Polarization of Separating Invariants

Jan Draisma[*], Gregor Kemper, and David Wehlau[†]

August 11, 2005

Abstract

We prove a characteristic free version of Weyl's theorem on polarization. Our result is an exact analogue of Weyl's theorem, the difference being that our statement is about separating invariants rather than generating invariants. For the special case of finite group actions we introduce the concept of *cheap polarization*, and show that it is enough to take cheap polarizations of invariants of just one copy of a representation to obtain separating vector invariants for any number of copies. This leads to upper bounds on the number and degrees of separating vector invariants of finite groups.

Introduction

We begin with a description of the standard invariant theory setting and recall the concepts of separating invariants and of polarization. Let K be any field and let V be a finite-dimensional vector space over K. We write $K[V]$ for the symmetric algebra of the dual space, V^*. If $\{x_1, \ldots, x_k\}$ is a basis of V^*, then $K[V]$ is the polynomial ring in the indeterminates $x_1, \ldots, x_k$.

Now suppose that G is any group acting linearly on V. Then there is a natural action of G on V^* which induces an action of G on $K[V]$. The **ring of invariants** is the subring $K[V]^G$ of $K[V]$ consisting of those polynomials fixed pointwise by all of G:

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

The main problem in invariant theory is to find a set of invariants $S \subset K[V]^G$ which generates $K[V]^G$ as a K-algebra. Such a set S is called a **generating** set.

Since generating sets are often very complicated, and in some cases no finite generating sets exist, the concept of a **separating** set of $K[V]^G$ has emerged as a useful weakening of a generating set. Loosely speaking, a separating set is a set of invariants that has the same capabilities of separating G-orbits as all the

[*]Supported by the Swiss National Science Foundation.
[†]Partially supported by NSERC and ARP.
2000 Mathematics Subject Classification: 13A50, 14L24

invariants from $K[V]^G$. More precisely, if for two points $x, x' \in V$ there exists an invariant in $K[V]^G$ taking different values on x and x', there should also exist an invariant from the separating set with this property. For more background on separating invariants we direct readers to Derksen and Kemper [2, Section 2.3.2] and to [3].

Polarization is an important classical technique used to describe invariants of certain representations. Before giving the general definition we illustrate the idea in a simpler setting. With V and G as above, take $a, b \in K$ arbitrary and consider the G-equivariant surjection

$$\begin{aligned} \varphi_{a,b}\colon V \oplus V &\to V \\ (u, v) &\mapsto au + bv, \end{aligned}$$

where G acts diagonally on $V \oplus V$. On the level of rings the map $\varphi_{a,b}$ induces a ring homomorphism $\Phi_{a,b}\colon K[V] \to K[V \oplus V]$ given by $(\Phi_{a,b}(f))(u, v) = f(au + bv)$. Since $\Phi_{a,b}$ is G-equivariant, it carries invariants to invariants: $\Phi_{a,b}\colon K[V]^G \to K[V \oplus V]^G$.

If we treat a and b as new indeterminates, rather than as elements of K, we obtain a ring homomorphism $\Phi\colon K[V] \to K[V \oplus V][a, b]$ and a corresponding homomorphism $\Phi\colon K[V]^G \to K[V \oplus V]^G[a, b]$ where G fixes a and b. Thus if $f \in K[V]^G$ then $\Phi(f) = \sum_i \sum_j f_{i,j} a^i b^j$. The coefficients, $f_{i,j}$, are invariants called the polarizations of f and we write $\mathrm{Pol}_1^2(f) := \{f_{i,j}\} \subset K[V \oplus V]^G$.

We will give a more general and formal definition of polarization at the end of the introduction and state Weyl's polarization theorem now. Note that in the standard situation one has $W = \{0\}$.

Theorem 0.1 (Weyl [8, II.5, Theorem 2.5A])**.** *Let G be a group acting linearly on two finite-dimensional vector spaces V and W over a field K of characteristic zero. Let n and m be positive integers such that $m \geq \min\{\dim(V), n\}$. If $S \subseteq K[V^m \oplus W]^G$ is a generating set of invariants, then $\mathrm{Pol}_m^n(S) \subseteq K[V^n \oplus W]^G$ is also generating.*

A proof of Theorem 0.1 can also be found in Kraft and Procesi [5, § 7.1].

The following examples show that the hypothesis that K has characteristic zero is necessary in Weyl's Theorem.

Example 0.2. (a) Let K be a field of characteristic 3 containing a primitive 4th root of unity ω. The invariant ring of the group $G \subset \mathrm{GL}_1(K)$ generated by ω is

$$K[V]^G = K[x]^G = K[x^4],$$

and the vector invariants of two copies are

$$K[V^2]^G = K[x, y]^G = K[x^4, x^3y, x^2y^2, xy^3, y^4],$$

where the given generating set is minimal. However,

$$\Phi(x^4) = (ax + by)^4 = x^4 \cdot a^4 + x^3y \cdot a^3b + xy^3 \cdot ab^3 + y^4 \cdot b^4,$$

so $\mathrm{Pol}_1^2(x^4) = \{x^4, x^3y, xy^3, y^4\}$. Thus polarization misses the necessary generator x^2y^2.

Apart from demonstrating the necessity of the hypothesis on the characteristic in Theorem 0.1, this example also shows that the paradigm that "any theorem in invariant theory of finite groups that holds in characteristic zero also holds in the case where $\mathrm{char}(K)$ does not divide the group order" does not carry over directly to Weyl's theorem. It should be mentioned, though, that a weaker form of Weyl's theorem does hold in this case; see Knop [4].

What we also see in this example is that the polarizations form a separating set. Indeed, we have

$$x^2y^2 = \frac{(x^3y)^2}{x^4},$$

so for any point in V^2 where x^4 takes a non-zero value, the value of x^2y^2 can be reconstructed from the values of x^3y and x^4; and for a point where x^4 vanishes, x^2y^2 also vanishes.

(b) Consider the two-dimensional indecomposable representation V of the cyclic group, G, of order p over a field K of characteristic p. It is known that $K[V^2]^G$ has five generators of degrees 1,1,2,p and p. Furthermore, Richman [6] showed that for $n \geq 3$, the ring $K[V^n]^G$ requires a generator, h, of degree $n(p-1)$. Since polarization preserves degree (see below), we see that h cannot be obtained from polarizations of the generators of $K[V^2]^G$. ◁

There are several known results which show that positive characteristic anomalies of invariant theory tend to disappear when the focus is shifted from generating to separating invariants (see [2, Section 2.3.2]). It is therefore natural to ask whether Theorem 0.1 holds in arbitrary characteristic if one replaces every instance of the word "generating" by "separating". In this paper we give an affirmative answer to this question.

In the first section we deal with the case where G may be infinite. In fact, we start by considering a more general setting which does not necessarily involve invariant theory. The key result is contained in Lemma 1.1, which leads to our characteristic free version of Weyl's theorem (Theorem 1.4 and Corollary 1.5). We find it remarkable that although the statements of Theorems 0.1 and 1.4 are in perfect analogy, the proofs are altogether different.

Section 2 deals with the case of finite groups. We introduce the concept of cheap polarization and prove that for G finite, the cheap polarizations of a separating set S of invariants in $K[V]^G$ yield a separating set of invariants in $K[V^n]^G$ for every n (see Theorem 2.4). In particular, $\mathrm{Pol}_1^n(S) \subseteq K[V^n]^G$ is separating (see Corollary 2.7). This result has no parallel in terms of generating invariants, even in characteristic zero. We conclude the paper by giving upper bounds on the degrees and number of separating invariants. In particular, for G finite, we obtain a bound on the number of separating invariants in $K[V^n]^G$

which is linear in n (see Corollary 2.12). We also show that no such bound can exist for generating invariants (see Theorem 2.13), again underscoring the benefit reaped from shifting focus from generating to separating invariants.

We finish the introduction by giving the general definition of polarization. Let V and W be finite-dimensional vector spaces over any field K, and write V^m for the direct sum of m copies of V. We write $K[V^m \oplus W]$ for the symmetric algebra of the dual $(V^m \oplus W)^*$. If $\{x_1, \ldots, x_k\}$ is a basis of V^* and $\{y_1, \ldots, y_l\}$ is a basis of W^*, then we obtain a basis $\{x_{i,\nu} \mid i = 1, \ldots, m, \nu = 1, \ldots, k\} \cup \{y_1, \ldots, y_l\}$ of $(V^m \oplus W)^*$ in the obvious way by defining $x_{i,\nu}(v_1, \ldots, v_m, w) := x_\nu(v_i)$ and $y_i(v_1, \ldots, v_m, w) := y_i(w)$ for $v_1, \ldots, v_m \in V$, $w \in W$. Then $K[V^m \oplus W]$ is a polynomial ring in the indeterminates $x_{i,\nu}$ $(i = 1, \ldots, m,\ \nu = 1, \ldots, k)$ and y_i $(i = 1, \ldots, l)$. Let n be a further positive integer, and for $i = 1, \ldots, m$ and $j = 1, \ldots, n$ let $a_{i,j}$ be an indeterminate. Form a homomorphism $\Phi\colon K[V^m \oplus W] \to K[V^n \oplus W][a_{1,1}, \ldots, a_{m,n}]$ of K-algebras by

$$\begin{array}{lcll} \Phi(x_{i,\nu}) & := & \sum_{j=1}^n a_{i,j} x_{j,\nu} & (i = 1, \ldots, m,\ \nu = 1, \ldots, k) \quad \text{and} \\ \Phi(y_i) & := & y_i & (i = 1, \ldots, l). \end{array} \tag{0.1}$$

So, pretending for a moment that the $a_{i,j}$ are elements of the field K, we obtain for every $f \in K[V^m \oplus W]$ and $v_1, \ldots, v_n \in V$, $w \in W$ that

$$(\Phi(f))(v_1, \ldots, v_n, w) = f\left(\sum_{j=1}^n a_{1,j} v_j, \sum_{j=1}^n a_{2,j} v_j, \ldots, \sum_{j=1}^n a_{m,j} v_j, w\right),$$

which connects the definition of Φ with what we said about the simpler situation above. Now we take the $a_{i,j}$ again for what they really are. For $f \in K[V^m \oplus W]$ let $\mathrm{Pol}_m^n(f) \subseteq K[V^n \oplus W]$ denote the set of all non-zero coefficients of $\Phi(f)$, considered as a polynomial in the "main" indeterminates $a_{i,j}$. It is easy to see that if f is a homogeneous polynomial, then $\deg(h) = \deg(f)$ for all polarizations $h \in \mathrm{Pol}_m^n(f)$. If $S \subseteq K[V^m \oplus W]$ is a set of polynomials, we write $\mathrm{Pol}_m^n(S) \subseteq K[V^n \oplus W]$ for the union of all sets $\mathrm{Pol}_m^n(f)$ for $f \in S$. Note that if S is a finite subset of $K[V^m \oplus W]$, then $\mathrm{Pol}_m^n(S)$ is a finite subset of $K[V^n \oplus W]$, which by construction has the following property: For any $f \in S$ and any m by n-matrix A with entries in K, inducing a natural linear map $\varphi_A\colon V^n \oplus W \to V^m \oplus W$, the polynomial $f \circ \varphi_A$ is a K-linear combination of $\mathrm{Pol}_m^n(S)$.

Now let G be any group acting linearly on V and W. G acts diagonally on $V^m \oplus W$. If we let G act trivially on the indeterminates $a_{i,j}$ in the above construction, then clearly $\Phi(\sigma(f)) = \sigma(\Phi(f))$ for $\sigma \in G$ and $f \in K[V^m \oplus W]$. It follows that for a subset $S \subseteq K[V^m \oplus W]^G$ of the invariant ring we have $\mathrm{Pol}_m^n(S) \subseteq K[V^n \oplus W]^G$.

A variation of the definition of polarization is given on page 8 before Corollary 1.5.

Acknowledgments. Most of the research for this paper was done during a visit of the second author to Queen's University at Kingston, Ontario. The

second author thanks Eddy Campbell, Ian Hughes, and David Wehlau for their hospitality. All authors thank Eddy Campbell for useful discussions.

1 Infinite groups

We start by considering a rather general situation. Let X and Y be sets and let F be a set of functions $f\colon X \to Y$. If $\rho \subseteq Y \times Y$ is a relation on Y we write

$$F^{-1}(\rho) := \{(x, x') \in X \times X \mid (f(x), f(x')) \in \rho \text{ for all } f \in F\} \subseteq X \times X$$

for the **preimage** of ρ under F. More specifically, let V be a vector space over a field K, n and m positive integers, and W any set. Put

$$X := V^n \times W \quad \text{and} \quad Y := V^m \times W.$$

For $A = (\alpha_{i,j}) \in K^{m\times n}$ an m by n matrix, define

$$\varphi_A\colon X \to Y,\ (v_1, \dots, v_n, w) \mapsto \left(\sum_{j=1}^{n} \alpha_{1,j} v_j, \dots, \sum_{j=1}^{n} \alpha_{m,j} v_j, w\right)$$

and set $F := \{\varphi_A\colon X \to Y \mid A \in K^{m\times n}\}$. Then for $\rho \subseteq Y \times Y$ we call $\mathrm{Pol}_m^n(\rho) := F^{-1}(\rho)$ the **polarization** of ρ. As we will see, this is closely related to the polarization of polynomials.

Lemma 1.1. *In the above situation let $\sim$ and $\approx$ be equivalence relations on X and Y, respectively, such that*

$$\sim \subseteq \mathrm{Pol}_m^n(\approx) \quad \textit{and} \quad \approx \subseteq \mathrm{Pol}_n^m(\sim). \tag{1.1}$$

If $m \geq \min\{\dim(V), n\}$, then

$$\sim = \mathrm{Pol}_m^n(\approx).$$

Remark 1.2. Before proving the lemma we make two remarks.

(a) Our main application of the lemma will be to the case that W has only one element (which amounts to saying that there is no set W) and two points $x_1, x_2 \in X = V^n$ are called equivalent if $f(x_1) = f(x_2)$ for all $f \in K[V^n]^G$, where G is a group acting linearly on V. In the same way, an equivalence relation is defined on $Y = V^m$. It is easy to see that the hypothesis (1.1) is satisfied in that situation.

(b) The following example shows that the hypothesis $m \geq \min\{\dim(V), n\}$ is essential. Let $V = K^2$, $n = 2$, $m = 1$, and let W have one element (meaning we can drop W). For $(v_1, v_2), (w_1, w_2) \in X = V^2$ we write

$$(v_1, v_2) \sim (w_1, w_2) \quad \text{if and only if} \quad \det(v_1, v_2) = \det(w_1, w_2).$$

Set $\approx = V \times V$, i.e., $v \approx w$ holds for all $v, w \in Y = V$. Then $\mathrm{Pol}_1^2(\approx) = X \times X$, so $\sim$ is a proper subset of $\mathrm{Pol}_1^2(\approx)$. Moreover, take $v, w \in V$ and $\varphi_B \in G$ arbitrary. With $B = (\beta_1, \beta_2)^T$ we have $\varphi_B(v) = (\beta_1 v, \beta_2 v)$, so $\det(\varphi_B(v)) = 0$. The same holds for $\varphi_B(w)$, so $\varphi_B(v) \sim \varphi_B(w)$. This shows that $\approx \subseteq \mathrm{Pol}_2^1(\sim)$, so the hypothesis (1.1) is satisfied. But $\sim \neq \mathrm{Pol}_1^2(\approx)$. ◁

Proof of Lemma 1.1. We need to show that $\mathrm{Pol}_m^n(\approx) \subseteq \sim$. To this end, take $x_1, x_2 \in X$ such that

$$\varphi_A(x_1) \approx \varphi_A(x_2) \quad \text{for all} \quad A \in K^{m\times n}. \tag{1.2}$$

We need to show $x_1 \sim x_2$. First consider the case $m \geq n$. In this case we can choose matrices $A \in K^{m\times n}$ and $B \in K^{n\times m}$ such that $BA = I_n$, the n by n identity matrix. This implies $\varphi_B \circ \varphi_A = \mathrm{id}_X$, so

$$x_1 = \varphi_B(\varphi_A(x_1)) \sim \varphi_B(\varphi_A(x_2)) = x_2,$$

where the equivalence "$\sim$" follows from (1.2) and (1.1).

Next we consider the case $m < n$, which implies $m \geq \dim(V)$. We interpret X as $(K^n \otimes V) \times W$ and for $i = 1, 2$ write $x_i = \left(\sum_{j=1}^k x_{i,j} \otimes v_{i,j}, w_i\right)$ with $x_{i,j} \in K^n$, $v_{i,j} \in V$, $w_i \in W$, and $k \leq \dim(V)$. Let $\widehat{U}_i \subseteq K^n$ be the span of $x_{i,1}, \ldots, x_{i,k}$. Then $\dim(\widehat{U}_i) \leq k \leq m$ and $x_i \in \left(\widehat{U}_i \otimes V\right) \times W$. Choose $\widetilde{U}_i \subseteq K^n$ with $\widehat{U}_i \subseteq \widetilde{U}_i$ and $\dim(\widetilde{U}_i) = m$. Set $Z := \widetilde{U}_1 \cap \widetilde{U}_2$ and let $U_i \subseteq \widetilde{U}_i$ be a complement of Z in $\widetilde{U}_i$. Then

$$\widetilde{U}_i \otimes V = (U_i \otimes V) \oplus (Z \otimes V).$$

Write $x_i = (u_i + z_i, w_i)$ with $u_i \in U_i \otimes V$ and $z_i \in Z \otimes V$. We have $\widetilde{U}_1 \cap U_2 = \widetilde{U_1} \cap \widetilde{U_2} \cap U_2 = Z \cap U_2 = \{0\}$. This and the fact that $\dim(\widetilde{U}_1) = m$ shows that there exists $A_1 \in K^{m\times n}$ such that the application of A_1 is injective on $\widetilde{U}_1$ and zero on U_2. Analogously, there is an $A_2 \in K^{m\times n}$ such that the application of A_2 is injective on $\widetilde{U}_2$ and zero on U_1. Thus for $i = 1, 2$ there exists $B_i \in K^{n\times m}$ such that $B_i A_i$ acts as the identity on $\widetilde{U}_i$. This yields

$$x_1 = \varphi_{B_1}(\varphi_{A_1}(x_1)) \sim \varphi_{B_1}(\varphi_{A_1}(x_2)) = \varphi_{B_1}(\varphi_{A_1}(u_2 + z_2, w_2)) = (z_2, w_2), \tag{1.3}$$

where the first equation holds since $x_1 \in \left(\widetilde{U}_1 \otimes V\right) \times W$, the equivalence "$\sim$" follows from (1.2) and (1.1), and the last equation holds since $B_1 A_1$ acts as the identity on $Z \subseteq \widetilde{U}_1$ and as 0 on U_2. Likewise, by using A_2 and B_2 we obtain

$$x_2 \sim (z_1, w_1). \tag{1.4}$$

Moreover, we have

$$\varphi_{A_1}(z_1, w_1) \approx \varphi_{A_1}(x_2) \approx \varphi_{A_1}(x_1) \approx \varphi_{A_1}(z_2, w_2), \tag{1.5}$$

where the fist equivalence follows from (1.4) and (1.1), the second from (1.2), and the third from (1.3) and (1.1). From this we see that

$$(z_1, w_1) = \varphi_{B_1}(\varphi_{A_1}(z_1, w_1)) \sim \varphi_{B_1}(\varphi_{A_1}(z_2, w_2)) = (z_2, w_2), \tag{1.6}$$

where both equations follow from $z_i \in Z \otimes V \subseteq \widetilde{U}_1 \otimes V$, and the equivalence follows from (1.5) and (1.1). Now (1.3), (1.6) and (1.4) yield $x_1 \sim x_2$, as required. □

Remark 1.3. Lemma 1.1 generalizes to infinite dimensions as follows: With V, N, and M vector spaces over K and W a set, form $X := (N \otimes V) \times W$ and $Y := (M \otimes V) \times W$, and let $F \subseteq \mathrm{Map}(X, Y)$ be the set of functions induced by all linear maps $N \to M$. For an equivalence relation $\approx$ on Y, we write $\mathrm{Pol}_M^N(\approx) := F^{-1}(\approx)$. Then Lemma 1.1 holds with the condition "$m \geq \min\{\dim(V), n\}$" replaced by

$$\dim(M) \geq \min\{\dim(V), \dim(N)\},$$

where the dimension of a vector space is either a non-negative integer or ∞, disregarding cardinalities. The original proof of Lemma 1.1 carries over almost word by word. ◁

Before formulating the main result, we recall that for any finite-dimensional vector space U with a linear G-action, a subset $S \subseteq K[U]^G$ is called **separating** if for all points $u, u' \in U$ we have that $f(u) = f(u')$ for all $f \in S$ implies $f(u) = f(u')$ for all $f \in K[U]^G$. In the following theorem, Pol_m^n, when applied to a set of polynomials, has the meaning defined at the end of the Introduction.

Theorem 1.4. *Let G be a group acting linearly on two finite-dimensional vector spaces V and W over a field K. Let n and m be positive integers such that $m \geq \min\{\dim(V), n\}$. If $S \subseteq K[V^m \oplus W]^G$ is a separating set of invariants, then $\mathrm{Pol}_m^n(S) \subseteq K[V^n \oplus W]^G$ is also separating.*

Remark. In Theorem 1.4, W might be the zero vector space. This yields the case that "there is no W" in Theorem 1.4. ◁

Proof of Theorem 1.4. Set $X := V^n \oplus W$ and define an equivalence relation $\sim$ on X by saying $x \sim x'$ for $x, x' \in X$ if $g(x) = g(x')$ for all $g \in K[X]^G$. An equivalence relation $\approx$ on $Y := V^m \oplus W$ is defined in the same way. We first show that $\sim \subseteq \mathrm{Pol}_m^n(\approx)$ and $\approx \subseteq \mathrm{Pol}_n^m(\sim)$.

For a matrix $A \in K^{m \times n}$, the map $\varphi_A \colon X \to Y$ is G-equivariant. Hence the same is true for the dual map $\varphi_A^* \colon Y^* \to X^*$ and also for its extension $\varphi_A^* \colon K[Y] \to K[X]$ as a homomorphism of algebras. It follows that $\varphi_A^*(f) \in K[X]^G$ for all $f \in K[Y]^G$. Let $x, x' \in X$ with $x \sim x'$. Then for $f \in K[Y]^G$ and $A \in K^{m \times n}$ we have

$$f(\varphi_A(x)) = (\varphi_A^*(f))(x) = (\varphi_A^*(f))(x') = f(\varphi_A(x')).$$

This shows that $\sim \subseteq \mathrm{Pol}_m^n(\approx)$. The inclusion $\approx \subseteq \mathrm{Pol}_n^m(\sim)$ is proved by reversing the roles of X and Y. Now Lemma 1.1 shows that $\sim = \mathrm{Pol}_m^n(\approx)$.

To prove that $\mathrm{Pol}_m^n(S) \subseteq K[X]^G$ is separating, take $x, x' \in X$ such that

$$g(x) = g(x') \quad \text{for all} \quad g \in \mathrm{Pol}_m^n(S). \tag{1.7}$$

We need to show that $x \sim x'$. By the above, this is equivalent to $\varphi_A(x) \approx \varphi_A(x')$ for all $A \in K^{m\times n}$. Since $S \subseteq K[Y]^G$ is separating by hypothesis, it is enough to show that $f(\varphi_A(x)) = f(\varphi_A(x'))$ for all $f \in S$. Write $A = (\alpha_{i,j}) \in K^{m\times n}$ and consider the homomorphism $\psi_A\colon K[X][a_{1,1},\dots,a_{m,n}] \to K[X]$, $a_{i,j} \mapsto \alpha_{i,j}$ with $a_{i,j}$ indeterminates. With $\Phi\colon K[Y] \to K[X][a_{1,1},\dots,a_{m,n}]$ defined by (0.1), we have $\psi_A(\Phi(f)) = \varphi_A^*(f)$ for all $f \in K[Y]$ (this is easily verified for the generators $x_{i,\nu}$ and y_i of $K[Y]$, and follows by homomorphic extension for all f). Since $\psi_A(\Phi(f))$ is a K-linear combination of $\mathrm{Pol}_m^n(f)$, it follows that for all $f \in S$ we have

$$f(\varphi_A(x)) = (\varphi_A^*(f))(x) = (\varphi_A^*(f))(x') = f(\varphi_A(x')),$$

where (1.7) is used for the middle equation. This completes the proof. □

There is a more general version of polarization, which we introduce now. Let $V_1,\dots,V_r$ be finite-dimensional vector spaces over K, each with a linear G-action. Let $m_1,\dots,m_r,n_1,\dots,n_r$ be positive integers. For a subset $S \subseteq K[V_1^{m_1}\oplus\cdots\oplus V_r^{m_r}]^G$ define $\mathrm{Pol}_{m_1,\dots,m_r}^{n_1,\dots,n_r}(S) \subseteq K[V_1^{n_1}\oplus\cdots\oplus V_r^{n_r}]^G$ by applying $\mathrm{Pol}_{m_1}^{n_1}, \mathrm{Pol}_{m_2}^{n_2},\dots,\mathrm{Pol}_{m_r}^{n_r}$ successively, where $\mathrm{Pol}_{m_i}^{n_i}$ is defined as the polarization operator Pol_m^n on page 4 with $V_i^{m_i}$ taking the role of V^m, and $V_1^{n_1}\oplus\cdots\oplus V_{i-1}^{n_{i-1}}\oplus V_{i+1}^{m_{i+1}}\oplus\cdots\oplus V_r^{m_r}$ taking the role of W. Using induction on r we obtain from Theorem 1.4:

Corollary 1.5. *Suppose that in the above setting we have $m_i \geq \min\{\dim(V_i), n_i\}$ for each $i = 1,\dots,r$. If $S \subseteq K[V_1^{m_1}\oplus\cdots\oplus V_r^{m_r}]^G$ is a separating set, then the same is true for $\mathrm{Pol}_{m_1,\dots,m_r}^{n_1,\dots,n_r}(S) \subseteq K[V_1^{n_1}\oplus\cdots\oplus V_r^{n_r}]^G$.*

If V is a finite-dimensional vector space over K with a linear G-action, we write $K[V]_d^G$ for the set of homogeneous invariants of degree d. Define

$$\beta_{\mathrm{sep}}\left(K[V]^G\right) := \min\left\{k \in \mathbb{N} \,\middle|\, \bigcup_{d=1}^{k} K[V]_d^G \text{ is separating}\right\}$$

with $\mathbb{N} := \{0,1,2,\dots\}$. By Derksen and Kemper [2, Theorem 2.3.15] there always exists a finite separating set, hence $\beta_{\mathrm{sep}}\left(K[V]^G\right)$ is a finite number. Clearly

$$\beta_{\mathrm{sep}}\left(K[V]^G\right) \leq \beta\left(K[V]^G\right) := \min\left\{k \in \mathbb{N} \,\middle|\, \bigcup_{d=1}^{k} K[V]_d^G \text{ generates } K[V]^G\right\} \in \mathbb{N}\cup\{\infty\}.$$

Corollary 1.6. *With the notation and hypotheses of Corollary 1.5 we have*

$$\beta_{\mathrm{sep}}\left(K[V_1^{n_1}\oplus\cdots\oplus V_r^{n_r}]^G\right) \leq \beta_{\mathrm{sep}}\left(K[V_1^{m_1}\oplus\cdots\oplus V_r^{m_r}]^G\right).$$

Proof. It is clear from the definition of polarization that for $f \in K[V^m]^G$ homogeneous of degree d, each polynomial in $\mathrm{Pol}_m^n(f)$ has degree d. This yields the result. □

2 Finite groups

In this section we consider actions of finite groups. Here the situation is much simpler. Indeed, if G is a finite group acting on a vector space V over a field K, then any two G-orbits in V can be separated by invariants from $K[V]^G$. For the convenience of the reader we present a short proof of this fact here.

Lemma 2.1. *Let G be a finite group acting linearly on a finite-dimensional vector space V. Then for each $v, w \in V$ with distinct G-orbits (i.e., $Gv \neq Gw$), there exists $f \in K[V]^G$ such that $f(v) \neq f(w)$.*

Proof. Write $K[V] = K[x_1, \dots, x_n]$. With additional indeterminates T and U, the polynomial

$$F(T,U) = \prod_{\sigma \in G} \left(T - \sum_{i=1}^{n} \sigma(x_i) U^{i-1} \right)$$

has coefficients in $K[V]^G$. Assume that $f(v) = f(w)$ for every $f \in K[V]^G$. Then this holds in particular for all coefficients of $F(T,U)$, so

$$\prod_{\sigma \in G} \left(T - \sum_{i=1}^{n} x_i \left(\sigma^{-1}(v)\right) U^{i-1} \right) = \prod_{\sigma \in G} \left(T - \sum_{i=1}^{n} x_i \left(\sigma^{-1}(w)\right) U^{i-1} \right).$$

Hence there exists a $\sigma \in G$ such that $x_i(w) = x_i\left(\sigma(v)\right)$ for all i, which implies $Gv = Gw$. □

It follows from Lemma 2.1 that a subset $S \subseteq K[V]^G$ is separating if any two G-orbits can be separated by invariants from S. The proof of Lemma 2.1 shows that the coefficients of $F(T,U)$ form a separating set of invariants.

In order to formulate the results of this section, we need to introduce the concept of cheap polarization. Let $V_1, \dots, V_r$ be finite-dimensional vector spaces over a field K. For $k = 1, \dots, r$ write $d_k := \dim(V_k)$, and let $x_1^{(k)}, \dots, x_{d_k}^{(k)}$ be a basis of the dual space V_k^*. Then $K[V_1 \oplus \cdots \oplus V_r]$ is a polynomial ring with the $x_j^{(k)}$ as indeterminates. Let $n_1, \dots, n_r$ be positive integers. Then $K[V_1^{n_1} \oplus \cdots \oplus V_r^{n_r}]$ is a polynomial ring with indeterminates $x_{i,j}^{(k)}$ ($k = 1, \dots, r$, $i = 1, \dots, n_k$, $j = 1, \dots, d_k$) defined in the obvious way (see page 4 in the Introduction). Let a be a further indeterminate and define a homomorphism

$$\Psi\colon K[V_1 \oplus \cdots \oplus V_r] \to K[V_1^{n_1} \oplus \cdots \oplus V_r^{n_r}][a],\ x_j^{(k)} \mapsto \sum_{i=1}^{n_k} a^{i-1} \cdot x_{i,j}^{(k)}$$

of K-algebras. To illustrate the effect of Ψ, we consider the standard case $r = 1$ and pretend for a moment that $a \in K$ is a scalar. Then for $f \in K[V]$ and $v_1, \dots, v_n \in V$ we have

$$\left(\Psi(f)\right)(v_1, \dots, v_n) = f(v_1 + a v_2 + \cdots + a^{n-1} v_n).$$

For a polynomial $f \in K[V_1 \oplus \cdots \oplus V_r]$, let $\mathrm{Pol}_{\mathrm{cheap}}^{n_1,\dots,n_r}(f) \subset K[V_1^{n_1} \oplus \cdots \oplus V_r^{n_r}]$ denote the set of all coefficients of $\Psi(f)$ as a polynomial in a. For $S \subseteq$

$K[V_1 \oplus \cdots \oplus V_r]$, let $\mathrm{Pol}^{n_1,\ldots,n_r}_{\mathrm{cheap}}(S)$ be the union of all $\mathrm{Pol}^{n_1,\ldots,n_r}_{\mathrm{cheap}}(f)$, $f \in S$. We call this set the **cheap polarization** of S. This construction is in particular interesting in the case $r = 1$. The main difference between cheap polarization and "ordinary" polarization is that there is only one additional indeterminate a involved in cheap polarization, which results in an easier computation and a smaller number of coefficients. This is why we use the word "cheap".

Example 2.2. Consider the case $r = 1$ and $\dim(V_1) = 1$, so $K[V_1] = K[x]$. For $n_1 = 3$ we have

$$\Psi(x^2) = (x_1 + ax_2 + a^2x_3)^2 = x_1^2 + 2x_1x_2 \cdot a + (2x_1x_3 + x_2^2) \cdot a^2 + 2x_2x_3 \cdot a^3 + x_3^2a^4,$$

so

$$\mathrm{Pol}^3_{\mathrm{cheap}}(x^2) = \left\{x_1^2, x_3^2, 2x_1x_2, 2x_2x_3, 2x_1x_3 + x_2^2\right\}.$$

On the other hand, "ordinary" polarization gives

$$\mathrm{Pol}^3_1(x^2) = \left\{x_1^2, x_3^2, 2x_1x_2, 2x_2x_3, 2x_1x_3, x_2^2\right\}.$$

◁

The following proposition gives some basic properties of cheap polarization. The first part compares cheap polarization with "ordinary" polarization.

Proposition 2.3. *In the above situation, let $S \subseteq K[V_1 \oplus \cdots \oplus V_r]$ be a set of polynomials.*

(a) Every element in $\mathrm{Pol}^{n_1,\ldots,n_r}_{\mathrm{cheap}}(S)$ can be expressed as a sum of elements from $\mathrm{Pol}^{n_1,\ldots,n_r}_{1,\ldots,1}(S)$.

(b) If d is an upper bound on the total degree of monomials occurring in polynomials from S, then it is also an upper bound on the total degree of monomials occurring in polynomials from $\mathrm{Pol}^{n_1,\ldots,n_r}_{\mathrm{cheap}}(S)$. If all polynomials in S are homogeneous of degree d, then the same is true for all polynomials in $\mathrm{Pol}^{n_1,\ldots,n_r}_{\mathrm{cheap}}(S)$.

(c) Let d be a bound as in part (b), let S be finite and set $n := \max\{n_1, \ldots, n_r\}$. Then

$$\left|\mathrm{Pol}^{n_1,\ldots,n_r}_{\mathrm{cheap}}(S)\right| \leq |S| \cdot (d \cdot (n-1) + 1).$$

(d) Let G be a group acting linearly on all V_k. If $S \subseteq K[V_1 \oplus \cdots \oplus V_r]^G$, then

$$\mathrm{Pol}^{n_1,\ldots,n_r}_{\mathrm{cheap}}(S) \subseteq K[V_1^{n_1} \oplus \cdots \oplus V_r^{n_r}]^G$$

(with G acting diagonally on $V_1 \oplus \cdots \oplus V_r$ and $V_1^{n_1} \oplus \cdots \oplus V_r^{n_r}$).

Proof. (a) Let $a_i^{(k)}$ be indeterminates ($k = 1, \ldots, r$, $i = 1, \ldots, n_k$) and define a homomorphism

$$\Phi\colon K[V_1 \oplus \cdots \oplus V_r] \to K[V_1^{n_1} \oplus \cdots \oplus V_r^{n_r}][a_1^{(1)}, \ldots, a_{n_r}^{(r)}],\ x_j^{(k)} \mapsto \sum_{i=1}^{n_k} a_i^{(k)} \cdot x_{i,j}^{(k)}$$

of K-algebras. Let $f \in K[V_1 \oplus \cdots \oplus V_r]$. Let $\mathcal{M}$ be the set of monomials in the $a_i^{(k)}$ occurring in $\Phi(f)$. Then

$$\Phi(f) = \sum_{t \in \mathcal{M}} f_t \cdot t$$

with $f_t \in K[V_1^{n_1} \oplus \cdots \oplus V_r^{n_r}]$, and

$$\mathrm{Pol}_{1,\ldots,1}^{n_1,\ldots,n_r}(f) = \{f_t \mid t \in \mathcal{M}\}.$$

Define a homomorphism Λ: $K[V_1^{n_1} \oplus \cdots \oplus V_r^{n_r}][a_1^{(1)}, \ldots, a_{n_r}^{(r)}] \to K[V_1^{n_1} \oplus \cdots \oplus V_r^{n_r}][a]$ of algebras over $K[V_1^{n_1} \oplus \cdots \oplus V_r^{n_r}]$ by $\Lambda(a_i^{(k)}) = a^{i-1}$. Then $\Psi = \Lambda \circ \Phi$. For $i \in \mathbb{N}$ a non-negative integer, set $\mathcal{M}_i := \{t \in \mathcal{M} \mid \Lambda(t) = a^i\}$. Then

$$\Psi(f) = \sum_{t \in \mathcal{M}} f_t \cdot \Lambda(t) = \sum_{i \in \mathbb{N}} \left(\sum_{t \in \mathcal{M}_i} f_t \right) a^i,$$

the latter sum being finite by the finiteness of $\mathcal{M}$. It follows that all elements in $\mathrm{Pol}_{\mathrm{cheap}}^{n_1,\ldots,n_r}(f)$ are sums of elements of $\mathrm{Pol}_{1,\ldots,1}^{n_1,\ldots,n_r}(f)$. This implies the statement (a).

(b) This is clear from the definition of cheap polarization (assign degree 0 to a).

(c) Let $f \in S$. Then clearly $d \cdot (n-1)$ is an upper bound on the degree of $\Psi(f)$ considered as a polynomial in a, and therefore

$$\left|\mathrm{Pol}_{\mathrm{cheap}}^{n_1,\ldots,n_r}(f)\right| \leq d \cdot (n-1) + 1.$$

The statement (c) follows.

(d) Extending the G-action to $K[V_1^{n_1} \oplus \cdots \oplus V_r^{n_r}][a]$ by $\sigma(a) = a$ for $\sigma \in G$, we see that the map Ψ is G-equivariant. This implies part (d). □

Theorem 2.4. *In the situation introduced at the beginning of this section, let G be a finite group acting linearly on all V_k ($k = 1, \ldots, r$). Let $S \subseteq K[V_1 \oplus \cdots \oplus V_r]^G$ and assume that at least one of the following hypotheses is satisfied:*

(a) S generates $K[V_1 \oplus \cdots \oplus V_r]^G$ as a K-algebra.

(b) S is separating and K has strictly more than $(\max\{n_1, \ldots, n_r\} - 1) \cdot |G|$ elements.

Then $\mathrm{Pol}_{\mathrm{cheap}}^{n_1,\ldots,n_r}(S) \subseteq K[V_1^{n_1} \oplus \cdots \oplus V_r^{n_r}]^G$ is separating.

Proof. We first assume that the hypothesis (b) holds. For $k = 1, \ldots, r$, let $v_1^{(k)}, \ldots, v_{n_k}^{(k)}, w_1^{(k)}, \ldots, w_{n_k}^{(k)} \in V_k$ be vectors such that for all $F \in \mathrm{Pol}_{\mathrm{cheap}}^{n_1,\ldots,n_r}(S)$ we have

$$F\left(v_1^{(1)}, \ldots, v_{n_1}^{(1)}, \ldots, v_1^{(r)}, \ldots, v_{n_r}^{(r)}\right) = F\left(w_1^{(1)}, \ldots, w_{n_1}^{(1)}, \ldots, w_1^{(r)}, \ldots, w_{n_r}^{(r)}\right). \tag{2.1}$$

We need to show that there exists a $\sigma \in G$ such that $w_i^{(k)} = \sigma(v_i^{(k)})$ for all $k \in \{1, \dots, r\}$, $i \in \{1, \dots, n_k\}$. For $\alpha \in K$ let $\eta_\alpha \colon K[V_1^{n_1} \oplus \cdots \oplus V_r^{n_r}][a] \to K[V_1^{n_1} \oplus \cdots \oplus V_r^{n_r}]$ be the map given by sending a to α, and set $\Psi_\alpha := \eta_\alpha \circ \Psi$ (with Ψ defined at the beginning of Section 2). Thus for $f \in K[V_1 \oplus \cdots \oplus V_r]$ and $u_i^{(k)} \in V_k$ $(k = 1, \dots, r,\ i = 1, \dots, n_k)$ we have

$$(\Psi_\alpha(f)) \left(u_1^{(1)}, \dots, u_{n_1}^{(1)}, \dots, u_1^{(r)}, \dots, u_{n_r}^{(r)} \right) = f \left(\sum_{i=1}^{n_1} \alpha^{i-1} u_i^{(1)}, \dots, \sum_{i=1}^{n_r} \alpha^{i-1} u_i^{(r)} \right). \quad (2.2)$$

Observe that $\Psi_\alpha(f)$ is a K-linear combination of elements of $\mathrm{Pol}_{\mathrm{cheap}}^{n_1,\dots,n_r}(f)$, so if $f \in S$, then (2.1) implies

$$(\Psi_\alpha(f)) \left(v_1^{(1)}, \dots, v_{n_1}^{(1)}, \dots, v_1^{(r)}, \dots, v_{n_r}^{(r)} \right) = (\Psi_\alpha(f)) \left(w_1^{(1)}, \dots, w_{n_1}^{(1)}, \dots, w_1^{(r)}, \dots, w_{n_r}^{(r)} \right), \quad (2.3)$$

which with (2.2) leads to

$$f \left(\sum_{i=1}^{n_1} \alpha^{i-1} v_i^{(1)}, \dots, \sum_{i=1}^{n_r} \alpha^{i-1} v_i^{(r)} \right) = f \left(\sum_{i=1}^{n_1} \alpha^{i-1} w_i^{(1)}, \dots, \sum_{i=1}^{n_r} \alpha^{i-1} w_i^{(r)} \right).$$

Since this holds for all $f \in S$, it follows by Lemma 2.1 that there exists a $\sigma \in G$ such that

$$\sum_{i=1}^{n_k} \alpha^{i-1} w_i^{(k)} = \sigma \left(\sum_{i=1}^{n_k} \alpha^{i-1} v_i^{(k)} \right) \quad \text{for all} \quad k \in \{1, \dots, r\}.$$

Since $\alpha \in K$ was chosen arbitrary, this means that for every $\alpha \in K$ there exists a $\sigma \in G$ such that

$$\sum_{i=1}^{n_k} \alpha^{i-1} \left(w_i^{(k)} - \sigma(v_i^{(k)}) \right) = 0 \quad \text{for all} \quad k \in \{1, \dots, r\}. \quad (2.4)$$

For $\sigma \in G$ let S_σ be the set of all $\alpha \in K$ such that (2.4) holds for α and σ. Thus $K = \bigcup_{\sigma \in G} S_\sigma$. By the hypothesis on the size of K there exists a $\sigma \in G$ such that $|S_\sigma| \geq \max\{n_1, \dots, n_r\}$. By using the Vandermonde determinant, we conclude from (2.4) that for this σ we have

$$w_i^{(k)} = \sigma(v_i^{(k)})$$

for all k and i. This completes the proof.

Now assume that (a) is satisfied, and let $\overline{K}$ be the algebraic closure of K. With $\overline{V_k} := \overline{K} \otimes_K V_k$ we have $\overline{K}[\overline{V_1} \oplus \cdots \oplus \overline{V_r}]^G \cong \overline{K} \otimes_K K[V_1 \oplus \cdots \oplus V_r]^G$, so S generates $\overline{K}[\overline{V_1} \oplus \cdots \oplus \overline{V_r}]^G$. In particular, S is $\overline{K}[\overline{V_1} \oplus \cdots \oplus \overline{V_r}]^G$-separating, and

the hypothesis (b) is satisfied. Therefore $\mathrm{Pol}_{\mathrm{cheap}}^{n_1,\dots,n_r}(S) \subseteq \overline{K}[\overline{V_1}^{n_1} \oplus \cdots \oplus \overline{V_r}^{n_r}]^G$ is separating. Since G is finite, this implies by Lemma 2.1 that any two points in $\overline{V_1}^{n_1} \oplus \cdots \oplus \overline{V_r}^{n_r}$ with distinct G-orbits can be separated by an element from $\mathrm{Pol}_{\mathrm{cheap}}^{n_1,\dots,n_r}(S)$, hence in particular this applies to points in $V_1^{n_1} \oplus \cdots \oplus V_r^{n_r}$. □

Remark 2.5. We know of no example which shows that the hypothesis on the size of K in Theorem 2.4(b) cannot be dropped. ◁

Example 2.6. Suppose $\mathrm{char}(K) \neq 2$, and let $G = \{\pm 1\} \subset \mathrm{GL}_1(K)$. Then $K[x]^G = K[x^2]$, so by Example 2.2 and Theorem 2.4 the set

$$\left\{x_1^2, x_3^2, x_1x_2, x_2x_3, 2x_1x_3 + x_2^2\right\} \subseteq K[x_1, x_2, x_3]^G$$

is separating. In other words, for $(\xi_1, \xi_2, \xi_3), (\eta_1, \eta_2, \eta_3) \in K^3$ we have

$$\begin{aligned}&(\xi_1, \xi_2, \xi_3) = \pm(\eta_1, \eta_2, \eta_3) \quad \Longleftrightarrow\\ &\xi_1^2 = \eta_1^2,\ \xi_3^2 = \eta_3^2,\ \xi_1\xi_2 = \eta_1\eta_2,\ \xi_2\xi_3 = \eta_2\eta_3, \quad \text{and} \quad 2\xi_1\xi_3 + \xi_2^2 = 2\eta_1\eta_3 + \eta_2^2.\end{aligned}$$

It is rather subtle to verify this equivalence without using Theorem 2.4. ◁

Corollary 2.7. *With the same situation and hypotheses as in Theorem 2.4,* $\mathrm{Pol}_{1,\dots,1}^{n_1,\dots,n_r}(S) \subseteq K[V_1^{n_1} \oplus \cdots \oplus V_r^{n_r}]^G$ *is also separating.*

Proof. This is a direct consequence of Theorem 2.4 and Proposition 2.3(a). □

As a consequence of Theorem 2.4 and Proposition 2.3(b) we obtain:

Corollary 2.8. *Let G be a finite group acting linearly on a finite-dimensional vector space V over a field K. Then for all positive integers n we have*

$$\beta_{\mathrm{sep}}\left(K[V^n]^G\right) \leq \beta\left(K[V]^G\right).$$

Remark 2.9. (a) Of course Corollary 2.8 holds in the more general situation of a linear action on several vector spaces $V_1, \dots, V_r$. Moreover, if K has more than $(n-1) \cdot |G|$ elements, one can use β_{sep} instead of β on the right hand side of the inequality. In fact, we have $\beta_{\mathrm{sep}}\left(K[V^n]^G\right) = \beta_{\mathrm{sep}}\left(K[V]^G\right)$ in this case.

(b) In the modular case (i.e., when the characteristic of K divides the group order), Corollary 2.8 stands in stark contrast to the results about generating invariants. In fact, we know from Richman [7] that for every faithful linear representation V of a finite group G with $\mathrm{char}(K) \mid |G|$ we have

$$\lim_{n\to\infty} \beta\left(K[V^n]^G\right) = \infty.$$

◁

Example 2.10. Corollary 2.8 contains new information in the characteristic zero case as well. In fact, whenever we have

$$\beta\left(K[V^n]^G\right) > \beta\left(K[V]^G\right), \tag{2.5}$$

Corollary 2.8 tells us that $\beta_{\mathrm{sep}}\left(K[V^n]^G\right) < \beta\left(K[V^n]^G\right)$. It is surprisingly hard to find an example in characteristic zero where (2.5) is satisfied. Using a rather extensive computer search, we found the subgroup $G \subseteq \mathrm{GL}_4(\mathbb{C})$ of order 32 generated by the four matrices

$$\begin{pmatrix} i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & 0 & 0 & -i \end{pmatrix}, \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

We used the computer algebra system Magma [1] for computing generating invariants for $\mathbb{C}[V]^G$ and $\mathbb{C}[V^2]^G$, where $V = \mathbb{C}^4$. The resulting beta-values are

$$\beta\left(\mathbb{C}[V]^G\right) = 6 \quad \text{and} \quad \beta\left(\mathbb{C}[V^2]^G\right) = 8.$$

So $\mathbb{C}[V^2]^G$ is an example of an invariant ring in characteristic zero where the degree bound on separating invariants is smaller than the one for generating invariants. ◁

Corollary 2.11. *Let G be a finite group acting linearly on a finite-dimensional vector space V over a field K. Moreover, let V_{reg} be the regular representation of G. Then*

$$\beta_{\mathrm{sep}}\left(K[V]^G\right) \leq \beta\left(K[V_{\mathrm{reg}}]^G\right).$$

Proof. Since V_{reg} is free of rank one as a module over the group ring KG, there exists a positive integer n and an epimorphism $V_{\mathrm{reg}}^n \to V^*$ of KG-modules, where V^* is the dual. Dualizing yields $V \hookrightarrow V_{\mathrm{reg}}^n$, which induces a degree-preserving, G-equivariant epimorphism $\pi\colon K[V_{\mathrm{reg}}^n] \to K[V]$ of K-algebras. By Theorem 2.8 there exists a separating set $S \subseteq K[V_{\mathrm{reg}}^n]^G$ consisting of homogeneous invariants of degree at most $\beta\left(K[V_{\mathrm{reg}}]^G\right)$. Moreover, by Derksen and Kemper [2, Theorem 2.3.16], $\pi(S)$ is a separating set of $K[V]^G$. This completes the proof. □

Theorem 2.4 also has consequences on the number of separating vector invariants. For stating them, we introduce the following notation. For a K-algebra R we write

$$\gamma(R) := \min\{k \in \mathbb{N} \mid R \text{ has a generating subset of size } k\} \in \mathbb{N} \cup \{\infty\}.$$

Moreover, if R consists of functions $f\colon X \to K$ from a set X to K, we write

$$\gamma_{\mathrm{sep}}(R) := \min\{k \in \mathbb{N} \mid \ R \text{ has an } R\text{-separating subset of size } k\} \in \mathbb{N} \cup \{\infty\}.$$

Corollary 2.12. *Let G be a finite group acting linearly on a finite-dimensional vector space V over a field K. Then for all positive integers n we have*

$$\gamma_{\mathrm{sep}}\left(K[V^n]^G\right) \leq \left((n-1)\cdot\beta\left(K[V]^G\right)+1\right)\cdot\gamma\left(K[V]^G\right).$$

Proof. This is a direct consequence of Proposition 2.3(c) and Theorem 2.4. □

Remark. Remark 2.9(a) applies to Corollary 2.12 in the same way as it does to Corollary 2.8. ◁

It is remarkable that the bound in Corollary 2.12 is linear in n. We will see in Theorem 2.13 that such a bound cannot hold for $\gamma\left(K[V^n]^G\right)$ unless G acts trivially. In fact, instead of a linear upper bound we have a quadratic lower bound. Before stating the theorem, we make a remark about multihomogeneity. The natural multigrading on $K[V^n] = K[x_{1,1},\ldots,x_{n,k}]$, defined by

$$\deg_{\mathrm{mult}}\left(\prod_{i=1}^{n}\prod_{\nu=1}^{k} x_{i,\nu}^{e_{i,\nu}}\right) = \left(\sum_{\nu=1}^{k} e_{1,\nu},\ldots,\sum_{\nu=1}^{k} e_{n,\nu}\right) \in \mathbb{Z}^n,$$

is inherited by $K[V^n]^G$. Observe that the invariants produced by cheap polarization are usually not multihomogeneous, whereas invariants produced by "ordinary" polarization always are. This can be seen as a shortcoming of cheap polarization. However, we will also show that no bound on the number of separating invariants as in Corollary 2.12 can be obtained when requiring the invariants to be multihomogeneous. In order to formulate our result we need another piece of notation. For a multigraded subalgebra $R \subseteq K[V^n]$ we write

$$\gamma_{\mathrm{sep,mult}}(R) := \min\left\{k\in\mathbb{N} \mid \text{ } R \text{ has an } R\text{-separating multihomogeneous subset of size } k\right\}.$$

Theorem 2.13. *Let G be a finite group acting linearly and non-trivially on a finite-dimensional vector space V over a field K. Then for all positive integers n we have*

$$\gamma\left(K[V^n]^G\right) \geq \gamma_{\mathrm{sep,mult}}\left(K[V^n]^G\right) \geq \binom{n+1}{2}.$$

Proof. We start by showing the first inequality. Set $k := \gamma\left(K[V^n]^G\right)$ and take generators $f_1,\ldots,f_k \in K[V^n]^G$. Write $I := K[V^n]^G_+$ for the ideal in $K[V^n]^G$ consisting of all invariants whose constant coefficient is 0. We can remove the constant coefficients from each f_i and thus assume $f_i \in I$. Clearly the $\overline{f_i} := f_i + I^2$ generate I/I^2 as a vector space over K. So

$$\dim_K\left(I/I^2\right) \leq k.$$

Observe that I/I^2 inherits a multigrading from $K[V^n]^G$. Thus there exist multihomogeneous invariants $g_i,\ldots,g_{k'} \in I$ with $k' := \dim_K\left(I/I^2\right)$ such that the $g_i + I^2$ form a basis of I/I^2. By the graded version of Nakayama's lemma (see

Derksen and Kemper [2, Lemma 3.5.1]), the g_i generate I (as an ideal), and also $K[V^n]^G$ (as a K-algebra). Therefore $k' = k$, and there exists a multihomogeneous system of generators of $K[V^n]^G$ of size k. Since every generating set is also separating, the first inequality follows.

For proving the second inequality, let $S \subseteq K[V^n]^G$ be a multihomogeneous separating set. We first claim that for all $1 \leq i < j \leq n$ there exists a non-zero $f \in S$ with

$$\deg_{\text{mult}}(f) = (0, \ldots, 0, d_i, 0, \ldots, 0, d_j, 0, \ldots, 0), \tag{2.6}$$

where d_i and d_j are positive and occur at the i-th and j-th position of the vector. Indeed, by the non-triviality of the action there exist $v \in V$ and $\sigma \in G$ such that $w := \sigma(v) \neq v$. Thus the vectors $(0, \ldots, 0, v, 0, \ldots, 0, v, 0, \ldots, 0) \in V^n$ (with v in the i-th and j-th position) and $(0, \ldots, 0, v, 0, \ldots, 0, w, 0, \ldots, 0) \in V^n$ (with v in the i-th and w in the j-th position) lie in distinct G-orbits. By Lemma 2.1, there exists $f \in S$ which takes different values at these two vectors. Let $\deg_{\text{mult}}(f) = (d_1, \ldots, d_n)$. Then $d_\nu = 0$ for $\nu \notin \{i, j\}$, since otherwise

$$f(0, \ldots, 0, v, 0, \ldots, 0, v, 0, \ldots, 0) = 0 = f(0, \ldots, 0, v, 0, \ldots, 0, w, 0, \ldots, 0).$$

Assume $d_i = 0$. Then

$$\begin{aligned} f(0, \ldots, 0, v, 0, \ldots, 0, v, 0, \ldots, 0) = f(0, \ldots, 0, 0, 0, \ldots, 0, v, 0, \ldots, 0) = \\ f(0, \ldots, 0, 0, 0, \ldots, 0, w, 0, \ldots, 0) = f(0, \ldots, 0, v, 0, \ldots, 0, w, 0, \ldots, 0), \end{aligned}$$

since $(0, \ldots, 0, 0, 0, \ldots, 0, v, 0, \ldots, 0)$ and $(0, \ldots, 0, 0, 0, \ldots, 0, w, 0, \ldots, 0)$ lie in the same G-orbit. Similarly, assuming $d_j = 0$ leads to

$$f(0, \ldots, 0, v, 0, \ldots, 0, v, 0, \ldots, 0) = f(0, \ldots, 0, v, 0, \ldots, 0, w, 0, \ldots, 0).$$

So indeed $\deg_{\text{mult}}(f)$ has the form claimed in (2.6). Our second claim is that for every $i \in \{1, \ldots, n\}$ there exists a non-zero $f \in S$ with

$$\deg_{\text{mult}}(f) = (0, \ldots, 0, d_i, 0, \ldots, 0), \tag{2.7}$$

where d_i is positive and occurs at the i-th position of the vector. Indeed, $(0, \ldots, 0, v, 0, \ldots, 0) \in V^n$ (with v in the i-th component) and $(0, \ldots, 0) \in V^n$ lie in distinct G-orbits. Thus there exists $f \in S$ with $f(0, \ldots, 0, v, 0, \ldots, 0) \neq f(0, \ldots, 0)$. This implies (2.7). Taking both claims together shows that S must contain at least $\binom{n}{2} + n = \binom{n+1}{2}$ distinct elements. □

Example 2.14. Let $G \leq GL_1(K)$ be the group of order 2 generated by -1, where $\text{char}(K) \neq 0$. Then $K[V^n]^G$ (with $V = K$) is minimally generated by all monomials of degree 2, so

$$\gamma\left(K[V^n]^G\right) = \binom{n+1}{2}.$$

This show that the bounds in Theorem 2.13 are sharp. ◁

References

[1] Wieb Bosma, John J. Cannon, Catherine Playoust, *The Magma Algebra System I: The User Language*, J. Symb. Comput. **24** (1997), 235–265.

[2] Harm Derksen, Gregor Kemper, *Computational Invariant Theory*, Encyclopaedia of Mathematical Sciences **130**, Springer-Verlag, Berlin, Heidelberg, New York 2002.

[3] Gregor Kemper, *Computing Invariants of Reductive Groups in Positive Characteristic*, Transformation Groups **8** (2003), 159–176.

[4] Friedrich Knop, *On Noether's and Weyl's bound in positive characteristic*, in: H. E. A. Eddy Campbell, David L. Wehlau, eds., *Invariant theory in all characteristics*, CRM Proceedings and Lecture Notes **35**, pp. 175–188, Amer. Math. Soc., Providence, RI 2004.

[5] Hanspeter Kraft, Claudio Procesi, *A Primer of invariant theory*, Notes by G. Boffi, Brandeis Lecture Notes 1. Updated version (2000) available at `http://www.math.unibas.ch/~kraft/Papers/KP-Primer.pdf`, 1982.

[6] David R. Richman, *On Vector Invariants over Finite Fields*, Adv. in Math. **81** (1990), 30–65.

[7] David R. Richman, *Invariants of Finite Groups over Fields of Characteristic p*, Adv. in Math. **124** (1996), 25–48.

[8] Hermann Weyl, *The Classical Groups. Their Invariants and Representations*, Princeton University Press, Princeton, N.J. 1939.

Jan Draisma
Mathematisches Institut der
Universität Basel
Rheinsprung 21
CH-4051 Basel
Switzerland
`jan.draisma@unibas.ch`

Gregor Kemper
Technische Universität München
Zentrum Mathematik - M11
Boltzmannstr. 3
85 748 Garching
Germany
`kemper@ma.tum.de`

David Wehlau
Department of Mathematics and Computer Science
Royal Military College
Kingston
Ontario K7K 7B4
Canada
`wehlau@rmc.ca`