_0^{p-2})\\
&= x_m^{p-1}\sum_{\sigma\in C_p} \sigma((z_0y_1y_2\cdots y_{m-1})^{p-1}y_0^{p-2})\\
&= x_m^{p-1}\operatorname{Tr}^{C_p}((z_0y_1y_2\cdots y_{m-1})^{p-1}y_0^{p-2})\\
&= x_m^{p-1}F_{m-1}\ .
\end{aligned}$$

Note that since $F_{m-1}\neq 0$, this shows that $F_m\neq 0$. Assume by way of contradiction that F_m is decomposable in $\mathbb{F}[U]^{C_p}$ and write $F_m=\sum_{i=1}^r f_ig_i$ for some multi-homogeneous invariants $f_i,g_i\in\mathbb{F}[U]_+^{C_p}$. Thus each of the products f_ig_i has multi-degree $(2p-3,p-1,p-1,\ldots,p-1)$. Let $x_m^{a_i}$ denote the largest power of x_m which divides f_i for $i=1,2,\ldots,r$. Similarly, let $x_m^{b_i}$ denote the largest power of x_m which divides g_i. Thus $a_i+b_i=p-1$ for all $i=1,2,\ldots,r$. Therefore, writing $f_i=x_m^{a_i}f_i'$ and $g_i=x_m^{b_i}g_i'$ we have $\pi(F_m)=x_m^{p-1}\sum_{i=1}^r f_i'g_i'$ with f_i' and g_i' in $\mathbb{F}[U']^{C_p}$. Hence $F_{m-1}:=\sum_{i=1}^r f_i'g_i'$. Since F_{m-1} is indecomposable, one of the f_i' or g_i', say f_i', must be a non-zero constant, c. Thus $\pi(f_i)=cx_m^{p-1}$ and therefore, f_i has multi-degree $(0,0,0,\ldots,0,p-1)$. Hence $f_i\in\mathbb{F}[y_m,x_m]_{p-1}^{C_p}=\mathbb{F}[y_m^p-x_m^{p-1}y_m,x_m]_{p-1}=\operatorname{span}_{\mathbb{F}}\{x_m^{p-1}\}$. But then $\pi(f_i)=0$ contradicting the fact that $f_i'=c$. This contradiction shows that F_m is indecomposable in $\mathbb{F}[U]^{C_p}$.

The proof that $\operatorname{Tr}^{C_p}((z_0y_1y_2\cdots y_m)^{p-1}y_0)$ is indecomposable in $\mathbb{F}[W]^{C_p}$ is similar. We proceed by induction as above using Lemma 7.7.2 and using Theorem 4.10.1 to handle the base case. □

Corollary 7.7.3. *Let V be a non-trivial reduced representation of C_p. Let s be the maximum dimension of an indecomposable summand of V. Suppose $s\geq 3$. Then*

$$\beta(V,C_p)\geq\begin{cases}(p-1)\dim(V^{C_p})+1, & \text{if } s=3;\\ (p-1)\dim(V^{C_p})+p-2, & \text{if } s\geq 4.\end{cases}$$

Proof. We treat the case $s=4$ first. Put $m:=\dim(V^{C_p})-1$ and $U:=V_4\oplus m\,V_2$. Decompose V as $V=V_{r_0}\oplus V_{r_1}\oplus\cdots\oplus V_{r_m}$ where $r_0\geq 4$ and $r_i\geq 2$ for $i=1,2,\ldots,m$. For each $i=0,1,2,\ldots,m$, choose a distinguished variable $X_i\in V_{r_i}^*$. We determine a C_p-equivariant surjection $\phi:V^*\to U^*$ by requiring that $\phi(X_0)=w_0$ and $\phi(X_i)=y_i$ for $i=1,2,\ldots,m$. This induces a C_p-equivariant surjection, which we also call ϕ mapping $\mathbb{F}[V]$ onto $\mathbb{F}[U]$. Thus $\phi(\Delta^j(X_i))=\Delta^j(\phi(X_i))$ for all i and j. Restricting to invariants we get an algebra map $\phi:\mathbb{F}[V]^{C_p}\to\mathbb{F}[U]^{C_p}$.

Since ϕ is an algebra homomorphism and since

$$\phi(\operatorname{Tr}^{C_p}((X_0X_1X_2\cdots X_m)^{p-1}\Delta(X_0)^{p-2}))=\operatorname{Tr}^{C_p}((w_0y_1y_2\cdots y_m)^{p-1}z_0^{p-2})$$

is indecomposable in $\mathbb{F}[U]^{C_p}$, it follows that

$$\mathrm{Tr}^{C_p}((X_0X_1X_2\cdots X_m)^{p-1}\Delta(X_0)^{p-2})$$

is indecomposable in $\mathbb{F}[V]^{C_p}$. Therefore, $\beta(V,C_p)\geq(m+1)(p-1)+p-2$.

The case $s=3$ is proved similarly. □

Our next goal is, of course, to show that any invariant whose degree exceeds the bound given in Corollary 7.7.3 must be decomposable.

Accordingly, we consider a degree a with $a\geq(m+2)(p-1)$. By the discussion preceding Theorem 7.3.2, we see that every $f\in\mathbb{F}[V]^G$ may be written as $f=f^\sharp+f^\flat$ where $f^\sharp\in\mathbb{F}[V]^\sharp_a$ and $f^\flat\in\mathbb{F}[V]^\flat_a$. Since $f^\sharp\in\mathbb{F}[V]^\sharp_a$, we may write $f^\sharp=\sum_{i=1}^{m+1}f_i\mathbf{N}^{C_p}(z_i)$ with $f_i\in\mathbb{F}[V]^{C_p}_{a-p}$. Therefore, f is indecomposable if and only if $f^\flat$ is indecomposable. Thus we concentrate our attention on $f^\flat$. Since $\mathbb{F}[V]^\flat_a$ is free, $(\mathbb{F}[V]^\flat_a)^{C_p}=\mathrm{Tr}^{C_p}(\mathbb{F}[V]_a)$. Thus to prove Theorem 7.7.1, it suffices to show that every homogeneous invariant of the form $\mathrm{Tr}^{C_p}(h)$ of degree at least $(m+2)(p-1)$ is decomposable.

In order to study the image of the transfer, we consider $\mathbb{F}[V]$ as an $\mathbb{F}[V]^{C_p}$-module. Since $\mathbb{F}[V]$ is integral over $\mathbb{F}[V]^{C_p}$, we know that $\mathbb{F}[V]$ is a finitely generated $\mathbb{F}[V]^{C_p}$-module, i.e., there exist $h_1,h_2,\ldots,h_t\in\mathbb{F}[V]$ such that $\mathbb{F}[V]=\mathbb{F}[V]^{C_p}h_1+\mathbb{F}[V]^{C_p}h_2+\cdots+\mathbb{F}[V]^{C_p}h_t$. By the graded Nakayama lemma 2.10.1, we may find a homogeneous minimal such set of generators $h_1,h_2,\ldots,h_t$ by lifting back to $\mathbb{F}[V]$ any homogeneous basis for the maximal homogeneous ideal in the ring of coinvariants $\mathbb{F}[V]_{C_p}=\mathbb{F}[V]/J$ where J is the Hilbert ideal. We define $\gamma(V)=\max\{d\mid(\mathbb{F}[V]_{C_p})_d\neq\{0\}\}$. Hence for a homogeneous minimal set of generators, $h_1,h_2,\ldots,h_t$, we have $\gamma(V)=\max\{\deg(h_i)\mid 1\leq i\leq t\}$, Thus γ is the degree of the Hilbert series $\mathcal{H}(\mathbb{F}[V]_{C_p},\lambda)$ (which is in fact a polynomial).

Given any homogeneous $h\in\mathbb{F}[V]$, we may write $h=f_1h_1+f_2h_2+\cdots+f_th_t$ with $f_1,f_2,\ldots,f_t\in\mathbb{F}[V]^{C_p}$ and $\deg(h)=\deg(f_i)+\deg(h_i)$ for $i=1,2,\ldots,t$. Therefore, $\mathrm{Tr}^{C_p}(h)=f_1\,\mathrm{Tr}_{C_P}(h_1)+f_2\,\mathrm{Tr}_{C_P}(h_2)+\cdots+f_t\,\mathrm{Tr}_{C_P}(h_t)$. This shows that if $\deg(h)>\gamma(V)$, then $\mathrm{Tr}^{C_p}(h)$ is a decomposable invariant.

Our next step is to get upper bounds (which will turn out to be sharp) for $\gamma(V)$. As above, we let s denote the maximum dimension of an indecomposable summand of V. We will show that

$$\gamma(V)=\begin{cases}(p-1)\dim(V^{C_p})+1, & \text{if } s=3;\\ (p-1)\dim(V^{C_p})+p-2, & \text{if } s\geq 4.\end{cases}$$

As we noted above, $\gamma(V)=\deg(\mathcal{H}(\mathbb{F}[V]_{C_p},\lambda)$. Thus

$$\begin{aligned}\gamma(V)&=\deg(\mathcal{H}(\mathbb{F}[V],\lambda)-\deg(\mathcal{H}(J,\lambda)\\&=\deg(\mathcal{H}(\mathbb{F}[V],\lambda)-\deg(\mathcal{H}(\mathrm{LT}(J),\lambda)\\&=\deg(\mathcal{H}(\mathbb{F}[V]/\,\mathrm{LT}(J),\lambda).\end{aligned}$$

This is true regardless of which monomial order is used to determine lead terms.

Thus $\gamma(V)$ is the highest degree of a monomial not lying in $\mathrm{LT}(J)$. We will use this characterization to determine $\gamma(V)$.

We consider the case $s = 3$ first. For this case, we may write $V = m_1V_2 \oplus m_2V_3$ where $m_2 \geq 1$. We choose distinguished variables $y_1, y_2, \ldots, y_{m_1}$ and $z_{m_1+1}, z_{m_1+2}, \ldots, z_{m_1+m_2}$. We define $y_i = \Delta(z_i)$ for $m_1 + 1 \leq i \leq m_1 + m_2$ and $x_i = \Delta(y_i)$ for $1 \leq i \leq m_1 + m_2$. Then we have $\mathbb{F}[V] = \mathbb{F}[z_{m_1+1}, z_{m_1+2}, \ldots, z_{m_1+m_2}, y_1, y_2, \ldots, y_{m_1+m_2}, x_1, x_2, \ldots, x_{m_1+m_2}]$. We fix a graded reverse lexicographic order with respect to these variables and satisfying $y_1 < z_i$ for $m_1 + 1 \leq i \leq m_1 + m_2$ and $x_i < y_i$ for all $1 \leq i \leq m_1 + m_2$.

It is easily verified that $d_i = y_i^2 - 2x_iz_i - x_iy_i$ is invariant with $\mathrm{LT}(d_i) = y_i^2$ for all $m_1 + 1 \leq i \leq m_1 + m_2$. Similarly, the polarizations of d_i given by $d_{ij} = y_iy_j - x_iz_j - z_jx_i - x_iy_j$ are also invariant and satisfy $\mathrm{LT}(d_{ij}) = y_iy_j$ for all $m_1 + 1 \leq i < j \leq m_1 + m_2$. The linear monomials $x_1, x_2, \ldots, x_{m_1+m_2}$ are all invariant. Finally, $z_i^p = \mathrm{LT}(\mathbf{N}^{C_p}(z_i)) \in \mathrm{LT}(J)$ for all $1 \leq i \leq m_1$ and $y_i^p = \mathrm{LT}(\mathbf{N}^{C_p}(y_i) \in \mathrm{LT}(J)$ for all $m_1 + 1 \leq i \leq m_1 + m_2$. From this it is clear that a monomial not lying in $\mathrm{LT}(J)$ cannot be divisible by any x_i, nor by any y_iy_j with $i, j \geq m_1 + 1$, nor by any z_i^p, nor by any y_i^p. This shows that the highest degree monomials which may lie outside $\mathrm{LT}(J)$ are those of the form $(y_1y_2 \cdots y_{m_1}z_{m_1+1} \cdots z_{m_1+m_2})^{p-1}y_j$ with $j > m_1$. Therefore, $\gamma(m_1V_2 \oplus m_2V_3) \leq (m_1+m_2)(p-1)+1 = (p-1)\dim(V^{C_p})+1$ which completes the proof for the case $s = 3$.

The proof for the case $s \geq 4$ is similar. In this case, we will need two small technical lemmas in order to bound $\gamma(V)$.

Before stating and proving these lemmas, we fix some notation. As above, we write $V = V_{d_1} \oplus V_{d_2} \oplus \cdots \oplus V_{d_m}$ and we choose a distinguished variable z_i in each $V_{d_i}^*$. We then consider the basis $B := \{z_{ij} \mid 1 \leq i \leq m, 0 \leq j < d_i\}$ of V^* where $z_{ij} := \Delta^j(z_i)$ for all $j = 0, 1, \ldots, d_i - 1$ and $i = 1, 2, \ldots, m$. We also define $B' := \{z_{ij} \mid 1 \leq i \leq m, 1 \leq j < d_i\} = B \setminus \{z_1, z_2, \ldots, z_m\}$ and the polynomial ring $R := \mathbb{F}[B'] = \mathbb{F}[z_{ij} \mid 1 \leq i \leq m, 1 \leq j < d_i]$. Let α be any monomial of degree p in R. Write $\alpha = u_1u_2 \ldots u_{p-1}$ and define $\alpha' = w_1w_2 \ldots w_{p-1}$ to be the monomial of degree $p - 1$ in $\mathbb{F}[V]$ where $w_s \in B$ and $\Delta(w_s) = u_s$ for $s = 1, 2, \ldots, p - 1$. We fix a graded reverse lexicographic order on $\mathbb{F}[V]$ (and R) with $z_{ij} < z_{ik}$ whenever $j > k$.

Let S be any subset of $\{1, 2, \ldots, p-1\}$. We will denote $\{1, 2, \ldots, p-1\} \setminus S$ by S'. We define $X_S := \prod s \in Sw_s$ and $X_{S'} := \prod s \in S'w_s$ so that for all subsets S, we have $X_SX_{S'} = \alpha'$. Fix a generator σ of C_p and consider the function

$$\begin{aligned} F &:= \sum_{t=0}^{p-1} \prod_{j=1}^{p-1} (w_j - \sigma^t(w_j)) \\ &= \sum_{\tau \in C_p} \prod_{j=1}^{p-1} (w_j - \tau(w_j)) \ . \end{aligned}$$

Lemma 7.7.4.

$$F = \sum_{S \subseteq \{1,2,\ldots,p-1\}} (-1)^{|S|} X_{S'} \operatorname{Tr}^{C_p}(X_S) \ .$$

Proof. For each t we have

$$\prod_{j=1}^{p-1}(w_j - \sigma^t(w_j)) = \sum_{S \subseteq \{1,2,\ldots,p-1\}} (-1)^{|S|} X_{S'} \sigma^t(X_S)$$

Thus

$$\begin{aligned} F &= \sum_{t=0}^{p-1} \prod_{j=1}^{p-1}(w_j - \sigma^t(w_j)) \\ &= \sum_{t=0}^{p-1} \sum_{S \subseteq \{1,2,\ldots,p-1\}} (-1)^{|S|} X_{S'} \sigma^t(X_S) \\ &= \sum_{S \subseteq \{1,2,\ldots,p-1\}} (-1)^{|S|} X_{S'} \sum_{t=0}^{p-1} \sigma^t(X_S) \\ &= \sum_{S \subseteq \{1,2,\ldots,p-1\}} (-1)^{|S|} X_{S'} \operatorname{Tr}^{C_p}(X_S) \end{aligned}$$

□

The function F defined above satisfies the following.

Lemma 7.7.5.

$$\operatorname{LT}(F) = -\alpha \ .$$

In particular, $\alpha \in \operatorname{LT}(J)$.

Proof. Since $\sigma^t(w_j) = w_j + t\Delta(w_j) + \binom{t}{2}\Delta^2(w_j) + \cdots = w_j + tu_j + \ldots$ we see that $\operatorname{LT}(w_j - \sigma^t(w_j)) = -tu_j$ for all $t \geq 1$ and all j. Therefore, if $1 \leq t \leq p-1$ we have $\operatorname{LT}(\prod_{j=1}^{p-1}(w_j - \sigma^t(w_j)) = \prod_{j=1}^{p-1} \operatorname{LT}(w_j - \sigma^t(w_j)) = \prod_{j=1}^{p-1} -tu_j = (-t)^{p-1}\alpha = \alpha$. For $t = 0$, we have $w_j - \sigma^0(w_j) = 0$. Therefore, F is the sum of $p-1$ non-zero summands, each of which has lead term α. Thus $\operatorname{LT}(F) = (p-1)\alpha = -\alpha$. □

Note that $z_i^p = \operatorname{LT}(\mathbf{N}^{C_p}(z_i)) \in \operatorname{LT}(J)$ for all $1 \leq i \leq m$. Thus we have shown that the monomials of largest degree which may lie outside $\operatorname{LT}(J)$ are those monomials of the form $(z_1 z_2 \cdots z_m)^{p-1}\alpha'$ where $\alpha' = u_1 u_2 \cdots u_{p-2}$ is some monomial in R of degree $p-2$. Thus $\gamma(V) \leq m(p-1) + p - 2$. This completes the proof of Theorem 7.7.1.

7.8 Hilbert Functions

Hughes and Kemper [54] provide an explicit formula for the calculation of the Hilbert series of an arbitrary representation of C_p in characteristic p as follows. We have that

$$V = V_{n_1+1} \oplus V_{n_2+1} \oplus \cdots \oplus V_{n_k+1}$$

where each $V_{n_\ell+1}$ is an indecomposable C_p-module of dimension $n_\ell + 1$, $0 \leq n_\ell < p$. We define a function $D_V(\xi, t) \in \mathbb{C}[t]$ as

$$D_V(\xi, t) = \prod_{\ell=1}^{k} \left(\frac{1 - \xi^{pn_\ell} t^p}{1 - t^p} \prod_{j=0}^{n_i} (1 - \xi^{n_\ell - 2j})^{-1} \right).$$

Theorem 7.8.1. *Let $\mathcal{M}(2p) \subset \mathbb{C}$ denote the set of $2p$-th roots of unity. Then, for V and D_V as just defined, the Hilbert series of $\mathbb{F}[V]^{C_p}$ is given by*

$$H(\mathbb{F}[V]^{C_p}, t) = \sum_{\xi \in \mathcal{M}(2p)} \frac{1+\xi}{2p} D_V(\xi, t).$$

If all of the n_ℓ are even, then the formula simplifies to

$$H(\mathbb{F}[V]^{C_p}, t) = \frac{1}{p} \sum_{\xi \in \mathcal{M}(p)} \prod_{\ell=1}^{k} \prod_{j=0}^{n_\ell} (1 - \xi^{n_\ell - 2j} t)^{-1}.$$

Furthermore, the multiplicity of V_n in $\mathbb{F}[V]$ as a C_p-module is given by

$$\sum_{\xi \in \mathcal{M}(2p)} \frac{\xi - \xi^{-1}}{2p} D_V(\xi, t), \quad \textit{if } n < p,$$

while the multiplicity of V_p in $\mathbb{F}[V]$ as a C_p-module is given by

$$\sum_{\xi \in \mathcal{M}(2p)} \frac{1+\xi}{2p} \xi^p D_V(\xi, t)\ .$$

Remark 7.8.2. Suppose that V is a sum of odd dimensional indecomposables

$$V_{i=1}^{k} := \oplus V_{n_i+1} \quad \text{where all} \quad n_i \text{ are even.}$$

Fix a primitive p-th root of unity $\xi \in \mathbb{C}$. Define the matrices

$$M_\ell = \operatorname{diag}(\xi^{n_\ell}, \xi^{n_\ell - 2}, \ldots, \xi, 1, \xi^{-1}, \ldots, \xi^{(n_\ell - 2)}, \xi^{-n_\ell}) \in \mathrm{GL}_{n_\ell+1}(\mathbb{C})$$

and set

$$M = \operatorname{diag}(M_1, M_2, \ldots, M_k) \in \mathrm{GL}_{\dim(V)}(\mathbb{C}).$$

We obtain a non-modular representation of C_p by mapping a generator to M. We may use Molien's theorem Theorem 3.7 to determine the Hilbert series of $\mathbb{C}[M]^{C_p}$. It is a remarkable fact that this Hilbert series is the same as the Hilbert series of $\mathbb{F}[V]^{C_p}$. Both Almkvist [2] and Hughes and Kemper [54] believe this to be a "combinatorial accident".

8

Polynomial Invariant Rings

In this chapter, we study representations of groups which have polynomial rings of invariants. In characteristic 0, this happens if and only if the group is a reflection group. Recall 1.5.1

Definition 8.0.1. *Suppose V is a representation of G defined over a field $\mathbb{K}$ of any characteristic. We say that $\sigma \in G$ is a reflection on V if the dimension of V^σ is $\dim(V) - 1$, that is, $e \neq \sigma$ fixes a hyperplane pointwise.*

A reflection $\sigma \in \mathrm{GL}(V)$ is said to be a *transvection* if it is not diagonalizable. If it is diagonalizable, it is said to be a *homology.*

Building on work of Shephard and Todd [101], and Chevalley [22], Serre proved

Theorem 8.0.2. *If $\mathbb{K}[V]^G$ is a polynomial ring, then G is generated by reflections.*

Lemma 8.0.3. *Suppose V is a representation of G defined over a field $\mathbb{F}$ of characteristic $p > 0$. If $\sigma \in G$ is a reflection whose order is p^r, then σ has order p.*

Proof. Choose any element $v \in V$ lying outside of the hyperplane V^σ. Then $\sigma(v) = \lambda v + w$ for some $\lambda \in \mathbb{F}$ and $w \in V^\sigma$. Now $\sigma^\ell(v) = \ell\lambda v + (\lambda^{\ell-1} + \lambda^{\ell-2} + \cdots + 1)w$. Therefore, $\sigma^{p^r}(v) = \lambda^{p^r} v + ((\lambda^{p^r-1} + \lambda^{p^r-2} + \cdots + 1)w = v$. In particular, $\lambda^{p^r} = 1$ and so as a $(p^r)^{\text{th}}$-root of unity, $\lambda = 1$. Thus $\sigma^p(v) = v$. Therefore, σ^p fixes all of V pointwise as claimed. □

Definition 8.0.4. *Let P be a p-subgroup of* $\mathrm{GL}(V)$ *where V is an n dimensional vector space over the field $\mathbb{F}$ of characteristic p. We say that P is a* Nakajima group (with respect to B) if there exists an ordered basis $B = \{x_1, x_2, \ldots, x_n\}$ of V^* such that

1. $\sigma(x_i) - x_i \in \mathrm{span}\{x_1, x_2, \ldots, x_{i-1}\}$ for all $i = 1, 2, \ldots, n$ and for all $\sigma \in P$, i.e., in the basis B, P is an upper triangular group, and

H.E.A. Eddy Campbell, D. Wehlau, *Modular Invariant Theory*,
Encyclopaedia of Mathematical Sciences 139,
DOI 10.1007/978-3-642-17404-9_8, © Springer-Verlag Berlin Heidelberg 2011

2. $P = P_nP_{n-1} \cdots P_1 := \{\sigma_n\sigma_{n-1} \cdots \sigma_1 \mid \sigma_i \in P_i\}$ where $P_i := \{\sigma \in P \mid \sigma(x_j) = x_j$ for all $j \neq i\}$.

If P is a Nakajima group, we call the basis B a *Nakajima basis.*

An example of a Nakajima basis is given in Example 8.0.8.

Lemma 8.0.5. *For $i < j$, P_i normalizes P_j, that is, $P_iP_j = P_jP_i$.* □

From this it follows that the set $P_nP_{n-1} \cdots P_2P_1$ is in fact a subgroup of P. It is also clear that $P_i \cap P_j = \{e\}$ for all $i \neq j$.

Lemma 8.0.6. *Suppose that $B = \{x_1, x_2, \ldots, x_n\}$ is a basis of V^* and that P is a finite subgroup of* GL(V)*. As above, let $P_k = \{\sigma \in P \mid \sigma(x_j) = x_j$ for all $j \neq k\}$. Then every element $\gamma \in P_nP_{n-1} \cdots P_1$ has a* unique *expression of the form $\gamma = \sigma_n\sigma_{n-1} \cdots \sigma_1$ with $\sigma_i \in P_i$ for $i = 1, 2, \ldots, n$.*

Proof. We proceed by induction on n with the case $n = 1$ being trivial. Suppose the result holds for $n - 1$ and assume $\sigma_n\sigma_{n-1} \cdots \sigma_1 = \tau_n\tau_{n-1} \cdots \tau_1$ where $\sigma_i, \tau_i \in P_i$ for $i = 1, 2, \ldots, n$. Let τ denote the element $\tau_n^{-1}\sigma_n = \tau_{n-1}\tau_{n-2} \cdots \tau_1\sigma_1^{-1}\sigma_2^{-1} \cdots \sigma_{n-1}^{-1}$. Then

$$\tau(x_i) = \begin{cases} \tau_n^{-1}\sigma_n(x_i) = x_i & \text{if } i < n; \\ \tau_{n-1}\tau_{n-2} \cdots \tau_1\sigma_1^{-1}\sigma_2^{-1} \cdots \sigma_{n-1}^{-1}(x_i) = x_n, & \text{if } i = n. \end{cases}$$

Thus $\tau \in P_1 \cap P_2 \cap \cdots \cap P_n = \{e\}$ and therefore $\tau_n = \sigma_n$. This implies $\sigma_{n-1}\sigma_{n-2} \cdots \sigma_1 = \tau_{n-1}\tau_{n-2} \cdots \tau_1$. Applying the induction hypothesis yields $\tau_i = \sigma_i$ for all $i = 1, 2, \ldots, n - 1$. □

Theorem 8.0.7. *Let P be a p-subgroup of* GL(V) *and suppose that $B = \{x_1, x_2, \ldots, x_n\}$ is an ordered basis of V^* with respect to which P is an upper triangular subgroup of* GL(V^*)*. Then G is a Nakajima group with Nakajima basis B if and only if $\mathbb{F}[V]^P = \mathbb{F}[N_1, N_2, \ldots, N_n]$ where $N_i = \mathbf{N}^P_{P_{x_i}}(x_i)$.*

Proof. First we assume that B is a Nakajima basis for P. Since

$$P = P_nP_{n-1} \cdots P_1,$$

Lemma 8.0.6 implies that

$$|P| = \prod_{k=1}^{n} |P_k|.$$

Since

$$\sigma_n\sigma_{n-1} \cdots \sigma_1(x_k) = \sigma_k(x_k)$$

where $\sigma_i \in P_i$ for $i = 1, 2, \ldots, n$, we see that $P \cdot x_k = P_k \cdot x_k$. Furthermore, if $\sigma_k, \tau_k \in P_k$ are such that $\sigma_k(x_k) = \tau_k(x_k)$, then $\sigma_k^{-1}\tau_k(x_k) = x_k$ and thus $\sigma_k^{-1}\tau_k = P_1 \cap P_2 \cap \cdots \cap P_n = \{e\}$. This shows that $|P_k \cdot x_k| = |P_k|$. Therefore,

$\deg N_k = |P\cdot x_k| = |P_k\cdot x_k| = |P_k|$ and thus $\prod_{k=1}^{n} \deg(N_k) = \prod_{k=1}^{n} |P_k| = |P|$. By Proposition 4.0.3, $N_1, N_2, \ldots, N_n$ is a homogeneous system of parameters for $\mathbb{F}[V]^P$. Therefore, by Corollary 3.1.6, $\mathbb{F}[V]^P = \mathbb{F}[N_1, N_2, \ldots, N_n]$.

The proof of the opposite direction follows immediately from Theorem 8.0.11 below. □

Example 8.0.8. 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 let G denote the subgroup of $\mathrm{GL}(V)$ given by these matrices as a, b, c and d vary over $\mathbb{F}_p$. We first saw this group in §1.1.1. The action of $\sigma(a,b,c,d)$ on V^* is given by the matrix

$$\begin{pmatrix} 1&0&0&-a&0&0&-d\\ 0&1&0&0&-b&0&-d\\ 0&0&1&0&0&-c&-d\\ 0&0&0&1&0&0&0\\ 0&0&0&0&1&0&0\\ 0&0&0&0&0&1&0\\ 0&0&0&0&0&0&1 \end{pmatrix}$$

with respect to the basis B dual to the standard basis. It is clear that $P_1 = P_2 = P_3 = \{e\}$ while

$$\begin{aligned} P_4 &= \{\sigma(a,0,0,0) \mid a \in \mathbb{F}_p\}\,,\\ P_5 &= \{\sigma(0,b,0,0) \mid b \in \mathbb{F}_p\}\,,\\ P_6 &= \{\sigma(0,0,c,0) \mid c \in \mathbb{F}_p\} \quad \text{and}\\ P_7 &= \{\sigma(0,0,0,d) \mid d \in \mathbb{F}_p\}\,. \end{aligned}$$

It is also clear that the Abelian group $G = P_7 P_6 \cdots P_1$, and so B is a Nakajima basis for G and thus $\mathbb{F}[V]^G$ is a polynomial ring.

By way of contrast in this next example, we return to the group H defined in §1.1.1 as the transpose of the G of the previous example acting on V.

Example 8.0.9. The transpose of $\sigma(a,b,c,d)$ has the form

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

We consider the group $H \subset \mathrm{GL}(V)$ given by these latter matrices as a, b, c and d vary over $\mathbb{F}_p$. Following our usual convention, we rewrite the action of H on V in lower triangular form. The matrix just given is written with respect to the standard basis $\{e_1, e_2, \ldots, e_7\}$. Re-ordering this basis as $\{e_4, e_5, e_6, e_7, e_1, e_2, e_3\}$ and re-writing the matrix accordingly we obtain

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

We will show that $\mathbb{F}[V]^H$ is not Cohen-Macaulay. We define $\alpha = \sigma(1,0,0,0)$, $\beta = \sigma(0,1,0,0)$, $\gamma = \sigma(0,0,1,0)$ and $\delta = \sigma(0,0,0,1)$.

Let K be the subgroup generated by α, β and γ. We label the basis dual to the basis $\{e_5, e_6, e_7, e_1, e_2, e_3, e_4\}$ as $\{x_1, x_2, x_3, x, y_1, y_2, y_3\}$ so that $(V^*)^G = \operatorname{span}\{x_1, x_2, x_3, x\}$. Define $N_i(t) = t^p - x_i^{p-1}t$. By Nakajima's Theorem 8.0.7, we have

$$\mathbb{F}[V]^K = \mathbb{F}[x_1, x_2, x_3, x, N_1(y_1), N_2(y_2), N_3(y_3)].$$

Now define $\Delta = \delta - 1$. The rest of this example flows from Example 4.0.4. We calculate

$$\Delta(N_i(y_i)) = x(x^{p-1} - x_i^{p-1}) = N_i(x)$$

and hence

$$u_{ij} := (x^{p-1} - x_i^{p-1})N_j - (x^{p-1} - x_j^{p-1})N_i$$

is H-invariant for $1 \le i < j \le 3$. It is easy to verify that

$$x_1u_{23} - x_2u_{13} + x_3u_{12} = 0$$

and that $\{x_1, x_2, x_3\}$ is a partial homogeneous system of parameters in $\mathbb{F}[V]$ and hence also in $\mathbb{F}[V]^H$. Assume by way of contradiction that $\mathbb{F}[V]^H$ is Cohen-Macaulay. Then x_1, x_2, x_3 is a regular sequence in $\mathbb{F}[V]^H$ and therefore, since $x_3u_{12} \equiv 0 \pmod{x_1, x_2}$, we have

$$u_{12} = f_1x_1 + f_2x_2$$

for some f_1, f_2 in $\mathbb{F}[V]^H$. Clearly we may assume that the f_i's are homogeneous of degree $2p-2$. As a K-invariant of this degree, we must have $f_1 = h_0 + h_1N_1(y_1) + h_2N_2(y_2) + h_3N_3(y_3)$, where $h_j \in \mathbb{F}[x_1, x_2, x_3, x]$ for $0 \leq j \leq 3$. Further, as a H-invariant, we have

$$0 = \Delta(f_1) = h_1\Delta(N_1(y_1)) + h_2\Delta(N_2(y_2)) + h_3N_3(y_2)) \ .$$

Rewriting this expression we have

$$(h_1 + h_2 + h_3)x^{p-1} = h_1x_1^{p-1} + h_2x_2^{p-1} + h_3x_3^{p-1}.$$

Expanding the right hand side as a linear combination of monomials we see that any non-zero term is a scalar multiple of a monomial from one of the h_i's times the monomial x_i^{p-1}. But any such term must be divisible by x^{p-1}. Therefore, the degree of h_i is at least $p-1$, and so $h_1 = h_2 = h_3 = 0$, from which we see $f_1 \in \mathbb{F}[x_1, x_2, x_3, x]$. The same argument applies to f_2. But this would mean $u_{12} \in \mathbb{F}[x_1, x_2, x_3, x]$, a contradiction.

Lemma 8.0.10. *Let S and T be subgroups of a group K. Then the cardinality of the* set $ST := \{st \mid s \in S, t \in T\}$ *is given by*

$$|ST| = \frac{|S||T|}{|S \cap T|} \ .$$

Proof. Let $\mathcal{S} := \{s_i \mid 1 \leq i \leq [S : S \cap T]\}$ be a set of left coset representatives for $S \cap T$ in S and let $\mathcal{T} := \{t_j \mid 1 \leq j \leq [T : S \cap T]\}$ be a set of left coset representatives for $S \cap T$ in T. Define the set $L := \{s_ixt_j \mid s_i \in \mathcal{S}, x \in S \cap T, t_j \in \mathcal{T}\}$. We see $ST = L$ since clearly $ST \subseteq L$ and $L \subseteq ST$. Suppose $s_ixt_j = s_{i'}x't_{j'}$ where $s_i, s_{i'} \in \mathcal{S}$, $t_j, t_{j'} \in \mathcal{T}$ and $x, x' \in S \cap T$. Therefore, $s_{i'}^{-1}s_ix = x't_{j'}t_j^{-1} \in S \cap T$. Therefore, $s_{i'}^{-1}s_i, t_{j'}t_j^{-1} \in S \cap T$ and thus $s_i = s_{i'}$ and $t_j = t_{j'}$. Hence $x = x'$. This shows that every element of ST can be written uniquely in the form s_ixt_j with $s_i \in \mathcal{S}$, $t_j \in \mathcal{T}$ and $x \in S \cap T$. Hence $|ST| = [S : S \cap T] \cdot |S \cap T| \cdot [T : S \cap T] = \frac{|S|}{|S \cap T|} \cdot |S \cap T| \cdot \frac{|T|}{|S \cap T|} = \frac{|S||T|}{|S \cap T|}$. □

The following Theorem is due to Yinglin Wu [114].

Theorem 8.0.11. *Let P be a p-subgroup of* $\mathrm{GL}(V)$. *Suppose $\mathbb{F}[V]^P$ is a polynomial ring on the norms of the elements of some basis $B = \{y_1, y_2, \ldots, y_n\}$ of V^*:*

$$\mathbb{F}[V]^P = \mathbb{F}[\mathbf{N}_{P_{y_1}}^P(y_1), \mathbf{N}_{P_{y_2}}^P(y_2), \ldots, \mathbf{N}_{P_{y_n}}^P(y_n)] \ .$$

Then there exists an ordering of B with respect to which P is a Nakajima group.

Proof. We begin by showing that

$$|P_{y_1}| \cdot |P_{y_2}| \cdots |P_{y_r}| = |W_1| \cdot |W_2| \cdots |W_{r-1}| \cdot |P_{y_1} \cap P_{y_2} \cap \cdots \cap P_{y_r}|$$

for all $r = 2, 3, \ldots, n$ where $W_i := (P_{y_1} \cap P_{y_2} \cap \cdots \cap P_{y_i})P_{y_{i+1}}$ for $1 \leq i \leq n-1$.

For $r = 2$, we need to show that $|P_{y_1}| \cdot |P_{y_2}| = |P_{y_1} P_{y_2}| \cdot |P_{y_1} \cap P_{y_2}|$. This is immediate from Lemma 8.0.10. Now we proceed by induction on r. Hence we assume that

$$|P_{y_1}| \cdot |P_{y_2}| \cdots |P_{y_r}| = |W_1| \cdot |W_2| \cdots |W_{r-1}| \cdot |P_{y_1} \cap P_{y_2} \cap \cdots \cap P_{y_r}| \ .$$

Multiplying both sides of this equation by $|P_{y_{r+1}}|$ gives

$$|P_{y_1}| \cdot |P_{y_2}| \cdots |P_{y_r}| \cdot |P_{y_{r+1}}| = |W_1| \cdot |W_2| \cdots |W_{r-1}| \cdot |P_{y_1} \cap P_{y_2} \cap \cdots \cap P_{y_r}| \cdot |P_{y_{r+1}}| \ .$$

Applying Lemma 8.0.10 with $S = P_{y_1} \cap P_{y_2} \cap \cdots \cap P_{y_r}$ and $T = P_{y_{r+1}}$ we get

$$|P_{y_1} \cap P_{y_2} \cap \cdots \cap P_{y_r}| \cdot |P_{y_{r+1}}| = |(P_{y_1} \cap P_{y_2} \cap \cdots \cap P_{y_r}) P_{y_{r+1}}| \cdot |P_{y_1} \cap P_{y_2} \cap \cdots \cap P_{y_{r+1}}|$$

and therefore,

$$|P_{y_1}| \cdot |P_{y_2}| \cdots |P_{y_{r+1}}| = |W_1| \cdot |W_2| \cdots |W_r| \cdot |P_{y_1} \cap P_{y_2} \cap \cdots \cap P_{y_{r+1}}| \ ,$$

as required.

In particular, since $P_{y_1} \cap P_{y_2} \cap \cdots \cap P_{y_n} = \{e\}$ we see that

$$|P_{y_1}| \cdot |P_{y_2}| \cdots |P_{y_n}| = |W_1| \cdot |W_2| \cdots |W_{n-1}| \ .$$

The hypothesis $\mathbb{F}[V]^P = \mathbb{F}[\mathbf{N}^P_{P_{y_1}}(y_1), \mathbf{N}^P_{P_{y_2}}(y_2), \ldots, \mathbf{N}^P_{P_{y_n}}(y_n)]$ implies that $\prod_{i=1}^n \deg(\mathbf{N}^P_{P_{y_i}}(y_i)) = |P|$ by Corollary 3.1.6. Since $\deg(\mathrm{N}(y_k)) = |P|/|P_{y_i}|$ for all i, we have that $\prod_{i=1}^n |P|/|P_{y_i}| = |P|$ and therefore, $\prod_{i=1}^n |P_{y_i}| = |P|^{n-1}$. Thus $\prod_{i=1}^{n-1} |W_i| = |P|^{n-1}$. Since each of the sets W_i is a subset of P, we must have $W_i = P$ for all $i = 1, 2, \ldots, n-1$. In particular, $(P_{y_1} \cap P_{y_2} \cap \cdots \cap P_{y_{n-1}}) P_{y_n} = P$, i.e., $P_n P_{y_n} = P$ and thus $|P_n P_{y_n}| = |P|$. But, again by Lemma 8.0.10, $|P_n P_{y_n}| = \frac{|P_n| \cdot |P_{y_n}|}{|P_n \cap P_{y_n}|} = |P_n| \cdot |P_{y_n}|$ and hence $|P_n| \cdot |P_{y_n}| = |P|$. Since B is unordered we similarly obtain $|P_k| \cdot |P_{y_k}| = |P|$ for all $k = 1, 2 \ldots, n$.

Therefore,

$$\prod_{k=1}^n |P_k| = \prod_{k=1}^n \frac{|P|}{|P_{y_k}|} = \frac{|P|^n}{\prod_{k=1}^n |P_{y_k}|} = \frac{|P|^n}{|P|^{n-1}} = |P| \ .$$

Applying Lemma 8.0.6 shows that $|P| = |P_n P_{n-1} \cdots P_1|$ and therefore, $P = P_n P_{n-1} \cdots P_1$. It only remains to show that there is some ordering of B with respect to which the group P is upper triangular.

Define $V^*_0 := \{0\}$ and $V^*_k := \{v \in V^* \mid (\sigma - 1)v \in V^*_{k-1} \text{ for all } \sigma \in P\}$ for all $k \geq 1$. Let s be minimal such that $V^*_s = V^*$. Choose an ordered basis $\{x_1, x_2, \ldots, x_n\}$ of V^* which is compatible with the flag $\{0\} = V^*_0 \subset V^*_1 \subset \cdots \subset V^*_s = V^*$. Thus for each k with $1 \leq k \leq s$, there exits a j such that $\{x_1, x_2, \ldots, x_j\}$ is a basis of V^*_k. Thus P is upper triangular with respect to the basis $\{x_1, x_2, \ldots, x_n\}$.

Write $x_i = \sum_{j=1}^n \alpha_{ij} \, y_j$ for $i = 1, 2, \ldots, n$. For $1 \leq k \leq s$, define

$$W_k := \operatorname{span}\{y_j \in B \mid \alpha_{ij} \neq 0 \text{ for some } i \text{ with } x_i \in V_k^*\}\ .$$

We claim that $W_k = V_k^*$ for all $k = 1, 2, \ldots, s$. Take $x_i \in V_k^*$. Then $x_i = \sum_{j=1}^n \alpha_{ij}\, y_j$ with $y_j \in W_k$ for all j with $\alpha_{ij} \neq 0$. Thus $x_i \in W_k$. Therefore $V_k^* \subseteq W_k$.

For the opposite inclusion, let $y_j \in W_k$. Then there exists some t such that $x_t \in V_k^*$ and $\alpha_{tj} \neq 0$. Take any $\sigma \in P$. We want to show that $(\sigma-1)y_j \in V_{k-1}^*$. Write $\sigma = \sigma_n\sigma_{n-1}\cdots\sigma_1$ with $\sigma_a \in P_a$ for $a = 1, 2, \ldots, n$. Since P_a normalizes P_b when $a < b$, we may write $\sigma = \sigma_k\sigma'_n\sigma'_n \cdots \sigma'_{j+1}\sigma_{j-1}\cdots\sigma_1$ where $\sigma'_a \in P_a$ for $a = j+1, j+2, \ldots, n$. Thus $(\sigma-1)y_j = \sigma_j\sigma'_n\sigma'_n\cdots\sigma'_{j+1}\sigma_{j-1}\cdots\sigma_1(y_j) - y_j = (\sigma_j - 1)y_j$. But $(\sigma_j - 1)x_t = (\sigma_j - 1)(\sum_{i=1}^n \alpha_{ti}y_i) = \alpha_{tj}(\sigma_j - 1)y_j$. Since $(\sigma_j - 1)x_t \in V_{k-1}$ and since $\alpha_{tj} \neq 0$, this shows that $(\sigma - 1)y_j = (\sigma_j - 1)y_j \in V_{k-1}$. Therefore, $y_j \in V_k$. Thus, $W_k \subseteq V_k^*$.

This shows that we may order the y_j in such a way that B is compatible with the flag $\{0\} = V_0^* \subset V_1^* \subset \cdots \subset V_s^* = V^*$. Using this ordered basis for B we see that P is upper triangular. Hence P is a Nakajima group with respect to this ordered basis. □

In 1980, H. Nakajima [83] working over the prime field proved that, if $\mathbb{F}_p[V]^P$ is polynomial, then P is a Nakajima group. In the next section, we give an example, due to Stong, which shows that Nakajima's result cannot be extended naively to other fields.

8.1 Stong's Example

Here we give an example due to R Stong that shows that Nakajima's Theorem does not extend to larger fields. We work over the field $\mathbb{F}_q$ with $q = p^3$. We may suppose the field $\mathbb{F}_q$ has basis over $\mathbb{F}_p$ consisting of $\{1, \omega, \mu\}$. Let H be the group generated by the matrices

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and let G be the group generated by H and the matrix

$$\sigma = \begin{pmatrix} 1 & \omega & \mu \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

with respect to the basis $\{x, y, z\}$ of V^*. We note that both groups are generated by reflections, but that G is not a Nakajima group, since we cannot choose a basis of V^* with respect to which each generating reflection is concentrated in a single column.

It is not hard to see that $\mathbb{F}_q[V]^H = \mathbb{F}_q[x, N(y), N(z)]$, where $N(y) = y^p - x^{p-1}y$ and $N(z) = z^p - x^{p-1}z$. We calculate $\sigma(N(y)) = N(y) - (\omega^p - \omega)x^p$, and $\sigma(N(z)) = N(z) - (\mu^p - \mu)x^p$. From here we can construct two G invariants $f_1 = (\mu^p - \mu)N(y) - (\omega^p - \omega)N(z)$ and $f_2 = N(y)^p - (\omega^p - \omega)^{(p-1)}N(y)x^{p(p-1)}$. Using Lemma 2.6.3 it is not hard to see that $\{x, f_1, f_2\}$ form a homogeneous system of parameters, and thus by Corollary 3.1.6 that $\mathbb{F}_q[V] = \mathbb{F}_q[x, f_1, f_2]$ is a polynomial ring.

We may also use Theorem 3.9.2 to see immediately that $\mathbb{F}_q[V]^G$ must be a polynomial ring.

8.2 A Counterexample

For non-modular groups, we have the characterization of Shephard and Todd which asserts that $\mathbb{F}[V]^G$ is a polynomial ring if and only if the action of G on V is generated by reflections. It is known that this characterization fails for modular representations. For example, the representation described in Example 11.0.3 is generated by reflections but its ring of invariants is not polynomial. One of the most important open questions in modular invariant theory is to give a geometric characterization of modular representations of finite groups having a polynomial ring of invariants.

In 1982, V. Kac [57] made the following conjecture.

Conjecture 8.2.1. Let $\mathbb{K}$ be an algebraically closed field. Then $\mathbb{K}[V]^G$ is a polynomial ring if and only if each isotropy group G_v for $v \in V$ is generated by reflections.

In this section, we describe an unpublished counter-example to Kac's conjecture due to Campbell, Hughes and Shank in 1995. We note that Example 11.0.3 is also an example of such a reflection group whose ring of invariants is not polynomial. In their classification of irreducible reflection groups, Kemper and Malle [63, Example 2.2] also gave a counter-example.

Let $\mathbb{F}$ be an algebraically closed field of characteristic p and consider

$$G = \left\{ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ a & c & 1 & 0 \\ c & b & 0 & 1 \end{pmatrix} \mid a, b, c \in \mathbb{F}_p \right\}.$$

G is a subgroup of $GL_4(\mathbb{F})$ isomorphic to C_p^3. Note that G is a reflection group since it is generated by the following three reflections:

$$\alpha^{-1} := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},\ \beta^{-1} := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix},\ \gamma^{-1} := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}.$$

Theorem 8.2.2. *For all $v \in V$, the isotropy group G_v is a reflection group.*

Proof. Let $v = (a_1, a_2, b_1, b_2) \in V$. It is clear that if $a_1 = a_2 = 0$, then $G_v = G$. If $a_1 = 0$ and $a_2 \neq 0$, then G_v is generated by α. If $a_1 \neq 0$ and $a_2 = 0$ then G_v is generated by β. Finally, if $a_1 \neq 0$ and $a_2 \neq 0$, then G_v is generated by the reflection $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{a_2}{a_1} & -1 & 1 & 0 \\ -1 & \frac{a_1}{a_2} & 0 & 1 \end{pmatrix}$.

Each of these subgroups is a reflection group and thus all of the isotropy subgroups are reflection groups. □

Remark 8.2.3. Since each proper non-trivial isotropy subgroup is generated by a single transvection, the ring of invariants for each of these groups is a polynomial algebra.

Theorem 8.2.4. $\mathbb{F}[V]^G$ *is not a polynomial algebra.*

Proof. Let $\{x_1, x_2, y_1, y_2\}$ denote the basis of V^* dual to the canonical basis of V. Thus $\alpha(y_1) = y_1 + x_1$ and $\beta(y_2) = y_2 + x_2$. If $\mathbb{F}[V]^G$ were a polynomial algebra, it would have four generators and the product of the degrees of the generators would be $|G| = p^3$. Observe that x_1 and x_2 are invariant and that $\dim_{\mathbb{K}} \mathbb{F}[V]^G_1 = 2$. Thus if $\mathbb{F}[V]^G$ is a polynomial algebra, it has one generator in degree p and one in degree p^2. We will show that $\mathbb{F}[V]^G$ does not have a generator in degree p.

Consider the subgroup H of G generated by α and β. Then H is a Nakajima group isomorphic to $C_p \times C_p$ and $\mathbb{F}[V]^H \cong \mathbb{F}[x_1, x_2, N_1, N_2]$ where $N_1 = y_1^p - x_1^{p-1} y_1$ and $N_2 = y_2^p - x_2^{p-1} y_2$. Since $\mathbb{F}[V]^G \subseteq \mathbb{F}[V]^H$, the degree p elements of $\mathbb{F}[V]^G$ are contained in the vector space spanned by N_1, N_2 and the monomials in x_1 and x_2 of degree p. Since $\gamma(N_1) = N_1 + (x_2^p - x_1^{p-1} x_2)$ and $\gamma(N_2) = N_2 + (x_1^p - x_2^{p-1} x_1)$ we see that γ does not fix any non-zero linear combination of N_2 and N_1. Hence $\mathbb{F}[V]^G$ does not have a generator in degree p and is not a polynomial algebra. □

Remark 8.2.5. We show in Example 10.0.11 that $\mathbb{F}[V]^G$ is a hypersurface ring with two generators x_1 and x_2 of degree 1, a generator $x_1 N_1 + x_2 N_2$ of degree $p + 1$ and two generators $\mathrm{N}_H^G(N_1)$ and $\mathrm{N}_H^G(N_2)$ of degree p^2. Note that Proposition 11.0.1 can also be used to show this.

8.3 Irreducible Modular Reflection Groups

In this section, we recall the work of Kemper and Malle [63]. Suppose G is represented on V over a field $\mathbb{F}$ of characteristic p dividing the order of G and that V is irreducible as a G-module. Of course, in modular representation theory, the irreducible representations are essentially few and far between, as we have just seen for C_p. Nevertheless, such representations have been classified. Using the classification, Kemper and Malle have proved the following

Theorem 8.3.1. *Suppose V is an irreducible representation of the modular group G. Then $\mathbb{F}[V]^G$ is a polynomial ring if and only if G is generated by reflections and if W is any non-trivial subspace of V, then $\mathbb{F}[V]^{G_W}$ has a polynomial ring of invariants.*

We note that $\mathbb{F}[V]^{G_W}$ polynomial implies that G_W is generated by reflections.

Remark 8.3.2. As observed by Kemper and Malle, the theorem is equivalent to the following. Suppose v is any point of $\overline{V} = \overline{\mathbb{F}} \otimes_{\mathbb{F}} V$, where $\overline{\mathbb{F}}$ denotes the algebraic closure of $\mathbb{F}$, and that G_v denotes the stabilizer of v. Then $\mathbb{F}[V]^G$ is polynomial if and only if G is generated by reflections and $\mathbb{F}[V]^{G_v}$ is polynomial.

Remark 8.3.3. Kemper and Malle, see [68], also proved that the invariant field, $\mathbb{F}(V)^G$, of an irreducible reflection group is purely transcendental.

It is not sufficient to assume that each G_v is a reflection group.

Example 8.3.4. Let $G = \Sigma_6$ act on its 4-dimensional irreducible representation V in characteristic 2. Then G_v is generated by reflections for all non-zero $v \in V$, but $\mathbb{F}[V]^G$ is not polynomial.

8.3.1 Reflection Groups

As noted above, a reflection $\sigma \in \mathrm{GL}(V)$ is said to be a *transvection* if it is not diagonalizable. If it is diagonizable, it is said to be a *homology.*

We denote by $\mathrm{RL}_n(\mathbb{F})$ the subgroup of $\mathrm{GL}_n(\mathbb{F})$ generated by its reflections. Similarly, we let $\mathrm{RU}_n(\mathbb{F})$ denote the subgroup of $\mathrm{GU}_n(\mathbb{F})$, the unitary group of dimension n over $\mathbb{F}$, generated by its reflections. Then $\mathrm{RL}_n(\mathbb{F})$ contains $\mathrm{SL}_n(\mathbb{F})$ as a subgroup of index 2. Similarly respectively $\mathrm{RU}_n(\mathbb{F})$ contains $\mathrm{SU}_n(\mathbb{F})$ as a subgroup of index 2.

We denote by $\mathrm{GO}_n(\mathbb{F})$ the general orthogonal group over $\mathbb{F}$ and by $\mathrm{SO}_n(\mathbb{F})$, the special orthogonal group over $\mathbb{F}$. And we denote by $\Omega_n^{\pm}(\mathbb{F})$ the commutator subgroup of $\mathrm{GO}_n(\mathbb{F})$. Finally, we denote by G_i the i-th group in the list of finite reflection groups given by Shephard and Todd [101] and by $W_p(G_i)$ the mod p reduction of its reflection representation. The irreducible subgroups of $\mathrm{GL}_n(\mathbb{F})$ generated by reflections were classified by Wagner [109] and by Zalesskiĭ, and Serežkin [115], for $n \geq 3$.

Theorem 8.3.5. *Let V be a vector space of dimension $n \geq 3$ over $\mathbb{F}$ of characteristic $p > 2$ and let G denote a subgroup of $\mathrm{GL}(\mathbb{F})$ generated by reflections such that V is irreducible and primitive as a G-module. Then one of the following holds:*

1. $G = \mathrm{RL}_n(\mathbb{F})$, *or* $G = \mathrm{RU}_n(\mathbb{F})$;
2. $\Omega_n^{\pm}(\mathbb{F}) \subsetneq G \subseteq \mathrm{GO}_n^{\pm}(\mathbb{F})$, $G \neq \mathrm{SO}_n^{\pm}(\mathbb{F})$;
3. $G = \Sigma_{n+1}$ *for* $p \nmid (n+1)$;

4. $G = \Sigma_{n+2}$ *for* $n \geq 5$ *and* $p \nmid (n+2)$*;*
5. $G = W_p(G_i)$ *for* $23 \leq i \leq 37$*,* $i \neq 25, 26, 32$ *and* $p \nmid |G|$ *or* $p \in \{3, 5, 7\}$ *corresponds to the columns marked p or np in [69, Table 6.3];*
6. $p = 5$*,* $n = 3$ *and* G *is the group* $3 \cdot A_7 \times 2$ *(sometimes denoted* $EJ_3(5)$*);*
7. $p = 3$*,* $n = 4$ *and* G *is the group* $4 \cdot L_3(4) : 2_2$ *(sometimes denoted* $J_4(4)$*).*

8.3.2 Groups Generated by Homologies of Order Greater than 2

These groups and representations were classified by Mitchell in characteristic 0 and by Wagner in positive characteristic, see [108].

Theorem 8.3.6. *Let* V *be a vector space of dimension* $n \geq 3$ *over* $\mathbb{F}$ *of characteristic* $p > 2$ *and let* G *denote a subgroup of* $\mathrm{GL}(\mathbb{F})$ *generated by pseudoreflections, containing homologies of order greater than* 2 *such that* V *is irreducible and primitive as a* G*-module. Then one of the following holds:*

1. $\mathrm{SL}_n(\mathbb{F}) \subset G \subset \mathrm{GL}_n(\mathbb{F})$ *or* $\mathrm{SU}_n(\mathbb{F} \subset G \subset \mathrm{GU}_n(\mathbb{F})$*,* $G \neq \mathrm{RL}_n(\mathbb{F}), \mathrm{RU}_n(\mathbb{F})$*;*
2. $G = W_p(G_{25})$ *and* $p \neq 3$*;*
3. $G = W_p(G_{26})$ *and* $p \neq 2, 3$*;*
4. $G = W_p(G_{32})$ *and* $p \neq 3$*.*

8.3.3 Groups Generated by Transvections

There is no analogue for these groups in characteristic 0. We refer to [58].

Theorem 8.3.7. *Let* V *be a vector space of dimension* n *over* $\mathbb{F}$*, a finite field of characteristic* p*. Let* G *denote a subgroup of* $\mathrm{GL}(\mathbb{F})$ *generated by transvections such that* V *is irreducible and primitive as a* G*-module. Then one of the following holds:*

1. $G = \mathrm{SL}_n(\mathbb{F}), \mathrm{Sp}_n(\mathbb{F}), \mathrm{SU}_n(\mathbb{F})$ *and* $p \neq 2$*,* $n \neq 3$ *;*
2. $G = \mathrm{SO}_n(\mathbb{F})$*,* $p = 2$ *and* $n \geq 4$ *is even;*
3. $G = \Sigma_{n+1}$ *or* Σ_{n+2}*,* $p = 2$ *and* $n \geq 6$ *is even;*
4. $G = \Sigma_{n+2}$ *for* $n \geq 5$ *and* $p \nmid (n+2)$*;*
5. $G = \mathrm{SL}_2(\mathbb{F})$*,* $p = 5$ *and* $n = 3$*;*
6. $p = 2$*,* $n = 3$ *and* G *is the group* $3 \cdot A_6$*;*
7. $p = 2$*,* $n = 6$ *and* G *is the group* $3 \cdot U_4(3) : 2_2$*.*

9

The Transfer

In this chapter, we consider in detail the transfer (also called the trace) map introduced in §1.2. Let H be a subgroup of the finite group G. Choose a set of left coset representatives for H in G. We denote this set of representatives by G/H. Thus $G = \sqcup_{\sigma \in G/H} \sigma H$ is a decomposition of G into left cosets. There is an extensive theory considering the relative versions of the results of this chapter, see Fleischmann [38] or Fleischmann and Shank [41].

As in §1.2, we define the relative transfer

$$\begin{array}{rcl} \mathrm{Tr}^G_H : \mathbb{K}[V]^H & \longrightarrow & \mathbb{K}[V]^G \\ f & \longmapsto & \sum\limits_{\sigma \in G/H} \sigma(f) \ . \end{array}$$

To see that the image of Tr^G_H lies in $\mathbb{K}[V]^G$, let $f \in \mathbb{K}[V]^H$ and take an arbitrary group element $\sigma_0 \in G$. Write $G/H = \{\sigma_1, \sigma_2, \ldots, \sigma_r\}$ where r is the index of H in G. Then $\{\sigma_0\sigma_1, \sigma_0\sigma_2, \ldots, \sigma_0\sigma_r\}$ is another set of left coset representatives for H in G. Therefore, there exist $\tau_1, \tau_2, \ldots, \tau_r \in H$ such that $\{\sigma_0\sigma_1, \sigma_0\sigma_2, \ldots, \sigma_0\sigma_r\} = \{\sigma_1\tau_1, \sigma_2\tau_2, \ldots, \sigma_r\tau_r\}$. Therefore, $\sigma_0(\mathrm{Tr}^G_H(f)) = \sigma_0(\sum_{i=1}^r \sigma_i(f)) = \sum_{i=1}^r \sigma_0\sigma_i(f) = \sum_{i=1}^r \sigma_i\tau_i(f) = \sum_{i=1}^r \sigma_i(f) = \mathrm{Tr}^G_H(f)$.

Lemma 9.0.1. *The map* Tr^G_H *is independent of the choice of coset representatives.* □

The most important of these maps is the map $\mathrm{Tr}^G_{\{e\}}$ which we denote by Tr^G.

In general, the transfer map does not behave well with respect to products. However, if $f_1 \in \mathbb{F}[V]^G$ and $f_2 \in \mathbb{F}[V]^H$ then we have

$$\begin{aligned} \mathrm{Tr}^G_H(f_1 f_2) &= \sum_{\sigma \in G/H} \sigma(f_1 f_2) = \sum_{\sigma \in G/H} \sigma(f_1)\sigma(f_2) \\ &= \sum_{\sigma \in G/H} f_1 \sigma(f_2) = f_1 \sum_{\sigma \in G/H} \sigma(f_2) = f_1 \, \mathrm{Tr}^G_H(f_2). \end{aligned}$$

H.E.A. Eddy Campbell, D. Wehlau, *Modular Invariant Theory*,
Encyclopaedia of Mathematical Sciences 139,
DOI 10.1007/978-3-642-17404-9_9, © Springer-Verlag Berlin Heidelberg 2011

Thus Tr_H^G is an $\mathbb{F}[V]^G$-module homomorphism. Also, note that the transfer map preserves degree.

Notice that if $f \in \mathbb{F}[V]^G$, then $\mathrm{Tr}_H^G(f) = \sum_{\sigma \in G/H} \sigma(f) = \frac{|G|}{|H|} f = [G : H]f$. Thus the composition

$$\mathbb{F}[V]^G \hookrightarrow \mathbb{F}[V]^H \xrightarrow{\mathrm{Tr}_H^G} \mathbb{F}[V]^G$$

is just multiplication by $[G : H]$. Therefore, if $[G : H]$ is invertible in $\mathbb{F}$, then

$$\frac{1}{[G:H]} \mathrm{Tr}_H^G : \mathbb{K}[V]^H \to \mathbb{K}[V]^G$$

is a projection operator. This projection onto the invariants is known as the *Reynolds Operator.* In particular, using the Reynolds Operator, we see that for non-modular representations, we have $\mathbb{F}[V]^H = \mathbb{F}[V]^G \oplus (\ker \mathrm{Tr}_H^G)$. For this reason, the transfer is well understood and accordingly less interesting in the non-modular case.

The following well-known lemma is very useful in studying the image of the transfer for modular groups.

Lemma 9.0.2. *Let $q = p^r$ be a prime power and suppose that ℓ is a positive integer. Then*

$$\sum_{c \in \mathbb{F}_q} c^\ell = \begin{cases} -1, & \textit{if } q-1 \textit{ divides } \ell; \\ 0, & \textit{if } q-1 \textit{ does not divide } \ell. \end{cases}$$

Proof. It is sufficient to sum over the non-zero elements of $\mathbb{F}_q$. Let η be a generator for the group of units of $\mathbb{F}_q$. Hence $\mathbb{F}_q \setminus \{0\} = \{\eta^0, \eta, \eta^2, \ldots, \eta^{q-2}\}$ and, therefore,

$$\sum_{c \in \mathbb{F}_q} c^\ell = \sum_{i=0}^{q-2} (\eta^i)^\ell = \sum_{i=0}^{q-2} (\eta^\ell)^i.$$

Every non-zero element of $\mathbb{F}_q$, including in particular η^ℓ, is a root of the polynomial $x^{q-1} - 1 = (x-1)(1+x+\ldots+x^{q-2})$. If $q-1$ divides ℓ, then $\eta^\ell = 1$ and the sum is $q-1$. If $q-1$ does not divide ℓ, then $\eta^\ell \neq 1$ and therefore, η^ℓ is a root of $1 + x + \ldots + x^{q-2}$. Hence if $q-1$ does not divide ℓ, then the sum is zero. □

Remark 9.0.3. In the above lemma, ℓ is a positive integer. However, it will be convenient to consider the case where $\ell = 0$. In that case, the sum, $\sum_{c \in \mathbb{F}_q} c^0$, involves 0^0 and so is not well-defined. In most of the applications we will use, the sum will most naturally be 0 and so we will consider 0^0 to be equal to 1. Thus even though $q-1$ does divide ℓ when $\ell = 0$, the sum in our applications will be 0.

Example 9.0.4. We consider the two dimensional irreducible representation, V_2, of the cyclic group of order p over a field $\mathbb{F}$ of characteristic p. We choose a basis $\{x, y\}$ for V_2^* and a generator σ of C_p such that $\sigma(y) = y + x$ and $\sigma(x) = x$.

Since the ring of invariants $\mathbb{F}[V_2]^{C_p} = \mathbb{F}[x, N = y^p - x^{p-1}y]$ is a polynomial subring, we have that $\mathbb{F}[V_2] = \mathbb{F}[x, y]$ is a free $\mathbb{F}[V_2]^{C_p}$-module. In fact, it is easy to see that $\mathbb{F}[V_2] = \oplus_{j=0}^{p-1}\mathbb{F}[V_2]^{C_p}\, y^j$. We use this Hironaka decomposition to study the transfer map:

$$\mathrm{Tr}^{C_p} : \mathbb{F}[V_2] = \oplus_{j=0}^{p-1}\mathbb{F}[V_2]^{C_p}\, y^j \to \mathbb{F}[V_2]^{C_p}\ .$$

We have $\mathrm{Tr}^{C_p} = \sum_{i=0}^{p-1} \sigma^i$ and therefore, $\mathrm{Tr}^{C_p}(f(x, y)) = \sum_{i=0}^{p-1} f(x, y + ix)$. It is easy to show that $\sigma^i(y) = y + ix$. Thus $\sigma^i(y^j) = (y + ix)^j$ and therefore

$$\begin{aligned}
\mathrm{Tr}^{C_p}(y^j) &= \sum_{i=0}^{p-1}(y + ix)^j \\
&= \sum_{i=0}^{p-1}\sum_{t=0}^{j}\binom{j}{t} y^{t-j}(ix)^t \\
&= \sum_{t=0}^{j}\binom{j}{t} y^{t-j}x^t(\sum_{i=0}^{p-1} i^t)\ .
\end{aligned}$$

Thus by Lemma 9.0.2, we see that $\mathrm{Tr}^{C_p}(y^j) = 0$ unless $j = p - 1$ in which case we have $\mathrm{Tr}^{C_p}(y^{p-1}) = \binom{p-1}{p-1}y^0x^{p-1}(\sum_{i=0}^{p-1} i^{p-1}) = -x^{p-1}$.

Given $f \in \mathbb{F}[V_2]$, we write $f = \sum_{j=0}^{p-1} f_j y^j$ where $f_j \in \mathbb{F}[V_2]^{C_p}$ for each $j = 0, 1, \ldots, p - 1$. Then $\mathrm{Tr}^{C_p}(f) = \sum_{j=0}^{p-1} \mathrm{Tr}^{C_p}(f_j y^j) = \sum_{j=0}^{p-1} f_j \mathrm{Tr}^{C_p}(y^j) = -f_{p-1}x^{p-1}$. Thus the image of Tr^{C_p} is the principal ideal of $\mathbb{F}[V_2]^{C_p}$ generated by x^{p-1}.

Example 9.0.5. Take G to be the symmetric group on 3 letters and $\mathbb{K}$ to be any field. Let G act on $V = \mathbb{K}^3$ by permuting the basis $\{e_1, e_2, e_3\}$ of V. Write $\{x, y, z\}$ for the dual basis of V^*. Then G also permutes $\{x, y, z\}$. Then $\mathbb{K}[V]^G = \mathbb{K}[s_1, s_2, s_3]$ where $s_1 = x + y + z$, $s_2 = xy + xz + yz$ and $s_3 = xyz$.

We have $G = \{e, (12), (13), (23), (123), (132)\}$ where for example $(123) \cdot ae_1 + be_2 + ce_3 = ae_2 + be_3 + ce_1$ and $(123) \cdot y = z$.

Then $\mathrm{Tr}^G(f) = f + (12) \cdot f + (13) \cdot f + (23) \cdot f + (123) \cdot f + (132) \cdot f$. Thus $\mathrm{Tr}^G(f(x, y, z)) = f(x, y, x) + f(y, x, z) + f(z, y, x) + f(x, z, y) + f(y, z, x) + f(z, x, y)$.

Let N denote the normal subgroup of index 2 in G, the alternating group on 3 letters. Thus $\mathrm{Tr}^N(f) = f + (123) \cdot f + (132) \cdot f = f(x, y, z) + f(y, z, x) + f(z, x, y)$.

The coset decomposition of G with respect to N is $G = N \sqcup (12)N = \{e, (123), (132)\} \sqcup \{(12), (13), (23)\}$. We have $\mathrm{Tr}_N^G(f) = f + (12) \cdot f$ for $f \in \mathbb{F}[V]^N$. Thus $\mathrm{Tr}_N^G(f(x, y, z)) = f(x, y, z) + f(y, x, z)$.

Now consider the subgroup $H = \{e, (12)\}$ of G of order 2 which fixes e_3 and z. Then $\mathrm{Tr}^H(f) = f + (12) \cdot f = f(x, y, z) + f(y, x, z)$.

Decomposing G as a union of left H-cosets gives

$$G = H \sqcup (123)H \sqcup (132)H = \{e, (12)\} \sqcup \{(123), (13)\} \sqcup \{(132), (23)\}.$$

Hence

$$\mathrm{Tr}^G_H(f) = f + (123) \cdot f + (132) \cdot f$$

for $f \in \mathbb{F}[V]^H$. Therefore, $\mathrm{Tr}^G_H(f(x, y, z)) = f(x, y, z) + f(y, z, x) + f(z, x, y)$.

To study the full transfer homomorphism, Tr^G, we exploit the block basis and corresponding Hironaka decomposition given in Proposition 6.1.1:

$$\mathbb{K}[V] \cong \mathbb{K}[V]^G \oplus \mathbb{K}[V]^G\, x \oplus \mathbb{K}[V]^G\, y \oplus \mathbb{K}[V]^G\, xy \oplus \mathbb{K}[V]^G\, x^2 \oplus \mathbb{K}[V]^G\, x^2 y\ .$$

Let $f \in \mathbb{K}[V]$ and write $f = f_0 + f_1 x + f_2 y + f_3 xy + f_4 x^2 + f_5 x^2 y$ where $f_i \in \mathbb{K}[V]^G$ for all $i = 0, 1, \ldots, 5$.

It is easy to compute that

$$\begin{aligned}
\mathrm{Tr}^G(1) &= |G| = 6,\\
\mathrm{Tr}^G(x) &= \mathrm{Tr}^G(y) = 2(x + y + z) = 2s_1,\\
\mathrm{Tr}^G(xy) &= 2(xy + xz + yz) = 2s_2,\\
\mathrm{Tr}^G(x^2) &= 2(x^2 + y^2 + z^2) = 2(s_1^2 - 2s_2),\ \text{and}\\
\mathrm{Tr}^G(x^2y) &= (x^2y + x^2z + xy^2 + xz^2 + y^2z + yz^2) = s_1 s_2 - 3s_3\ .
\end{aligned}$$

Thus

$$\begin{aligned}
\mathrm{Tr}^G(f) &= \mathrm{Tr}^G(f_0 + f_1 x + f_2 y + f_3 xy + f_4 x^2 + f_5 x^2 y)\\
&= \mathrm{Tr}^G(f_0) + \mathrm{Tr}^G(f_1 x) + \mathrm{Tr}^G(f_2 y) + \mathrm{Tr}^G(f_3 xy) + \mathrm{Tr}^G(f_4 x^2)\\
&\quad + \mathrm{Tr}^G(f_5 x^2 y)\\
&= f_0 \mathrm{Tr}^G(1) + f_1 \mathrm{Tr}^G(x) + f_2 \mathrm{Tr}^G(y) + f_3 \mathrm{Tr}^G(xy) + f_4 \mathrm{Tr}^G(x^2)\\
&\quad + f_5 \mathrm{Tr}^G(x^2 y)\\
&= 6f_0 + 2f_1 s_1 + 2f_2 s_1 + 2f_3 s_2 + 2f_4(s_1^2 - 2s_2)\\
&\quad + f_5(s_1 s_2 - 3s_3)\ .
\end{aligned}$$

Lemma 9.0.6. *For G a finite group, if $N \le H \le G$ is a sequence of subgroups of G, then $\mathrm{Tr}^G_N = \mathrm{Tr}^G_H \circ \mathrm{Tr}^H_N$.*

We may use right cosets to define another useful map. If H is a subgroup of G, choose a set Ω of right coset representatives for H in G. and define $\widehat{\mathrm{Tr}^G_H} : \mathbb{K}[V] \to \mathbb{K}[V]$ by $\widehat{\mathrm{Tr}^G_H}(f) = \sum_{\sigma \in \Omega} \sigma(f)$. This map depends upon the choice of coset representatives used. However, for all possible choices we have the following factorization of Tr^G_N.

Lemma 9.0.7. *Let N be a normal subgroup of the finite group G and suppose H is another subgroup of G with $N \le H \le G$. Then $\mathrm{Tr}^G_N = \mathrm{Tr}^H_N \circ \widehat{\mathrm{Tr}^G_H}$.*

Proof. Write $\Omega = \sigma_1, \sigma_2, \ldots, \sigma_r$ for the right coset representatives of H in G chosen to define the map $\widehat{\mathrm{Tr}}^G_H$. Then we have the decomposition $G = \sqcup_{i=1}^r H\sigma_i$ of G into right H-cosets.

Let $H = \sqcup_{j=1}^s \tau_j N$ be a decomposition of H into left N-cosets. Then $G = \sqcup_{i=1}^r \sqcup_{j=1}^s \tau_j N \sigma_i$. Since N is normal in G, this gives $G = \sqcup_{i=1}^r \sqcup_{j=1}^s \tau_j \sigma_i N$.

Thus for any $f \in \mathbb{K}[V]^N$,

$$\mathrm{Tr}^G_N(f) = \sum_{i=1}^r \sum_{j=1}^s \tau_j \sigma_i(f) = \sum_{j=1}^s \sum_{i=1}^r \tau_j \sigma_i(f) = \sum_{j=1}^s \tau_j \widehat{\mathrm{Tr}}^G_H(f) \ .$$

We claim that $\sum_{j=1}^s \tau_j \widehat{\mathrm{Tr}}^G_H(f) = \mathrm{Tr}^H_N(\widehat{\mathrm{Tr}}^G_H(f))$. In order to see this, we only need show that $\widehat{\mathrm{Tr}}^G_H(f)$ lies in the domain of Tr^G_N, i.e., that $\widehat{\mathrm{Tr}}^G_H(f)$ is N-invariant. To see this, let $n \in N$ be arbitrary and note that f is N-invariant. Thus writing $n_i = \sigma_i^{-1} n \sigma_i \in N$ we have $n \cdot \widehat{\mathrm{Tr}}^G_H(f) = n \cdot \sum_{i=1}^r \sigma_i(f) = \sum_{i=1}^r (n\sigma_i)(f) = \sum_{i=1}^r \sigma_i n_i(f) = \sum_{i=1}^r \sigma_i(f) = \widehat{\mathrm{Tr}}^G_H(f)$, as required. □

The following corollary is immediate.

Corollary 9.0.8. $\mathrm{Tr}^G_N(\mathbb{K}[V]^N) \subseteq \mathrm{Tr}^H_N(\mathbb{K}[V]^N)$.

Example 9.0.9. We continue with Example 9.0.5. Examples of factorizations of Tr^G include $\mathrm{Tr}^G(f) = \mathrm{Tr}^G_N(\mathrm{Tr}^N(f)) = \mathrm{Tr}^G_H(\mathrm{Tr}^H(f))$ and $\mathrm{Tr}^G(f) = \mathrm{Tr}^H(\widehat{\mathrm{Tr}}^G_H(f))$ for $f \in \mathbb{K}[V]$. This latter factorization gives $\mathrm{Tr}^G(f(x,y,z)) = \mathrm{Tr}^H(h(x,y,z)) = h(x,y,z) + h(y,x,z)$ where for one choice of coset representatives, we have $h(x,y,z) = \widehat{\mathrm{Tr}}^G_H(f(x,y,z)) = f(x,y,z) + f(y,z,x) + f(z,y,x)$.

As we noted above, the transfer map Tr^G_H is a homogeneous $\mathbb{F}[V]^G$-module homomorphism. In particular, its image, $\mathrm{Tr}^G_H(\mathbb{F}[V]^H)$, is a graded $\mathbb{F}[V]^G$-submodule of $\mathbb{F}[V]^G$, that is to say, the image of the transfer, $\mathrm{Tr}^G_H(\mathbb{F}[V]^H)$, is a homogeneous ideal of $\mathbb{F}[V]^G$. We will be interested in describing the algebraic subset, X, of $V/\!\!/G$ (more precisely, its lift $\pi^{-1}_{V,G}(X) \subset V$) corresponding to this ideal.

We will assume throughout the rest of this chapter that the base field has positive characteristic p unless stated otherwise.

Theorem 9.0.10. *Let N be a normal subgroup of the finite group G and suppose $\mathbb{F}$ has characteristic p. Then*

$$\mathcal{V}_{\overline{V}}(\mathrm{Tr}^G_N(\mathbb{F}[V]^N)) = \bigcup_{\mathrm{order}(\sigma N)=p} \overline{V}^{\sigma}$$

where the union is over all quotient group elements $\sigma N \in G/N$ of order p.

Proof. First, we prove

$$\mathcal{V}(\mathrm{Tr}^G_N(\mathbb{F}[V]^N)) \supseteq \cup_{\mathrm{order}(\sigma N)=p}\overline{V}^\sigma .$$

Let $\mathbf{v} \in \cup_{\mathrm{order}(\sigma N)=p}\overline{V}^\sigma$, i.e., suppose there exists $\sigma \in G \setminus N$, with $\sigma^p \in N$ and $\sigma\mathbf{v} = \mathbf{v}$.

Let H denote the subgroup of G generated by N and σ. Thus $[H : N] = p$. Take $f \in \mathbb{F}[V]^N$. Then

$$\mathrm{Tr}^H_N(f) = \sum_{i=0}^{p-1} \sigma^i(f)$$

and

$$(\mathrm{Tr}^H_N(f))(\mathbf{v}) = \sum_{i=0}^{p-1} f(\sigma^{-i}\mathbf{v}) = \sum_{i=0}^{p-1} f(\mathbf{v}) = pf(\mathbf{v}) = 0.$$

By the previous corollary, $(\mathrm{Tr}^G_N(f))(\mathbf{v}) = 0$ also. Therefore,

$$\mathbf{v} \in \mathcal{V}_{\overline{V}}\left(\mathrm{Tr}^G_N(\mathbb{F}[V]^N)\right).$$

Next, we prove the opposite inclusion, $\mathcal{V}(\mathrm{Tr}^G_N(\mathbb{F}[V]^N)) \subseteq \cup_{\mathrm{order}(\sigma N)=p}\overline{V}^\sigma$. Take $\mathbf{v} \in V \setminus \cup_{\mathrm{order}(\sigma N)=p}\overline{V}^\sigma$. By Corollary 2.1.3, there exists $h \in \mathbb{F}[V]$ such that

$$h(w) = \begin{cases} 1 & \text{if } w \in N\mathbf{v}, \\ 0 & \text{if } w \in G\mathbf{v} \setminus N\mathbf{v}. \end{cases}$$

Notice that it can happen that $\sigma \notin N$ but $\sigma\mathbf{v} \in N\mathbf{v}$, i.e., $\sigma\mathbf{v} = \tau\mathbf{v}$ for some $\tau \in N$ and some $\sigma \in G \setminus N$. In this case, $\tau^{-1}\sigma \in G_{\mathbf{v}}$ and therefore, $\sigma \in N \cdot G_{\mathbf{v}}$ where since N is normal in G, the set $N \cdot G_{\mathbf{v}}$ is a subgroup of G. Conversely, if σ is any element of $N \cdot G_{\mathbf{v}}$, then there exist $\tau \in N$ and $\sigma_1 \in G_{\mathbf{v}}$ such that $\sigma = \tau\sigma_1$ and thus $\sigma\mathbf{v} = \tau\mathbf{v} \in N\mathbf{v}$. In conclusion, for $\sigma \in G$, we have $\sigma\mathbf{v} \in N\mathbf{v}$ if and only if $\sigma \in N \cdot G_{\mathbf{v}}$.

Let $f := \prod_{\tau \in N} \tau(h) \in \mathbb{F}[V]^N$. For all $\tau \in N$, we have $\tau^{-1}w \in N\mathbf{v}$ if and only if $w \in N\mathbf{v}$ and thus

$$f(w) = \prod_{\tau \in N} h(\tau^{-1}w) = \begin{cases} 1 & \text{if } w \in N\mathbf{v}, \\ 0 & \text{if } w \in G\mathbf{v} \setminus N\mathbf{v}. \end{cases}$$

Let $G = \sqcup_{i=1}^t \sigma_i N$ be a decomposition of G into left N-cosets. Then

$$\begin{aligned}(\mathrm{Tr}^G_N(f))(\mathbf{v}) &= \sum_{i=1}^t f(\sigma_i^{-1}\mathbf{v}) = |\{i \mid \sigma_i^{-1}\mathbf{v} \in N\mathbf{v}\}| \\ &= |\{i \mid \sigma_i^{-1} \in N \cdot G_{\mathbf{v}}\}| = |\{i \mid \sigma_i \in N \cdot G_{\mathbf{v}}\}| .\end{aligned}$$

Consider a right coset $N\sigma$ of N and suppose this right coset meets $N \cdot G_{\mathbf{v}}$. Then we have $\tau_1\sigma = \tau\sigma_1$ where $\tau_1, \tau \in N$ and $\sigma_1 \in G_{\mathbf{v}}$. Now let $\tau_2\sigma$ be any other element of the coset $N\sigma$. Then $\tau_2\sigma = \tau_2\tau_1^{-1}\tau\sigma_1$ is also in $N \cdot G_{\mathbf{v}}$. Thus if the coset $N\sigma$ meets $N \cdot G_{\mathbf{v}}$, it is contained in $N \cdot G_{\mathbf{v}}$. Since N is normal, every right coset is also a left coset: $N\sigma = \sigma N$, and thus we see that

$$N \cdot G_{\mathbf{v}} = \bigsqcup_{\sigma_i \in N \cdot G_{\mathbf{v}}} \sigma_i N \ .$$

Therefore,

$$\begin{aligned}(\mathrm{Tr}_N^G(f))(\mathbf{v}) &= |\{i \mid \sigma_i \in N \cdot G_{\mathbf{v}}\}| = |\frac{N \cdot G_{\mathbf{v}}}{N}| \\ &= |\frac{G_{\mathbf{v}}}{N \cap G_{\mathbf{v}}}| \text{ by the second isomorphism theorem for groups} \\ &= |\frac{G_{\mathbf{v}}}{N_{\mathbf{v}}}| \ .\end{aligned}$$

Thus it remains to prove that p does not divide the index of $N_{\mathbf{v}}$ in $G_{\mathbf{v}}$. Assume by way of contradiction that p does divide $[G_{\mathbf{v}} : N_{\mathbf{v}}]$. Then there exists $\sigma_1 \in G_{\mathbf{v}}$ such that $\sigma_1 N_{\mathbf{v}}$ has order p in $\frac{G_{\mathbf{v}}}{N_{\mathbf{v}}}$, i.e., $\sigma_1 \in G_{\mathbf{v}} \setminus N_{\mathbf{v}}$ and $\sigma_1^p \in N_{\mathbf{v}}$. Since $N_{\mathbf{v}} = G_{\mathbf{v}} \cap N$, we see that $\sigma_1 \notin N$. Thus $\sigma_1 N$ has order p as an element of G/N. Also, $\mathbf{v} \in \overline{V}^{\sigma_1}$ since $\sigma_1 \in G_{\mathbf{v}}$. This contradicts our original assumption on $\mathbf{v}$ and this contradiction shows that p does divide $|\frac{G_{\mathbf{v}}}{N_{\mathbf{v}}}|$. Therefore $(\mathrm{Tr}_N^G(f))(\mathbf{v}) \neq 0$ and hence $\mathbf{v} \notin \mathcal{V}_{\overline{V}}(\mathrm{Tr}_N^G(\mathbb{F}[V]^N))$. □

Of course, the most important case of the above Corollary is when $N = \{e\}$. In that case we have

$$\mathcal{V}_{\overline{V}}\left(\mathrm{Tr}^G(\mathbb{F}[V])\right) = \bigcup_{\mathrm{order}(\sigma)=p} \overline{V}^{\sigma}.$$

Thus the subvariety of V corresponding to the ideal of $\mathbb{F}[V]$ generated by the elements in the image of the transfer consists of precisely those points x in V for which p divides the order of the isotropy group G_x.

Example 9.0.11. We continue with Example 9.0.9. We have seen that the image of Tr^{Σ_3} is the ideal in $\mathbb{K}[V]^{\Sigma_3} = \mathbb{K}[s_1, s_2, s_3]$ generated by $6, 2s_1, 2s_2, 2(s_1^2 - 2s_2), s_1s_2 - 3s_3$. Thus if the characteristic of $\mathbb{K}$ is not 2 nor 3, then Tr^{Σ_3} is surjective.

If $\mathbb{K}$ has characteristic 2, then the image of the transfer is the principal ideal generated by

$$\begin{aligned}s_1s_2 - s_3 &= (x+y+z)(xy+xz+yz) - xyz \\ &= x^2y + x^2z + xy^2 + xz^2 + y^2z + yz^2 \\ &= x^2(y+z) + x(y+z)^2 + yz(y+z) \\ &= (y+z)(x^2 + xy + xz + yz) \\ &= (y+z)(x(x+y) + z(x+y)) \\ &= (y+z)(x+y)(x+z)\end{aligned}$$

Thus the image of the transfer vanishes precisely on the union of planes: $\overline{V}^{(12)} \cup \overline{V}^{(13)} \cup \overline{V}^{(23)} = \{(a,a,c) \in \overline{V} \mid a,c \in \overline{\mathbb{K}}\} \cup \{(a,b,a) \in \overline{V} \mid a,b \in \overline{\mathbb{K}}\} \cup \{(a,b,b) \in \overline{V} \mid a,b \in \overline{\mathbb{K}}\}$.

Now we consider the situation when the characteristic of $\mathbb{K}$ is 3. In that case, the image of the transfer is generated by the two invariants s_1 and s_2. Taking the two equations $x+y+z=0$ and $xy+xz+yz=0$ and substituting $z=-(x+y)$ into the second gives

$$\begin{aligned}0 &= xy - x(x+y) - y(x+y)\\ &= -(x^2+xy+y^2)\\ &= -(x^2-2xy+y^2)\\ &= -(x-y)^2\end{aligned}$$

Thus if both s_1 and s_2 vanish at a point $v=(a,b,c) \in \overline{V}$, we must have $a=b$ and similarly, $b=c$. Therefore, we find that in characteristic 3, the ideal $\mathrm{Tr}^{\Sigma_3}(\mathbb{K}[V])$ cuts out the line $\{(a,a,a) \in \overline{V} \mid a \in \overline{\mathbb{K}}\}$. Of course, this line is precisely the set of points of $\overline{V}$ fixed by the subgroup $N=\{e,(123),(132)\}$ of order $p=3$.

For any $\sigma \in G$, we denote by $\mathcal{P}_\sigma$ the ideal of $\mathbb{F}[V]$ generated by the set $(\sigma-1)V^* := \{\sigma \cdot x - x \mid x \in V^*\}$.

Lemma 9.0.12. *Let $\sigma \in G$. Then*

$$\mathcal{V}_{\overline{V}}(\mathcal{P}_\sigma) = \overline{V}^\sigma.$$

In fact,

$$\mathcal{P}_\sigma = \mathcal{I}_{\mathbb{F}[V]}(\overline{V}^\sigma).$$

Proof. First, let $v \in \overline{V}^\sigma$ and take $x \in V^*$. Then $((\sigma-1)\cdot x)(v) = (\sigma \cdot x)(v) - x(v) = x(\sigma^{-1}v) - x(v) = x(v) - x(v) = 0$. Thus $v \in \mathcal{V}_{\overline{V}}(\mathcal{P}_\sigma)$.

For the opposite inclusion, take $v \in \mathcal{V}_{\overline{V}}(\mathcal{P}_\sigma)$ and consider a basis

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

of V^*. Write $v=(v_1,v_2,\dots,v_n)$ and

$$w := \sigma^{-1}v = (w_1, w_2, \dots, w_n)$$

in terms of (the dual of) this basis. Then

$$\begin{aligned}0 &= ((\sigma-1)\cdot x_i)(v) = (\sigma \cdot x_i)(v) - x_i(v) = x_i(\sigma^{-1}v) - x_i(v)\\ &= x_i(w) - x_i(v) = w_i - v_i\end{aligned}$$

for all $i=1,2,\dots,n$. Therefore $v = w = \sigma^{-1}v$ and thus $v \in \overline{V}^\sigma$.

Having shown $\mathcal{V}_{\overline{V}}(\mathcal{P}_\sigma) = \overline{V}^\sigma$, the final statement in the lemma is equivalent to the assertion that $\mathcal{P}_\sigma$ is a radical ideal. Since $\mathcal{P}_\sigma$ is generated by homogeneous elements of degree 1, it is in fact a (homogeneous) prime ideal and thus is certainly radical. □

Since $\mathcal{P}_\sigma$ is a prime ideal, by the “lying over” Theorem 2.5.2 (1), $\mathcal{P}_\sigma \cap \mathbb{F}[V]^G$ is also a prime ideal. Let $f, h \in \mathbb{F}[V]$. Then $(\sigma-1)(fh) = \sigma(f)\sigma(h) - fh = \sigma(f)\sigma(h) - \sigma(f)h + \sigma(f)h - fh = \sigma(f)\cdot(\sigma-1)(h) + h\cdot(\sigma-1)(f) \in \mathcal{P}_\sigma$. Hence $(\sigma-1)\mathbb{F}[V] \subset \mathcal{P}_\sigma$.

Specializing of the result of Theorem 9.0.10 to the case $N = \{e\}$, we have the following equality of ideals:

$$\sqrt{\operatorname{Im}\operatorname{Tr}^G} = (\cap_{\operatorname{order}(\sigma)=p}\mathcal{P}_\sigma) \cap \mathbb{F}[V]^G.$$

Proposition 9.0.13. *Let G be a permutation group. Then* $\operatorname{Im}\operatorname{Tr}^G$ *is a radical ideal.*

Proof. Recall that $\{\mathcal{O}_G(m) \mid m \text{ is a monomial of } \mathbb{F}[V]\}$ is a vector space basis for $\mathbb{F}[V]^G$. Since for any monomial, m, $\operatorname{Tr}^G(m) = |G_m|\mathcal{O}_G(m)$, we see that the set $\{\mathcal{O}_G(m) \mid m \text{ is a monomial of } \mathbb{F}[V], p \text{ does not divide } |G_m|\}$ is a vector space basis for $\operatorname{Im}\operatorname{Tr}^G$.

Suppose that $f \in \sqrt{\operatorname{Im}\operatorname{Tr}^G}$ and write $f^r \in \operatorname{Im}\operatorname{Tr}^G$. Express the invariant f as a linear combination of orbit sums: $f = \sum_{i=1}^t c_i\mathcal{O}_G(m_i)$ where $c_i \in \mathbb{F}$ and m_i is a monomial in $\mathbb{F}[V]$ for all $i = 1, 2, \ldots, t$. Choose $m \in \mathrm{N}$ large enough that $p^m \geq r$. Then $f^{p^m} = \sum_{i=1}^t c_i^{p^m}\mathcal{O}_G(m_i^{p^m}) \in \operatorname{Im}\operatorname{Tr}^G$. Since V is a permutation representation, it follows that $G_m = G_{m^s}$ for all monomials m of $\mathbb{F}[V]$ and for all positive integers s. Thus $\sum_{i=1}^t c_i^{p^m}\mathcal{O}_G(m_i^{p^m}) \in \operatorname{Im}\operatorname{Tr}^G$ implies that $c_i^{p^m} = 0$ for all i for which p divides $|G_{m_i}|$. Hence $f = \sum_{i=1}^t c_i\mathcal{O}_G(m_i)$ expresses f as a linear combination of elements of $\operatorname{Im}\operatorname{Tr}^G$ and thus $f \in \operatorname{Im}\operatorname{Tr}^G$. □

Suppose now that $G = \langle\tau\rangle$ is a cyclic permutation group with p dividing $|G|$. Define $\sigma := \tau^{|G|/p}$. Then the subgroup generated by σ contains all the elements of order p in G. Every element of order p in G is of the form σ^r for some r with $1 \leq r < p$. Clearly, $\overline{V}^\sigma \subseteq \overline{V}^{\sigma^r}$. If we choose integers $s \geq 1$ and $t \leq 0$ such that $rs + pt = 1$, then $(\sigma^r)^s = \sigma(\sigma^p)^{-t} = \sigma$ and we see that $\overline{V}^{\sigma^r} \subseteq \overline{V}^\sigma$. Therefore, $\mathcal{V}_{\overline{V}}\left(\operatorname{Tr}^G(\mathbb{F}[V])\right) = \bigcup_{r=1}^{p-1}\overline{V}^{\sigma^r} = \overline{V}^\sigma$. which is a subspace of $\overline{V}$ and hence is an irreducible variety. Thus $\operatorname{Im}\operatorname{Tr}^G = \sqrt{\operatorname{Im}\operatorname{Tr}^G}$ is a prime ideal.

We may give the same argument algebraically as follows. Consider such an element of G of order p which we again write as σ^r where $1 \leq r < p$. From the factorization $\sigma^r - 1 = (1 + \sigma + \sigma^2 + \cdots + \sigma^{r-1})(\sigma - 1)$, we see that $(\sigma^r - 1)V^* \subseteq (\sigma-1)V^*$ and therefore, $\mathcal{P}_{\sigma^r} \subseteq \mathcal{P}_\sigma$. Again, writing $(\sigma^r)^s = \sigma$ for an appropriate positive integer s shows that $\mathcal{P}_\sigma = \mathcal{P}_{(\sigma^r)^s} \subseteq \mathcal{P}_{\sigma^r}$. Hence $\mathcal{P}_\sigma = \mathcal{P}_{\sigma^r}$ for all $r = 1, 2, \ldots, p-1$. Therefore, $\operatorname{Im}\operatorname{Tr}^G = \sqrt{\operatorname{Im}\operatorname{Tr}^G} = (\cap_{r=1}^{p-1}\mathcal{P}_{\sigma^r}) \cap \mathbb{F}[V]^G = \mathcal{P}_\sigma \cap \mathbb{F}[V]^G$ is a prime ideal.

We record this result as

Proposition 9.0.14. *If G is a modular cyclic permutation group, then*

$$\operatorname{Im}\operatorname{Tr}^G$$

is a prime ideal.

Example 9.0.15. Consider the regular representation V_p of $C_p = \langle\sigma\rangle$ defined over $\mathbb{F}_p$. As usual, we choose a triangular basis $\{x_1, x_2, \ldots, x_p\}$ for V_p^* dual to the basis $\{e_1, e_2, \ldots, e_p\}$ of V_p. Then $V_p^\sigma = V_p^{C_p} = \operatorname{span}_{\mathbb{F}_p}\{e_p\}$ is a one dimensional space. By definition, the ideal $\mathcal{P}_\sigma$ of $\mathbb{F}[V_p]$ is generated by the degree 1 elements $(\sigma - 1)x$ as x varies over V_p^*. We have

$$(\sigma - 1)x_i = \sigma(x_i) - x_i = \begin{cases} x_{i-1} & \text{if } 2 \le i \le p; \\ 0 & \text{if } i = 1. \end{cases}$$

Thus $\mathcal{P}_\sigma = (x_1, x_2, \ldots, x_{p-1})$ is the ideal generated by the linear functionals which vanish on the fixed line $\overline{V}^{C_p}$ and $\operatorname{Im}\operatorname{Tr}^{C_p} = (x_1, x_2, \ldots, x_{p-1}) \cap \mathbb{F}[V_p]^{C_p}$.

Next, we will show that for any modular representation of G, the ideal $\operatorname{Tr}^G(\mathbb{F}[V])$ is a non-zero proper subset of the irrelevant ideal, $\mathbb{F}[V]^G_+ := \oplus_{d=1}^\infty \mathbb{F}[V]^G_d$, of $\mathbb{F}[V]^G$.

If f is invariant, then $\operatorname{Tr}^G(f) = |G|f = 0$ since p divides $|G|$. In particular, this means that in degree 0, $\operatorname{Tr}^G(\mathbb{F}[V])$ contains only 0. This proves that $\operatorname{Tr}^G(\mathbb{F}[V])$ is a subset of the irrelevant ideal.

In fact, it is always a proper subset of the irrelevant ideal since by Theorem 9.0.10, it cuts out the variety of points v whose isotopy group G_v contains an element of order p. Thus $\mathcal{V}_{\overline{V}}(\operatorname{Tr}^G(\mathbb{F}[V])) = \{v \in V \mid p \text{ divides } |G_v|\}$. Therefore, $\operatorname{Tr}^G(\mathbb{F}[V]) = \mathbb{F}[V]^G_+$ if and only if $V^\sigma = \{0\}$ for every element of order $p \in G$. But by Lemma 4.0.1, we see that the subspace V^σ is never zero if the order of σ is p. Of course, since G is a modular group, there will exist elements in G of order p.

The following lemma which is a consequence of Artin's Theorem on Characters (see, for example, [75, VI Theorem 4.1]) will be used to show that Tr^G is never the zero map.

Lemma 9.0.16. *Let K be any field and let $\Gamma := \{\phi_\alpha \mid \alpha \in A\}$ be any set of (distinct) field automorphisms of K. Then Γ is a linearly independent subset of the vector space of all maps from K to K.*

Proof. Assume by way of contradiction that there is a linear relation

$$c_1\phi_{\alpha_1} + c_2\phi_{\alpha_2} + \cdots + c_t\phi_{\alpha_t} = 0$$

among the elements of Γ. Furthermore, we assume that t is minimal so that any subset of Γ containing $t - 1$ or fewer elements is linearly independent. Thus each of $c_1, c_2, \ldots, c_t$ must be non-zero. Note that $t \ge 2$ since the zero map is not a field automorphism.

Let $a \in K$ be arbitrary. Evaluating the above linear relation at $x \in K$ and multiplying by $\phi_{\alpha_1}(a)$ yields the equation

$$c_1\phi_{\alpha_1}(a)\phi_{\alpha_1}(x) + c_2\phi_{\alpha_1}(a)\phi_{\alpha_2}(x) + \cdots + c_t\phi_{\alpha_1}(a)\phi_{\alpha_t}(x) = 0$$

which is valid for all $x \in K$. On the other hand, if we evaluate the linear relation at the point ax we get

$$c_1\phi_{\alpha_1}(a)\phi_{\alpha_1}(x) + c_2\phi_{\alpha_2}(a)\phi_{\alpha_2}(x) + \cdots + c_t\phi_{\alpha_t}(a)\phi_{\alpha_t}(x) = 0$$

for all $x \in K$. Subtracting we find

$$\begin{aligned} &c_2(\phi_{\alpha_1}(a) - \phi_{\alpha_2}(a))\phi_{\alpha_2}(x) \\ &\quad + c_3(\phi_{\alpha_1}(a) - \phi_{\alpha_3}(a))\phi_{\alpha_3}(x) + \ldots \\ &\quad + c_t(\phi_{\alpha_1}(a) - \phi_{\alpha_t}(a))\phi_{\alpha_t}(x) = 0 \end{aligned}$$

which holds for all $x \in K$. By the minimality of t, we must have $\phi_{\alpha_1}(a) - \phi_{\alpha_i}(a) = 0$ for all $i = 2, 3, \ldots, t$. In particular, $\phi_{\alpha_1}(a) = \phi_{\alpha_2}(a)$. But since $a \in K$ was arbitrary, this implies $\phi_{\alpha_1} = \phi_{\alpha_2}$. This contradiction shows that Γ is linearly independent. □

Corollary 9.0.17. *Let V be a representation of the finite group G. The image of the transfer $\mathrm{Tr}^G : \mathbb{F}[V] \to \mathbb{F}[V]^G$ is non-zero.*

Proof. Each $\sigma \in G$ defines an automorphism, ϕ_σ of the ring $\mathbb{F}[V]$ given by $\phi_\sigma(f) = \sigma(f)$. We may extend these ring automorphisms to field automorphisms of $\mathbb{F}(V)$, the quotient field of $\mathbb{F}[V]$. We do this by defining $\overline{\phi}_\sigma(f/h) := \phi_\sigma(f)/\phi_\sigma(h)$. Then, by the preceding lemma, $\sum_{\sigma\in G}\overline{\phi}_\sigma \neq 0$, i.e., there exists $f, h \in \mathbb{F}[V]$ such that $\sum_{\sigma\in G}\overline{\phi}_\sigma(f/h) \neq 0$. Now define $h' := \prod_{\sigma\in G}\sigma(h)$ and $f' := f\prod_{\sigma\in G\setminus\{e\}}\sigma(h)$. Then $f' \in \mathbb{F}[V]$ and $h' \in \mathbb{F}[V]^G$ and $f/h = f'/h'$. Therefore $\sum_{\sigma\in G}\overline{\phi}_\sigma(f'/h') = \sum_{\sigma\in G}(\phi_\sigma(f')/\phi_\sigma(h')) = (\sum_{\sigma\in G}\phi_\sigma(f'))/h' = \mathrm{Tr}^G(f')/h'$ is non-zero. Hence $\mathrm{Tr}^G(f') \neq 0$. □

Note that by Theorem 9.0.10, this corollary shows that $\cup_{\mathrm{order}(\sigma)=p}\overline{V}^\sigma \neq \overline{V}$, i.e., there must always be some points of $\overline{V}$ whose isotropy group is nonmodular.

Proposition 9.0.18. *Let $N \leq H \leq G$ be a sequence of subgroups of G where N is normal in G and suppose that p does not divide the index of H in G. Then $\mathrm{Im}\,\mathrm{Tr}^G_N = \mathrm{Im}\,\mathrm{Tr}^H_N \cap \mathbb{F}[V]^G$.*

Proof. First, suppose that $f \in \mathrm{Im}\,\mathrm{Tr}^H_N \cap \mathbb{F}[V]^G$ and write $f = \mathrm{Tr}^H_N(h)$ where $h \in \mathbb{F}[V]^N$. Since $f \in \mathbb{F}[V]^G$, we have $\mathrm{Tr}^G_H(f) = [G : H]f$ and thus

$$\mathrm{Tr}^G_N(h/[G:H]) = \mathrm{Tr}^G_H(\mathrm{Tr}^H_N(h/[G:H])) = \mathrm{Tr}^G_H(f/[G:H]) = f\ .$$

For the opposite inclusion, we consider the factorization from Lemma 9.0.7: $\mathrm{Tr}^G_N = \mathrm{Tr}^H_N \circ \widehat{\mathrm{Tr}}^G_H$. Suppose that $f \in \mathrm{Im}\,\mathrm{Tr}^G_N$ and write $f = \mathrm{Tr}^G_N(h)$ where $h \in \mathbb{F}[V]^N$. Then $f = \mathrm{Tr}^H_N(\widehat{\mathrm{Tr}}^G_H(h)) \in \mathrm{Im}\,\mathrm{Tr}^H_N$. □

Corollary 9.0.19. *Let P be a p-Sylow subgroup of G and let $N \subset P$ be a normal subgroup of G. Then the two ideals* $\operatorname{Im}\operatorname{Tr}_N^P$ *and* $\operatorname{Im}\operatorname{Tr}_N^G$ *have the same height.*

Proof. Since $\mathbb{F}[V]^P$ is integral over $\mathbb{F}[V]^G$, we may apply "going-up" and "going-down" (Theorem 2.5.2). Since the above lemma shows that $\operatorname{Im}\operatorname{Tr}_N^P$ lies over $\operatorname{Im}\operatorname{Tr}_N^G$, the result follows. □

9.1 The Transfer for Nakajima Groups

We want to describe the image of the transfer for a Nakajima group . We begin with some preliminaries. The following lemma is known as Lucas' Lemma.

Lemma 9.1.1. *Let p be prime and let $m = m_0 + m_1p + m_2p^2 + \cdots + m_sp^s$ and $i = i_0 + i_1p + i_2p^2 + \cdots + i_sp^s$ be the p-adic expansions of two non-negative integers m and i. Then*

$$\binom{m}{i} \equiv \binom{m_0}{i_0}\binom{m_1}{i_1}\cdots\binom{m_s}{i_s} \pmod{p}.$$

Proof. Considering the expression $(1+y)^m$ as a polynomial with coefficients in $\mathbb{F}_p$ we have

$$\begin{aligned}
\sum_{j=0}^{m}\binom{m}{j}y^j &= \prod_{j=0}^{s}(1+y)^{m_jp^j} \\
&= \prod_{j=0}^{s}(1+y^{p^j})^{m_j} \\
&= \prod_{j=0}^{s}(\sum_{k=0}^{m_j}\binom{m_j}{k}y^{kp^j}) \\
&= \prod_{j=0}^{s}(\sum_{k=0}^{p-1}\binom{m_j}{k}y^{kp^j}) \\
&= \sum_{k_0=0}^{p-1}\sum_{k_1=0}^{p-1}\cdots\sum_{k_s=0}^{p-1}\binom{m_0}{k_0}\binom{m_1}{k_1}\cdots\binom{m_s}{k_s}y^{k_0+k_1p+\ldots k_sp^s} .
\end{aligned}$$

Comparing the coefficient of y^i in the first and last expressions gives the result. □

In fact, the following generalization is true.

Lemma 9.1.2. *Let $q = p^r$ be a prime power and let $m = m_0+m_1q+\ldots+m_sq^s$ be the q-adic expansion of m and $i_r = i_{r,0} + i_{r,1}q + \ldots + i_{r,s}q^s$ is the q-adic expansion of i_r. Then*

$$\binom{m}{i_1,\ldots,i_k} \equiv \binom{m_0}{i_{1,0},\ldots,i_{k,0}} \cdot \binom{m_1}{i_{1,1},\ldots,i_{k,1}} \cdots \binom{m_s}{i_{1,s},\ldots,i_{k,s}} \pmod{p}.$$

Remark 9.1.3. The preceding lemma is essentially due to Dickson — see [28]. There he also proves that the highest power of p dividing $\binom{m}{I}$ is p^e where e is the total amount carried when the addition $i_1 + \cdots + i_k = m$ is performed base p.

Definition 9.1.4. *Let m be a non-negative integer. We denote by $\alpha_p(m)$ the sum of the digits in the p-adic expansion of m: $\alpha_p(m) := m_0 + m_1 + \cdots + m_s$ where $m = \sum_{i=0}^{s} m_i p^i$ and $0 \leq m_i < p$ for all $i = 0, 1, \ldots, s$.*

Let W be a subset of V^*. If W is an $\mathbb{F}_p$ subspace, i.e., if W is closed under addition, we define $d(W) := \prod_{x \in W \setminus \{0\}} x$.

Now we prove a generalization of Lemma 9.0.2.

Lemma 9.1.5. *Let V be an $\mathbb{F}$ vector space and let W be a finite subset of V^* which is closed under addition, so that W is a vector space over $\mathbb{F}_p$. Put $m := \dim_{\mathbb{F}_p}(W)$. Then in $\mathbb{F}[V]$ we have*

1. $\displaystyle\sum_{x \in W} x^s = 0$ *unless $p-1$ divides s and $\alpha_p(s) \geq m(p-1)$.*
2. $\displaystyle\sum_{x \in W} x^{p^m - 1} = d(W)$.

Proof. We use the convention that $0^0 = 1$. Thus when $s = 0$, we have $\sum_{x \in W} x^0 = p^m \cdot 1 = 0$. We assume for the following that $s \geq 1$.

The proof is by induction on m. Let $\{x_1, x_2, \ldots, x_m\}$ be an $\mathbb{F}_p$ basis of W. For $m = 0$, the result is clear. For $m = 1$, we have $\sum_{x \in W} x^s = \sum_{c \in \mathbb{F}_p} (cx_1)^s = x_1^s \sum_{c \in \mathbb{F}_p} c^s$. Now let η be a generator for the cyclic group of non-zero elements of $\mathbb{F}_p$. Then $\sum_{c \in \mathbb{F}_p} c^s = \sum_{c \in \mathbb{F}_p \setminus \{0\}} c^s = \sum_{j=0}^{p-2} (\eta^j)^s$. If $p-1$ divides s, then $\eta^{sj} = 1$ for all j and $\sum_{c \in \mathbb{F}_p} c^s = p - 1 = -1$. Since every non-zero element of $\mathbb{F}_p$ is a root of the polynomial $x^p - 1 = (x-1)(x^{p-2} + x^{p-3} + \cdots + x + 1)$, if $p-1$ does not divide s, then $\eta^s \neq 1$ and thus η^s is a root of $(x^{p-2} + x^{p-3} + \cdots + x + 1)$. Thus if $p-1$ does not divide s, then $\sum_{c \in \mathbb{F}_p} c^s = 0$. Therefore, for $m = 1$ we have

$$\sum_{x \in W} x^s = \begin{cases} -x_1^s, & \text{if } p-1 \text{ divides } s; \\ 0, & \text{if } p-1 \text{ does not divide } s. \end{cases}$$

Suppose $m > 1$ and let W' denote the $\mathbb{F}_p$ vector space spanned by $\{x_1, x_2, \ldots, x_{m-1}\}$. Then

$$\begin{aligned} \sum_{x \in W} x^s &= \sum_{x \in W'} \sum_{c \in \mathbb{F}_p} (x + cx_m)^s \\ &= \sum_{x \in W'} \sum_{c \in \mathbb{F}_p} \sum_{j=0}^{s} \binom{s}{j} x^j (cx_m)^{s-j} \end{aligned}$$

$$= \sum_{j=0}^{s} \binom{s}{j} x_m^{s-j} \sum_{c\in\mathbb{F}_p} c^{s-j} \sum_{x\in W'} x^j$$
$$= \sum_{j=1}^{s-1} \binom{s}{j} x_m^{s-j} \sum_{c\in\mathbb{F}_p} c^{s-j} \sum_{x\in W'} x^j$$

since $\sum_{c\in\mathbb{F}_p} c^0 = 0$ and $\sum_{x\in W'} x^0 = 0$. This shows that $\sum_{x\in W} x^s$ is non-zero if and only if there is some value of j with $1 \le j \le s-1$ such that $\sum_{c\in\mathbb{F}_p} c^{s-j} \sum_{x\in W'} x^j \neq 0$. By induction, $\sum_{x\in W'} x^j = 0$ unless $p-1$ divides j. By the proof of the $m = 1$ case, $\sum_{c\in\mathbb{F}_p} c^{s-j} = 0$ unless $p-1$ divides $s-j$. Thus we may concentrate on the values of s and j which are both divisible by $p-1$.

Suppose $\binom{s}{j} \sum_{c\in\mathbb{F}_p} c^{s-j} \sum_{x\in W'} x^j \neq 0$ and let $s = a_0 + a_1 p + \cdots + a_t p^t$ and $j = b_0 + b_1 p + \cdots + b_t p^t$ be the p-adic expansions of s and j. Since $\binom{s}{j} \neq 0$, by Lucas' Lemma, we must have $b_r \le a_r$ for all $r = 0, 1, \ldots, t$. Since $j < s$, there is some value of r with $b_r < a_r$. Therefore, $\alpha_p(j) < \alpha_p(s)$. Since $\sum_{x\in W'} x^j \neq 0$, by induction, we must have $\alpha_p(j) \ge (m-1)(p-1)$ and therefore, $\alpha_p(s) > (m-1)(p-1)$.

Since $p^r \equiv 1 \pmod{p-1}$ for all r, we see that $\alpha_p(i) \equiv i \pmod{p-1}$ for all non-negative integers i. Therefore, $\alpha_p(s) \equiv 0 \pmod{p-1}$ and $\alpha_p(s) > (m-1)(p-1)$ which implies that $\alpha_p(s) \ge m(p-1)$.

Now we prove 2. Take $W_0 \subset W \setminus \{0\}$ such that W_0 contains exactly one element of every one dimensional $\mathbb{F}_p$ subspace of W (thus $|W_0| = (p^m-1)/(p-1)$). Thus every non-zero element $x \in W$ can be written uniquely as $x = cx_0$ where $c \in \mathbb{F}_p \setminus \{0\}$ and $x_0 \in W_0$ and $d(W) = \prod_{x_0\in W_0}(\prod_{c\in\mathbb{F}_p\setminus\{0\}} cx_0) = (-1)^{|W_0|} \prod_{x_0\in W_0} x_0 = (-1)^m \prod_{x_0\in W_0} x_0$.

The above discussion shows that x_m^{p-1} divides $\sum_{x\in W} x^s$. Since $\sum_{x\in W} x^s$ is a $\mathrm{GL}(W)$-invariant, this shows that y^{p-1} divides $\sum_{x\in W} x^s$ for all $y \in W$. Since the elements of W_0 are pairwise relatively prime and since $\mathbb{F}[V]$ is a unique factorization domain, this shows that $\prod_{x_0\in W_0} x_0$ divides $\sum_{x\in W} x^s$. Therefore, $d(W)$ divides $\sum_{x\in W} x^s$.

To complete the proof, we consider $\sum_{x\in W} x^{p^m-1}$ and $d(W)$ as polynomials in x_m with coefficients in $\mathbb{F}[W']$. We use the lexicographic order on $\mathbb{F}[W]$ with $x_1 < x_2 < \cdots < x_m$ and compare $\mathrm{LT}(\sum_{x\in W} x^{p^m-1})$ and $\mathrm{LT}(d(W))$. The least value of j such that $p-1$ divides j and $\alpha_p(j) \ge (m-1)(p-1)$ is $j = p^{m-1}-1$. Therefore, the lead term of $\sum_{x\in W} x^{p^m-1}$ is

$$\binom{p^{m-1}-1}{p^{m-1}-1} \sum_{c\in\mathbb{F}_p} c^{p^m-p^{m-1}} (\sum_{x\in W'} x^{p^{m-1}-1}) x_m^{p^m-p^{m-1}}.$$

By induction, $\sum_{x\in W'} x^{p^{m-1}-1} = d(W')$ and thus the lead term of

$$\sum_{x\in W} x^{p^m-1}$$

is

$$-d(W')x_m^{p^m-p^{m-1}}.$$

On the other hand,

$$\begin{aligned}\mathrm{LT}(d(W)) &= \mathrm{LT}(\prod_{x\in W\setminus\{0\}} x) = \mathrm{LT}((\prod_{x\in W\setminus W'} x)(\prod_{x\in W'\setminus\{0\}} x))\\ &= \mathrm{LT}(\prod_{x\in W\setminus W'} x)d(W') = d(W')\prod_{x\in W\setminus W'} \mathrm{LT}(x)\\ &= d(W')(\prod_{c\in\mathbb{F}_p} cx_m)^{p^{m-1}} = d(W')(-x_m^{p-1})^{p^{m-1}}\\ &= -d(W')x_m^{p^m-p^{m-1}}.\end{aligned}$$

Therefore, $d(W)$ and $\sum_{x\in W} x^{p^m-1}$ have the same degree and the same lead term. Since the former divides the latter, they must be equal as claimed. □

Proposition 9.1.6. *We take P to be a Nakajima group with Nakajima basis $\{x_1, x_2, \ldots, x_n\}$ for V^*. Then*

1. *The set $W_i := \{(\sigma-1)x_i \mid \sigma\in P\}$ is an $\mathbb{F}_p$ vector space.*
2. *$\mathrm{Im}\,\mathrm{Tr}^P$ is the principal ideal of $\mathbb{F}[V]^P$ generated by $\prod_{i=1}^n d(W_i)$.*

Proof. Let $P_i := \{\sigma\in P \mid \sigma(x_j) = x_j \text{ for all } j\neq i\}$. Since P is a Nakajima group , $P = P_nP_{n-1}\ldots P_1$ and the P-orbit of x_i is the same as the P_i-orbit of x_i. Since each non-identity element of P_i is a transvection, every such element has order p by Lemma 8.0.3 and thus P_i is an elementary Abelian p-group. Let $\sigma,\tau\in P_i$ and write $\sigma(x_i) = x_i + y_\sigma$ and $\tau(x_i) = x_i + y_\tau$ where $y_\sigma, y_\tau \in \mathrm{span}\{x_1, x_2, \ldots, x_{i-1}\}$. Then $\sigma\tau(x_i) = \sigma(\tau(x_i)) = \sigma(x_i + y_\tau) = \sigma(x_i) + \sigma(y_\tau) = x_i + y_\sigma + y_\tau$ and thus $(\sigma\tau-1)x_i = (\sigma-1)x_i + (\tau-1)x_i$. This shows that W_i is closed under addition and is therefore an $\mathbb{F}_p$-vector space.

For the second statement, let d_i denote the $\mathbb{F}_p$ dimension of W_i. Then $\mathbb{F}[V]^P = \mathbb{F}[\mathrm{N}(x_1), \mathrm{N}(x_2), \ldots, \mathrm{N}(x_n)]$ where $\deg(\mathrm{N}_i) = p^{d_i}$. By Lemma 6.2.1, the monomial factors of $m := \prod_{i=1}^n x_i^{p^{d_i}-1}$ form a block basis for $\mathbb{F}[V]$ over $\mathbb{F}[V]^P$. Thus if $f\in\mathbb{F}[V]$, we may write (uniquely) $f = \sum_{\alpha \text{ divides } m} f_\alpha\alpha$ where each $f_\alpha\in\mathbb{F}[V]^P$. Hence

$$\mathrm{Tr}^P(f) = \sum_{\alpha \text{ divides } m} \mathrm{Tr}^P(f_\alpha\alpha) = \sum_{\alpha \text{ divides } m} f_\alpha \mathrm{Tr}^P(\alpha).$$

Thus to find $\mathrm{Im}\,\mathrm{Tr}^P$, it suffices to compute $\mathrm{Tr}^P(\alpha)$ for all monomials α which divide $m = \prod_{i=1}^n x_i^{p^{d_i}-1}$.

Suppose $x^A = x_1^{a_1}x_2^{a_2}\cdots x_n^{a_n}$ divides m, i.e., suppose $a_i \le p^{d_i}-1$ for all $i = 1, 2, \ldots, n$. Since $\sigma_i(x_i)\in\mathrm{span}\{x_1, x_2, \ldots, x_i\}$ and $(\sigma_i - 1)x_j = 0$ for $\sigma_i\in P_i$ and $i\neq j$, we see that

$$\begin{aligned}(\sigma_n\sigma_{n-1}\cdots\sigma_1)\big(x^A\big) &= (\sigma_n\sigma_{n-1}\cdots\sigma_2)\big(\sigma_1(x_1^{a_1})x_2^{a_2}\cdots x_n^{a_n}\big)\\ &= (\sigma_n\sigma_{n-1}\cdots\sigma_3)\big(\sigma_1(x_1^{a_1})\sigma_2(x_2^{a_2})x_3^{a_3}\cdots x_n^{a_n}\big)\\ &= (\sigma_n\sigma_{n-1}\cdots\sigma_4)\big(\sigma_1(x_1^{a_1})\sigma_2(x_2^{a_2})\sigma_3(x_3^{a_3})x_4^{a_4}\cdots x_n^{a_n}\big)\\ &\ \ \vdots\\ &= \sigma_1(x_1^{a_1})\sigma_2(x_2^{a_2})\cdots\sigma_n(x_n^{a_n})\ .\end{aligned}$$

Therefore,

$$\begin{aligned}\mathrm{Tr}^P(x_1^{a_1}x_2^{a_2}\cdots x_n^{a_n}) &= \sum_{\sigma\in P}\sigma(x^A)\\ &= \sum_{\sigma_n\in P_n}\sum_{\sigma_{n-1}\in P_{n-1}}\cdots\sum_{\sigma_1\in P_1}(\sigma_n\sigma_{n-1}\cdots\sigma_1)(x^A)\\ &= \sum_{\sigma_n\in P_n}\sum_{\sigma_{n-1}\in P_{n-1}}\cdots\sum_{\sigma_1\in P_1}\sigma_1(x_1^{a_1})\sigma_2(x_2^{a_2})\cdots\sigma_n(x_n^{a_n})\\ &= \prod_{i=1}^n\sum_{\sigma_i\in P_i}\sigma_i(x_i)^{a_i}\\ &= \prod_{i=1}^n\sum_{y\in W_i}(x_i+y)^{a_i}\ .\end{aligned}$$

Now,

$$\sum_{y\in W_i}(x_i+y)^{a_i} = \sum_{y\in W_i}\sum_{j=0}^{a_i}\binom{a_i}{j}x_i^{a_i-j}y^j = \sum_{j=0}^{a_i}\binom{a_i}{j}x_i^{a_i-j}\sum_{y\in W_i}y^j.$$

By Lemma 9.1.5, $\sum_{y\in W_i}y^j$ equals 0 unless $p-1$ divides j and $\alpha_p(j)\geq d_i(p-1)$. Since $j\leq a_i<d_i$ for all $i=1,2,\ldots,n$ we have, $\mathrm{Tr}^P(x_1^{a_1}x_2^{a_2}\cdots x_n^{a_n})=0$ unless $a_i=d_i-1$ for all $i=1,2,\ldots,n$. By the second statement in Lemma 9.1.5, we have

$$\begin{aligned}\mathrm{Tr}^P(x_1^{d_1-1}x_2^{d_2-1}\cdots x_n^{d_n-1}) &= \prod_{i=1}^n\sum_{y\in W_i}(x_i+y)^{d_i-1}\\ &= \prod_{i=1}^n\sum_{y\in W_i}y^{d_i-1}\\ &= \prod_{i=1}^n d(W_i).\end{aligned}$$

□

We remind the reader that for $V=\mathbb{F}^n$, we showed in §3.4 that the ring of invariants $\mathbb{F}[V]^{\mathrm{U}_n}$ is the polynomial ring on the invariants $h_1,h_2,\ldots,h_n$.

Lemma 9.1.7. *We have that* $\operatorname{Im}\operatorname{Tr}^{U_n(\mathbb{F}_q)}$ *is the principal ideal generated by* $(h_1^{n-1}h_2^{n-2}\cdots h_{n-1})^{q-1}$.

Proof. By Proposition 9.1.6, $\operatorname{Im}\operatorname{Tr}^{U_n(\mathbb{F}_q)}$ is generated by $\prod_{i=1}^n d(W_i)$ where W_i is the $\mathbb{F}_q$ vector space spanned by $\{x_1, x_2, \dots, x_{i-1}\}$.

We will show that $d(W_t) = -h_{t-1}^{q-1}d(W_{t-1})$ for all $t = 1, 2, \dots, n$. By convention, we take $d(W_0) = -1$ and $h_0 = 1$. We proceed by induction on t. For $t = 1$, we have $W_1 = \{0\}$ and thus $d(W_1) = 1 = h_0^{q-1}d(W_0)$.

Assume then that $d(W_t) = h_{t-1}^{q-1}d(W_{t-1})$ and consider $d(W_{t+1})$. We have

$$\begin{aligned} d(W_{t+1}) &= \prod_{x\in W_{t+1}\setminus\{0\}} x = \prod_{x\in W_{t+1}\setminus W_t} x \prod_{y\in W_t\setminus\{0\}} y \\ &= \prod_{x\in W_{t+1}\setminus W_t} xd(W_t) \\ &= d(W_t) \prod_{c\in\mathbb{F}_q\setminus\{0\}} \prod_{z\in W_t} cx_{t+1} + z \\ &= d(W_t) \prod_{c\in\mathbb{F}_q\setminus\{0\}} \prod_{z\in W_t} c(x_{t+1} + z/c) \\ &= d(W_t) \prod_{c\in\mathbb{F}_q\setminus\{0\}} \left(c \prod_{z'\in W_t} x_{t+1} + z' \right) \\ &= d(W_t) \left(\prod_{c\in\mathbb{F}_q\setminus\{0\}} c \right) \left(\prod_{z'\in W_t} x_{t+1} + z' \right)^{q-1} \\ &= d(W_t)(-1) \prod_{z'\in W_t} (x_{t+1} + z')^{q-1} = -d(W_t)h_t\ . \end{aligned}$$

From this we get $d(W_t) = (-1)^t(h_1h_2\cdots h_{t-1})^{q-1}$. Thus $\operatorname{Im}\operatorname{Tr}^{U_n(\mathbb{F}_q)}$ is generated by $\prod_{t=1}^n d(W_t) = \pm\prod_{t=1}^n (h_1h_2\cdots h_{t-1})^{q-1} = \pm(h_1^{n-1}h_2^{n-2}\cdots h_{n-1})^{q-1}$. □

Corollary 9.1.8. *Let V be an n dimensional vector space over $\mathbb{F}_q$. Then* $\operatorname{Im}\operatorname{Tr}^{\mathrm{GL}(V)}$ *is the principal ideal of* $\mathbb{F}[V]^{\mathrm{GL}(V)}$ *generated by* $d(V)^{n-1}$.

Proof. Suppose that $f \in \operatorname{Im}\operatorname{Tr}^{\mathrm{GL}(V)}$. By Proposition 9.0.18,

$$\operatorname{Im}\operatorname{Tr}^{\mathrm{GL}(V)} = \operatorname{Im}\operatorname{Tr}^{U_n(\mathbb{F}_q)} \cap \mathbb{F}_q[V]^{\mathrm{GL}(V)}.$$

By the preceding lemma, $\operatorname{Im}\operatorname{Tr}^{U_n}$ is the principle ideal generated by

$$(h_1^{n-1}h_2^{n-2}\cdots h_{n-1})^{q-1}.$$

Therefore,

$$f = (h_1^{n-1}h_2^{n-2}\cdots h_{n-1})^{q-1}h$$

for some $h \in \mathbb{F}_q[V]^{U_n(\mathbb{F}_q)}$. Since $\prod_{c\in\mathbb{F}_q\setminus\{0\}} cx_1 = -x_1^{q-1}$ and $d(V) = \prod_{y\in V^*\setminus\{0\}} y$, we see that x_1^{q-1} divides $d(V)$ but x_1^q does not divide $d(V)$. Since $h_1 = x_1$, we see that $x_1^{(n-1)(q-1)}$ divides f. Take any non-zero element $y \in V^*$. Since x_1 and y lie in the same $\mathrm{GL}(V)$-orbit, and since f is $\mathrm{GL}(V)$-invariant, we see that $y^{(n-1)(q-1)}$ also divides f. Putting these facts together we get that $d(V)^{n-1}$ divides f.

For the opposite inclusion, we must show that $d(V)^{n-1}$ lies in $\operatorname{Im}\operatorname{Tr}^{\mathrm{GL}(V)}$. But

$$d(V) = (-1)^n(h_1h_2\dots h_n)^{q-1}$$

and thus using the preceding lemma, $d(V)^{n-1}$ lies in $\operatorname{Im}\operatorname{Tr}^{U_n(\mathbb{F}_q)}$. But clearly, from its definition, $d(V) \in \mathbb{F}_q[V]^{\mathrm{GL}(V)}$. Therefore,

$$d(V)^{n-1} \in \operatorname{Im}\operatorname{Tr}^{U_n(\mathbb{F}_q)} \cap \mathbb{F}_q[V]^{\mathrm{GL}(V)} = \operatorname{Im}\operatorname{Tr}^{\mathrm{GL}(V)} \ .$$

□

9.2 Cohen-Macaulay Invariant Rings of *p*-Groups

Hochster and Eagon [53] have shown that non-modular invariant rings are always Cohen-Macaulay. A very important question in modular invariant theory is whether or not the invariant ring is Cohen-Macaulay. Campbell, Hughes and Pollack [18] show that a sufficient condition is that the p-Sylow subgroup have a Cohen-Macaulay invariant ring. In this section, we give a necessary and restrictive condition for the invariant ring of a p-group to be Cohen-Macaulay, Theorem 9.2.2. This result is the culmination of a sequence of papers beginning with Ellingsrud and Skjelbred [35] and leading to Campbell, Geramita, Hughes, Shank, Wehlau [17], and Kemper [65]. In particular, this result can be viewed as a generalization of the calculation given in Example 4.0.4. The proof we give here of Theorem 9.2.2 avoids the use of group cohomology.

Definition 9.2.1. *A element $\sigma \in \mathrm{GL}(V)$ is called a* bi-reflection *if the endomorphism $\sigma - I_V$ of V has rank less than or equal to 2.*

Here $\sigma - I_V$ is a linear endomorphism of V and thus its rank is just the rank of the corresponding matrix. Thus σ is a bi-reflection if and only if $\dim_{\mathbb{F}}(V^\sigma) \geq \dim_{\mathbb{F}}(V) - 2$.

Kemper's proof of the following result uses group cohomology.

Theorem 9.2.2. *Let $G \leq \mathrm{GL}(V)$ be a p-group and suppose that $\mathbb{F}[V]^G$ is a Cohen-Macaulay ring. Then G is generated by bi-reflections.*

Proof. Suppose that $\mathbb{F}[V]^G$ is Cohen-Macaulay. Let H be the subgroup of G generated by all the bi-reflections in G. Assume, by way of contradiction, that

$H \neq G$. By Lemma 1.10.3, there exists a maximal proper subgroup N of G such that $H \leq N$, $N \triangleleft G$ and $[G : N] = p$. Take $\sigma \in G \setminus N$. By Theorem 9.0.10,

$$\mathcal{I}_{\mathbb{F}[V]^G}(\mathcal{V}_{\overline{V}}(\mathrm{Tr}^G_N(\mathbb{F}[V]^N)) = \mathcal{I}_{\mathbb{F}[V]^G}(\cup_{\mathrm{order}(\sigma N)=p}\overline{V}^{\sigma})$$

and thus

$$\sqrt{\mathrm{Tr}^G_N(\mathbb{F}[V]^N)} = \bigcap_{\mathrm{order}(\sigma N)=p} \mathcal{I}_{\mathbb{F}[V]^G}(\overline{V}^{\sigma}) \ .$$

Here we are writing $\mathcal{I}_{\mathbb{F}[V]^G}(X)$ for $X \subseteq \overline{V}$ to denote the ideal of functions in $\mathbb{F}[V]^G$ which vanish on $\pi_{V,G}(X) \subset \overline{V}/\!\!/G$, i.e., $\mathcal{I}_{\mathbb{F}[V]^G}(X)$ denotes $\mathcal{I}_{\mathbb{F}[V]^G}(\pi_{V,G}(X)) = \mathcal{I}_{\mathbb{F}[V]}(X) \cap \mathbb{F}[V]^G$.

Now $\mathcal{I}_{\mathbb{F}[V]^G}(\overline{V}^{\sigma}) = \mathbb{F}[V]^G \cap \mathcal{I}_{\mathbb{F}[V]}(\overline{V}^{\sigma})$. Therefore by the going-up and going-down Theorem 2.5.2 (2) and 2.5.2 (3),

$$\mathrm{height}(\mathcal{I}_{\mathbb{F}[V]^G}(\overline{V}^{\sigma})) = \mathrm{height}(\mathcal{I}_{\mathbb{F}[V]}(\overline{V}^{\sigma})).$$

But $\overline{V}^{\sigma}$ is just a subspace of $\overline{V}$ and clearly $\mathrm{height}(\mathcal{I}_{\mathbb{F}[V]}(\overline{V}^{\sigma})) = \mathrm{rank}(\sigma - I_{\overline{V}})$. Since an ideal and its radical have the same height, we see that

$$\begin{aligned}\mathrm{height}(\mathrm{Tr}^G_N(\mathbb{F}[V]^N)) &= \min\{\mathrm{rank}(\sigma - I_V) \mid \mathrm{order}(\sigma N) = p\} \\ &\geq 3, \text{ by definition of } H \text{ and } N \ .\end{aligned}$$

But if R is any Noetherian ring and I is an ideal of R of height m, then (by [79][Theroem 24 14.F] say) there exist $a_1, a_2, \ldots, a_m \in I$ such that the ideal $(a_1, a_2, \ldots, a_m)$ also has height m. Therefore, there exist $a_1, a_2, a_3 \in \mathrm{Tr}^G_N(\mathbb{F}[V]^N)$ such that a_1, a_2, a_3 is a partial homogeneous system of parameters in $\mathbb{F}[V]^G$. Since $\mathbb{F}[V]^G$ is Cohen-Macaulay, a_1, a_2, a_3 is a regular sequence in $\mathbb{F}[V]^G$.

Write $a_i = \mathrm{Tr}^G_N(f_i)$ with $f_i \in \mathbb{F}[V]^N$ for $i = 1, 2, 3$. Now the map $\mathrm{Tr}^G_N = 1 + \sigma + \sigma^2 + \cdots + \sigma^{p-1} = (\sigma - 1)^{p-1}$ and thus we have $a_i = (\sigma - 1)^{p-1}(f_i)$. Define b_i by $b_i := (\sigma - 1)^{p-2}(f_i)$ (so $b_i = f_i$ if $p = 2$) for $i = 1, 2, 3$. Then $(\sigma - 1)(b_i) = a_i$. Also, note that for any $h \in \mathbb{F}[V]^N$ and any $n \in N$, we have $n \cdot (\sigma - 1)(h) = n\sigma(h) - nh = \sigma n'(h) - h = (\sigma - 1)h$ where $n' = \sigma^{-1}n\sigma \in N$. Thus $b_i = (\sigma - 2)^{p-1}(f_i) \in \mathbb{F}[V]^N$ for all $i = 1, 2, 3$.

Expanding

$$\det\left(\begin{pmatrix} a_1 & a_2 & a_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}\right) = 0$$

along the first row, we have $a_1(a_2b_3 - a_3b_2) + a_2(a_3b_1 - a_1b_3) = a_3(a_2b_1 - a_1b_2)$. It is easily verified that $(\sigma - 1)(a_ib_j - a_jb_i) = 0$ and thus $a_ib_j - a_jb_i \in \mathbb{F}[V]^G$ for all $1 \leq i \neq j \leq 3$. Therefore, by the definition of a regular sequence, there must exist $h_1, h_2 \in \mathbb{F}[V]^G$ such that $a_2b_1 - a_1b_2 = h_1a_1 + h_2a_2$. Thus $a_2(b_1 - h_2) = a_1(h_1 + b_2)$. Again, by the definition of a regular sequence, this implies that there exist $t \in \mathbb{F}[V]^G$ such that $b_1 - h_2 = ta_1$ and thus $b_1 \in \mathbb{F}[V]^G$. But then $a_1 = (\sigma - 1)b_1 = 0$. This contradiction shows that H must be equal to G, i.e., that G is generated by its bi-reflections. □

More general versions of this result and related material may be found in Kemper [65, §3], and Lorenz [77, §8].

Corollary 9.2.3. *Let V be a faithful modular representation of a p group G defined over a field $\mathbb{F}$. Then $\mathbb{F}[3\,V]^G$ is not Cohen-Macaulay.*

Proof. Consider $e \neq \sigma \in G$. Then $\dim V^\sigma \leq \dim V - 1$. Therefore $\dim(3\,V)^\sigma \leq 3(\dim V - 1) = \dim(3\,V) - 3$. Thus no non-trivial element of G can act on $3\,V$ as a bi-reflection. □

Example 4.0.4 is a simple example of Corollary 9.2.3. Historically it was study of this example which lead to the development of the proof of Theorem 9.2.2.

Theorem 3.9.2 provides a partial converse to Theorem 9.2.2, by giving a class of representations for which all elements are bi-reflections and for which the ring of invariants is always Cohen-Macaulay. Note that the converse of Theorem 9.2.2 is not valid. For example, at the end of §1.1.1, we exhibited an example of a group generated by reflections whose ring of invariants is not Cohen-Macaulay.

Let H be a subgroup of G such that H contains a p-Sylow subgroup of G. We have already seen in Proposition 9.0.18 that this condition means there is closer relation between $\mathbb{F}[V]^H$ and $\mathbb{F}[V]^G$ than in general. The next result shows there is another similarity between these two rings.

Proposition 9.2.4. *Suppose that H is a subgroup of G such that $[G:H]$ is invertible in $\mathbb{F}$. If $\mathbb{F}[V]^H$ is Cohen-Macaulay, then $\mathbb{F}[V]^G$ is also Cohen-Macaulay.*

Proof. Choose a homogeneous system of parameters $f_1, f_2, \ldots, f_n$ for $\mathbb{F}[V]^G$. Then by Corollary 3.0.6, this is also a homogeneous system of parameters for $\mathbb{F}[V]^H$ and $\mathbb{F}[V]$. Let r denote the index of H in G. Recall that $1/r\,\mathrm{Tr}^G_H : \mathbb{F}[V]^H \to \mathbb{F}[V]^G$ is the Reynolds operator and accordingly we have the $\mathbb{F}[V]^G$-module decomposition $\mathbb{F}[V]^H = \mathbb{F}[V]^G \oplus U$ where $U := \ker \mathrm{Tr}^G_H$.

Let $B := \mathbb{F}[f_1, f_2, \ldots, f_n]$ and let I denote the ideal of $\mathbb{F}[V]^H$ generated by $(f_1, f_2, \ldots, f_n)$ and let $J := I \cap \mathbb{F}[V]^G$. The above decomposition implies that

$$\mathbb{F}[V]^H/I \cong \mathbb{F}[V]^G/J \oplus U/K$$

where $K = f_1U + f_2U + \cdots + f_nU$. Choose a basis

$$\{\overline{h}_1, \overline{h}_2, \ldots, \overline{h}_r\}$$

for $\mathbb{F}[V]^G/J$ and another basis

$$\{\overline{h}_{r+1}, \overline{h}_{r+2}, \ldots, \overline{h}_s\}$$

for U/K. Then $\{\overline{h}_1, \overline{h}_2, \ldots, \overline{h}_s\}$ is a basis for $\mathbb{F}[V]^H/I$. Lift each $\overline{h}_i$ to an element $h_i \in \mathbb{F}[V]^H$. Then

$$\mathbb{F}[V]^H = \sum_{i=1}^{s} Bh_i$$

and $\mathbb{F}[V]^G = \sum_{i=1}^{r} Bh_i$. But since $\mathbb{F}[V]^H$ is Cohen-Macaulay, we have in fact that $\mathbb{F}[V]^H = \oplus_{i=1}^{s} Bh_i$ and $\mathbb{F}[V]^G = \oplus_{i=1}^{r} Bh_i$. Thus, since it is a free B-module, $\mathbb{F}[V]^G$ is Cohen-Macaulay. □

Remark 9.2.5. Let V be a fixed non-trivial modular representation of the group G. Then Gregor Kemper has proved that for sufficiently large m, the ring of vector invariants $\mathbb{F}[mV]^G$ is not Cohen-Macaulay. See [65]. It remains an open question whether $m = 3$ is sufficiently large. Suppose G is a p-group generated by reflections and $\mathbb{F}[V]^G$ is Cohen-Macaulay. There are examples that show that sometimes, but not always, $\mathbb{F}[2V]^G$ is Cohen-Macaulay. It is an open problem to characterize those V for which $\mathbb{F}[2V]^G$ is Cohen-Macaulay. See the work of Chuai [23] and Shank and Wehlau [99].

9.3 Differents in Modular Invariant Theory

Here we recall work of A. Broer [13], closely following his exposition. His work is best understood in the context of §2.5. That is, suppose that $R \subset S$ are connected graded commutative rings with unit, that S is integral over R, both of which are domains and integrally closed in their quotient fields $\mathrm{Quot}(R)$ and $\mathrm{Quot}(S)$, and that the extension $\mathrm{Quot}(R) \subset \mathrm{Quot}(R)$ is separable. Then the trace map $\mathrm{Tr} : \mathrm{Quot}(S) \to \mathrm{Quot}(R)$ is surjective, respects the grading, and $\mathrm{Tr}(S) \subseteq R$.

We suppose that, as a R-module, S has homogeneous generators $s_1, \ldots, s_m$, and we define an epimorphism of algebras

$$R[a_1, a_2, \ldots, a_m] \overset{\rho}{\twoheadrightarrow} S \text{ by the rule}$$
$$\rho(a_\ell) = s_\ell$$

where $\deg(a_\ell) = \deg(s_\ell)$ with kernel I. We refer to elements of I as relations in S and to I as the ideal of relations. Now S integral over R implies that there is a monic polynomial

$$f_\ell(t) = \sum r_{j,\ell} t^j, \text{ where } r_{j,\ell} \in R \text{ and } f_\ell(s_\ell) = 0 \ .$$

As before, we write R_+ for the ideal of positive degree elements of R and we consider the Hilbert ideal R_+S of S. The quotient algebra $S/(R_+S)$ is a finite dimensional algebra. In the event that $R = S^G$, we call $S/(R_+)^G$ the ring of coinvariants of G.

9.3.1 Construction of the Various Different Ideals

The Noether Different Ideal

The enveloping algebra of the extension S over R is the algebra $S^e = S \otimes_R S$. We define J to be the kernel of the multiplication map

$$\begin{aligned} S \otimes_R S &\xrightarrow{\mu} S \\ s \otimes s' &\to ss' . \end{aligned}$$

Then J is generated by $s_\ell \otimes 1 - 1 \otimes s_\ell$, $1 \leq \ell \leq m$. We define $\mathfrak{K}$ to be the annihilator of J in S^e and the *Noether different ideal* as the image $\mu(\mathfrak{K}) = \mathfrak{D}^N_{S/R} = \mathfrak{D}^N \subset S$. In other words, $s \in \mathfrak{D}^N$ if and only if there is a $t \in S^e$ such that $\mu(t) = s$ and for all $s' \in S$ we have $(1 \otimes s')t = (s' \otimes 1)t$.

The Galois Different Ideal

Consider the natural map associated to an element $\sum t_\ell \otimes t'_\ell \in S^e$

$$\begin{aligned} \omega : S \otimes S &\to \operatorname{End}_R(S) \\ [\omega(t_\ell \otimes t'_\ell)](s) &= \sum t_\ell \operatorname{Tr}(t'_\ell s) \ . \end{aligned}$$

Then ω is a S^e-module map under the natural S^e-module structure on $\operatorname{End}_R(S)$, and the image $\omega(S^e) \subset \operatorname{End}_R(S)$ is the cyclic S^e-module generated by Tr. Let

$$\begin{aligned} \operatorname{End}_R(S) &\xrightarrow{\varepsilon} S \\ \varepsilon(\eta) &= \eta(1_S) \ . \end{aligned}$$

We define the *Galois different ideal* to be

$$\mathfrak{D}^{\mathrm{Gal}}_{S/R} = \mathfrak{D}^{\mathrm{Gal}} = \{\varepsilon(\eta) \mid \eta \in \omega(S^e) \cap \operatorname{End}_R R\} = \varepsilon(\omega(S^e) \cap \operatorname{End}_R R \ .$$

We note that $\operatorname{End}_R R$ is the collection of maps $R \to R$ given by multiplication by the elements of R, the elements of the Galois different are the maps $s \to \sum t_\ell \operatorname{Tr}(t_\ell s)$.

There is an alternate description of $\mathfrak{D}^{\mathrm{Gal}}$ in the case that $R = S^G$. We consider the twisted group ring $(SG, \cdot)$ which is additively the free S-module on generators $\{e_\sigma \mid \sigma \in G\}$ with multiplication determined by the formula

$$s_1 e_\sigma \cdot s_2 e_\tau = s_1 \sigma(s_2) e_{\sigma\tau}.$$

The unit of the group is the unit of the twisted group ring, and $(SG, \cdot)$ is naturally a graded module over S^e. There is also a natural inclusion of associative rings $(SG, \cdot) \to \operatorname{End}_R(S)$ taking $\alpha = \sum s_\ell e_{\sigma_\ell}$ to the R-linear map $\rho_\alpha(s) = \sum s_\ell \sigma_\ell(s)$. Of course, the element $\sum_{\sigma \in G} e_\sigma$ determines the trace over G. The Galois different is equal to the intersection of the two-sided ideal generated by $\sum_{\sigma \in G} e_\sigma$ with the subring S (that is $S[\mathrm{Id}_G]$). Broer notes that an extension $S^G \subset S$ is said to be Galois if the Galois different ideal is the trivial ideal (1_S), providing a rationale for the language.

The Inverse Dedekind Different Ideal

The *inverse Dedekind different ideal* is defined as

$$(\mathfrak{D}^D)^{-1} = (\mathfrak{D}^D_{S/R})^{-1} = \{r \in \mathrm{Quot}(R) \mid \text{ for all } s \in S, \mathrm{Tr}(rs) \in R\}$$

We note that, for every R-linear map $\eta : S \to R$, there is a unique $r \in (\mathfrak{D}^D)^{-1}$ such that $\eta(s) = \mathrm{Tr}(rs)$ for all $s \in S$. Therefore, we define the *Dedekind different ideal* to be

$$\mathfrak{D}^D = \mathfrak{D}^D_{S/R} = \left\{r \in \mathrm{Quot}(R) \mid r(\mathfrak{D}^D)^{-1} \subset R\right\} \ .$$

Under our hypotheses, the Dedekind different is a divisorial graded ideal in S, that is, the height of any one of its primary components is one. If, in addition, S is a unique factorization domain, then the Dedekind different ideal is a homogeneous principal ideal. In this case, we denote a generator by θ.

The Twisted Trace Different Ideal

Here we define the twisted trace different ideal or T-different ideal to be

$$\mathfrak{D}^T = \mathfrak{D}^T_{S/R} = \mathfrak{D}^D(\mathrm{Tr}((\mathfrak{D}^D)^{-1})).$$

When $\mathfrak{D}^D = (\theta)$, we define the twisted trace to be the R-module homomorphism given by

$$S \overset{\mathrm{Tr}_\theta}{\to} S$$
$$\mathrm{Tr}_\theta(s) = \theta\, \mathrm{Tr}(s\theta) \ .$$

In this case, since $\frac{S}{\theta} = (\mathfrak{D}^D_{S/R})^{-1}$, we do have $\mathrm{Tr}_\theta(s\theta) \in \theta R \subset S$. Then we have that

$$\mathfrak{D}^D = S\, \mathrm{Tr}_\theta(S).$$

The Kähler Different Ideal

Recall that J is defined to be the kernel of the multiplication map $S^e = S \otimes S \to S$. The graded $S = S^e/J$-module of *Kähler differentials* is defined to be

$$\Omega = \Omega_{S/R} = J/J^2.$$

Recall the ideal of relations I determined by

$$I \subset R[a_1, a_2, \ldots, a_m] \overset{\rho}{\to} S \ .$$

We suppose that I is generated by $\{f_1, f_2, \ldots, f_k\}$. For each such f_ℓ, we define

$$df_\ell = \frac{\partial f_\ell}{\partial t_1} dt_1 + \frac{\partial f_\ell}{\partial t_2} dt_2 + \cdots + \frac{\partial f_\ell}{\partial t_m} dt_m$$

where

$$\frac{\partial f_\ell}{\partial t_j} = \mu(\frac{\partial f_\ell}{\partial a_j}).$$

Then Ω is generated by the $dt_\ell = t_\ell \otimes 1 - 1 \otimes t_\ell + J^2$, $1 \leq \ell \leq m$ subject to the relations just given and we obtain a presentation of Ω by free graded S-modules

$$F_1 \to F_0 \to \Omega_{S/R}$$

where F_1 is of rank k and F_0 is of rank m. The S-linear map between F_1 and F_0 is associated to the $m \times k$ matrix $\mathcal{A}$ with entries

$$\mathcal{A}_{i,j} = \frac{\partial f_j}{\partial t_i} \ .$$

The *Kähler different ideal* or *Jacobian ideal*, denoted $\mathfrak{D}^K = \mathfrak{D}^K_{S/R}$, is the annihilator $\mathrm{Ann}_S(\Omega)$, in other words, the zero-th Fitting ideal of Ω. We note that $\mathfrak{D}^K$ is generated by the determinants of all $m \times m$-minors of $\mathcal{A}$. The elements of this theory show that the Kähler different ideal does not depend upon the choices of presentation or homogeneous bases.

For any m relations $\{f_1, f_2, \ldots, f_m\} \subset R[a_1, a_2, \ldots, a_m]$, the *Jacobian determinant* is defined to be

$$\mathrm{Jac}(f_1, f_2, \ldots, f_m) = \det(\frac{\partial f_j}{\partial t_i}) \in S$$

and the Kähler different ideal is the ideal generated by the Jacobians associated to the $\binom{k}{m}$ m-tuples of generators of I.

Now suppose that $S = \mathbb{F}[x_1, x_2, \ldots, x_n]$ is a graded polynomial algebra on algebraically independent generators x_ℓ. We consider the presentation of R-algebras given by

$$I \to R[a_1, a_2, \ldots, a_n] \xrightarrow{\rho} S$$

given by $\rho(a_\ell) = x_\ell$. Each element $r \in R$ can be written as a polynomial in the x_ℓ's, that is, $r = r(x_1, x_2, \ldots, x_n)$, and we can construct an element $r(a) = r(a_1, a_2, \ldots, a_n) \in R[a_1, a_2, \ldots, a_n]$. On the other hand, we can simply consider r as a constant in $R[a_1, a_2, \ldots, a_n]$. Then the difference $r(a) - r \in I$ is a relation. Therefore, for any $1 \leq \ell \leq n$ we have

$$\frac{\partial(r(a) - r)}{\partial x_\ell} = \mu(\frac{\partial(r(a) - r)}{\partial a_\ell}) = \mu(\frac{\partial(r(a))}{\partial a_\ell}) = \frac{\partial(r(x))}{\partial x_\ell}.$$

This latter expression is the usual partial derivative of R as an element of $\mathbb{F}[x_1, x_2, \ldots, x_n]$. It follows in this case that

$$\mathrm{Jac}(\frac{\partial(r(a) - r)}{\partial x_\ell}) = \mathrm{Jac}(\frac{\partial(r(a))}{\partial x_\ell}),$$

the classical Jacobian. Further, since the relations $r(a) - r$ generate I, we see that $\mathfrak{D}^K$ is generated by the Jacobians of the m-tuples of elements from R.

Differents and Properties of Modular Invariant Rings

Broer has determined the inclusions of the various differents as follows.

Theorem 9.3.1. *1. We have $\mathfrak{D}^K \subseteq \mathfrak{D}^N \subseteq \mathfrak{D}^{Gal} \subseteq \mathfrak{D}^D$ and $\mathfrak{D}^N \subseteq \operatorname{Ann} \Omega$;*
2. We have $\mathfrak{D}^T \subset \mathfrak{D}^D$ and if $\mathfrak{D}^T$ is a principal ideal, then $\mathfrak{D}^{Gal} \subseteq \mathfrak{D}^T$;
3. If S is free as a graded R-module, then $\mathfrak{D}^N = \mathfrak{D}^{Gal} = \mathfrak{D}^T = \mathfrak{D}^D$.

□

Broer's techniques allow him to analyze how various properties of modular invariant rings are related to properties of the differents. We recall here two of his theorems.

First, recall that $S/(R_+)$ is said to be a Poincaré duality algebra if there exists a non-degenerate symmetric bilinear mapping

$$S/(R_+) \otimes S/(R_+) \to S/(R_+)$$

Theorem 9.3.2. *The following statements for $R \subset S$ are equivalent:*

1. *$S/(R_+)$ is a Poincaré duality algebra;*
2. *$\mathfrak{D}^N$ is a principal ideal;*
3. *$\mathfrak{D}^{Gal}$ is a principal ideal;*
4. *$\mathfrak{D}^D$ is a principal ideal and $\mathfrak{D}^N = \mathfrak{D}^{Gal} = \mathfrak{D}^D$.*

□

In the case of most interest for this chapter, Broer proves

Theorem 9.3.3. *We assume that $S = \mathbb{F}[x_1, \dots, x_n]$ is a graded polynomial algebra with $\deg(x_\ell) > 0$ for all $\ell = 1, 2, \dots, n$. We consider the extension $R \subset S$ as in this section. Then the Dedekind different $\mathfrak{D}^D = (\theta)$ is principal, and the following statements are equivalent.*

1. *The ring R is also a polynomial algebra;*
2. *S is free as graded R-module;*
3. *$S/(R_+)$ is a Poincarduality algebra;*
4. *Either $\mathfrak{D}^K$ or $\mathfrak{D}^N$ or $\mathfrak{D}^{Gal}$ equals $\mathfrak{D}^D$;*
5. *We have $\mathfrak{D}^K = \mathfrak{D}^N = \mathfrak{D}^{Gal} = \mathfrak{D}^T = \mathfrak{D}^D$;*
6. *There exists $\{f_1, f_2, \dots, f_n\}$ such that $\sum \deg(f_\ell) - \deg(x_\ell) = \deg(\theta)$ and $0 \neq \operatorname{Jac}(f_1, f_2, \dots, f_n)$;*
7. *There exists $\{f_1, f_2, \dots, f_n\}$ such that $R = \mathbb{F}[f_1, f_2, \dots, f_n]$.*

□

10

Invariant Rings via Localization

In some circumstances, it is possible to study a ring of invariants through a well-chosen localization. In this chapter, we give some instances of this technique.

Let R be a finitely generated algebra which is a domain. Recall that R_f denotes the localization of R with respect to the multiplicative set generated by f.

Proposition 10.0.1. *Suppose that B is a subalgebra of R and that f_1, f_2 is a regular sequence in B such that $B_{f_1} = R_{f_1}$ and $B_{f_2} = R_{f_2}$. Then $B = R$.*

Proof. Take $h \in R$. Since $R \subseteq R_{f_i} = B_{f_i}$ we may write $h = b_1/f_1^n$ and $h = b_2/f_2^m$ for some $b_1, b_2 \in B$ and $n, m \in \mathbb{N}$. Therefore, $b_1 f_2^m = b_2 f_1^n$. Since f_1, f_2 is a regular sequence in B, so also is f_1^n, f_2^m. This implies that $b_2 = b f_2^m$ for some $b \in B$. Thus $h = b f_2^m / f_2^m = b$ lies in B. □

Lemma 10.0.2. *Suppose that B is a subalgebra of R and I is an ideal of R. If $I \subseteq B$ and $f \in \sqrt{I} \cap B$, then $B_f = R_f$.*

Proof. There exists $m \in \mathbb{N}$ such that $f^m \in I$. Take $h \in R_f$ and write $h = r/f^k$ with $r \in R$. Then $rf^m = hf^{k+m} \in I \subseteq B$ and $h = rf^m/f^{k+m} \in B_f$. □

Theorem 10.0.3. *Suppose B is a subalgebra of $\mathbb{F}[V]^G$ containing the image of the transfer, $\operatorname{Im}\operatorname{Tr}^G \subset B$. Let $f_1, f_2 \in \sqrt{\operatorname{Im}\operatorname{Tr}^G} \cap B$ be two elements such that f_1, f_2 is a regular sequence in B. Then $B = \mathbb{F}[V]^G$.*

Remark 10.0.4. Suppose that B is a subalgebra of $\mathbb{F}[V]^G$ containing the image of the transfer. Further, suppose that f_1 and f_2 are non-associate primes of B lying in $\sqrt{\operatorname{Im}\operatorname{Tr}^G}$. A relatively routine calculation shows that f_1, f_2 is a regular sequence in B and so $B = \mathbb{F}[V]^G$.

Remark 10.0.5. Let $\overline{\mathbb{F}}$ denote the algebraic closure of $\mathbb{F}$ and consider $\overline{V} := \overline{\mathbb{F}} \otimes V$. By Theorem 9.0.10, we have that $\sqrt{\operatorname{Im}\operatorname{Tr}^G}$ consists of those invariant

H.E.A. Eddy Campbell, D. Wehlau, *Modular Invariant Theory*,
Encyclopaedia of Mathematical Sciences 139,
DOI 10.1007/978-3-642-17404-9_10, © Springer-Verlag Berlin Heidelberg 2011

polynomials in $\mathbb{F}[V]$ which vanish on the subvariety $\mathcal{V}$ of $\overline{V}$ defined by $\mathcal{V} = \cup_{\sigma\in\Sigma}\overline{V}^{\sigma}$ where Σ consists of all the elements σ of G of order p. Thus given an element $f \in \mathbb{F}[V]^G$, we may check that $f \in \sqrt{\operatorname{Im}\operatorname{Tr}^G}$ by verifying that f vanishes on $\overline{V}^{\sigma}$ for every element $\sigma \in G$ of order p.

We illustrate Theorem 10.0.3 by using it to compute (again) the ring of C_p invariants, $\mathbb{F}[2\,V_2]^{C_p}$. Let σ be a generator for C_p. Choose a triangular basis y_2, x_2, y_1, x_1 for $(2\,V_2)^*$ with $\sigma(y_i) = y_i + x_i$ and $\sigma(x_i) = x_i$ for $i = 1, 2$. We use the graded reverse lexicographic order on $\mathbb{F}[2\,V_2]$ with $x_1 < y_1 < x_2 < y_2$. We take $N_i = \mathbf{N}(y_i) = y_i^p - x_i^{p-1}y_i$ for $i = 1, 2$ and $u = x_2y_1 - x_1y_2$. Let B denote the ring $B = \mathbb{F}[x_1, x_2, u, N_1, N_2]$. We will use Theorem 10.0.3 to show that $\mathbb{F}[2\,V_2]^{C_p} = B$.

We begin by verifying that $\operatorname{Im}\operatorname{Tr}^{C_p} \subset B$.

Lemma 10.0.6. *Let $0 \le a, b \le p-1$ and put $e := a + b - (p-1)$.*

$$\operatorname{Tr}^{C_p}(y_1^a y_2^b) = \begin{cases} 0, & \textit{if } e < 0;\\ -u^{p-1} - (x_1x_2)^{p-1}, & \textit{if } a = b = p-1;\\ -\binom{a}{e} x_1^{p-1-b} x_2^{p-1-a} u^e, & \textit{otherwise.}\end{cases}$$

Proof.

$$\begin{aligned}\operatorname{Tr}^{C_p}(y_1^a y_2^b) &= \sum_{k=0}^{p-1} (y_1 + kx_1)^a (y_2 + kx_2)^b\\ &= \sum_{k=0}^{p-1}\sum_{i=0}^{a}\binom{a}{i} k^i x_1^i y_1^{a-i} \sum_{j=0}^{b}\binom{b}{j} k^j x_2^j y_2^{b-j}\\ &= \sum_{i=0}^{a}\sum_{j=0}^{b}\binom{a}{i}\binom{b}{j} x_1^i y_1^{a-i} x_2^j y_2^{b-j} \sum_{k=0}^{p-1} k^{i+j}\ .\end{aligned}$$

By Lemma 9.0.2, $\sum_{k=0}^{p-1} k^{i+j}$ is zero unless $i + j = p - 1$ or $i + j = 2(p-1)$ in which case this sum is -1. Therefore, $\operatorname{Tr}^{C_p}(y_1^a y_2^b) = 0$ if $a+b < p-1$. Moreover, if $p - 1 \le a + b < 2(p-1)$, then

$$\begin{aligned}\operatorname{Tr}^{C_p}(y_1^a y_2^b) &= -\sum_{i=0}^{a}\binom{a}{i}\binom{b}{p-1-i} x_1^i y_1^{a-i} x_2^{p-1-i} y_2^{b-p+1+i}\\ &= -\sum_{i=p-1-b}^{a}\binom{a}{i}\binom{b}{p-1-i} x_1^i y_1^{a-i} x_2^{p-1-i} y_2^{b-p+1+i}\\ &= -\sum_{t=0}^{a+b-p+1}\binom{a}{p-1-b+t}\binom{b}{b-t} x_1^{p-1-b+t} y_1^{a+b-p+1-t} x_2^{b-t} y_2^t\end{aligned}$$

$$= -x_1^{p-1-b}x_2^{p-a-1}\left(\sum_{t=0}^{a+b-p+1}\binom{a}{p-1-b+t}\binom{b}{b-t}\right.$$
$$\left. x_1^t y_1^{a+b-p+1-t} x_2^{b-t-p+1+a} y_2^t \right)$$
$$= -x_1^{p-1-b}x_2^{p-a-1}\sum_{t=0}^{e}\binom{a}{p-1-b+t}\binom{b}{b-t}x_1^t y_1^{e-t}x_2^{e-t}y_2^t\ .$$

Wilson's Theorem asserts that $(p-1)! = -1 \pmod p$ and Lucas' Lemma gives $\binom{p-1}{s} = (-1)^s \pmod p$. Using these two facts we get

$$\begin{aligned}\binom{a}{p-1-b+t}\binom{b}{t} &= \frac{a!}{(p-1-b+t)!\,(a+b-p+1-t)!}\ \frac{b!}{t!\,(b-t)!}\\ &= \frac{a!}{(e-t)!\,t!}\ \frac{b!}{(p-1-b+t)!\,(b-t)!}\\ &= -\frac{a!\,b!}{(e-t)!\,t!}\ \frac{(p-1)!}{(p-1-b+t)!\,(b-t)!}\\ &= -\frac{a!\,b!}{(e-t)!\,t!}(-1)^{b-t}\\ &= -\frac{a!\,b!}{(e-t)!\,t!}(-1)^{b}(-1)^{t}\\ &= -(-1)^t\frac{a!\,b!}{(e-t)\,!t!}\ \frac{(p-1)!}{(p-1-b)!\,b!}\\ &= (-1)^t\frac{a!\,b!}{(e-t)!\,t!}\ \frac{1}{(p-1-b)!\,b!}\\ &= (-1)^t\frac{a!}{(p-1-b)!}\ \frac{b!}{b!(e-t)!\,t!}\\ &= (-1)^t\frac{a!}{(p-1-b)!\,e!}\ \frac{e!}{(e-t)!\,t!}\\ &= (-1)^t\binom{a}{e}\binom{e}{t}\ .\end{aligned}$$

Therefore, for $p-1 \le a+b < 2(p-1)$ we have

$$\begin{aligned}\mathrm{Tr}^{C_p}(y_1^a y_2^b) &= -x_1^{p-1-b}x_2^{p-a-1}\sum_{t=0}^{e}(-1)^t\binom{e}{t}x_1^t y_1^{e-t}x_2^{e-t}y_2^t\\ &= -\binom{a}{e}x_1^{p-1-b}x_2^{p-a-1}(x_2y_1 - x_1y_2)^e\\ &= -\binom{a}{e}x_1^{p-1-b}x_2^{p-a-1}u^e\ .\end{aligned}$$

For $a = b = p-1$, we get one extra term corresponding to $i+j = 2(p-1)$ which does not vanish. Thus $\mathrm{Tr}^{C_p}(y_1^{p-1}y_2^{p-1}) = -u^{p-1} - (x_1x_2)^{p-1}$. □

Using Lemma 10.0.6, we can see that $\mathrm{Tr}^{C_p}(\mathbb{F}[2\,V_2]) \subset \mathbb{F}[x_1, x_2, u, N_1, N_2]$ as follows. The four invariants x_1, x_2, N_1, N_2 form a homogeneous system of parameters for $\mathbb{F}[2\,V_2]$. Taking $H = \mathbb{F}[x_1, x_2, N_1, N_2]$ we have the Hironaka decomposition $\mathbb{F}[2\,V_2] = \oplus_{a=0}^{p-1} \oplus_{b=0}^{p-1} H y_1^a y_2^b$. Since Tr^{C_p} is H-linear, we see that $\mathrm{Tr}^{C_p}(\mathbb{F}[2\,V_2]) \subset \sum_{a=0}^{p-1}\sum_{b=0}^{p-1} H \cdot \mathrm{Tr}^{C_p}(y_1^a y_2^b) \subset \mathbb{F}[x_1, x_2, u, N_1, N_2]$.

Lemma 10.0.6 also shows that $x_1, x_2 \in \sqrt{\mathrm{Im}\,\mathrm{Tr}^{C_p}}$ since $\mathrm{Tr}(y_1^{p-1}) = x_1^{p-1}$ and $\mathrm{Tr}(y_2^{p-1}) = x_2^{p-1}$.

To use Theorem 10.0.3 to show that $\mathbb{F}[2\,V_2]^{C_p} = \mathbb{F}[x_1, x_2, u, N_1, N_2]$, it only remains to show that x_2, x_1 is a regular sequence in $B = \mathbb{F}[x_1, x_2, u, N_1, N_2]$. We begin by noting that x_1, x_2, u, N_1, N_2 is a SAGBI basis for B. This is clear since $\mathrm{LT}(x_1) = x_1$, $\mathrm{LT}(x_2) = x_2$, $\mathrm{LT}(u) = x_2 y_1$, $\mathrm{LT}(N_1) = y_1^p$ and $\mathrm{LT}(N_2) = y_2^p$. Thus the only non-trivial tête-a-tête difference is $u^p - x_2^p N_1$ which subducts to zero via $u^p - x_2^p N_1 + x_1^p N_2 - (x_1 x_2)^{p-1} u = 0$. In fact, we already found this SAGBI basis in Example 5.1.9.

To see that x_2, x_1 is a regular sequence in B, we suppose that $x_1 g = x_2 h$ for some $g, h \in B$. Then x_1 divides $\mathrm{LT}(h)$. We need to show that x_1 divides h in B. This follows from the following lemma.

Lemma 10.0.7. *Suppose $\mathcal{B}$ is a SAGBI basis for a ring $B \subset \mathbb{F}[x_1, x_2, \ldots, x_n]$ with respect to a graded reverse lexicographic term order with $x_1 < x_i$ for all $i = 2, 3, \ldots, n$. Further, suppose that $x_1 \in \mathcal{B}$ and that x_1 is the unique element of $\mathcal{B}$ whose lead term is divisible by x_1. Let $f \in B$. If x_1 divides $\mathrm{LT}(f)$, then x_1 divides f in B.*

Proof. Assume that the result is false. Then there exist elements $f \in B$ whose lead term is divisible by x_1 but for which $f/x_1 \notin B$. Choose such an f with minimal lead term. Since $\mathcal{B}$ is a SAGBI basis for B, we may write $\mathrm{LT}(f) = \mathrm{LT}(x_1)^{e_0}\,\mathrm{LT}(g_1)^{e_1}\,\mathrm{LT}(g_2)^{e_2} \cdots \mathrm{LT}(g_s)^{e_s}$ where $x_1, g_1, g_2, \ldots, g_s \in \mathcal{B}$ and $e_1, e_2, \ldots, e_s$ are non-negative integers. Since x_1 divides $\mathrm{LT}(f)$ and x_1 does not divide $\mathrm{LT}(g_i)$ for $i = 1, 2, \ldots, s$, we must have $e_0 \geq 1$. Define $f' = f - x_1^{e_0} g_1^{e_1} g_2^{e_2} \cdots g_s^{e_s}$. Then $\mathrm{LT}(f') < \mathrm{LT}(f)$. Since we are using a graded reverse lexicographic ordering with x_1 small, this implies that x_1 divides $\mathrm{LT}(f')$. Then by the minimality of f we see that $h' = f'/x_1 \in B$. But then $f = x_1(x_1^{e_0-1} g_1^{e_1} g_2^{e_2} \cdots g_s^{e_s} + h')$ is divisible by x_1 in B. □

An important application of the above lemma is in combination with the following proposition.

Proposition 10.0.8. *Let $A = \mathbb{F}[f_1, f_2, \ldots, f_r] \subseteq \mathbb{F}[V]^G$. Suppose*

1. *$\mathbb{F}[V]$ is integral over A;*
2. *$\mathrm{Quot}(A) = \mathbb{F}(V)^G$;*
3. *There exists $h \in A$ such that hA is a prime ideal of A and A_h is a unique factorization domain.*

Then $A = \mathbb{F}[V]^G$.

Proof. By [6, Lemma 6.3.1, p. 73], hypothesis (3) implies that A is itself a unique factorization domain and is therefore by Proposition 3.0.2 integrally closed. Let $f \in \mathbb{F}[V]^G$ be arbitrary. By (2), f lies in $\mathrm{Quot}(A)$. Then (1) implies that f is integral over A and since A integrally closed, $f \in A$. Therefore $A = \mathbb{F}[V]^G$. □

Verifying the first hypotheses of Proposition 10.0.8 is usually easy as all we need to do is ensure that A contains a homogeneous system of parameters. Verifying the second hypothesis is a little harder but the following result due to Kemper, [64, Cor 1.8], is often applicable or by the use of Theorem 4.3.4.

Proposition 10.0.9. *Let $f_1, f_2, \ldots, f_n$ be n algebraically independent homogeneous polynomials over a field $\mathbb{K}$. Then the degree of the field extension $[\mathbb{K}(x_1, x_2, \ldots, x_n) : \mathbb{K}(f_1, f_2, \ldots, f_n)]$ is at most $\prod_{i=1} \deg(f_i)$.*

Corollary 10.0.10. *Let $f_1, f_2, \ldots, f_n \in \mathbb{K}[V]^G$ be algebraically independent homogeneous invariants. If $\prod_{i=1} \deg(f_i) < 2|G|$, then $\mathbb{K}(f_1, f_2, \ldots, f_n) = \mathbb{K}(V)^G$.*

Proof. Since $\mathbb{K}(f_1, f_2, \ldots, f_n) \subseteq \mathbb{K}(V)^G \subseteq \mathbb{K}(x_1, x_2, \ldots, x_n)$, we have

$$\begin{aligned} 2|G| &> \prod_{i=1} \deg(f_i) \\ &\geq [\mathbb{K}(x_1, x_2, \ldots, x_n) : \mathbb{K}(f_1, f_2, \ldots, f_n)] \\ &= [\mathbb{K}(x_1, x_2, \ldots, x_n) : \mathbb{K}(V)^G] \cdot [\mathbb{K}(V)^G : \mathbb{K}(f_1, f_2, \ldots, f_n)] \\ &= |G| \cdot [\mathbb{K}(V)^G : \mathbb{K}(f_1, f_2, \ldots, f_n)]. \end{aligned}$$

Thus $[\mathbb{K}(V)^G : \mathbb{K}(f_1, f_2, \ldots, f_n)] < 2$. □

Verifying the third hypothesis of Proposition 10.0.8 is generally much harder than verifying the first two hypotheses. In particular, it is hard to show that the element h is prime in the ring A which is usually given by specifying a set of generators for A. One way to do this however is when the invariant h is x_1. If a SAGBI basis for A with respect to grevlex, with x_1 small, is known, then Lemma 10.0.7 can be used to show that the principal ideal $x_1 A$ is a prime ideal of A. We illustrate this with a simple example.

Example 10.0.11. We return to the 4 dimensional representation of ${C_p}^3$ considered in Section 8.2. This representation is given by

$$G = \left\{ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ a & c & 1 & 0 \\ c & b & 0 & 1 \end{pmatrix} \mid a, b, c \in \mathbb{F}_p \right\}.$$

Let $\{x_1, x_2, y_1, y_2\}$ be the basis of V^* dual to the standard basis of V. We will use the graded reverse lexicographic ordering with $x_1 < x_2 < y_1 < y_2$.

One may verify that the following functions are G-invariants:

$$\begin{aligned}
&x_1\\
&x_2\\
&h = x_2(y_2^p - x_2^{p-1}y_2) + x_1(y_1^p - x_1^{p-1}y_1) = x_2y_2^p + \dots\\
&f_1 = \mathbf{N}^G_{G_{y_1}}(y_1) = y_1^{p^2} + \dots\\
&f_2 = \mathbf{N}^G_{G_{y_2}}(y_2) = y_2^{p^2} + \dots
\end{aligned}$$

We will show that $\mathbb{F}[V]^G = \mathbb{F}[x_1, x_2, h, f_1, f_2]$ by applying Lemma 10.0.7 and Proposition 10.0.8. We seek a SAGBI basis for $\mathbb{F}[V]^G$. Accordingly, we consider the lone tête-a-tête among x_1, x_2, h, f_1, f_2 which is $h^p - x_2^p f_2$. We perform the subduction of this tête-a-tête and obtain

$$h^p - x_2^p f_2 - x_1^p f_1 - h\sum_{j=1}^{p}(x_1^{p-j+1}x_2^j)^{p-1} = 0\ .$$

Thus $\{x_1, x_2, h, f_1, f_2\}$ is a SAGBI basis for the algebra $A := \mathbb{F}[x_1, x_2, h, f_1, f_2]$ which they generate. Since $\mathrm{LT}(x_1) = x_1$, $\mathrm{LT}(x_2) = x_2$, $\mathrm{LT}(f_1) = y_1^{p^2}$ and $\mathrm{LT}(f_2) = y_2^{p^2}$, we see by Lemma 6.2.1 that x_1, x_2, f_1, f_2 forms a homogeneous system of parameters and so $\mathbb{F}[V]$ is integral over A.

Since the lead terms of x_1, x_2, h, f_1 are algebraically independent, so are these 4 polynomials themselves. Applying Corollary 10.0.10 we see that $\mathbb{F}(x_1, x_2, h, f_1) = \mathbb{F}(V)^G$.

By Lemma 10.0.7, we know that Ax is a prime ideal of A. Furthermore, from the subduction of the tête-a-tête above, we see that $f_1 \in \mathbb{F}[x_1, x_2, h, f_2]_{x_1}$ and therefore, $A_{x_1} = \mathbb{F}[x_1, x_2, h, f_1, f_2]_{x_1} = \mathbb{F}[x_1, x_2, h, f_2]_{x_1}$. Since $\mathbb{F}[x_1, x_2, h, f_2]$ is a polynomial ring, this implies that A_{x_1} is a unique factorization domain.

Applying Proposition 10.0.8 we see that $\mathbb{F}[V]^G = A = \mathbb{F}[x_1, x_2, h, f_1, f_2]$.

11

Rings of Invariants which are Hypersurfaces

As we have seen, we seek to characterize those representations V of groups G whose rings of invariants are well-behaved. The best behaved rings of invariants, $\mathbb{K}[V]^G$, are those which are polynomial rings, that is, $\mathbb{K}[V]^G$ is generated by $\dim(V)$ many invariants. A slightly less well behaved class of examples is provided by those rings of invariants which are hypersurfaces, that is, $\mathbb{K}[V]^G$ is generated by $\dim(V)+1$ many invariants. Those representations with this property have been extensively studied in characteristic 0 by Nakajima [84]. Less is known for modular groups. When G is a Nakajima group with maximal proper subgroup H, the following proposition shows that the ring of H-invariants is a hypersurface (or a polynomial) ring.

Proposition 11.0.1. *Let R be an integral domain of characteristic p and suppose the finite group G acts faithfully on R. Suppose $H \leq G$ is a maximal proper subgroup of index $k \leq p$. Let $\sigma \in G \setminus H$. If there exists $y \in R^H$ such that $x := (\sigma - 1)y$ lies in R^G and such that the set $(\sigma-1)R^H$ lies in the principal ideal of R generated by x, then $R^H = R^G[y]$.*

Proof. Since the group H is a proper subgroup of G, the field $\mathrm{Quot}(R)^G$ is a proper subfield of $\mathrm{Quot}(R)^H$ (in fact, the degree of the extension $\mathrm{Quot}(R)^G \subset \mathrm{Quot}(R)^H$ is $|G|/|H|$) and thus the ring R^G is a proper subring of R^H. Thus there exists $f \in R^H \setminus R^G$. Since G is generated by H together with σ, we see that $(\sigma-1)f$ is a non-zero element of R which is divisible by x. This shows that $x \neq 0$ and hence that y is not G-invariant. Therefore, $y \notin \mathrm{Quot}(R)^G$. By the maximality of H in G, this implies, using Galois Theory, that

$$\mathrm{Quot}(R)^H = \mathrm{Quot}(R)^G[y] = \oplus_{i=0}^{k-1} \mathrm{Quot}(R)^G y^i \ .$$

We will will show, using induction, that for every non-negative integer m,

$$R^H \cap \sum_{i=0}^{m} \mathrm{Quot}(R)^G y^i = R^H \cap \sum_{i=0}^{m} R^G y^i \ .$$

H.E.A. Eddy Campbell, D. Wehlau, *Modular Invariant Theory*, Encyclopaedia of Mathematical Sciences 139,

DOI 10.1007/978-3-642-17404-9_11,

For $m = 0$, we need to show that $R^H \cap \mathrm{Quot}(R)^G = R^H \cap R^G$, i.e., that $R^H \cap \mathrm{Quot}(R)^G = R^G$. Clearly, $R^G \subseteq R^H \cap \mathrm{Quot}(R)^G$. Conversely, if $h \in R^H \cap \mathrm{Quot}(R)^G$, then h is an element of R which is fixed by G and thus $h \in R^G$.

For the general case, $m \geq 1$, we choose a set of left coset representatives, denoted G/H, for H in G. Consider the polynomial

$$F(t) = \prod_{g \in G/H} (t - g(y)) = t^k + f_{k-1}t^{k-1} + \cdots + f_0 \in R^G[t] \ .$$

One of the roots of $F(t)$ is the element y since one of the coset representatives lies in H. Therefore, $y^k \in \sum_{i=0}^{k-1} R^G y^i$.

First, we consider the cases $m < k$. Take $f \in R^H \cap \sum_{i=0}^{m} \mathrm{Quot}(R)^G y^i$. Write $f = \sum_{i=0}^{m} f_i y^i$ where $f_i \in \mathrm{Quot}(R)^G$. Since $(\sigma - 1)f$ lies in the principal ideal of R generated by x, we may write $(\sigma - 1)f = xf'$ where $f' \in R$. Thus

$$\begin{aligned} xf' &= (\sigma - 1)f = (\sigma - 1)\sum_{i=0}^{m} f_i y^i = \sum_{i=0}^{m} \sigma(f_i y^i) - \sum_{i=0}^{m} f_i y^i \\ &= \sum_{i=0}^{m} f_i (y + x)^i - \sum_{i=0}^{m} f_i y^i \\ &= m f_m y^{m-1} x + \sum_{j=0}^{m-2} \binom{m}{j} f_m y^j x^{m-j} + \sum_{i=0}^{m-1} f_i (y+x)^i - \sum_{i=0}^{m-1} f_i y^i \ . \end{aligned}$$

Thus $xf' = mf_m y^{m-1}x + w$ where $w \in \sum_{i=0}^{m-2} \mathrm{Quot}(R)^G y^i$. Since $x \in R^G$, this shows that $f' = mf_m y^{m-1} + w/x$ lies in $\sum_{i=0}^{m-1} \mathrm{Quot}(R)^G y^i$. Therefore, $f' \in \mathrm{Quot}(R)^G[y] \cap R = \mathrm{Quot}(R)^H \cap R = R^H$. Hence $f' \in R^H \cap \sum_{i=0}^{m-1} \mathrm{Quot}(R)^G y^i$ and thus by the induction hypothesis, $f' \in \sum_{i=0}^{m-1} R^G y^i$.

Since

$$\mathrm{Quot}(R)^H = \oplus_{i=0}^{k-1} \mathrm{Quot}(R)^G y^i$$

and since $m < k$, we see that $\{1, y, y^2, \ldots, y^{m-1}\}$ is linearly independent over $\mathrm{Quot}(R)^G$ and thus

$$\sum_{i=0}^{m-1} R^G y^i = \oplus_{i=0}^{m-1} R^G y^i.$$

Therefore, $mf_m y^{m-1} = f' - w/x \in R^G y^{m-1}$ and thus $mf_m \in R^G$. But $m < k \leq p$ implies that m is invertible in R and thus $f_m \in R^G$.

Now the induction hypothesis applied to the element

$$f - f_m y^m = \sum_{i=0}^{m-1} f_i y^i$$

show that it lies in

$$R^H \cap \sum_{i=0}^{m-1} R^G y^i$$

and therefore, $f \in \sum_{i=0}^{m} R^G y^i$, as required.

Thus we have shown that $R^H \cap \operatorname{Quot}(R)^G[y] = R^H \cap R^G[y]$. Combining this with $\operatorname{Quot}(R)^G[y] = \operatorname{Quot}(R)^H$ gives $R^H \cap \operatorname{Quot}(R)^H = R^H \cap R^G[y]$, i.e., $R^H = R^G[y]$. This completes the cases $m < k$.

For the cases $m \geq k$, we utilize the fact $y^k \in \sum_{i=0}^{k-1} R^G y^i$ which we showed above. Using this we have $\sum_{i=0}^{m} R^G y^i = \sum_{i=0}^{k-1} R^G y^i$ and $\sum_{i=0}^{m} \operatorname{Quot}(R)^G y^i = \sum_{i=0}^{k-1} \operatorname{Quot}(R)^G y^i$. Thus for $m \geq k$, using the induction hypothesis, we are done. □

The prototypical example of Proposition 11.0.1 is the action of C_p on $2\,V_2$ given in §1.12.

Proposition 11.0.2. *The hypotheses of Proposition 11.0.1 imply that H is a normal subgroup of index p in G.*

Proof. Define $z := y^p - x^{p-1}y$. It is easy to verify that z is fixed by σ and thus $z \in R^G$. Consider the polynomial $F(t) := t^p - x^{p-1}t - z \in R^G[t]$. Note that for every $k \in \mathbb{F}_p$, we have $F(y + kx) = (y + kx)^p - x^{p-1}(y + kx) - z = y^p + k^p x^p - x^{p-1}y - kx^p - z = y^p - x^{p-1}y - z + kx^p - kx^p = 0$. Thus $F(t) = \prod_{k=0}^{p-1}(t - (y + kx))$. Since $F(t)$ is G-invariant (having declared that G fixes t), we see that each element $g \in G$ must permute the factors of F and hence also the roots of F. Thus for each $g \in G$, there exists $k \in \mathbb{F}_p$ such that $g(y) = y + kx$.

We define $\phi : G \to \mathbb{F}_p$ by $\phi(g) = (g(y) - y)/x$. Note that for $g_1, g_2 \in G$, if $g_1(y) = y + k_1x$ and $g_2(y) = y + k_2x$, then $(g_2g_1)(y) = g_2(g_1(y)) = g_2(y + k_1x) = g_2(y) + k_1x = y + k_2x + k_1x$. Thus $\phi(g_2g_1) = \phi(g_2) + \phi(g_1)$ and thus ϕ is a group homomorphism.

Since $y \notin R^G$, we know that $\phi(\sigma) \neq 0$ and thus ϕ is a surjective group homomorphism. Furthermore, since $y \in R^H$, the subgroup H lies in the kernel of ϕ. But since H is a maximal proper subgroup of G, this means H is the kernel of ϕ. Therefore, H is a normal subgroup. Finally, since the index of the kernel of ϕ equals the cardinality of the image of ϕ, we see that the index of H is $|\mathbb{F}_p| = p$. □

Example 11.0.3. This example illustrates the use of Proposition 11.0.1. It is also interesting as a simple example of a reflection group whose ring of invariants fails to be a polynomial ring. Let $\mathbb{F}$ denote a field of characteristic p containing an element ϵ satisfying $\epsilon^p - \epsilon = 1$. For example, if $\epsilon^p - \epsilon = 1$, we could take $\mathbb{F} = \mathbb{F}_p[\epsilon]$, a field of order p^p.

We consider the representation generated by the three reflections

$$\alpha^{-1} = \begin{pmatrix} 1&0&0&0\\ a&1&0&0\\ 0&0&1&0\\ 0&0&0&1 \end{pmatrix}, \beta^{-1} = \begin{pmatrix} 1&0&0&0\\ 0&1&0&0\\ 0&0&1&0\\ 0&0&1&1 \end{pmatrix}, \gamma^{-1} = \begin{pmatrix} 1&0&0&0\\ 0&1&0&0\\ 0&0&1&0\\ 1&0&\epsilon&1 \end{pmatrix}$$

with respect to the standard basis of V. We let $\{y_2, x_2, y_1, x_1\}$ denote the dual basis of V^*.

We denote by H the group generated by $\{\alpha, \beta, \gamma\}$ and define G to be the group generated by $\{\alpha, \beta, \gamma, \sigma\}$ where

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

Then G is a Nakajima group and we will use Proposition 11.0.1 to show that H has a hypersurface as a ring of invariants.

Further, we let K denote the group generated by $\{\alpha, \beta\}$. Then $K \cong C_p \times C_p$ and as in §1.12, we have $\mathbb{F}[V]^K = \mathbb{F}[x_1, y_1^p - x_1^{p-1} y_1, x_2, y_2^p - x_2^{p-1} y_2]$.

We define $Y_1 := y_1^p - x_1^{p-1} y_1$, $X_1 := (\gamma - 1) Y_1 = (\epsilon^p - \epsilon) x_1^p = x_1^p$, $Y_2 := y_2^p - x_2^{p-1} y_2$ and $X_2 := (\gamma - 1) Y_2 = x_1^p - x_2^{p-1} x_1$. Then $\mathbb{F}[V]^K = \mathbb{F}[Y_2, x_2, Y_1, x_1]$.

We will compute $\mathbb{F}[V]^H = (\mathbb{F}[V]^K)^{H/K}$ by applying Proposition 11.0.1 to $R = \mathbb{F}[V]^K$ and the index p subgroup H/K of G/K.

Since σ fixes x_1, y_1 and x_2 and $\sigma(y_2) = y_2 + x_1$, we see that σ fixes X_1, Y_1 and X_2 and $\sigma(Y_2) = Y_2 + (x_1^p - x_2^{p-1} x_1) = Y_2 + X_2$. Therefore, applying $(\sigma - 1)$ to an arbitrary monomial $m = x_1^{a_1} Y_1^{b_1} x_2^{a_2} Y_2^{b_2}$ of R gives

$$\begin{aligned} (\sigma_1 - 1) m &= x_1^{a_1} x_2^{a_2} Y_1^{b_1} (\sigma_1 - 1) Y_2^{b_2} \\ &= x_1^{a_1} x_2^{a_2} y_2^{b_2} (Y_2 + X_2)^{b_2} - Y_2^{b_2} \\ &= x_1^{a_1} x_2^{a_2} y_2^{b_2} \sum_{j=1}^{b_2} \binom{b_2}{j} Y_2^{b_2 - j} X_2^j \ . \end{aligned}$$

Thus $(\sigma_1 - 1) m$ is always divisible by X_2 and therefore X_2 divides $(\sigma_1 - 1) f$ for all $f \in R$.

Defining τ by $\gamma = \tau \sigma$ gives

$$\gamma^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \epsilon & 1 \end{pmatrix} .$$

Then τ fixes y_2, x_2 and x_1 and $\tau(y_1) = y_1 + \epsilon x_1$. Therefore, τ fixes Y_2, X_2 and X_1 and $\tau(Y_1) = Y_1 + (\epsilon^p - \epsilon) x_1^p = Y_1 + x_1^p = Y_1 + X_1$.

Working with $\tau^{-1} = \tau^{p-1}$ we find as above that $(\tau^{-1} - 1) m$ is divisible by X_1 and thus X_1 divides $(\tau^{-1} - 1) f$ for all $f \in R$.

Let $f \in R^H$ be arbitrary. Then $\tau \sigma(f) = f$. Thus $\sigma(f) = \tau^{-1}(f)$ and therefore $(\sigma - 1) f = (\tau^{-1} - 1) f$. Hence X_2 divides $(\sigma - 1) f$ in R and also X_1 divides $(\sigma - 1) f = (\tau^{-1} - 1) f$ in R. Since R is a unique factorization domain, this implies that the least common multiple $X_1 X_2 / x_1 =$

$x_1^p(x_1^{p-1} - x_2^{p-1})$ of X_1 and X_2 divides $(\sigma - 1)f$ in R. Thus all the hypotheses of Proposition 11.0.1 are satisfied and we may apply the Proposition and conclude that $R^{H/K} = R^{G/K}[h]$ where h is any element of R such that $(\sigma - 1)h = x_1^p(x_1^{p-1} - x_2^{p-1})$. For example, we may take $h = (X_2Y_1 - X_1Y_2)/x_1 = (x_1^{p-1} - x_2^{p-1})(y_1^p - x_1^{p-1}y_1) - x_1^{p-1}(y_2^p - x_2^{p-1}y_2)$. Thus $\mathbb{F}[V]^H = (\mathbb{F}[V]^K)^{H/K} = R^{G/K}[h] = (\mathbb{F}[V]^K)^{G/K}[h] = \mathbb{F}[V]^G[h]$. Hence we have reduced to computing $\mathbb{F}[V]^G$.

We may easily find the ring of G-invariants since G is a Nakajima group. We find that $\mathbb{F}[V]^G = \mathbb{F}[x_1, x_2, f_1, f_2]$ where $f_1 = \mathbf{N}(y_1) = Y_1^p - X_1^{p-1}Y_1 = y_1^{p^2} - x_1^{p^2-p}y^p - x_1^{p^2-p}y_1^p + x_1^{p^2-1}y_1$, and $f_2 = \mathbf{N}(y_2) = Y_2^p - X_2^{p-1}Y_2 = y_2^{p^2} - x_2^{p^2-p}y_2^p - (x_1^p - x_1x_2^{p-1})^{p-1}(y_2^p - x_2^{p-1}y_2)$. Therefore, $\mathbb{F}[V]^H = \mathbb{F}[x_1, x_2, f_1, f_2, h]$ where $h = x_1^{p-1}y_1^p - x_2^{p-1}y_1^p - x_1^{2p-2}y_1^p + x_1^{p-1}x_2^{p-1}y_1 - x_1^{p-1}y_2^p + x_1^{p-1}x_2^{p-1}y_2$.

Furthermore, from the C_p representation $2\,V_2$ we know that $U^p - X_2^pN_1 + X_1^pN_2 - (X_1X_2)^{p-1}U = 0$ where $U = X_2Y_1 - X_1Y_2 = hx_1$, $N_1 = f_1$ and $N_2 = f_2$. Therefore,

$$\begin{aligned}
0 &= U^p - X_2^p f_1 + X_1^p f_2 - (X_1X_2)^{p-1}U \\
&= (hx_1)^p - (x_1^p - x_1x_2^{p-1})^p f_1 + (x_1^p)^p f_2 - (x_1^p)^{p-1}(x_1^p - x_1x_2^{p-1})^{p-1}(hx_1) \\
&= h^p - (x_1^{p-1} - x_2^{p-1})^p f_1 + x_1^{p^2-p} f_2 - x_1^{p^2-p}(x_1^{p-1} - x_2^{p-1})^{p-1} h\ .
\end{aligned}$$

12

Separating Invariants

The original and possibly most important use of invariants is to detect whether two mathematical objects are equivalent under some transformation. For example, given two matrices, we may wish to decide whether or not they are conjugate. If their eigenvalues differ, then they cannot be conjugate. In other words, the eigenvalues serve to partially distinguish non-conjugate matrices. Separating invariants typically play a similar role.

We note that much of the material in this chapter is valid in characteristic 0. This topic has generated much recent interest, see the book of Derksen and Kemper, [26, §2.3.2, p. 54 and §3.9.4, p. 119], and thesis and papers of Dufresne, [30] and [31], the papers of Kemper, [66], and Draisma, Kemper and Wehlau, [29].

Definition 12.0.1. *Let $S \subset \mathbb{K}[V]$ and let $u, v \in \overline{V}$. We will say that u and v are S-*equivalent *if $f(u) = f(v)$ for all $f \in S$.*

It is clear that S-equivalence is an equivalence relation on $\overline{V}$ and also (via restriction) on V. A trivial example is obtained by taking $S = \mathbb{K}[V]$ in which case the equivalence classes in V are the individual points of V. At the opposite end of the spectrum, if we take $S = \mathbb{K} \subset \mathbb{K}[V]^G$, then all of $\overline{V}$ is one S-equivalence class.

Remark 12.0.2. Note that if S is any subset of $\mathbb{K}[V]$ and $R = \mathbb{K}[S]$ is the subalgebra of $\mathbb{K}[V]$ generated by S, then S-equivalence and R-equivalence are the same equivalence relation.

For our purposes, the most important example is $S = \mathbb{K}[V]^G$.

Theorem 12.0.3. *Let G be a finite subgroup of* $\mathrm{GL}(V)$ *where V is an n dimensional vector space over the field $\mathbb{K}$. Then $u, v \in V$ are $\mathbb{K}[V]^G$-equivalent if and only if $u \in G \cdot v$.*

Proof. Of course, if $f(u) \neq f(v)$ for any $f \in \mathbb{K}[V]^G$, then u and v cannot lie in the same G orbit (since if $v = \sigma \cdot u$ then $f(v) = f(\sigma u) = (\sigma^{-1}(f))(u) = f(u)$).

H.E.A. Eddy Campbell, D. Wehlau, *Modular Invariant Theory*,
Encyclopaedia of Mathematical Sciences 139,
DOI 10.1007/978-3-642-17404-9_12, © Springer-Verlag Berlin Heidelberg 2011

Conversely, suppose $u, v \in V$ are $\mathbb{K}[V]^G$-equivalent. Assume, by way of contradiction, that $u \notin Gv$. Then by Corollary 2.1.3, there exists $f \in \mathbb{K}[V]$ such that $f(u) = 0$ and $f(\sigma v) = 1$ for all $\sigma v \in Gv$. Define $h := \mathbf{N}^G(f) \in \mathbb{K}[V]^G$. Since f divides h, we have $h(u) = 0$. Thus $0 = h(v) = \prod_{\sigma \in G}(\sigma f)(v)$. Hence there exists $\sigma \in G$ with $0 = (\sigma f)(v) = f(\sigma^{-1}v)$ contradicting the definition of f. □

Note that the hypothesis that G be finite is required in the above theorem. For example, if we consider the subgroup $\mathbb{C}^* = \mathrm{GL}(1, \mathbb{C})$ acting on the line $V = \mathbb{C}$, then it is clear that $\mathbb{C}[V]^{\mathbb{C}^*} = \mathbb{C}$. Thus there is a single $\mathbb{C}[V]^{\mathbb{C}^*}$-equivalence class. However, there are two orbits in V: one consisting of the origin and a second consisting of all other points of V.

Remark 12.0.4. If G is a reductive group over an algebraically closed field $\mathbb{K}$, then it can be shown that u and v are $\mathbb{K}[V]^G$-equivalent if and only if the topological closures of their orbits $G \cdot u$ and $G \cdot v$ in V have non-empty intersection.

Theorem 12.0.3 shows that we may detect whether two points $v, w \in V$ lie in the same G orbit by evaluating invariants on the two points. Indeed, by Remark 12.0.2, if $f_1, f_2, \ldots, f_r$ is a generating set for $\mathbb{K}[V]^G$ then u and v lie in the same G orbit if and only if $f_i(u) = f_i(v)$ for all $i = 1, 2, \ldots, r$.

Suppose now that $f_1, f_2, \ldots, f_r$ is a minimal generating set for $\mathbb{K}[V]^G$. The question arises whether it is necessary to verify all r of the equations $f_i(u) = f_i(v)$ for $i = 1, 2, \ldots, r$ in order to be certain that $v \in G \cdot u$. The answer to this question is almost always no.

Example 12.0.5. Let $\mathbb{K}$ be any field and suppose that the integer $n \geq 3$ is invertible in $\mathbb{K}$. Let ξ be a primitive n^{th} root of unity in $\mathbb{K}$ and let G denote the subgroup of $\mathrm{GL}(V)$ generated by $\sigma^{-1} = \begin{pmatrix} \xi & 0 \\ 0 & \xi \end{pmatrix}$. Clearly, G is the cyclic group of order n. If $\{x, y\}$ is the basis of V^* dual to the standard basis of V then $\sigma(x) = \xi^{-1}x$ and $\sigma(y) = \xi^{-1}y$. It is easy to see that $\mathbb{K}[V]^G$ is minimally generated by the $n+1$ monomials $x^n, x^{n-1}y, \ldots xy^{n-1}, y^n$. However, the values of the three monomials $x^n, x^{n-1}y, y^n$ alone suffice to determine whether two points lie in the same orbit.

To see this, consider any two points $u = (u_1, u_2)$ and $v = (v_1, v_2) \in V$. Suppose that $x^n(u) = x^n(v)$, $x^{n-1}y(u) = x^{n-1}y(v)$ and $y^n(u) = y^n(v)$. We consider the two possibilities: $x^n(u) = 0$ and $x^n(u) \neq 0$. If $x^n(u) = 0$, then u lies on the x-axis which consists of two orbits: the origin and the set $\{(a, 0) \mid a \neq 0\}$. The value $y^n(u) = u_2^n = y^n(v)$ distinguishes in which of these two orbits the points u and v lie. Thus we have seen that if $x^n(u) = 0$, then the pair of invariant monomials x^n and y^n suffice to determine whether u and v lie in the same G-orbit.

Next, consider the case where $x^n(u) \neq 0$. In this case, $u_1 \neq 0$ and thus

$$\begin{aligned} x^i y^{n-i}(u) &= u_1^i u_2^{n-i} = \frac{(u_1^{n-1}u_2)^i}{(u_1^n)^{n-i}} \\ &= \frac{((x^{n-1}y)(u))^i}{x^n(u)} = \frac{((x^{n-1}y)(v))^i}{x^n(v)} = x^i y^{n-i}(v) \end{aligned}$$

for all $i = 2, 3, \dots, n-1$. Therefore, by Theorem 12.0.3, u and v lie in the same G-orbit.

Definition 12.0.6. *Let G be a subgroup of $\mathrm{GL}(V)$ where V is an n dimensional vector space over the field. We say that a subset $S \subseteq \mathbb{K}[V]^G$* separates *if S-equivalence and $\mathbb{K}[V]^G$-equivalence are the same equivalence relation on the points of $\overline{V}$.*

It is easily seen that the following is an equivalent definition.

Definition 12.0.7. *A subset $S \subset \mathbb{K}[V]^G$* separates *if for all $u, v \in \overline{V}$, we have the following: if there exists an invariant $f \in \mathbb{K}[V]^G$ with $f(u) \neq f(v)$, then there exists $h \in S$ with $h(u) \neq h(v)$.*

Theorem 12.0.8. *Let G be an arbitrary (possibly infinite) group of automorphisms of $\mathbb{K}[V]$. There is a finite separating set $S \subset \mathbb{K}[V]^G$.*

Proof. We consider the ideal

$$J := (f \otimes 1 - 1 \otimes f \mid f \in \mathbb{K}[V]^G) \subset \mathbb{K}[V] \otimes_{\mathbb{K}} \mathbb{K}[V].$$

Since the ring $\mathbb{K}[V] \otimes_{\mathbb{K}} \mathbb{K}[V] \cong \mathbb{K}[V \oplus V]$ is Noetherian, every generating set for J contains a finite generating set. Thus there exist

$$f_1, f_2, \dots, f_m \in \mathbb{K}[V]^G$$

such that

$$J = (f_1 \otimes 1 - 1 \otimes f_1, f_2 \otimes 1 - 1 \otimes f_2, \dots, f_m \otimes 1 - 1 \otimes f_m).$$

We will show that $S = \{f_1, f_2, \dots, f_m\}$ is a separating set. Let $u, v \in \overline{V}$ and suppose there exists $f \in \mathbb{K}[V]^G$ with $f(u) \neq f(v)$. Write $f \otimes 1 - 1 \otimes f = \sum_{i=1}^m g_i(f_i \otimes 1 - 1 \otimes f_i)$ with $g_i \in \mathbb{K}[V] \otimes_{\mathbb{K}} \mathbb{K}[V]$.

Define an algebra homomorphism

$$\begin{aligned} \pi : \mathbb{K}[V] \otimes_{\mathbb{K}} \mathbb{K}[V] &\longrightarrow \mathbb{K} \\ g \otimes h &\mapsto g(u)h(v)\ . \end{aligned}$$

Then

$$\begin{aligned} 0 \neq f(u) - f(v) &= \pi(f \otimes 1 - 1 \otimes f) \\ &= \pi\left(\sum_{i=1}^m g_i(f_i \otimes 1 - 1 \otimes f_i)\right) \\ &= \sum_{i=1}^m \pi(g_i)(f_i(u) - f_i(v))\ . \end{aligned}$$

Therefore, there exists i such that $f_i(u) \neq f_i(v)$. □

Example 12.0.9. Consider the action of $C_p = \langle\sigma\rangle$ on $V = m\,V_2$. As usual, we choose a triangular basis $\{x_1, y_1, x_2, y_2, \ldots, x_m, y_m\}$ for V^* with $\sigma(y_i) = y_i + x_i$ and $\sigma(x_i) = x_i$ for all $i = 1, 2, \ldots, m$. Generators for this ring of invariants are given in Theorem 7.4.1. Here we will show that S is a separating subalgebra where S is the subalgebra 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. $u_{ij} = x_j y_i - x_i y_j$ for $1 \le i < j \le m$.

Suppose then that $v, w \in \overline{V}$ and that $f(v) = f(w)$ for all $f \in S$. Write $x_i(v) = a_i$, $y_i(v) = b_i$, $x_i(w) = \alpha_i$ and $y_i(w) = \beta_i$ for $i = 1, 2, \ldots, m$.

First we consider the case where $x_i(v) = x_i(w) = 0$ for all $i = 1, 2, \ldots, m$. Then $N_i(v) = N_i(w)$ implies that $b_i^p = \beta_i^p$ for all i. Thus $0 = b_i^p - \beta_i^p = (b_i - \beta_i)^p$ and therefore $y_i(v) = y_i(w)$ for all $i = 1, 2, \ldots, m$. Thus $v = w$.

Thus we may assume that $x_k(v) = x_k(w) \neq 0$ for some k with $1 \le k \le m$. Let W denote the k^{th} copy of V_2 (with basis dual to $\{x_k, y_k\}$). Since $\mathbb{F}[W]^{C_p} = \mathbb{F}[x_k, \mathbf{N}^{C_p}(y_k)]$, we see that the two invariants x_k and $\mathbf{N}^{C_p}(y_k)$ separate C_p orbits on W. In particular, since $a_k = \alpha_k$ and $b_k = \beta_k$, there must exist ℓ with $0 \le \ell \le p-1$ such that $(a_k, b_k) = \sigma^\ell \cdot (\alpha_k, \beta_k) = (\alpha_k, \beta_k - \ell\alpha_k)$.

Now $u_{jk}(v) = u_{jk}(w)$ for all $j \neq k$ (where for ease of notation, we define $u_{jk} := -u_{kj}$ if $j > k$). Thus $a_j b_k - a_k b_j = \alpha_j \beta_k - \alpha_k \beta_j$. But using $b_k = \beta_k - \ell\alpha_k$ we see that $a_j(\beta_k - \ell\alpha_k) - a_k b_j = \alpha_j\beta_k - \alpha_k\beta_j$. Thus $\ell\alpha_j\alpha_k + \alpha_k b_j = \alpha_k\beta_j$. Therefore, $\alpha_k(b_j - \beta_j + \ell\alpha_j) = 0$. Since $\alpha_k \neq 0$ this yields $b_j = \beta_j - \ell\alpha_j$ for all $j \neq k$. Therefore, $v = \sigma^\ell(w) \in C_p w$.

If G is finite, we can give a constructive proof of Theorem 12.0.8. Suppose G is a finite group of automorphisms of $\mathbb{K}[V]$.

Consider the polynomial

$$F(t,z) := \prod_{\sigma\in G} (t - \sum_{i=1}^{n} \sigma(x_i) z^i) \in \mathbb{K}[V]^G[t,z]$$

Theorem 12.0.10. *The coefficients of $F(t,z)$ form a finite separating set.*

Proof. Suppose that each coefficient of $F(t,z)$ takes the same value at u as it does at v. Then $F(t,z)(u) = F(t,z)(v)$ in the unique factorization domain $\mathbb{K}[t,z]$. Since $t - \sum_{i=1}^n x_i(u) z^i$ is a factor of $F(t,z)(u)$, it must also be a factor of $F(t,z)(v)$. Therefore, there exists $\tau \in G$ such that $t - \sum_{i=1}^n x_i(u) z^i = t - \sum_{i=1}^n (\tau \cdot x_i)(v) z^i$. Hence $\sum_{i=1}^n x_i(u) z^i = \sum_{i=1}^n (\tau\cdot x_i)(v) z^i$ and thus $x_i(u) = x_i(\tau^{-1}\cdot v)$ for all $i = 1, 2, \ldots, n$. Hence $u = \tau^{-1}(v) \in Gv$. □

Corollary 12.0.11. *If G is a finite subgroup of* $\mathrm{GL}(V)$ *then there exists a finite separating set $S = \{f_1, f_2, \ldots, f_m\}$ with* $\deg(f_i) \le |G|$ *for all $i = 1, 2, \ldots, m$.*

Proof. The coefficients of $F(t,z)$ separate and have degree at most $|G|$. □

12.1 Relation Between $\mathbb{K}[V]^G$ and Separating Subalgebras

In this section, we prove a number of results that show that each separating subalgebra is closely related to the full ring of invariants.

In [87, Lemma 3.4.2], the following result is proved for G any geometrically reductive group. Here we suppose G is finite.

Lemma 12.1.1. *Let G be a finite group and let R be a $\mathbb{K}$-algebra equipped with a G-action. Suppose $f_1, f_2, \ldots, f_s \in R^G$ and consider the ideal $I = (f_1, f_2, \ldots, f_s)$ of R^G. If $f \in (f_1, f_2, \ldots, f_s)R \cap R^G$, then there exists a positive integer t such that $f^t \in I$.*

Proof. The proof is by induction on s. If $s = 1$, then we may write $f = f_1 h$ for some $h \in R$. Since both $f_1, f \in R^G$, it follows that $f = f_1\sigma(h)$ for all $\sigma \in G$. Thus $\prod_{\sigma\in G} f = \prod_{\sigma\in G} f_1\sigma(h)$ and therefore $f^{|G|} = f_1^{|G|}\mathbf{N}(h) \in f_1 R^G$.

Now assume $s \geq 2$ and that the result is true when I has $s-1$ generators. Write $\overline{R} = R/f_s R$ and for $h \in R$, denote by $\overline{h}$ the image of h in $\overline{R}$. Take $f \in (f_1, f_2, \ldots, f_s)R \cap R^G$. Then $\overline{f} \in (\overline{f_1}, \overline{f_2}, \ldots, \overline{f_{s-1}})\overline{R} \cap \overline{R}^G$. Thus by our induction hypothesis, there exists a positive integer t such that $\overline{f}^t \in \overline{I} = (\overline{f_1}, \overline{f_2}, \ldots, \overline{f_{s-1}})\overline{R}^G$. Hence we may write $f^t = \sum_{i=1}^{s} f_i h_i$ where $h_1, h_2, \ldots, h_s \in R$ and $\overline{h_1}, \overline{h_2}, \ldots, \overline{h_{s-1}} \in \overline{R}^G$. Therefore $f^t - \sum_{i=1}^{s-1} f_i h_i = f_s h_s$ and $f^t - \sum_{i=1}^{s-1} f_i\sigma(h_i) = f_s\sigma(h_s)$ for all $\sigma \in G$. Hence $\prod_{\sigma\in G}(f^t - \sum_{i=1}^{s-1} f_i\sigma(h_i)) = \prod_{\sigma\in G} f_s\sigma(h_s)$. Thus $f^{t|G|} - f_s^{|G|}\mathbf{N}(h_s) \in (f_1, f_2, \ldots, f_{s-1})R \cap R^G$. Applying the induction hypothesis again, there is a positive integer v such that $(f^{t|G|} - f_s^{|G|}\mathbf{N}(h_s))^v \in (f_1, f_2, \ldots, f_{s-1})R^G$. Therefore $f^{t|G|v} \in (f_1, f_2, \ldots, f_s)R^G$. □

Theorem 12.1.2. *Suppose that $S \subset \mathbb{K}[V]^G$ is a graded separating subalgebra. Then $\mathbb{K}[V]^G$ is integral over S.*

Proof. Let $I = S_+ \cdot \mathbb{K}[V]$ denote the ideal of $\mathbb{K}[V]$ generated by S_+ and let $J = \mathbb{K}[V]^G_+ \cdot \mathbb{K}[V]$ denote the Hilbert ideal. Take any $v \in \mathcal{V}(I)$. Since S is a separating set and since $h(v) = 0$ for $h \in S_+$, it follows that $f(v) = 0$ for all $f \in \mathbb{K}[V]^G_+$. Thus $v \in \mathcal{V}(J)$. Hence $\mathcal{V}(I) \subseteq \mathcal{V}(J)$. Therefore $\sqrt{J} \subseteq \sqrt{I}$. In particular, $\mathbb{K}[V]^G_+ \subseteq \sqrt{I} \cap \mathbb{K}[V]^G$. Take $f \in \mathbb{K}[V]^G_+$. Then, by the above, $f \in \sqrt{I}$. Therefore, there exists a positive integer m such that $f^m \in I \cap \mathbb{K}[V]^G$. By Lemma 12.1, there is a positive integer t with $f^{mt} \in S_+\mathbb{K}[V]^G$. This implies that the ring $\mathbb{K}[V]^G/S_+\mathbb{K}[V]^G$ has Krull dimension 0 and thus this ring is a finite dimensional $\mathbb{K}$ vector space. Therefore by the graded Nakayama lemma 2.10.1, we see that $\mathbb{K}[V]^G$ is a finitely generated S-module. This implies that $\mathbb{K}[V]^G$ is integral over S. □

In particular, we have the following result.

Corollary 12.1.3. *If $S \subset \mathbb{K}[V]^G$ is a graded separating subalgebra, then $\dim S = \dim \mathbb{K}[V]^G$.*

Remark 12.1.4. It is possible to show (see [30][Corollary 3.2.5]) that the hypothesis that S be graded is not required for the proof of Corollary 12.1.3 provided that G is reductive.

Note that the converse of Theorem 12.1.2 is not true as the following example shows.

Example 12.1.5. Consider the two dimensional representation of the trivial group, $G = \{e\}$ over a field $\mathbb{K}$. Take $S = \mathbb{K}[x, y^2, x^2y] \subset \mathbb{K}[V]^G = \mathbb{K}[x, y]$. Since $y = (x^2y)/(x)^2 \in \mathrm{Quot}(S)$ and since y satisfies the monic equation $T^2 - y^2 \in S[T]$, we see that y lies in the integral closure of S. Hence $\mathbb{K}[x, y] = \mathbb{K}[V]^G$ is the integral closure of S. Consider the two points $u = (0, 1)$ and $v = (0, -1)$. Since $x(u) = 0 = x(v)$ and $y^2(u) = 1 = y^2(v)$ and $x^2y(u) = 0 = x^2y(v)$ we see that S does not separate u from v and thus S is not a separating subalgebra.

Theorem 12.1.6. *Suppose G is finite and $S \subset \mathbb{K}[V]^G$ is a graded separating subalgebra. Then the field $\mathrm{Quot}(\mathbb{K}[V]^G)$ is a finite purely inseparable extension of $\mathrm{Quot}(S)$.*

Proof. Dual to the inclusion $\phi : S \hookrightarrow \mathbb{K}[V]^G$ is the morphism $\phi^* : V/\!\!/G \to \mathrm{Spec}(S)$. In terms of ideals, this map is given by $\phi^*(P) = P \cap S$ for a prime ideal $P \in V/\!\!/G = \mathrm{Spec}(\mathbb{K}[V]^G)$. Since ϕ is injective, ϕ^* is dominant, i.e., its image is dense in $\mathrm{Spec}(S)$. We claim that ϕ^* is also injective.

First, we show it is injective on maximal ideals. Take two different maximal ideals I and J of $\mathbb{K}[V]^G$. The ideals I and J have height $n = \dim V$. Hence both $\mathcal{V}_{\overline{V}}(I)$ and $\mathcal{V}_{\overline{V}}(J)$ are zero dimensional subvarities of $\overline{V}$, i.e., both these varieties consist of a (non-empty) finite set of points of $\overline{V}$. Write $\mathcal{V}_{\overline{V}}(I) = \{a_1, a_2, \ldots, a_s\}$ and $\mathcal{V}_{\overline{V}}(J) = \{b_1, b_2, \ldots, b_t\}$. Since S separates, there exists $f_i \in S$ with $f_i(a_i) \neq f_i(b_1)$ for all $i = 1, 2, \ldots, s$. Define $f := \prod_{i=1}^{s}(f_i - f_i(a_i))$. Then $f \in S$ with $f(a_i) = 0$ for all i and $f(b_1) \neq 0$. Therefore, $f \in S \cap I = \phi^*(I)$ but $f \notin S \cap J = \phi^*(J)$ and $\phi^*(I) \neq \phi^*(J)$.

Now we use this to show that ϕ^* is injective. Take any two different prime ideals P and Q of $\mathbb{K}[V]^G$. Assume, by way of contradiction, that $\phi^*(P) = \phi^*(Q)$. Since the prime ideals P and Q are different, there exists (without loss of generality) a maximal ideal I of $\mathbb{K}[V]^G$ containing P but not containing Q (see for example [80][Theorem 5.5]).

Since $\phi^*(P) = \phi^*(Q)$ and $P \subset I$ we have the chain of prime ideals $\phi^*(Q) \subset \phi^*(I)$ in S. By Theorem 12.1.2, the "Going Up" Theorem (Theorem 2.5.2 (2)) applies here. Thus there exists a prime ideal J of $\mathbb{K}[V]^G$ such that $Q \subset J$ and $J \cap S = \phi^*(I)$. By "Going Up" and "Going Down" (Theorem 2.5.2), since I is maximal in $\mathbb{K}[V]^G$, the ideal $I \cap S = J \cap S$ is maximal in S and this implies that J is maximal in $\mathbb{K}[V]^G$. But then $\phi^*(I) = \phi^*(J)$ implies that $I = J$ since ϕ^* is injective on maximal ideals. However, $Q \subset J$ and $Q \not\subseteq I$. This contradiction shows that ϕ^* is indeed injective.

By [56, Theorem 4.6], since ϕ^* is a dominant injective morphism, we have that $\mathrm{Quot}(\mathbb{K}[V]^G)$ is a finite purely inseparable extension of $\phi(\mathrm{Quot}(S))$ hence also of $\mathrm{Quot}(S)$. □

Remark 12.1.7. The above theorem is true even without the hypothesis that S be graded. For a proof of this see [26, Proposition 2.3.10] or [30].

Definition 12.1.8. *Let S be a subalgebra of $\mathbb{K}[V]$. If $\mathbb{K}$ has characteristic p then we define the* inseparable closure *or* p-root closure *of S in $\mathbb{K}[V]$, denoted $\widehat{S}$ by*

$$\widehat{S} := \{f \in \mathbb{K}[V] \mid \exists\, t \in \mathbb{N} \text{ with } f^{p^t} \in S\}.$$

For ease of notation, we define $\widehat{S} = S$ when $\mathbb{K}$ has characteristic 0.

Proposition 12.1.9. *Let S be a subalgebra of $\mathbb{K}[V]^G$ and suppose that $\widehat{S}$ separates. Then S also separates.*

Proof. If $\mathbb{K}$ has characteristic 0 then there is nothing to prove. Thus we suppose that the characteristic of $\mathbb{K}$ is $p > 0$. Take $u, v \in \overline{V}$ and suppose that $h(u) = h(v)$ for all $h \in S$. Then $0 = (h(u) - h(v))^{p^r} = h^{p^r}(u) - h^{p^r}(v)$ for $h \in S$ and all $r \in \mathbb{N}$. Therefore $f(u) = f(v)$ for all $f \in \widehat{S}$. Since $\widehat{S}$ separates this implies that u and v are $\mathbb{K}[V]^G$-equivalent. Thus S separates. □

Theorem 12.1.10. *Let S be a graded separating subalgebra of $\mathbb{K}[V]^G$. Then $\widehat{\widetilde{S}} = \mathbb{K}[V]^G$.*

Proof. Since $\mathbb{K}[V]^G$ is integrally closed, we know that $\widehat{\widetilde{S}} \subseteq \mathbb{K}[V]^G$. For the opposite inclusion, take $f \in \mathbb{F}[V]^G$. Then Theorem 12.1.6 implies that $f^{p^r} \in \mathrm{Quot}(S)$ for some $r \in \mathbb{N}$ (use $r = 0$ and $f^{p^r} = f$ if $\mathbb{K}$ has characteristic 0). Furthermore, by Theorem 12.1.2, f^{p^r} is integral over S, and thus $f^{p^r} \in \widetilde{S}$. Thus $f \in \widehat{\widetilde{S}}$. □

Proposition 12.1.11. *Let S be a subalgebra of $\mathbb{K}[V]$. Then $\widehat{\widetilde{S}} = \widetilde{\widehat{S}}$.*

Proof. Take $f \in \widetilde{\widehat{S}}$. Then $f \in \mathrm{Quot}(\widehat{S})$ and f satisfies some equation $z^t + c_{t-1}z^{t-1} + \cdots + c_0 = 0$ where $c_0, c_1, \ldots, c_t \in S$. We must show that there exists $r \in \mathbb{N}$ with $f^{p^r} \in \widetilde{S}$. Write $f = a/b$ where $a, b \in \widehat{S}$. Then there exists $\ell, m \in \mathbb{N}$ with $a^{p^\ell}, b^{p^m} \in S$. Take $r \geq \max\{\ell, m\}$. Then $a^{p^r}, b^{p^r} \in S$ and $f^{p^r} = a^{p^r}/b^{p^r} \in \mathrm{Quot}(S)$. Furthermore, $0 = (f^t + c_{t-1}f^{t-1} + \cdots + c_0)^{p^r} = (f^t)^{p^r} + c_{t-1}^{p^r}(f^{t-1})^{p^r} + \cdots + c_0^{p^r} = (f^{p^r})^t + c_{t-1}^{p^r}(f^{p^r})^{t-1} + \cdots + c_0^{p^r}$ with $c_0^{p^r}, c_1^{p^r}, \ldots, c_{t-1}^{p^r} \in S$. Thus $f^{p^r} \in \widetilde{S}$ which means that $f \in \widehat{\widetilde{S}}$.

For the opposite inclusion, take $f \in \widehat{\widetilde{S}}$. Then there exists $r \in \mathbb{N}$ with $h = f^{p^r} \in \widetilde{S}$. Write $h = a/b$ with $a, b \in S$. Then $bf^{p^r} = a \in S$. Thus $(bf)^{p^r} = b^{p^r-1}(bf^{p^r}) = b^{p^r-1}a \in S$ and hence $bf \in \widehat{S}$. Thus $f = bf/b \in \mathrm{Quot}(\widehat{S})$.

Furthermore, if h satisfies the monic equation $z^t + c_{t-1}z^{t-1} + \cdots + c_0 = 0$ with $c_0, c_1, \ldots, c_t \in S \subset \widehat{S}$, then f satisfies the monic equation $z^{tp^r} + c_{t-1}z^{(t-1)p^r} + \cdots + c_0 = 0$. Therefore, $f \in \widetilde{\widehat{S}}$. □

Remark 12.1.12. In fact, it can be shown that if S is a graded subalgebra of $\mathbb{F}[V]^G$ where $\mathbb{F}$ is a field of characteristic p, then S separates if and only if $\widehat{S} = \mathbb{F}[V]^G$. For details see [50] or [27, Remark 1.3].

12.2 Polynomial Separating Algebras and Serre's Theorem

The ideal $J := (f \otimes 1 - 1 \otimes f \mid f \in \mathbb{K}[V]^G) \subset \mathbb{K}[V] \otimes_{\mathbb{K}} \mathbb{K}[V]$ plays a central role in the proof of Theorem 12.0.8. We consider this ideal more carefully now. This ideal corresponds to the reduced scheme having the same underlying topological space as $V \times_{V/\!\!/G} V$. Following Dufresne we call this the *separating scheme* and denote it by $\mathcal{S}_G$. Thus

$$\mathcal{S}_G = \operatorname{Spec}(\mathbb{K}[V] \otimes_{\mathbb{K}[V]^G} \mathbb{K}[V]) \ .$$

We consider the map $\delta : \mathbb{K}[V] \to \mathbb{K}[V] \otimes_{\mathbb{K}} \mathbb{K}[V]$ defined by $\delta(f) = f \otimes 1 - 1 \otimes f$. For any subalgebra $B \subseteq \mathbb{K}[V]$, we have

$$\mathbb{K}[V] \otimes_B \mathbb{K}[V] \cong \frac{\mathbb{K}[V] \otimes_{\mathbb{K}} \mathbb{K}[V]}{(\delta(B))}$$

and in particular,

$$\mathbb{K}[V] \otimes_{\mathbb{K}[V]^G} \mathbb{K}[V] \cong \frac{\mathbb{K}[V] \otimes_{\mathbb{K}} \mathbb{K}[V]}{(\delta(\mathbb{K}[V]^G))} = \frac{\mathbb{K}[V] \otimes_{\mathbb{K}} \mathbb{K}[V]}{J} \ .$$

Theorem 12.2.1 ([30] Theorem 3.2.1). *Let B be a subalgebra of $\mathbb{K}[V]^G$. Then B is a separating subalgebra if and only if $\sqrt{(\delta(B))} = \sqrt{(\delta(\mathbb{K}[V]^G)}$.*

Proof. We will exploit the fact that the radical of any ideal I is obtained by taking the intersection of all maximal ideals containing I. Thus we need to consider the maximal ideals of $\mathbb{K}[V] \otimes_{\mathbb{K}} \mathbb{K}[V]$. These are precisely the ideals of the form $\mathcal{I}_{\overline{V} \times \overline{V}}\{(u,v)\} \cap (\mathbb{K}[V] \otimes_{\mathbb{K}} \mathbb{K}[V])$. (Recall that $\mathcal{I}_{\overline{V} \times \overline{V}}\{(u,v)\}$ denotes the maximal ideal of $\overline{\mathbb{K}[V]} \otimes_{\overline{\mathbb{K}}} \overline{\mathbb{K}[V]}$ corresponding to a geometric point $(u,v) \in \overline{V} \times \overline{V}$.)

The assertion that B is a separating subalgebra means that every pair of points $u, v \in \overline{V}$ can be separated by invariants if and only if they can be separated by elements of B. In terms of ideals, this can be stated as

$$\mathcal{I}_{\overline{V} \times \overline{V}}(u,v) \cap (\mathbb{K}[V] \otimes_{\mathbb{K}} \mathbb{K}[V]) \supseteq \delta(\mathbb{K}[V]^G)$$

if and only if

$$\mathcal{I}_{\overline{V}\times\overline{V}}(u,v)\cap(\mathbb{K}[V]\otimes_{\mathbb{K}}\mathbb{K}[V])\supseteq\delta(B)$$

for all $u, v \in \overline{V}$. Thus B is a separating subalgebra if and only if $\sqrt{(\delta(B))} = \sqrt{(\delta(\mathbb{K}[V]^G))}$. □

If G is a finite group, then the separating scheme is quite simple; it is a union of $|G|$ subspaces each of dimension $n = \dim V$ lying in $V \times V$. Indeed, take $\tau \in G$ and let H_τ denote the subspace $H_\tau := \{(v, \tau(v)) \mid v \in V\}$ of $V \times V$.

Lemma 12.2.2. *If G is a finite group, then*

$$\mathcal{S}_G = \bigcup_{\tau\in G} H_\tau \ .$$

Proof. Since G is finite, the invariants will separate points not in the same G-orbit by Theorem 12.0.3. Thus $(u,v) \in \mathcal{S}_G$ if and only if $\mathcal{I}_{\overline{V}\times\overline{V}}\{(u,v)\} \cap (\mathbb{K}[V] \otimes_{\mathbb{K}} \mathbb{K}[V]) \supseteq \delta(\mathbb{K}[V]^G)$, if and only if $v = \tau(u)$ for some $\tau \in G$, if and only if $(u, v) \in H_\tau$ for some $\tau \in G$. □

A natural question is whether we may find separating subalgebras which are better behaved than the full ring of invariants. In particular, when may we find a polynomial subalgebra which separates? The following theorem answers this question in characteristic 0 and gives a necessary answer in positive characteristic.

Theorem 12.2.3. *Let G be a finite group. If $\mathbb{K}[V]^G$ contains a graded separating algebra which is a polynomial ring, then the action of G on V is generated by reflections.*

Proof. Suppose $B \subseteq \mathbb{K}[V]^G$ is a graded polynomial separating algebra. By Corollary 12.1.3, the algebra B has dimension n. Write $B = \mathbb{K}[f_1, f_2, \ldots, f_n]$. Then $(\delta(B)) = (\delta(f_1), \delta(f_2), \ldots, \delta(f_n))$ has n generators and cuts out a variety $\mathcal{S}_G = \cup_{\tau\in G} H_\tau$ of codimension n in $V \times V$. Thus $\delta(f_1), \delta(f_2), \ldots, \delta(f_n)$ is a partial homogeneous system of parameters. Therefore

$$\operatorname{Spec}(\mathbb{K}[V]\otimes_B \mathbb{K}[V]) \cong \operatorname{Spec}\left(\frac{\mathbb{K}[V]\otimes_{\mathbb{K}}\mathbb{K}[V]}{(\delta(B))}\right)$$

is a complete intersection and hence is also Cohen-Macaulay. Therefore, since $\mathbb{K}[V] \otimes_{\mathbb{K}} \mathbb{K}[V]$ is Noetherian, Hartshorne's Connectedness Theorem, see [51][Corollary 2.4], applies and we conclude from it that $\operatorname{Spec}(\mathbb{K}[V] \otimes_B \mathbb{K}[V])$ is connected in co-dimension 1. Since $\operatorname{Spec}(\mathbb{K}[V] \otimes_B \mathbb{K}[V])$ and $\mathcal{S}_G$ correspond to the same underlying topological space, this means that $\mathcal{S}_G$ is connected in co-dimension 1.

Consider the two components H_e and H_σ of $\mathcal{S}_G$ corresponding to the identity and an arbitrary element σ of G. Since $\mathcal{S}_G$ is connected in co-dimension 1, there exists a sequence of components $H_e = H_{\tau_0}, H_{\tau_1}, \ldots, H_{\tau_s} = H_\sigma$

such that $H_{\tau_{i-1}} \cap H_{\tau_i}$ has dimension $n-1$ (i.e., codimension 1 in $H_{\tau_{i-1}}$ and in H_{τ_i}) for all $i = 1, 2, \ldots, s$. Now $(u, v) \in H_{\tau_{i-1}} \cap H_{\tau_i}$ if and only if $\tau_{i-1}(u) = v = \tau_i(u)$, if and only if $\tau_{i-1}^{-1}\tau_i(u) = u$ and $v = \tau_i(u)$, if and only if $u \in V^{\tau_{i-1}^{-1}\tau_i}$ and $v = \tau_i(u)$. Therefore, $H_{\tau_{i-1}} \cap H_{\tau_i} \cong V^{\tau_{i-1}^{-1}\tau_i}$. This means that $\dim V^{\tau_{i-1}^{-1}\tau_i} = \dim H_{\tau_{i-1}} \cap H_{\tau_i} = n-1$, i.e., $\tau_{i-1}^{-1}\tau_i$ is a reflection. Therefore, $\sigma = \tau_s = \tau_0(\tau_0^{-1}\tau_1)(\tau_1^{-1}\tau_2)\cdots(\tau_{s-1}^{-1}\tau_s)$ expresses σ as a product of reflections in G. □

Remark 12.2.4. By Remark 12.1.4 we may omit the hypothesis of graded from Theorem 12.2.3.

Corollary 12.2.5 (Serre's Theorem). *Let G be a finite group. If $\mathbb{K}[V]^G$ is a polynomial ring, then the action of G on V is generated by reflections.*

In characteristic zero, Theorem 12.1.10 shows that if S is a graded polynomial separating subalgebra, then $S = \mathbb{K}[V]^G$. In positive characteristic, suppose $\mathbb{F}[V]^G = \mathbb{F}[f_1, f_2, \ldots, f_n]$ is a polynomial ring. Then using Proposition 12.1.9, we see that $S = \mathbb{F}[f_1^{p^{r_1}}, f_2^{p^{r_2}}, \ldots, f_n^{p^{r_n}}]$ is a polynomial separating subalgebra for all $r_1, r_2, \ldots, r_n \in \mathbb{N}$. These two results together with Theorem 12.2.3 suggest that perhaps whenever there is a polynomial separating subalgebra, the ring of invariants itself is polynomial. The following example, due to Dufresne [30] shows that this is not the case.

Example 12.2.6. In this example, we show that for positive characteristic, the existence of a polynomial separating subalgebra, $S \subset \mathbb{F}[V]^G$, does not necessarily imply that $\mathbb{F}[V]^G$ is a polynomial ring. To see this, we consider Example 11.0.3 from the point of view of separating invariants. We maintain the notation from Example 11.0.3, although we will now denote the group of order p^3 by G. In that example, we computed the ring of invariants of a four dimensional representation V of G, the elementary Abelian group of order p^3 over a field $\mathbb{F}$. We found that

$$\mathbb{F}[V]^G = \mathbb{F}[x_1, f_1, x_2, f_2, h]$$

where

$$\begin{aligned} f_1 &= y_1^{p^2} - x_1^{p^2-p} y^p - x_1^{p^2-p} y_1^p + x_1^{p^2-1} y_1, \\ f_2 &= y_2^{p^2} - x_2^{p^2-p} y_2^p - (x_1^p - x_1 x_2^{p-1})^{p-1}(y_2^p - x_2^{p-1} y_2) \text{ and} \\ h &= (x_1^{p-1} - x_2^{p-1})(y_1^p - x_1^{p-1} y_1) - x_1^{p-1}(y_2^p - x_2^{p-1} y_2) \\ &= x_1^{p-1} y_1^p - x_2^{p-1} y_1^p - x_1^{2p-2} y_1^p + x_1^{p-1} x_2^{p-1} y_1 - x_1^{p-1} y_2^p + x_1^{p-1} x_2^{p-1} y_2 \end{aligned}$$

with

$$h^p - (x_1^{p-1} - x_2^{p-1})^p f_1 + x_1^{p^2-p} f_2 - x_1^{p^2-p}(x_1^{p-1} - x_2^{p-1})^{p-1} h = 0 .$$

We define $\widetilde{f_2}$, a perturbation of f_2, as follows,

$$\widetilde{f_2} := f_2 - (x_1^{p-1} - x_2^{p-1})^{p-1} h$$

and take $S = \mathbb{F}[x_1, x_2, f_1, \widetilde{f_2}] = \mathbb{F}[x_1, x_2, f_1, f_2]$. Note that $\widetilde{f_2}$ is homogeneous of the same degree as f_2.

Then $\mathbb{F}[V]^G = \mathbb{F}[x_1, f_1, x_2, \widetilde{f_2}, h] = S[h]$ with

$$h^p - (x_1^{p-1} - x_2^{p-1})^p f_1 + (x_1^{p-1})^p \widetilde{f_2} = 0 \ .$$

In particular, $h^p \in S$ and therefore $h \in \widehat{S}$. This shows that $\widehat{S} = \mathbb{F}[V]^G$ and thus by Proposition 12.1.9, S is a separating subalgebra.

Since $x_1, x_2, f_1, \widetilde{f_2}$ forms a homogeneous system of parameters, these elements are algebraically independent and thus S is a polynomial subalgebra of $\mathbb{F}[V]^G$ which is also a separating algebra.

12.3 Polarization and Separating Invariants

Recall the polarization operators defined in §1.9. In this section, we prove a surprising theorem relating separating invariants and polarization. This material comes from the work of Draisma, Kemper and Wehlau [29].

Theorem 12.3.1. *Suppose S is a separating subset of $\mathbb{K}[V]^G$ and let $m \geq 1$. Then $\mathcal{P}\mathrm{ol}^m(S)$ is a separating subset of $\mathbb{K}[m\,V]^G$. More generally, suppose S is a separating subset of $\mathbb{K}[W_1 \oplus W_2 \oplus \cdots \oplus W_r]^G$ and that $m_1, m_2, \ldots, m_r$ are positive integers. Then $\mathcal{P}\mathrm{ol}^{m_1,m_2,\ldots,m_r}(S)$ is a separating subset of $\mathbb{K}[m_1\,W_1 \oplus m_2\,W_2 \oplus \cdots \oplus m_r W_r]^G$.*

Proof. We will prove the second assertion. Consider two points

$$\begin{aligned}\mathbf{v} &= (v_1^{(1)}, v_2^{(1)}, \ldots, v_{m_1}^{(1)}, v_1^{(2)}, v_2^{(2)}, \ldots, v_{m_2}^{(2)}, \ldots, v_1^{(r)}, v_2^{(r)}, \ldots, v_{m_r}^{(r)}),\\ \mathbf{w} &= (w_1^{(1)}, w_2^{(1)}, \ldots, w_{m_1}^{(1)}, w_1^{(2)}, w_2^{(2)}, \ldots, w_{m_2}^{(2)}, \ldots, w_1^{(r)}, w_2^{(r)}, \ldots, w_{m_r}^{(r)})\end{aligned}$$

of $\overline{V}$ where $V = m_1\,W_1 \oplus m_2\,W_2 \oplus \cdots \oplus m_r W_r$. Suppose that $g(\mathbf{v}) = g(\mathbf{w})$ for all

$$g \in \mathcal{P}\mathrm{ol}^{m_1,m_2,\ldots,m_r}(S) = \cup_{f \in S}\, \mathcal{P}\mathrm{ol}^{m_1,m_2,\ldots,m_r}(f).$$

We need to show that this assumption implies that $\mathbf{v}$ and $\mathbf{w}$ lie in the same G-orbit.

Take $f \in S$ and for $k = 1, 2, \ldots, r$ and $i = 1, 2, \ldots, m_k$, let $\lambda_i^{(k)}$ denote an indeterminate. Consider

$$f\Big(\sum_{i=1}^{m_1} \lambda_i^{(1)} v_i^{(1)}\ , \sum_{i=1}^{m_2} \lambda_i^{(2)} v_i^{(2)}\ , \ldots,\ \sum_{i=1}^{m_r} \lambda_i^{(r)} v_i^{(r)}\Big) = \sum_J \lambda^{\mathbf{J}} f_{\mathbf{J}}(v_1^{(1)}, v_2^{(1)}, \ldots, v_{m_r}^{(r)})$$

where each $f_{\mathbf{J}}$ lies in $\mathcal{P}\mathrm{ol}^{m_1,m_2,\ldots,m_r}(S)$. Therefore

$$f_{\mathbf{J}}(v_1^{(1)}, v_2^{(1)}, \ldots, v_{m_r}^{(r)}) = f_{\mathbf{J}}(w_1^{(1)}, w_2^{(1)}, \ldots, w_{m_r}^{(r)})$$

for all J and

$$f\Big(\sum_{i=1}^{m_1}\lambda_i^{(1)}v_i^{(1)}\ ,\sum_{i=1}^{m_2}\lambda_i^{(2)}v_i^{(2)}\ ,\dots,\ \sum_{i=1}^{m_r}\lambda_i^{(r)}v_i^{(r)}\Big) = f\Big(\sum_{i=1}^{m_1}\lambda_i^{(1)}w_i^{(1)}\ ,\sum_{i=1}^{m_2}\lambda_i^{(2)}w_i^{(2)}\ ,\dots,\ \sum_{i=1}^{m_r}\lambda_i^{(r)}w_i^{(r)}\Big).$$

Now pick $\alpha \in \overline{\mathbb{K}}$ and substitute in the above $\lambda_i^{(k)}$ by α^{i-1} for all $k = 1, 2, \dots, r$. Thus

$$f\Big(\sum_{i=1}^{m_1}\alpha^{i-1}v_i^{(1)}\ ,\sum_{i=1}^{m_2}\alpha^{i-1}v_i^{(2)}\ ,\dots,\ \sum_{i=1}^{m_r}\alpha^{i-1}v_i^{(r)}\Big) = f\Big(\sum_{i=1}^{m_1}\alpha^{i-1}w_i^{(1)}\ ,\sum_{i=1}^{m_2}\alpha^{i-1}w_i^{(2)}\ ,\dots,\ \sum_{i=1}^{m_r}\alpha^{i-1}w_i^{(r)}\Big).$$

Since this holds for all f in the separating set S, there must exist $\sigma_\alpha \in G$ such that

$$\sigma_\alpha\Big(\sum_{i=1}^{m_1}\alpha^{i-1}v_i^{(1)}\ ,\sum_{i=1}^{m_2}\alpha^{i-1}v_i^{(2)}\ ,\dots,\ \sum_{i=1}^{m_r}\alpha^{i-1}v_i^{(r)}\Big) = \Big(\sum_{i=1}^{m_1}\alpha^{i-1}w_i^{(1)}\ ,\sum_{i=1}^{m_2}\alpha^{i-1}w_i^{(2)}\ ,\dots,\ \sum_{i=1}^{m_r}\alpha^{i-1}w_i^{(r)}\Big).$$

Now to each $\alpha \in \overline{\mathbb{K}}$, we have associated an element σ_α of G. Let $t := \max\{m_1, m_2, \dots, m_r\}$. Choose $\tau \in G$ such that there exist $\alpha_1, \alpha_2, \dots, \alpha_t$ distinct elements of $\overline{\mathbb{K}}$ with $\sigma_{\alpha_s} = \tau$ for all $s = 1, 2, \dots, t$.

Thus we have $\tau\Big(\sum_{i=1}^{m_k}\alpha_s^{i-1}v_i^{(k)}\Big) = (\sum_{i=1}^{m_k}\alpha_s^{i-1}w_i^{(k)}\Big)$ for all $1 \le s \le t$ and $1 \le k \le r$. Writing this in matrix terms we have

$$\begin{pmatrix} 1 & \alpha_1 & \alpha_1^2 & \dots & \alpha_1^{m_k-1} \\ 1 & \alpha_2 & \alpha_2^2 & \dots & \alpha_2^{m_k-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \alpha_{m_k} & \alpha_{m_k}^2 & \dots & \alpha_{m_k}^{m_k-1} \end{pmatrix} \begin{pmatrix} \tau(v_1^{(k)}) - w_1^{(k)} \\ \tau(v_2^{(k)}) - w_2^{(k)} \\ \vdots \\ \tau(v_{m_k}^{(k)}) - w_{m_k}^{(k)} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

for all $1 \le k \le r$. Since the square matrix is a Vandermonde matrix with determinant $\prod_{1 \le i < j \le j}(\alpha_j - \alpha_i) \neq 0$, we see that $\tau(v_i^{(k)}) - w_i^{(k)} = 0$ for all $1 \le i \le m_k$ and $1 \le k \le r$. Thus $\tau(\mathbf{v}) = \mathbf{w}$. □

Example 12.3.2. We consider the representation $3\,V_2$ of C_3 over a field $\mathbb{F}$ of characteristic 3. By Theorem 1.11.2, we know that $\mathbb{F}[V_2]^{C_3} = \mathbb{F}[N, x]$ where V_2^* has triangular basis $\{x, y\}$ and $N = \mathbf{N}(y) = y^3 - x^2y$. Let $\{x_1, y_1, x_2, y_2, x_3, y_3\}$ be a triangular basis for $(3\,V_2)^*$ and consider the $\mathcal{P}\mathrm{ol}^3(\{x, N\})$. Clearly, $\mathcal{P}\mathrm{ol}(x) = \{x_1, x_2, x_3\}$. To compute $\mathcal{P}\mathrm{ol}(N)$, we consider

$$
\begin{aligned}
&(\lambda_1 y_1 + \lambda_2 y_2 + \lambda_3 y_3)^3 - (\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3)^2 (\lambda_1 y_1 + \lambda_2 y_2 + \lambda_3 y_3)\\
&\qquad = \lambda_1^3 N_1 + \lambda_2^3 N_2 + \lambda_3^3 N_3 + \lambda_1^2 \lambda_2 x_1 u_{12} + \lambda_1^2 \lambda_3 x_1 u_{23} - \lambda_1 \lambda_2^2 x_2 u_{12}\\
&\qquad\quad - \lambda_1 \lambda_3^2 x_3 u_{13} + \lambda_2^2 \lambda_3 x_2 u_{23} - \lambda_2 \lambda_3^2 x_3 u_{23} - \lambda_1 \lambda_2 \lambda_3 (x_2 u_{13} + x_1 u_{23})
\end{aligned}
$$

where $N_i = \mathbf{N}(y_i)$ and $u_{ij} = x_i y_j - x_j y_i$ for $1 \leq i < j \leq 3$. Thus $\mathcal{P}\mathrm{ol}(N) = \{N_1, N_2, N_3\} \cup \{x_k u_{ij} \mid k = 1, 2, 3, 1 \leq i < j \leq 3\}$. Therefore by Theorem 12.3.1, we see that the set $A = \{x_1, x_2, x_3, N_1, N_2, N_3\} \cup \{x_k u_{ij} \mid k = 1, 2, 3, 1 \leq i < j \leq 3\}$ is a separating set for the action of C_3 on $3\,V_2$.

Note that $u_{12}^3 = x_2^3 N_1 - x_1^3 N_2 + (x_1 x_2)^2 u_{12}$ and thus we see that $u_{12} \in \widehat{S}$ where S is the $\mathbb{F}$-algebra generated by the separating set A. Similarly, $u_{13}, u_{23} \in \widehat{S}$.

Comparing Example 1.13.3 with Theorem 12.1.10 we see that the four other generators of $\mathbb{F}[3\,V_2]^{C_3}$,

$$\mathrm{Tr}^{C_3}(y_1^2 y_2^2 y_3), \mathrm{Tr}^{C_3}(y_1^2 y_2 y_3^2), \mathrm{Tr}^{C_3}(y_1 y_2^2 y_3^2)$$

and

$$\mathrm{Tr}^{C_3}(y_1^2 y_2^2 y_3^2)$$

also lie in $\widehat{\widehat{S}}$.

13

Using SAGBI Bases to Compute Rings of Invariants

In [98], Shank constructed generating sets, in fact, SAGBI bases, for the rings $\mathbb{F}[V_4]^{C_p}$ and $\mathbb{F}[V_5]^{C_p}$ for all primes $p \geq 5$. Of course, for the primes $p = 2, 3$, the corresponding actions are actions of C_{p^2} or C_{p^3}, not C_p. The rings of invariants $\mathbb{F}[V_4]^{C_4}$, $\mathbb{F}[V_4]^{C_9}$, $\mathbb{F}[V_5]^{C_8}$ and $\mathbb{F}[V_5]^{C_9}$ are all easily computed by computer, for example, using MAGMA.

We explain here the method used in [98]. This is a method used to confirm that a conjectural set of generators does indeed generate the full ring of invariants. In theory, this method will work for any ring of invariants $\mathbb{F}[V]^G$ provided we can compute a closed form for its Hilbert series, $\mathcal{H}(\mathbb{F}[V]^G, \lambda)$. Currently, for modular groups, there is only a formula for $\mathcal{H}(\mathbb{F}[V]^G, \lambda)$ for those groups G for which p^2 does not divide $|G|$ (and for permutation representations). This formula is described in [55].

We begin with a simple example to illustrate the main ideas.

Example 13.0.1. Let $G = C_p$, $V = V_2 \oplus V_2$ and let $\mathbb{F}$ be any field of characteristic p. We choose a generator σ of C_p and a basis $\{x_1, y_1, x_2, y_2\}$ for V^* such that $\sigma(x_i) = x_i$ and $\sigma(y_i) = y_i + x_i$ for $i = 1, 2$. We want to show that $\mathbb{F}[V]^G = \mathbb{F}[x_1, x_2, u, N_1, N_2]$ where $u = x_2y_1 - x_1y_2$, and $N_i = y_i^p - x_i^{p-1}y_i$ for $i = 1, 2$.

The method of Hughes and Kemper gives

$$\mathcal{H}(\mathbb{F}[V]^G, \lambda) = \frac{(1-\lambda^2)^p}{(1-\lambda^2)(1-\lambda)^2(1-\lambda^p)^2}.$$

We declare that $x_1 < y_1 < x_2 < y_2$ and work with the resulting graded reverse lexicographic ordering on the monomials of $\mathbb{F}[V] = \mathbb{F}[x_1, y_1, x_2, y_2]$. With this ordering we have $\mathrm{LM}(x_i) = x_i$, $\mathrm{LM}(N_i) = y_i^p$ for $i = 1, 2$ and $\mathrm{LM}(u) = x_2y_1$.

We define $T := \mathbb{F}[x_1, N_1, x_2, N_2, u]$. It is easily verified that these five functions are all invariants and therefore, $T \subset \mathbb{F}[V]^G$. We want to show that T is the full ring of invariants. We consider the monomial algebras $Q := \mathbb{F}[x_1, y_1^p, x_2, y_2^p, x_2y_1] \subseteq \mathrm{LT}(T)$ and $A := \mathbb{F}[x_1, y_1^p, x_2, y_2^p]$. Since x_1, y_1^p, x_2, y_2^p is

H.E.A. Eddy Campbell, D. Wehlau, *Modular Invariant Theory*,
Encyclopaedia of Mathematical Sciences 139,
DOI 10.1007/978-3-642-17404-9_13, © Springer-Verlag Berlin Heidelberg 2011

a homogeneous system of parameters for $\mathbb{F}[V]$, we see that Q must be a finite A-module.

It is not too difficult to see that $Q = \sum_{i=0}^{p-1} A\,(x_2y_1)^i$, but rather than prove this directly we consider $M := \sum_{i=0}^{p-1} A\,(x_2y_1)^i$ as an A-submodule of Q. It is clear that M is a free A-module: $M = \oplus_{i=0}^{p-1} A\,(x_2y_1)^i$. Thus

$$\begin{aligned}\mathcal{H}(M,\lambda) &= \sum_{i=0}^{p-1} \mathcal{H}(A\,(x_2y_1)^i,\lambda)\\ &= \sum_{i=0}^{p-1} \mathcal{H}(A,\lambda)\lambda^{2i}\\ &= \frac{1}{(1-\lambda)^2(1-\lambda^p)^2}\sum_{i=0}^{p-1}\lambda^{2i}\\ &= \mathcal{H}(\mathbb{F}[V]^G,\lambda).\end{aligned}$$

Thus we have the chain of inclusions $M \subseteq Q \subseteq \mathrm{LT}(T) \subseteq \mathrm{LT}(\mathbb{F}[V]^G)$. But since $\mathcal{H}(\mathbb{F}[V]^G,\lambda) = \mathcal{H}(\mathrm{LT}(\mathbb{F}[V]^G),\lambda)$, we have $\mathcal{H}(M,\lambda) = \mathcal{H}(\mathbb{F}[V]^G,\lambda)$, and therefore, $M = Q = \mathrm{LT}(T) = \mathrm{LT}(\mathbb{F}[V]^G)$. In particular, the monomials in M form a vector space basis for the vector space of lead terms of $\mathbb{F}[V]^G$. Since the monomials in M are generated multiplicatively by x_1, x_2, $y_1^p = \mathrm{LT}(N_1)$, $y_2^p = \mathrm{LT}(N_2)$ and $x_2y_1 = \mathrm{LT}(u)$, we have, in fact, proved that $\{x_1, x_2, N_1, N_2, u\}$ is a SAGBI basis (and hence a generating set) for $\mathbb{F}[V]^G$.

There are some features of Example 13.0.1 that make the computation simpler than it is in general. These include the fact that the natural minimal algebra generating set is already a SAGBI basis and the fact that monomial algebra Q is a free A-module in the example. In particular, this second fact made the calculation of $\mathcal{H}(M,\lambda)$ particularly easy. We will have to account for less well behaved algebras in the general case.

Having seen this first example we proceed to describe the method in general. Suppose then that we have a closed expression for $\mathcal{H}(\mathbb{F}[V]^G,\lambda)$. The first step is to somehow guess a finite set of invariants $\mathcal{B}$ which is a SAGBI basis for $\mathbb{F}[V]^G$. The method then gives a way to use the knowledge of $\mathcal{H}(\mathbb{F}[V]^G,\lambda)$ to prove that $\mathcal{B}$ is indeed a SAGBI basis for $\mathbb{F}[V]^G$.

Consider the algebra T generated by the elements of $\mathcal{B}$ and the algebra Q generated by the set of monomials $\{\mathrm{LM}(f) \mid f \in \mathcal{B}\}$. Clearly $Q \subseteq \mathrm{LT}(T)$. Let n denote the Krull dimension of $\mathbb{F}[V]$. We choose n invariants $f_1, f_2, \ldots, f_n$ in $\mathbb{F}[V]^G$ such that $a_1, a_2, \ldots, a_n$ forms a homogeneous system of parameters for $\mathbb{F}[V]$ where $a_i = \mathrm{LM}(f_i)$ for $i = 1, 2, \ldots, n$. Then Q (and T) is a finitely generated A-module and so there exists a finite set $\Gamma_0 \subset Q$ such that

$$Q = \sum_{h\in\Gamma_0} A \cdot h\ .$$

We wish to find such a set Γ_0. We choose a finite subset Γ of Q and consider the A-module, M, generated by Γ:

$$M := \sum_{h \in \mathit{\Gamma}} A \cdot h \ .$$

Thus we have $M \subseteq Q \subseteq \mathrm{LT}(T) \subseteq \mathrm{LT}(\mathbb{F}[V]^G$. If we can show $M = \mathrm{LT}(\mathbb{F}[V]^G)$, then we will have shown that $\mathcal{B}$ and also $\mathit{\Gamma} \cup \{a_1, a_2, \ldots, a_n\}$ is a SAGBI basis for $\mathbb{F}[V]^G$. Since $M \subseteq \mathrm{LT}(\mathbb{F}[V]^G)$, it suffices to show that $\mathcal{H}(M,\lambda) = \mathcal{H}(\mathrm{LT}(\mathbb{F}[V]^G),\lambda)$. But since $\mathcal{H}(\mathrm{LT}(\mathbb{F}[V]^G),\lambda) = \mathcal{H}(\mathbb{F}[V]^G,\lambda)$, our goal is to prove that $\mathcal{H}(M,\lambda) = \mathcal{H}(\mathbb{F}[V]^G,\lambda)$.

Since we have a closed expression for $\mathcal{H}(\mathbb{F}[V]^G,\lambda)$, we need to compute $\mathcal{H}(M,\lambda)$ from knowledge of $\mathit{\Gamma}$. This computation is at the heart of this method. We exploit the fact that M is an A-module and A is generated by monomials. We first decompose M into a direct sum of certain A-submodules and then proceed to compute the Hilbert series for each of these submodules.

We illustrate the entire method by computing a generating set for $\mathbb{F}[V_4]^{C_{11}}$. This example displays essentially all of the details involved in computing $\mathbb{F}[V_4]^{C_p}$ for general p.

Example 13.0.2. We write $R := \mathbb{F}[V_4]^{C_{11}}$ where $\mathbb{F}$ is a field of characteristic 11. We take the triangular basis $\{x, y, z, w\}$ of V_4 where $\sigma(x) = x$, $\sigma(y) = y + x$, $\sigma(z) = z + y$ and $\sigma(w) = w + z$. Again, we use the reverse lexicographic ordering with $x < y < z < w$. Thus by Theorem 5.2.4, there is indeed a finite SAGBI basis for $\mathbb{F}[V_4]^{C_{11}}$.

We begin by choosing $h_1 = x$, $h_2 = d = y^2 - 2xz - xy$, $h_3 = \mathrm{Tr}(w^{p-1}) = \mathrm{Tr}(w^{10}) = z^{10} + \ldots$ and $h_4 = \mathrm{N}(w) = w^{11} + \ldots$. Then the corresponding lead terms $a_1 = x$, $a_2 = y^2$, $a_3 = z^{10}$, $a_4 = w^{11}$ do indeed form a homogeneous system of parameters for $\mathbb{F}[V]$. There are two more integral invariants:

$$\begin{aligned} e &:= y^3 - 3xyz + x^2 \quad \text{and} \\ f &:= 3z^2y^2 - 6wy^3 - 3zy^3 - y^4 - 8z^3x + 18wzyx + 6z^2yx + 9wy^2x + 9zy^2x \\ &\quad + 2y^3x - 9w^2x^2 - 18wzx^2 - 12z^2x^2 - 3wyx^2 - 6zyx^2 - y^2x^2 . \end{aligned}$$

We make the “guess” that the set $\mathcal{B}$ comprised of the following twenty-four invariants is a SAGBI basis for R.

(0) h_1, h_2, h_3, h_4
(1) e, f
(2) $\mathrm{Tr}(z^i w^{10})$ for $1 \le i \le 9$
(3) $\mathrm{Tr}(z^i w^5)$ for $3 \le i \le 9$
(4) $\mathrm{Tr}(w^i)$ for $7 \le i \le 9$
(5) $\mathrm{Tr}(z^2 w^i)$ for $7 \le i \le 9$

In reality, the corresponding guess for general p was motivated by computer experiments with small values of p.

Using Lemma 9.0.2 it is not too hard to show that the corresponding lead monomials are

(0) x, y^2, z^{10}, w^{11}
(1) y^3, y^2z^2
(2) z^{10+i} for $1 \leq i \leq 9$
(3) yz^{8+i} for $3 \leq i \leq 9$
(4) $y^{10-i}z^{2i-10}$ for $7 \leq i \leq 9$
(5) $y^{10-j}z^{2j-8}$ for $7 \leq j \leq 9$

We take $A := \mathbb{F}[x, y^2, z^{10}, w^{11}]$ and consider R as an A-module. Now we wish to guess our set Γ of A-module generators for R. We take Γ_1 to be the above listed monomials excepting x, y^2, z^{10}, w^{11} which, of course, lie in A. Thus Γ_1 consists of 24 monomials. To form Γ, we add four more monomials:

(6) 1
(7) $yz^{19} = \mathrm{LM}(\mathrm{Tr}(zw^{10})\,\mathrm{Tr}(w^9)) = \mathrm{LM}(\mathrm{Tr}(zw^{10}))\,\mathrm{LM}(\mathrm{Tr}(w^9))$
(8) $y^4z^4 = \mathrm{LM}(f^2) = \mathrm{LM}(f)^2$
(9) $y^5z^2 = \mathrm{LM}(ef) = \mathrm{LM}(e)\,\mathrm{LM}(f)$

Note that each of these four extra monomials lies in $LT(\mathbb{F}[\mathcal{B}])$.

We let M denote the A-module generated by the twenty-eight monomials of Γ. We will show, using their Hilbert series, that $M = \mathrm{LT}(\mathbb{F}[V_4]^{C_{11}})$ and thus that $\Gamma \cup \{x, y^2, z^{10}, w^{11}\}$ generate $\mathrm{LT}(\mathbb{F}[V_4]^{C_{11}})$ as an algebra.

We are faced with the problem of calculating $\mathcal{H}(M, \lambda)$. In Example 13.0.1, this was easily done because M was a free A-module. Here this is definitely not the case. We proceed by decomposing M into a direct sum of simpler A-modules as follows. For $a, b \in \mathrm{N}$ we define

$$M_{(a,b)} := \mathrm{span}_F\{x^iy^jz^kw^\ell \mid j \equiv a \pmod 2 \text{ and } k \equiv b \pmod{10}\} \ .$$

Then each $M_{(a,b)}$ is an A-module and

$$M = \oplus_{a=0}^{1} \oplus_{b=0}^{9} M_{(a,b)} \ ,$$

and therefore,

$$\mathcal{H}(M, \lambda) = \sum_{a=0}^{1}\sum_{b=0}^{9} \mathcal{H}(M_{(a,b)}, \lambda) \ . \tag{13.0.1}$$

Thus we have reduced to finding the Hilbert series of the simpler A-modules, $M_{(a,b)}$.

Writing each of the twenty-eight module generators for M in the form $x^iy^jz^kw^\ell$, we note that i and ℓ are always zero. It is clear that for an A-module generator, i must be zero since $x \in A$. There are deeper reasons why we always find $\ell = 0$. This is related to our choice of w as the greatest variable in our triangular basis.

We begin by sorting these twenty-eight A-module generators of M according to the parity of j and the residue class of k modulo 10.

		k mod 10									
		0	1	2	3	4	5	6	7	8	9
j mod 2	0	1	z^{11}	y^2z^2 z^{12}	z^{13}	y^4z^4 z^{14}	z^{15}	y^2z^6 z^{16}	z^{17}	y^2z^8 z^{18}	z^{19}
	1	y^3 yz^{10}	yz^{11}	y^5z^2 yz^{12}	yz^{13}	y^3z^4 yz^{14}	yz^{15}	y^3z^6 yz^{16}	yz^{17}	yz^8	yz^{19}

Thus $M_{(j,k)}$ is cyclic for $j = 0, 1$ and k odd. Therefore $\mathcal{H}(M_{(j,k)}, \lambda) = \mathcal{H}(A, \lambda)\lambda^{10+k+j}$ for $j = 0, 1$ and k odd. Also $\mathcal{H}(M_{(0,0)}, \lambda) = \mathcal{H}(A, \lambda)$ and $\mathcal{H}(M_{(1,8)}, \lambda) = \mathcal{H}(A, \lambda)\lambda^9$. For the other values of (j, k), the A-module $M_{(j,k)}$ has exactly two generators and is not a free A-module.

Suppose $M_{(j,k)}$ is minimally generated by the two monomials $m_1 = y^{a_1}z^{b_1}$ and $m_1 = y^{a_2}z^{b_2}$. Note that m_1 cannot divide m_2 nor can m_2 divide m_1 since both are required as generators of $M_{(j,k)}$. For concreteness, we suppose that $a_1 < a_2$ (and then $b_1 > b_2$). We can resolve $M_{(j,k)}$ by free A-modules by the following exact sequence:

$$0 \longrightarrow A\beta \xrightarrow{\Psi_0} A\alpha_1 \oplus A\alpha_2 \xrightarrow{\Psi_1} M_{(j,k)} \longrightarrow 0$$

with $\Psi_0(\beta) = y^{a_2-a_1}\alpha_1 - z^{b_1-b_2}\alpha_2$ and $\Psi_1(\alpha_1) = m_1$ and $\Psi_1(\alpha_2) = m_2$. We make this exact sequence graded by declaring that α_1 has the same degree as m_1, that α_2 has the same degree as m_2 and that β has the same degree as $y^{a_2}z^{b_1}$. Then by the additivity of Hilbert series $\mathcal{H}(\cdot, \lambda)$ we find that

$$\begin{aligned}\mathcal{H}(M_{(j,k)}, \lambda) &= -\mathcal{H}(A\beta, \lambda) + \mathcal{H}(A\alpha_1 \oplus A\alpha_2, \lambda)\\ &= -\mathcal{H}(A, \lambda)\lambda^{a_2+b_1} + \mathcal{H}(A, \lambda)\lambda^{a_1+b_1} + \mathcal{H}(A, \lambda)\lambda^{a_2+b_2}\\ &= \mathcal{H}(A, \lambda)(-\lambda^{a_2+b_1} + \lambda^{a_1+b_1} + \lambda^{a_2+b_2})\end{aligned}$$

As an example, consider $M_{(1,2)}$ which is generated by $m_1 = yz^{12}$ and $m_2 = y^5z^2$. We can resolve $M_{(1,2)}$ by free A-modules as follows:

$$0 \longrightarrow A\beta \xrightarrow{\Psi_0} A\alpha_1 \oplus A\alpha_2 \xrightarrow{\Psi_1} M_{(1,2)} \longrightarrow 0$$

with $\Psi_0(\beta) = y^4\alpha_1 - z^{10}\alpha_2$ and $\Psi_1(\alpha_1) = yz^{12}$ and $\Psi_1(\alpha_2) = y^5z^2$.

Thus $\mathcal{H}(M_{(1,2)}, \lambda) = \mathcal{H}(A, \lambda)(-\lambda^{17} + \lambda^{13} + \lambda^7)$.

This argument yields the following Hilbert series:

$$\begin{aligned}\mathcal{H}(M_{(1,0)}, \lambda) &= \mathcal{H}(A, \lambda)(-\lambda^{13} + \lambda^{11} + \lambda^3),\\ \mathcal{H}(M_{(0,2)}, \lambda) &= \mathcal{H}(A, \lambda)(-\lambda^{14} + \lambda^{12} + \lambda^4),\\ \mathcal{H}(M_{(1,2)}, \lambda) &= \mathcal{H}(A, \lambda)(-\lambda^{17} + \lambda^{13} + \lambda^7),\\ \mathcal{H}(M_{(0,4)}, \lambda) &= \mathcal{H}(A, \lambda)(-\lambda^{18} + \lambda^{14} + \lambda^8),\\ \mathcal{H}(M_{(1,4)}, \lambda) &= \mathcal{H}(A, \lambda)(-\lambda^{17} + \lambda^{15} + \lambda^7),\end{aligned}$$

$$\begin{aligned}\mathcal{H}(M_{(0,6)},\lambda) &= \mathcal{H}(A,\lambda)(-\lambda^{18}+\lambda^{16}+\lambda^{8}),\\ \mathcal{H}(M_{(1,6)},\lambda) &= \mathcal{H}(A,\lambda)(-\lambda^{19}+\lambda^{17}+\lambda^{9}),\\ \mathcal{H}(M_{(0,8)},\lambda) &= \mathcal{H}(A,\lambda)(-\lambda^{20}+\lambda^{18}+\lambda^{10}).\end{aligned}$$

Using Equation (13.0.1) we have the following calculation of $\mathcal{H}(M,\lambda)$. First, let

$$q(\lambda) = (1+\lambda^3+\lambda^4+2\lambda^7+2\lambda^8+2\lambda^9+\lambda^{10}+2\lambda^{11}+2\lambda^{12}+\lambda^{13}+\lambda^{14}+2\lambda^{15}+2\lambda^{16})\,.$$

Then

$$\begin{aligned}\mathcal{H}(M,\lambda) &= q(\lambda)\mathcal{H}(A,\lambda)\\ &= \frac{q(\lambda)}{(1-\lambda)(1-\lambda^2)(1-\lambda^{10})(1-\lambda^{11})}\,.\end{aligned}$$

Simplifying this yields

$$\begin{aligned}\mathcal{H}(M,\lambda) &= \frac{1+\lambda^3+\lambda^7+\lambda^8+2\lambda^9+2\lambda^{10}}{(1-\lambda)(1-\lambda^2)(1-\lambda^4)(1-\lambda^{11})}\\ &= \frac{1-2\lambda+2\lambda^2-\lambda^3+\lambda^4-2\lambda^5+2\lambda^6}{(1-\lambda)^3(1-\lambda^{11})}\\ &= \mathcal{H}(\mathbb{F}[V_4]^{C_{11}},\lambda).\end{aligned}$$

Thus the monomials generated by the elements of $\Gamma\cup\{x,y^2,z^{10},w^{11}\}$ form a vector space basis for $LT(\mathbb{F}[V_4]^{C_{11}})$ and so $\mathcal{B}$ is a SAGBI basis (and thus a generating set) for $\mathbb{F}[V_4]^{C_{11}}$.

As stated above, this method works entirely similarly to verify a SAGBI basis for $\mathbb{F}[V_4]^{C_p}$ for general $p\geq 5$. The reader may see all the details in [98] from which this example was adapted.

14

Ladders

We consider a group G with a normal subgroup N. If $\sigma \in G$ and $\tau \in N$, then $\tau\sigma = \sigma\tau'$ for some $\tau' \in N$ by normality. Therefore for $f \in \mathbb{F}[V]^N$, we have $\tau \cdot (\sigma \cdot f) = \tau\sigma \cdot f = \sigma\tau' \cdot f = \sigma \cdot f$ and thus $\sigma \cdot f \in \mathbb{F}[V]^N$. This shows that G acts on $\mathbb{F}[V]^N$. Clearly $(\mathbb{F}[V]^N)^G = \mathbb{F}[V]^G$. Since N acts trivially on $\mathbb{F}[V]^N$, in fact, G/N acts on $\mathbb{F}[V]^N$ and $(\mathbb{F}[V]^N)^{G/N} = \mathbb{F}[V]^G$. We have seen this in detail in Lemma 1.10.1.

This description suggests that we may compute $\mathbb{F}[V]^G$ by first computing $\mathbb{F}[V]^N$ and then computing the G/N invariants in the ring of $\mathbb{F}[V]^N$. Thus we may work with the two smaller groups N and G/N. In characteristic zero, this method is very powerful. See for example the work of Littelmann [76], Popov [89], Wehlau [110].

A major difficulty with this technique is that we must find the invariants of the G/N action on a ring which is usually not a polynomial ring. If G/N is a linearly reductive group, then this difficulty can be surmounted by replacing $\mathbb{F}[V]^N$ by a polynomial ring as follows. The space of decomposable invariants, $D := (\mathbb{F}[V]^N_+)^2$ is a G/N-stable ideal in $\mathbb{F}[V]^N$. If G/N is linearly reductive, we may choose a G/N stable complement, Q to D:

$$\mathbb{F}[V]^N = D \oplus Q.$$

A set $\{f_1, \ldots, f_r\}$ of N-invariants is a generating set for $\mathbb{F}[V]^N$ if and only if the images $\{\overline{f_1}, \ldots, \overline{f_r}\}$ span $\mathbb{F}[V]^N/D$. Thus we may choose generators for $\mathbb{F}[V]^N$ by taking a basis $\{f_1, \ldots, f_m\}$ of Q. We then consider the dual of Q, Q^* as a G/N representation space. More precisely, we introduce indeterminates $y_1, \ldots, y_m$ of degree 1 and define a G/N action on $\mathbb{F}[y_1, \ldots, y_m]$ as follows. For $\sigma \in G/N$ we put

$$\sigma \cdot y_i = \sum_{j=1}^{m} \alpha^{\sigma}_{i,j} y_j$$

where $\alpha^{\sigma}_{i,j} \in \mathbb{F}$ is defined by

DOI 10.1007/978-3-642-17404-9_14, © Springer-Verlag Berlin Heidelberg 2011

$$\sigma \cdot f_i = \sum_{j=1}^{m} \alpha_{i,j}^{\sigma} f_j.$$

The point here is that $\mathbb{F}[y_1, \ldots, y_m]$ is a polynomial ring of Krull dimension m whereas there are usually algebraic relations among the functions $f_1, \ldots, f_m$. There is a natural G/N-equivariant algebra surjection

$$\rho : \mathbb{F}[W] \cong \mathbb{F}[y_1, \ldots, y_m] \to \mathbb{F}[f_1, \ldots, f_m] = \mathbb{F}[V]^N$$

carrying y_i to f_i where $W \cong Q$ is the m dimensional G/N-representation dual to $\operatorname{span}_{\mathbb{F}}\{y_1, \ldots, y_m\}$. Since G/N is linearly reductive, ρ restricts to the algebra surjection

$$\rho^{G/N} : \mathbb{F}[W]^{G/N} \cong \mathbb{F}[y_1, \ldots, y_m]^{G/N} \to \mathbb{F}[f_1, \ldots, f_m]^{G/N} = \mathbb{F}[V]^G.$$

Thus to compute $\mathbb{F}[V]^G$, it suffices to compute $\mathbb{F}[V]^N$ and $\mathbb{F}[W]^{G/N}$.

If G/N is not reductive, the above programme requires two modifications. The first change is not too difficult. The construction of W must be done more carefully. It no longer suffices to lift a minimal generating set $f_1, \ldots, f_m$ of $\mathbb{F}[V]^N$. We require a set of generators $h_1, \ldots, h_r$ whose span is a G/N stable vector space. To do this, we begin with a set of generators (which need not be minimal), $\{f_1, \ldots, f_m\}$ for $\mathbb{F}[V]^N$. We define $\widetilde{Q}$ to be the vector space

$$\widetilde{Q} := \operatorname{span}_{\mathbb{F}}\{\sigma \cdot f_i \mid \sigma \in G, i = 1, \ldots, m\}$$

and choose a vector space basis $h_1, \ldots, h_r$ for $\widetilde{Q}$. Then we use indeterminates $y_1, \ldots, y_r$ as before: for $\sigma \in G$, we define

$$\sigma \cdot y_i = \sum_{j=1}^{r} \alpha_{i,j}^{\sigma} y_j$$

where

$$\sigma \cdot h_i = \sum_{j=1}^{r} \alpha_{i,j}^{\sigma} h_j.$$

Thus $\mathbb{F}[W] = \mathbb{F}[y_1, \ldots, y_r]$ where $W \cong \widetilde{Q}$. Since this construction produces a larger G/N-module W than in the reductive case, the computation of $\mathbb{F}[W]^{G/N}$ may be expected to be correspondingly harder.

The second difficulty we must overcome, if G/N is not reductive, is more serious. Although

$$\rho : \mathbb{F}[y_1, \ldots, y_m] \to \mathbb{F}[f_1, \ldots, f_m] = \mathbb{F}[V]^N$$

is an algebra surjection, the restriction

$$\rho^{G/N} : \mathbb{F}[y_1, \ldots, y_m]^{G/N} \to \mathbb{F}[f_1, \ldots, f_m]^{G/N} = \mathbb{F}[V]^G$$

is usually not surjective.

Before discussing how to deal with this, we give an example of an algebra surjection whose restriction is not surjective.

Example 14.0.1. We take $\mathbb{F}$ a field of characteristic p and $G = C_p$ with generator σ. With $\mathbb{F}[V_2] = \mathbb{F}[x, y]$ and $\mathbb{F}[V_1] = \mathbb{F}[u]$ we have $\sigma \cdot y = y + x$, $\sigma \cdot x = x$ and $\sigma \cdot u = u$. Consider the surjective algebra homomorphism

$$\rho : \mathbb{F}[V_2] \to \mathbb{F}[V_1]$$

given by $\rho(y) = u$ and $\rho(x) = 0$.

This surjection is C_p-equivariant since

$$\begin{aligned} \sigma \cdot \rho(y) &= \sigma \cdot u = u = \rho(y + x) = \rho(\sigma \cdot y) \quad \text{and} \\ \sigma \cdot \rho(x) &= \sigma \cdot 0 = 0 = \rho(\sigma \cdot x). \end{aligned}$$

However, at the level of invariants, the map is not surjective:

$$\rho^{C_p} : \mathbb{F}[V_2]^{C_p} = \mathbb{F}[y^p - x^{p-1}y, x] \to \mathbb{F}[V_1] = \mathbb{F}[u]$$

where $\rho^{C_p}(y^p - x^{p-1}y) = u^p$ and $\rho^{C_p}(x) = 0$.

Now we return to the question of overcoming this difficulty. We use group cohomology to measure the extent to which $\rho^{G/N}$ fails to be surjective.

14.1 Group Cohomology

A good reference for group cohomology may be found in Benson's book [8]. We will very briefly recall the definitions needed for cyclic groups. Let $G = C_m$ be a cyclic group of order m with generator σ. Define two maps, Tr and Δ by $\mathrm{Tr}(x) = x + \sigma(x) + \sigma^2(x) + \cdots + \sigma^{m-1}(x)$ and $\Delta(x) = \sigma(x) - x$. With this notation we construct a periodic projective resolution of the trivial C_m-module, $\mathbb{F}$ as follows. Put $F_i \cong \mathbb{F}G$ for all i and define $\partial_i : F_i \to F_{i-1}$ by

$$\partial_i = \begin{cases} \Delta & \text{if } i \text{ is odd,} \\ \mathrm{Tr} & \text{if } i \text{ is even.} \end{cases}$$

Thus we have

$$\cdots \xrightarrow{\mathrm{Tr}} \mathbb{F}G \xrightarrow{\Delta} \mathbb{F}G \xrightarrow{\mathrm{Tr}} \mathbb{F}G \xrightarrow{\Delta} \mathbb{F}G \xrightarrow{\mathrm{Tr}} \mathbb{F}G \xrightarrow{\Delta} \mathbb{F}G$$

with $F_0/\mathrm{Im}(\partial_1) = \mathbb{F}G/\mathrm{Im}(\Delta) \cong \mathbb{F}$.

Applying the functor $\mathrm{Hom}_{\mathbb{F}G}(\cdot, M)$ and taking the cohomology of the resulting complex we get

$$\begin{aligned} H^0(C_m; M) &= \mathrm{Kernel}(\Delta|_M) = M^{C_m}, \\ H^1(C_m; M) &= \frac{\mathrm{Kernel}(\mathrm{Tr}\,|_M)}{\mathrm{Im}(\,\Delta|_M\,)}, \\ H^2(C_m; M) &= \frac{\mathrm{Kernel}(\,\Delta|_M\,)}{\mathrm{Im}(\mathrm{Tr}\,|_M)}, \end{aligned}$$

and, for $i > 0$, $H^{2i+1}(C_m; M) = H^1(C_m; M)$ and $H^{2i}(C_m; M) = H^2(C_m; M)$.

Since the cohomology functor is additive, we see that a C_m-module decomposition of M gives a vector space decomposition of $H^i(C_m; M)$ in terms of $H^i(C_m; V)$ with V indecomposable.

We now specialize to the case where $m = p$, is the characteristic of $\mathbb{F}$. By the above, we see that it is important to understand $H^i(C_p; V_n)$. The reader may wish to refer back to §7.1. Consider any triangular basis $\{e_1, e_2, \ldots, e_n\}$ of V_n.

Clearly $H^0(C_p; V_n)$ is a one dimensional vector space spanned by e_1. Using the fact that $\Delta^{p-1} = (\sigma - 1)^{p-1} = \sum_{\ell=0}^{p-1} \sigma^\ell = \mathrm{Tr}$, we see that

$$H^1(C_p; V_p) = H^2(C_p; V_p) = 0$$

while, for $n < p$, we have

$$\begin{aligned} H^1(C_p; V_n) &= \{\overline{e_n}\} \quad \text{and} \\ H^2(C_p; V_n) &= \{\overline{e_1}\} \,. \end{aligned}$$

The first equality follows since $\Delta(V_n) = V_{n-1}$ and $\mathrm{Tr}(V_n) = \{0\}$. The second equality follows similarly since e_1 spans Kernel(Δ) and $\mathrm{Im}(\mathrm{Tr}\,|_{V_n}) = 0$.

14.2 Cohomology and Invariant Theory

Recall the situation at the beginning of this chapter with G/N acting on $\mathbb{F}[V]^N$. Choose a set of generators $\{f_1, f_2, \ldots, f_m\}$ for $\mathbb{F}[V]^N$ and a polynomial algebra $A = \mathbb{F}[y_1, y_2, \ldots, y_m]$ equipped with a G/N-action such that the map $\rho : A \to \mathbb{F}[V]^N$ given by $\rho(y_i) = f_i$ is G/N equivariant. We recall the notation W for the representation of G/N given by $\{y_1, y_2, \ldots, y_m\}$ so that $A = \mathbb{F}[W]$.

We let J denote the ideal in A which is the kernel of ρ. The short exact sequence of A-modules

$$0 \to J \xrightarrow{\gamma} A \xrightarrow{\rho} \mathbb{F}[V]^N \to 0$$

induces a long exact sequence in group cohomology

$$0 \to J^{G/N} \overset{\gamma^{G/N}}{\to} A^{G/N} \overset{\rho^{G/N}}{\to} \mathbb{F}[V]^G \overset{\delta}{\to} H^1(G/N; J) \overset{\gamma^1}{\to} H^1(G/N; A) \overset{\rho^1}{\to} \ldots .$$

Of course, all the maps in this long exact sequence are $A^{G/N}$-module maps.

From the sequence, we see that

$$\frac{\mathbb{F}[V]^G}{\ker(\delta)} \cong \mathrm{Im}(\delta)$$

and therefore, we have the following isomorphisms of vector spaces.

$$\mathbb{F}[V]^G \cong \ker(\delta) \oplus \mathrm{Im}(\delta) \cong \mathrm{Im}(\rho^{G/N}) \oplus \ker(\gamma^1).$$

This shows that $\mathbb{F}[V]^G$ is generated by module generators of the two $\mathbb{F}[W]^{G/N}$-modules: $\mathrm{Im}(\rho^{G/N})$ and $\ker(\gamma^1)$.

Therefore, in order to find a generating set for $\mathbb{F}[V]^G$ in this situation, we have the additional work of computing $\mathbb{F}[W]^{G/N}$-module generators of $\ker(\gamma^1)$. In practice, in our experience, doing so is a hard problem. Indeed, even computing $H^1(G/N; J)$ and $H^1(G/N; \mathbb{F}[W])$ seems impractical unless G/N is cyclic.

We finish this section by giving three examples.

Example 14.2.1. We let $\mathbb{F}$ denote the finite field of order p^2 and we choose an element $\omega \in \mathbb{F}$ with ω not in the prime field $\mathbb{F}_p$. Thus w generates $\mathbb{F}$ over $\mathbb{F}_p$. We consider the subgroup of $\mathrm{GL}(3, \mathbb{F})$ generated by the three matrices.

$$\sigma^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \qquad \tau^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \omega & 0 & 1 \end{pmatrix}, \qquad \mu^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix},$$

Then G is a group of order p^3. Let V denote the three dimensional $\mathbb{F}$ vector space on which G is acting and let $\{x, y, z\}$ denote the basis of V^* dual to the ordinary basis.

Let H be the subgroup of order p^2 generated by σ and μ. Since the index of H in G is p, H is normal in G. Note that H is a Nakajima group and thus we have

$$\mathbb{F}[V]^H = \mathbb{F}[x, \mathrm{N}^H(y), \mathrm{N}^H(z)] = \mathbb{F}[x, y, z^{p^2} - x^{p^2-1}z].$$

Hence the space of indecomposables $Q(\mathbb{F}[V]^H) = (\mathbb{F}[V]^H_+)/(\mathbb{F}[V]^H_+)^2$ is a three dimensional vector space with basis the images of

$$\left\{x, y, z^{p^2} - x^{p^2-1}z\right\}$$

in the quotient.

Now we consider the action of $G/H \cong C_p$ on $Q(\mathbb{F}[V]^H)$. The image of τ generates G/H, and we write $\Delta := \tau - 1$. Then $\Delta(x) = 0$, $\Delta(y) = x$ and $\Delta(\mathrm{N}^H(z)) = y^{p^2} - x^{p^2-1}y$. Hence Q is not C_p stable and so we extend Q to the four dimensional space $\widetilde{Q}$ with basis

$$\{y, x, y^{p^2} - x^{p^2-1}y, z^{p^2} - x^{p^2-1}z\}.$$

Dualizing and lifting we take $W = V_2 \oplus V_2$.

We define $A := \mathbb{F}[W] = \mathbb{F}[X_1, Y_1, X_2, Y_2]$ where $\Delta(X_i) = 0$ and $\Delta(Y_i) = X_i$ for $i = 1, 2$. Define $\rho : A \to \mathbb{F}[V]^H$ by $\rho(Y_1) = y$, $\rho(X_1) = x$, $\rho(Y_2) = \mathrm{N}^H(z)$ and $\rho(X_2) = \Delta(\rho(Y_2)) = y^{p^2} - x^{p^2-1}y$. Then ρ is a C_p-equivariant algebra surjection.

The short exact sequence of A-modules

$$0 \to J \xrightarrow{\gamma} A \xrightarrow{\rho} \mathbb{F}[V]^H \to 0$$

induces a long exact sequence in group cohomology

$$0 \to J^{G/H} \overset{\gamma^{G/H}}{\to} A^{G/H} \overset{\rho^{G/H}}{\to} \mathbb{F}[V]^G \xrightarrow{\delta} H^1(G/H; J) \xrightarrow{\gamma^1} H^1(G/H; A) \xrightarrow{\rho^1} \dots.$$

We will show that $\rho^{G/N}$ is surjective. By exactness, this is equivalent to showing that $\gamma^1 : H^1(G/H; J) \to H^1(G/H; A)$ is injective.

Define $N := Y_1^{p^2} - X_1^{p^2-1} Y_1$. Then the kernel of ρ, J, is the principal ideal generated by $Y_1^{p^2} - X_1^{p^2-1}Y_1 - X_2 = N - X_2$. We write $[f]_A$ to denote the cohomology class in $H^1(G/H; A)$ represented by $f \in A$. Similarly, $[f]_J$ denotes the cohomology class in $H^1(G/H; J)$ represented by $f \in J$.

Suppose $[f'(X_2 - N)]_J \in \ker \gamma^1$. Since $[f'(X_2 - N)]_A = 0$, we have $f'(X_2 - N) \in \Delta(A)$ and so we may write $f'(X_2 - N) = \Delta(f)$ for some $f \in A$. Dividing f by N considered as a polynomial in Y_1 we have $f = qN + r$ where $\deg_{Y_1}(r) < p^2$. Similarly, $f' = q'N + r'$ where $\deg_{Y_1}(r') < p^2$. Thus

$$\begin{aligned} &f'(N - X_2) = \Delta(f) = \Delta(q)N + \Delta(r) \quad \text{and} \\ &f'(N - X_2) = (q'N + r')(N - X_2) = N(q'N + r' - X_2) - r'X_2. \end{aligned}$$

Thus $q'N + r' - X_2 = \Delta(q) \in \Delta(A)$. Hence

$$\begin{aligned} f' = q'N + r' = \Delta(q + Y_2) \in \Delta(A) \quad \text{and thus,} \\ f'(N - X_2) = \Delta((q + Y_2)(N - X_2)) \in \Delta(J). \end{aligned}$$

Therefore, $[f'(N - X_2)]_J = 0$ and thus γ^1 is injective.

This shows that $\rho^{G/N} : A^{C_p} \to \mathbb{F}[V]^G$ is surjective. We define

$$\begin{aligned} f_1 &= x, \\ f_2 &= \mathrm{N}^G(y) = y^p - x^{p-1}y, \\ f_3 &= \rho(X_2 Y_1 - X_1 Y_2) \\ &= x(z^{p^2} - x^{p^2-1}z) - (y^{p^2} - x^{p^2-1}y)y \\ &= y^{p^2+1} - xz^{p^2} - x^{p^2-1}y^2 + x^{p^2}z, \quad \text{and} \\ f_4 &= \mathrm{N}^G(z) = z^{p^3} + \dots \end{aligned}$$

Then $\mathbb{F}[V]^G$ is generated by

$$\begin{aligned} &\rho(X_1) = x = f_1, \qquad \rho(X_2) = y^{p^2} - x^{p^2-1}y = f_2^p + f_1^{p^2-p} f_2, \\ &\rho(\mathrm{N}^{C_p}(Y_1)) = f_2, \quad \rho(X_1Y_2 - X_2Y_1) = f_3, \quad \text{and} \\ &\rho(\mathrm{N}^{C_p}(Y_2)) = f_4. \end{aligned}$$

Hence $\mathbb{F}[V]^G = \mathbb{F}[f_1, f_2, f_3, f_4]$.

Example 14.2.2. Consider the matrix

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

acting on the three dimensional vector space V_3 over $\mathbb{F}$ of characteristic 2.

It is easy to see that σ generates a cyclic group, G, of order 4, and we write $\mathbb{F}[V_3] = \mathbb{F}[x, y, z]$. Then

$$\begin{aligned} \sigma \cdot x &= x, \\ \sigma \cdot y &= y + x, \quad \text{and} \\ \sigma \cdot z &= z + y. \end{aligned}$$

We consider the normal subgroup N of G generated by σ^2. We see immediately that as $\mathbb{F}N$-modules, $V_3 \cong V_2 \oplus V_1$, where V_2 has basis $\{x, z\}$ and V_1 has basis $\{y\}$. Therefore, as $\mathbb{F}N$-modules, we have

$$\mathbb{F}[V_3] = \mathbb{F}[V_2] \otimes \mathbb{F}[V_1] = \mathbb{F}[x, z] \otimes \mathbb{F}[y].$$

Therefore,

$$\mathbb{F}[V_3]^N \cong \mathbb{F}[V_2 \oplus V_1]^N = \mathbb{F}[x, z^2 - xz, y].$$

Thus the space of indecomposables $Q(\mathbb{F}[V_3]^N) = \mathbb{F}[V_3]^N_+/(\mathbb{F}[V_3]^N_+)^2$ is the three dimensional vector space with basis $\{x, y, z^2 + xz\}$. Next, we consider the action of $G/N \cong C_2$ on $Q(\mathbb{F}[V_3]^N)$. Let τ denote a generator of G/N. Then

$$\begin{aligned} \tau \cdot x &= x, \\ \tau \cdot y &= y + x, \qquad \text{and} \\ \tau \cdot (z^2 + xz) &= z^2 + xz + y^2 + xy. \end{aligned}$$

Hence the abstract vector space $\operatorname{span}_{\mathbb{F}}\{x, y, z^2+xz, z^2+xz\}$ is not G/N stable and we must instead consider the abstract 4 dimensional space, $\widetilde{Q}$, with basis

$$\{x, y, z^2 + xz, y^2 + xy\}$$

which is G/N-stable since $\tau \cdot (y^2+xy) = y^2+xy$. We dualize and lift to obtain $W \cong V_2 \oplus V_2$ and we write

$$\mathbb{F}[W] = \mathbb{F}[a_1, b_1, a_2, b_2]$$

where

$$\begin{aligned} \tau \cdot a_i &= a_i, \quad \text{and} \\ \tau \cdot b_i &= b_i + a_i \end{aligned}$$

Define

$$\rho : \mathbb{F}[W] \to \mathbb{F}[V_3]^N$$

by

$$\begin{aligned}\rho(a_1) &= x,\\ \rho(b_1) &= y,\\ \rho(a_2) &= y^2 + xy, \quad \text{and}\\ \rho(b_2) &= z^2 + xz.\end{aligned}$$

Then ρ is a G/N-equivariant surjection.

As usual, the short exact sequence of $A := \mathbb{F}[W]$-modules

$$0 \to J \xrightarrow{\gamma} A \xrightarrow{\rho} \mathbb{F}[V_3]^N \to 0$$

induces a long exact sequence in group cohomology

$$0 \to J^{G/N} \overset{\gamma^{G/N}}{\to} A^{G/N} \overset{\rho^{G/N}}{\to} \mathbb{F}[V_3]^G \overset{\delta}{\to} H^1(G/N; J) \overset{\gamma^1}{\to} H^1(G/N; A) \overset{\rho^1}{\to} \dots .$$

We claim that in this case, the map $\rho^{G/N}$ is surjective. By exactness, we may prove this by showing that

$$\gamma^1 : H^1(G/N; J) \to H^1(G/N; \mathbb{F}[W])$$

is injective.

Since $p = 2$, we have that $\mathrm{Tr}^{G/N} = \tau + \mathrm{Id} = \tau - \mathrm{Id} = \Delta$. Therefore

$$\begin{aligned} H^1(G/N; M) &= \frac{\mathrm{Kernel}(\mathrm{Tr}\,|_M)}{\mathrm{Im}(\,\Delta|_M\,)},\\ &= \frac{\mathrm{Kernel}(\Delta|_M)}{\mathrm{Im}(\,\mathrm{Tr}\,|_M\,)}\end{aligned}$$

As usual, for $f \in \mathbb{F}[W]^{C_p}$ we let $[f]_{\mathbb{F}[W]}$ denote the cohomology class in $H^1(G/N; \mathbb{F}[W])$ represented by f. Similarly if $f \in J^{C_p}$, we denote by $[f]_J$ the cohomology class f represents in $H^1(G/N; J)$. To see that γ^1 is injective, consider an element $f \in J^{C_p}$ such that $\gamma^1([f]_{\mathbb{F}[W]}) = 0$. Thus $f \in \mathrm{Tr}^G_N(\mathbb{F}[W])$. Write $f = \mathrm{Tr}^G_N(f')$ for some $f' \in \mathbb{F}[W]$.

Note that $\mathbb{F}[W]$ has Krull dimension 4 and that $\mathbb{F}[V]^N$ has Krull dimension 3. We note that

$$\rho(a_2) = \rho(b_1^2 + a_1 b_1),$$

and so J is the principal ideal generated by the element $r := a_2 + b_1^2 + a_1 b_1$. Since $f \in J$, we may write $f = f'r$ for some $f' \in \mathbb{F}[W]$. Since $f \in \mathrm{Im}\,\mathrm{Tr}$, it must vanish on

$$\overline{W}^{G/N} = \{(\beta_2, 0, \beta_1, 0) \mid \beta_2, \beta_1 \in \overline{\mathbb{F}}\}.$$

Thus

$$0 = f'(\beta_2, 0, \beta_1, 0) = f'(\beta_2, 0, \beta_1, 0) \cdot r(\beta_2, 0, \beta_1, 0) = f'(\beta_2, 0, \beta_1, 0)\beta_1^2.$$

Therefore f' must vanish on the set

$$\{(\beta_2, 0, \beta_1, 0) \mid \beta_1 \neq 0, \beta_2, \beta_1 \in \overline{\mathbb{F}}\}.$$

Hence, by continuity, f' must vanish on

$$\{(\beta_2, 0, \beta_1, 0) \mid \beta_2, \beta_1 \in \overline{\mathbb{F}}\} = \overline{W}^{G/N}.$$

Therefore $f' \in \sqrt{\operatorname{Im} \operatorname{Tr}_N^G}$. Since W is equivalent to a permutation representation, Proposition 9.0.13 implies that $f' \in \operatorname{Im} \operatorname{Tr}_N^G$. Thus there exists an element $f'' \in \mathbb{F}[W]$ with $\operatorname{Tr}_N^G(f'') = f'$. This shows that $f = f'r = \operatorname{Tr}_N^G(f'')r = \operatorname{Tr}_N^G(f''r) \in \operatorname{Tr}_N^G(J)$. Therefore $[f]_J = 0$. Thus we have shown that γ^1 is injective and so $\rho^{G/N}$ is surjective. Now

$$\mathbb{F}[W]^{G/N} = \mathbb{F}[a_1, b_1, a_2b_1 + a_1b_2, b_1^2 + a_1b_1, b_2^2 + a_2b_2]$$

and

$$\begin{aligned}\mathbb{F}[V_3]^G = \rho(\mathbb{F}[W]^{G/N}) &= \mathbb{F}[f_1, f_2, f_3, f_4] \quad \text{where}\\ f_1 &= x,\\ f_2 &= y^2 + xy,\\ f_3 &= y^3 + xy^2 + xz^2 + x^2z, \quad \text{and}\\ f_4 &= z^4 + x^2z^2 + y^2z^2 + xyz^2 + xy^2z + x^2yz.\end{aligned}$$

In particular, we see that $\mathbb{F}[V_3]^G$ is a hypersurface. Furthermore, it is straightforward to verify that the relation among the generators is given by

$$f_2^3 + f_3^2 + f_1f_2f_3 + f_1^2f_4 = 0\ .$$

Example 14.2.3. Let $\mathbb{F}_4 = \{0, 1, \omega, \omega^2\}$ be the field of order 4 where ω satisfies $\omega^2 + \omega + 1 = 0$. Consider the subgroup G of $\mathrm{GL}(3, \mathbb{F}_4)$ generated by the 2 matrices

$$\sigma^{-1} = \begin{pmatrix}1&0&0\\1&1&0\\0&1&1\end{pmatrix}, \qquad \tau^{-1} = \begin{pmatrix}1&0&0\\0&1&0\\\omega&0&1\end{pmatrix}$$

Let V denote the three dimensional $\mathbb{F}_4$ vector space on which G acts with $\{x, y, z\}$ the basis of V^* dual to the standard basis. Notice that the subgroup N generated by σ is the group we considered in Example 14.2.2. Thus N has order 4 and

$$\mathbb{F}_4[V]^N = \mathbb{F}_4[f_1, f_2, f_3, f_4]$$

where

$$\begin{aligned}f_1 &= x,\\ f_2 &= y^2 + xy,\\ f_3 &= y^3 + xy^2 + xz^2 + x^2z,\\ f_4 &= z^4 + x^2z^2 + y^2z^2 + xyz^2 + xy^2z + x^2yz.\end{aligned}$$

Furthermore, $\tau^2 = I_3$ and thus G has order 8 and N is normal in G.

We will use a ladder to study the ring of G invariants: $\mathbb{F}_4[V]^G = (\mathbb{F}_4[V]^N)^{G/N}$ where $G/N \cong C_2$. Write $\Delta = \tau - 1$. The action of G/N on $\mathbb{F}_4[V]^N$ is given by

$$\begin{aligned}\Delta(f_1) &= 0,\\ \Delta(f_2) &= 0,\\ \Delta(f_3) &= f_1^3, \quad \text{and}\\ \Delta(f_4) &= f_1^2(f_2 + f_1^2)\ .\end{aligned}$$

Thus we define $W = 2\,V_1 \oplus 2\,V_2$ where W^* has basis $\{a_1, a_2, a_3, b_3, a_4, b_4\}$ with

$$\begin{aligned}\tau(a_i) &= a_i \quad \text{for } i = 1,2,3,4 \quad \text{and}\\ \tau(b_i) &= b_i + a_i \quad \text{for} i = 3,4\end{aligned}$$

We define $\rho : A := \mathbb{F}_4[W] \to \mathbb{F}_4[V]^N$ by

$$\begin{aligned}\rho(a_1) &= f_1,\\ \rho(a_2) &= f_2,\\ \rho(b_3) &= f_3,\\ \rho(a_3) &= \Delta(f_3) = f_1^3,\\ \rho(b_4) &= f_4 \quad \text{and}\\ \rho(a_4) &= \Delta(f_4) = f_1^4 + f_1^2 f_2.\end{aligned}$$

Thus ρ is a G/N-equivariant surjection.

Again the short exact sequence of A-modules

$$0 \to J \xrightarrow{\gamma} A \xrightarrow{\rho} \mathbb{F}_4[V]^N \to 0$$

induces a long exact sequence in group cohomology

$$0 \to J^{G/N} \overset{\gamma^{G/N}}{\to} A^{G/N} \overset{\rho^{G/N}}{\to} \mathbb{F}_4[V]^G \xrightarrow{\delta} H^1(G/N; J) \xrightarrow{\gamma^1} H^1(G/N; A) \xrightarrow{\rho^1} \dots .$$

In this case, we will see that $\rho^{G/N}$ is not surjective. We define

$$f_5 := y^5 + xz^4 + x^3y^2 + x^4z.$$

Since $\Delta(f_5) = x(\omega x)^4 + x^4(\omega x) = 0$, we see that f_5 is G-invariant.

It is easy to see that f_5 does not lie in the image of $\rho^{G/N}$ as follows. We have

$$\begin{aligned}\rho^{G/N}(\mathbb{F}_4[W]^{G/N}) &= \rho^{G/N}(\mathbb{F}_4[a_1, a_2, b_3^2 + a_3b_3, b_4^2 + a_4b_4, a_3b_4 + a_4b_3])\\ &= \mathbb{F}_4[f_1, f_2, f_3^2 + f_1^3 f_3, f_4^2 + f_1^4 f_4 + f_1^2 f_2 f_4, f_1^3 f_4 + f_1^4 f_3 + f_1^2 f_2 f_3].\end{aligned}$$

Note that f_i has degree i for all $i = 1, 2, \ldots, 5$ and thus degree considerations show that if $f_5 \in \rho^{G/N}(\mathbb{F}_4[W]^{G/N})$, then $f_5 \in \mathbb{F}_4[f_1, f_2]_5$. Since every element of $\mathbb{F}_4[f_1, f_2]_5$ is divisible by $f_1 = x$ and the lead monomial of f_5 is y^5 we see that $f_5 \notin \mathbb{F}_4[f_1, f_2]_5$.

The ideal J is generated by the 3 relations

$$\begin{aligned}&a_2^3 + b_3^2 + a_1a_2b_3 + a_1^2b_4\\ &a_3 + a_1^3 \quad \text{and}\\ &a_4 + a_1^2a_2 + a_1^4.\end{aligned}$$

Consider

$$f = (a_2 + a_1^2)(a_3 + a_1^3) + a_1(a_4 + a_1^2a_2 + a_1^4) = a_2a_3 + a_1^2a_3 + a_1a_4 \in J.$$

Note that $f = \Delta(f') \in \Delta(\mathbb{F}_4[W])$ where $f' = a_2b_3 + a_1^2b_3 + a_1b_4$. However, $f \notin \Delta(J)$. Therefore f represents a non-zero cohomology class $[f]_J \in H^1(G/N; J)$ with $\gamma^1([h]_J) = 0 \in H^1(G/N; \mathbb{F}_4[W])$. This corresponds to the invariant

$$\rho(g) = f_2f_3 + f_1^2f_3 + f_1f_4 = f_5 \in \mathbb{F}_4[W]^G.$$

With a little more work the reader may show that the $\mathbb{F}_4[W]^{G/N}$-module, $\ker\gamma^1$ is generated by the element just considered $[a_2a_3 + a_1^2a_3 + a_1a_4]_J$. Thus $\mathbb{F}_4[V]^G$ is generated by $\rho(\mathbb{F}_4[W]^H)$ together with f_5. From this it easily follows that

$$\mathbb{F}_4[V]^G = \mathbb{F}_4[f_1, f_2, f_5, f_8]$$

where

$$\begin{aligned}f_8 &= \rho(\mathrm{N}^{G/H}(b_4))\\ &= z^8 + y^4z^4 + xy^7 + x^2y^2z^4 + x^3y^4z + x^3y^5x^5y^2z + x^6z^2.\end{aligned}$$

Remark 14.2.4. Note that the representation considered in Example 14.2.3 is the same representation as that considered in Example 14.2.1 when $p = 2$. However, the choice of ladder (i.e., the tower of normal subgroups used) is different in the two examples.

References

1. William W. Adams and Philippe Loustaunau, *An introduction to Gröbner bases*, Graduate Studies in Mathematics, vol. 3, American Mathematical Society, Providence, RI, 1994. MR 1287608 (95g:13025)
2. Gert Almkvist, *Representations of* $\mathbf{Z}/p\mathbf{Z}$ *in characteristic* p *and reciprocity theorems*, J. Algebra **68** (1981), no. 1, 1–27. MR 604290 (82k:14047)
3. Gert Almkvist and Robert Fossum, *Decomposition of exterior and symmetric powers of indecomposable* $\mathbf{Z}/p\mathbf{Z}$*-modules in characteristic* p *and relations to invariants*, Séminaire d'Algèbre Paul Dubreil, 30ème année (Paris, 1976–1977), Lecture Notes in Math., vol. 641, Springer, Berlin, 1978, pp. 1–111. MR 499459 (81b:14024)
4. Emil Artin, *Galois theory*, second ed., Dover Publications Inc., Mineola, NY, 1998, Edited and with a supplemental chapter by Arthur N. Milgram. MR 1616156 (98k:12001)
5. M. F. Atiyah and I. G. Macdonald, *Introduction to commutative algebra*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1969. MR 39:4129
6. D. J. Benson, *Polynomial invariants of finite groups*, London Mathematical Society Lecture Note Series, vol. 190, Cambridge University Press, Cambridge, 1993. MR 94j:13003
7. ———, *Representations and cohomology. I*, second ed., Cambridge Studies in Advanced Mathematics, vol. 30, Cambridge University Press, Cambridge, 1998, Cohomology of groups and modules. MR 1634407 (99f:20001b)
8. ———, *Representations and cohomology. II*, second ed., Cambridge Studies in Advanced Mathematics, vol. 31, Cambridge University Press, Cambridge, 1998, Cohomology of groups and modules. MR 1634407 (99f:20001b)
9. Marie-José Bertin, *Anneaux d'invariants d'anneaux de polynomes, en caractéristique* p, C. R. Acad. Sci. Paris Sér. A-B **264** (1967), A653–A656. MR 0215826 (35 #6661)
10. W. Bosma, J. Cannon, and C. Playous, *The magma algebra system. i. the user language*, J. Symbolic Comput. **24** (1997), no. 3–4, 235–265.
11. N. Bourbaki, *Éléments de mathématique. I: Les structures fondamentales de l'analyse. Fascicule XI. Livre II: Algèbre. Chapitre 4: Polynomes et fractions rationnelles. Chapitre 5: Corps commutatifs*, Deuxième édition.

H.E.A. Eddy Campbell, D. Wehlau, *Modular Invariant Theory*,
Encyclopaedia of Mathematical Sciences 139,
DOI 10.1007/978-3-642-17404-9, © Springer-Verlag Berlin Heidelberg 2011

Actualités Scientifiques et Industrielles, No. 1102, Hermann, Paris, 1959. MR 0174550 (30 #4751)

12. Abraham Broer, *Remarks on invariant theory of finite groups*, Unpublished manuscript, 1997.
13. ———, *Differents in modular invariant theory*, Transform. Groups **11** (2006), no. 4, 551–574. MR 2278139 (2007h:13007)
14. Winfried Bruns and Jürgen Herzog, *Cohen-Macaulay rings*, Cambridge Studies in Advanced Mathematics, vol. 39, Cambridge University Press, Cambridge, 1993. MR 1251956 (95h:13020)
15. Roger M. Bryant and Gregor Kemper, *Global degree bounds and the transfer principle for invariants*, J. Algebra **284** (2005), no. 1, 80–90. MR 2115005 (2005i:13007)
16. H. E. A. Campbell and Jianjun Chuai, *On the invariant fields and localized invariant rings of p-groups*, Quarterly Journal of Mathematics **10** (2007), no. 10.1093/qmath/ham011, 1–7.
17. H. E. A. Campbell, A. V. Geramita, I. P. Hughes, R. J. Shank, and D. L. Wehlau, *Non-Cohen-Macaulay vector invariants and a Noether bound for a Gorenstein ring of invariants*, Canad. Math. Bull. **42** (1999), no. 2, 155–161. MR 2000b:13005
18. H. E. A. Campbell, I. Hughes, and R. D. Pollack, *Rings of invariants and p-Sylow subgroups*, Canad. Math. Bull. **34** (1991), no. 1, 42–47. MR 92h:13008
19. H. E. A. Campbell and I. P. Hughes, *Vector invariants of $U_2(\mathbf{F}_p)$: a proof of a conjecture of Richman*, Adv. Math. **126** (1997), no. 1, 1–20. MR 98c:13007
20. H. E. A. Campbell and P. S. Selick, *Polynomial algebras over the Steenrod algebra*, Comment. Math. Helv. **65** (1990), no. 2, 171–180. MR 1057238 (91f:55006)
21. H. E. A. Campbell, R. J. Shank, and David L. Wehlau, *Vector invariants for the two dimensional modular representation of a cyclic group of prime order*, Advances in Math. **225** (2010), no. 2, 1069–1094.
22. Claude Chevalley, *Invariants of finite groups generated by reflections*, Amer. J. Math. **77** (1955), 778–782. MR 17,345d
23. Jianjun Chuai, *Two-dimensional vector invariant rings of abelian p-groups*, J. Algebra **266** (2003), no. 1, 362–373. MR 2004e:13010
24. H. S. M. Coxeter, *The product of the generators of a finite group generated by reflections*, Duke Math. J. **18** (1951), 765–782. MR 0045109 (13,528d)
25. Charles W. Curtis and Irving Reiner, *Representation theory of finite groups and associative algebras*, Wiley Classics Library, John Wiley & Sons Inc., New York, 1988, Reprint of the 1962 original, A Wiley-Interscience Publication. MR 1013113 (90g:16001)
26. Harm Derksen and Gregor Kemper, *Computational invariant theory*, Invariant Theory and Algebraic Transformation Groups, I, Springer-Verlag, Berlin, 2002, Encyclopaedia of Mathematical Sciences, 130. MR 2003g:13004
27. ———, *Computing invariants of algebraic groups in arbitrary characteristic*, Adv. Math. **217** (2008), no. 5, 2089–2129. MR 2388087 (2009a:13005)
28. L. E. J. Dickson, *Theorems on the residues of multinomial coefficients with respect to a prime modulus*, Q. J. Pure Appl. Math. **33** (1902), 378–384.
29. Jan Draisma, Gregor Kemper, and David Wehlau, *Polarization of separating invariants*, Canad. J. Math. **60** (2008), no. 3, 556–571. MR 2414957 (2009c:13011)
30. Emilie Dufresne, *Separating invariants*, Ph.D. thesis, Queen's University, Kingston, Ontario, Canada, 2008, defended.

31. ______, *Separating invariants of finite reflections groups*, Advances in Math. **221** (2009), no. 6, 1979–1989.
32. David S. Dummit and Richard M. Foote, *Abstract algebra*, third ed., John Wiley & Sons Inc., Hoboken, NJ, 2004. MR 2286236 (2007h:00003)
33. David Eisenbud, *Commutative algebra*, Graduate Texts in Mathematics, vol. 150, Springer-Verlag, New York, 1995. MR 97a:13001
34. ______, *Commutative algebra*, Graduate Texts in Mathematics, vol. 150, Springer-Verlag, New York, 1995, With a view toward algebraic geometry. MR 1322960 (97a:13001)
35. Geir Ellingsrud and Tor Skjelbred, *Profondeur d'anneaux d'invariants en caractéristique p*, Compositio Math. **41** (1980), no. 2, 233–244. MR 82c:13015
36. Jonathan Elmer and Peter Fleischmann, *On the depth of modular invariant rings for the groups $C_p \times C_p$*, Progress in Mathematics **278** (2009), 45–62.
37. Charles S. Fisher, *The death of a mathematical theory: a study in the sociology of knowledge*, Arch. History Exact Sci. **3** (1966), no. 2, 137–159. MR 0202546
38. Peter Fleischmann, *Relative trace ideals and Cohen-Macaulay quotients of modular invariant rings*, Computational methods for representations of groups and algebras (Essen, 1997), Progr. Math., vol. 173, Birkhäuser, Basel, 1999, pp. 211–233. MR 1714612 (2000j:13007)
39. ______, *The Noether bound in invariant theory of finite groups*, Adv. Math. **156** (2000), no. 1, 23–32. MR 1800251 (2001k:13009)
40. Peter Fleischmann, Gregor Kemper, and Chris Woodcock, *Homomorphisms, localizations and a new algorithm to construct invariant rings of finite groups*, J. Algebra **309** (2007), no. 2, 497–517. MR 2303190
41. Peter Fleischmann and R. James Shank, *The relative trace ideal and the depth of modular rings of invariants*, Arch. Math. (Basel) **80** (2003), no. 4, 347–353. MR 2004e:13012
42. John Fogarty, *On Noether's bound for polynomial invariants of a finite group*, Electron. Res. Announc. Amer. Math. Soc. **7** (2001), 5–7 (electronic). MR 1826990 (2002a:13002)
43. Robert M. Fossum and Phillip A. Griffith, *Complete local factorial rings which are not Cohen-Macaulay in characteristic p*, Ann. Sci. École Norm. Sup. (4) **8** (1975), no. 2, 189–199. MR 0382257 (52 #3142)
44. Manfred Göbel, *Computing bases for rings of permutation-invariant polynomials*, J. Symbolic Comput. **19** (1995), no. 4, 285–291. MR 96f:13006
45. ______, *A constructive description of SAGBI bases for polynomial invariants of permutation groups*, J. Symbolic Comput. **26** (1998), no. 3, 261–272. MR 1633927 (99f:13002)
46. ______, *The optimal lower bound for generators of invariant rings without finite SAGBI bases with respect to any admissible order*, Discrete Math. Theor. Comput. Sci. **3** (1999), no. 2, 65–70 (electronic). MR 1695195 (2001b:13008)
47. ______, *Rings of polynomial invariants of the alternating group have no finite SAGBI bases with respect to any admissible order*, Inform. Process. Lett. **74** (2000), no. 1–2, 15–18. MR 1761193 (2001d:13003)
48. P. Gordan, *Beweis dass jede covariante und invariante einer bindren form eine ganze function mit numerischen coefficienten solcher formen ist*, J. fur reine u. angew. Math. **69** (1868), 323–354.
49. Paul Gordan, *Vorlesungen über Invariantentheorie*, second ed., Chelsea Publishing Co., New York, 1987, Erster Band: Determinanten. [Vol. I: Determi-

nants], Zweiter Band: Binäre Formen. [Vol. II: Binary forms], Edited by Georg Kerschensteiner. MR 917266 (89g:01034)

50. F. D. Grosshans, *Vector invariants in arbitrary characteristic*, Transform. Groups **12** (2007), no. 3, 499–514. MR 2356320
51. Robin Hartshorne, *Complete intersections and connectedness*, Amer. J. Math. **84** (1962), 497–508. MR 0142547 (26 #116)
52. David Hilbert, *Ueber die Theorie der algebraischen Formen*, Math. Ann. **36** (1890), no. 4, 473–534. MR 1510634
53. M. Hochster and John A. Eagon, *Cohen-Macaulay rings, invariant theory, and the generic perfection of determinantal loci*, Amer. J. Math. **93** (1971), 1020–1058. MR 46:1787
54. Ian Hughes and Gregor Kemper, *Symmetric powers of modular representations, Hilbert series and degree bounds*, Comm. Algebra **28** (2000), no. 4, 2059–2088. MR 2001b:13009
55. ———, *Symmetric powers of modular representations for groups with a Sylow subgroup of prime order*, J. Algebra **241** (2001), no. 2, 759–788. MR 1843324 (2002e:13012)
56. James E. Humphreys, *Linear algebraic groups*, Springer-Verlag, New York, 1975, Graduate Texts in Mathematics, No. 21. MR 0396773 (53 #633)
57. Victor Kac and Keiichi Watanabe, *Finite linear groups whose ring of invariants is a complete intersection*, Bull. Amer. Math. Soc. (N.S.) **6** (1982), no. 2, 221–223. MR 640951 (83h:14042)
58. William M. Kantor, *Subgroups of classical groups generated by long root elements*, Trans. Amer. Math. Soc. **248** (1979), no. 2, 347–379. MR 522265 (80g:20057)
59. Irving Kaplansky, *Commutative rings*, revised ed., The University of Chicago Press, Chicago, Ill.-London, 1974. MR 0345945 (49 #10674)
60. Deepak Kapur and Klaus Madlener, *A completion procedure for computing a canonical basis for a k-subalgebra*, Computers and mathematics (Cambridge, MA, 1989), Springer, New York, 1989, pp. 1–11. MR 1005954 (90g:13001)
61. D. B. Karagueuzian and P. Symonds, *The module structure of a group action on a polynomial ring: examples, generalizations, and applications*, Invariant theory in all characteristics, CRM Proc. Lecture Notes, vol. 35, Amer. Math. Soc., Providence, RI, 2004, pp. 139–158. MR 2066462 (2005g:13011)
62. Dikran B. Karagueuzian and Peter Symonds, *The module structure of a group action on a polynomial ring: a finiteness theorem*, J. Amer. Math. Soc. **20** (2007), no. 4, 931–967 (electronic). MR 2328711
63. G. Kemper and G. Malle, *The finite irreducible linear groups with polynomial ring of invariants*, Transform. Groups **2** (1997), no. 1, 57–89. MR 98a:13012
64. Gregor Kemper, *A constructive approach to Noether's problem*, Manuscripta Math. **90** (1996), no. 3, 343–363. MR 1397662 (97d:13005)
65. ———, *On the Cohen-Macaulay property of modular invariant rings*, J. Algebra **215** (1999), no. 1, 330–351. MR 2000d:13008
66. ———, *Computing invariants of reductive groups in positive characteristic*, Transform. Groups **8** (2003), no. 2, 159–176. MR 2004b:13006
67. ———, *A course in commutative algebra*, first ed., Graduate Texts in Mathematics, vol. 256, Springer, to appear, 2010.
68. Gregor Kemper and Gunter Malle, *Invariant fields of finite irreducible reflection groups*, Math. Ann. **315** (1999), no. 4, 569–586. MR MR1731462 (2001c:13006)

69. ———, *Invariant fields of finite irreducible reflection groups*, Math. Ann. **315** (1999), no. 4, 569–586. MR 2001c:13006

70. Hanspeter Kraft and Claudio Procesi, *A primer of invariant theory, notes by G. Boffi*, http://www.math.unibas.ch/~kraft/Papers/KP-Primer.pdf **1** (1982, rev. 2000), 125.

71. Ernst Kunz, *Introduction to commutative algebra and algebraic geometry*, Birkhäuser Boston Inc., Boston, MA, 1985, Translated from the German by Michael Ackerman, With a preface by David Mumford. MR 789602 (86e:14001)

72. Shigeru Kuroda, *The infiniteness of the SAGBI bases for certain invariant rings*, Osaka J. Math. **39** (2002), no. 3, 665–680. MR 1932287 (2003k:13033)

73. Peter S. Landweber and Robert E. Stong, *The depth of rings of invariants over finite fields*, Number theory (New York, 1984–1985), Lecture Notes in Mathematics, vol. 1240, Springer, Berlin, 1987, pp. 259–274. MR 88k:13004

74. Serge Lang, *Algebra*, second ed., Addison-Wesley Publishing Company Advanced Book Program, Reading, MA, 1984. MR 783636 (86j:00003)

75. ———, *Algebra*, third ed., Addison-Wesley Publishing Company Advanced Book Program, Reading, MA, 1993.

76. Peter Littelmann, *Koreguläre und äquidimensionale Darstellungen*, J. Algebra **123** (1989), no. 1, 193–222. MR 1000484 (90e:20039)

77. Martin Lorenz, *Multiplicative invariant theory*, Encyclopaedia of Mathematical Sciences, vol. 135, Springer-Verlag, Berlin, 2005, Invariant Theory and Algebraic Transformation Groups, VI. MR 2131760 (2005m:13012)

78. I. G. Macdonald, *Symmetric functions and Hall polynomials*, second ed., Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 1995, With contributions by A. Zelevinsky, Oxford Science Publications. MR 1354144 (96h:05207)

79. Hideyuki Matsumura, *Commutative algebra*, second ed., Mathematics Lecture Note Series, vol. 56, Benjamin/Cummings Publishing Co., Inc., Reading, Mass., 1980. MR 575344 (82i:13003)

80. ———, *Commutative ring theory*, Cambridge Studies in Advanced Mathematics, vol. 8, Cambridge University Press, Cambridge, 1986. MR 88h:13001

81. Takehiko Miyata, *Invariants of certain groups. I*, Nagoya Math. J. **41** (1971), 69–73. MR 0272923 (42 #7804)

82. H. Mui, *Modular invariant theory and cohomology algebras of symmetric groups*, J. of Fac. Sci. Univ. Tokyo Sect. 1A Math. **22** (1975), 319–369.

83. Haruhisa Nakajima, *Modular representations of p-groups with regular rings of invariants*, Proc. Japan Acad. Ser. A Math. Sci. **56** (1980), no. 10, 469–473. MR 82a:20016

84. ———, *Rings of invariants of finite groups which are hypersurfaces*, J. Algebra **80** (1983), no. 2, 279–294. MR 85e:20036

85. Mara D. Neusel, *Invariant theory*, Student Mathematical Library, vol. 36, American Mathematical Society, Providence, RI, 2007. MR 2280491 (2007m:13007)

86. Mara D. Neusel and Larry Smith, *Invariant theory of finite groups*, Mathematical Surveys and Monographs, vol. 94, American Mathematical Society, Providence, RI, 2002. MR 1869812 (2002k:13012)

87. P. E. Newstead, *Introduction to moduli problems and orbit spaces*, Tata Institute of Fundamental Research Lectures on Mathematics and

Physics, vol. 51, Tata Institute of Fundamental Research, Bombay, 1978. MR 546290 (81k:14002)

88. E. Noether, *Der endlichkeitssatz der invarianten endlicher linearer gruppen der characteristik p*, Nachr. v. d. Ges. d. Wiss. zu Göttingen (1926), 28–35.
89. V. L. Popov, *Constructive invariant theory*, Young tableaux and Schur functors in algebra and geometry (Toruń, 1980), Astérisque, vol. 87, Soc. Math. France, Paris, 1981, pp. 303–334. MR 646826 (83i:14040)
90. Claudio Procesi, *Lie groups*, Universitext, Springer, New York, 2007, An approach through invariants and representations. MR 2265844 (2007j:22016)
91. Zinovy Reichstein, *SAGBI bases in rings of multiplicative invariants*, Comment. Math. Helv. **78** (2003), no. 1, 185–202. MR 1966757 (2004c:13005)
92. David R. Richman, *On vector invariants over finite fields*, Adv. Math. **81** (1990), no. 1, 30–65. MR 91g:15020
93. Lorenzo Robbiano and Moss Sweedler, *Subalgebra bases*, Commutative algebra (Salvador, 1988), Lecture Notes in Math., vol. 1430, Springer, Berlin, 1990, pp. 61–87. MR 1068324 (91f:13027)
94. Michael Roberts, *On the covariants of a binary quantic of the nth degree*, Quarterly Journal of Pure and Applied Mathematics **4** (1861), 168–178.
95. Jean-Pierre Serre, *Groupes finis d'automorphismes d'anneaux locaux réguliers*, Colloque d'Algèbre (Paris, 1967), Exp. 8, Secrétariat mathématique, Paris, 1968, p. 11. MR 38:3267
96. Müfit Sezer and R. James Shank, *On the coinvariants of modular representations of cyclic groups of prime order*, J. Pure Appl. Algebra **205** (2006), no. 1, 210–225. MR 2193198 (2006k:13015)
97. R. J. Shank, *Classical covariants and modular invariants*, Invariant theory in all characteristics, CRM Proc. Lecture Notes, vol. 35, Amer. Math. Soc., Providence, RI, 2004, pp. 241–249. MR 2005d:13012
98. R. James Shank, *S.A.G.B.I. bases for rings of formal modular seminvariants*, Comment. Math. Helv. **73** (1998), no. 4, 548–565. MR 2000a:13016
99. R. James Shank and David L. Wehlau, *Computing modular invariants of p-groups*, J. Symbolic Comput. **34** (2002), no. 5, 307–327. MR 2003j:13006
100. ______, *Noether numbers for subrepresentations of cyclic groups of prime order*, Bull. London Math. Soc. **34** (2002), no. 4, 438–450. MR 2003a:13005
101. G. C. Shephard and J. A. Todd, *Finite unitary reflection groups*, Canadian J. Math. **6** (1954), 274–304. MR 15,600b
102. Larry Smith, *Polynomial invariants of finite groups*, Research Notes in Mathematics, vol. 6, A K Peters Ltd., Wellesley, MA, 1995. MR 1328644 (96f:13008)
103. ______, *Polynomial invariants of finite groups. A survey of recent developments*, Bull. Amer. Math. Soc. (N.S.) **34** (1997), no. 3, 211–250. MR 98i:13009
104. Richard P. Stanley, *Invariants of finite groups and their applications to combinatorics*, Bull. Amer. Math. Soc. (N.S.) **1** (1979), no. 3, 475–511. MR 81a:20015
105. Bernd Sturmfels, *Algorithms in invariant theory*, Texts and Monographs in Symbolic Computation, Springer-Verlag, Vienna, 1993. MR 94m:13004
106. Peter Symonds, *On the castelnuovo-mumford regularity of rings of polynomial invariants*, preprint (2009), 1–11.
107. N. M. Thiéry and S. Thomassé, *Convex cones and SAGBI bases of permutation invariants*, Invariant theory in all characteristics, CRM Proc. Lecture Notes, vol. 35, Amer. Math. Soc., Providence, RI, 2004, pp. 259–263. MR 2066473 (2005e:13006)

108. A. Wagner, *Collineation groups generated by homologies of order greater than* 2, Geom. Dedicata **7** (1978), no. 4, 387–398. MR 512113 (81e:20055)
109. Ascher Wagner, *Determination of the finite primitive reflection groups over an arbitrary field of characteristic not* 2. *I*, Geom. Dedicata **9** (1980), no. 2, 239–253. MR 578199 (81g:20096)
110. David Wehlau, *Equidimensional representations of* 2*-simple groups*, J. Algebra **154** (1993), no. 2, 437–489. MR 1206131 (93k:14064)
111. David L. Wehlau, *The Noether number in invariant theory*, C. R. Math. Acad. Sci. Soc. R. Can. **28** (2006), no. 2, 39–62. MR 2257602 (2007h:13008)
112. Hermann Weyl, *The classical groups*, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, NJ, 1997. MR 98k:01049
113. Clarence Wilkerson, *A primer on the Dickson invariants*, Proceedings of the Northwestern Homotopy Theory Conference (Evanston, Ill., 1982) (Providence, RI), Contemp. Math., vol. 19, Amer. Math. Soc., 1983, pp. 421–434. MR 85c:55017
114. Yinglin Wu, *Rings of invariants of certain modular groups*, Ph.D. thesis, Queen's University, Kingston, Ontario, Canada, 2009, defended.
115. A. E. Zalesskiĭ and V. N. Serežkin, *Finite linear groups generated by reflections*, Izv. Akad. Nauk SSSR Ser. Mat. **44** (1980), no. 6, 1279–1307, 38. MR 603578 (82i:20060)
116. Oscar Zariski and Pierre Samuel, *Commutative algebra. Vol. 1*, Springer-Verlag, New York, 1975. MR 52:5641
117. ______, *Commutative algebra. Vol. II*, Springer-Verlag, New York, 1975. MR 52:10706

Index

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