On Noether's and Weyl's bound in positive characteristic

Friedrich Knop[1]
Department of Mathematics, Rutgers University, New Brunswick NJ 08903, USA
knop@math.rutgers.edu

Abstract: In this note we generalize several well known results concerning invariants of finite groups from characteristic zero to positive characteristic not dividing the group order. The first is Schmid's relative version of Noether's theorem. That theorem compares the degrees of generators of a group with those of a subgroup. Then we prove a suitable positive characteristic version of Weyl's theorem on vector invariants: polarization works in small degrees. Using that we show that the regular representation has the "most general" ring of invariants, thereby generalizing theorems of Schmid and Smith.

1. Introduction

Let G be a finite group acting linearly on a k-vector space V. If $k = \mathbb{C}$, Noether showed in [N] that the ring $k[V]^G$ of G-invariants is generated by invariants of degree $\le |G|$. This is called *Noether's bound.* An inspection of her proof reveals that it stays valid for any field k of characteristic 0 or characteristic $p > |G|$. It has been a long standing question whether Noether's bound holds under the weaker assumption that p does not divide $|G|$. This has been proved only recently by Fleischmann [Fl1] and Fogarty [Fo].

Meanwhile, Schmid showed in [Sch] that Noether's bound is sharp only for cyclic groups. Again, her proof is valid if either $p = 0$ or $p > |G|$. The only reason why Schmid's proof doesn't carry over to characteristics not dividing $|G|$ is a certain refinement of Noether's theorem which compares degrees of generators of G-invariants with those of H-invariants where H is any subgroup of G. The purpose of this paper is to provide this tool for $p \not| \, |G|$ (see Theorem 3.3 for a precise statement).

After a first draft of this note has been completed it came to my attention that this has also been done independently by Fleischmann, [Fl1], and Sezer, [Se]. Nevertheless, we

---

[1] Supported by an NSA-grant.

find it still worthwhile to include a proof. First, our proof is an adaptation of Fogarty's (as opposed to Fleischmann's) proof of Noether's bound. Secondly, we introduce another numerical invariant of independent interest. More precisely, in section 2 we look at the zero fiber of the quotient map $V \to V/G$ (the null cone) which is the spectrum of a local Artinian ring. We study the degree of nilpotence of its maximal ideal in Theorem 2.1 and show that it is closely related to degrees of generating invariants.

In a further section, we study degree bounds for algebras over arbitrary ground rings k. We show that it suffices to consider for k a field and even there the problem depends only on the characteristic of k with $\operatorname{char} k = 0$ giving the lowest bounds.

The second major thread of this paper is Weyl's theorem on vector invariants. It asserts that one can obtain all invariants on any number of copies of a representation V from invariants which live on only $\dim V$ copies. The process allowing this is called polarization and works in that generality only in characteristic zero. Nevertheless, in positive characteristic it is plausible that polarization still works in "low" degrees. Theorem 5.1 makes this statement precise. On the other hand, generating invariants for finite groups have "low" degree by the theorem of Noether-Fleischmann-Fogarty. This yields Weyl type theorems for vector invariants in positive characteristic. In particular, we strengthen theorems of Schmid and Smith according to which the regular representation has the most general ring of invariants.

It should be added that even in characteristic zero our approach is not without interest. It yields a completely elementary proof of Weyl's theorem for the (special but most important) case that G is linearly reductive. In particular, we avoid the use of Capelli's identity and representation theory of $GL(n)$.

Acknowledgment: The author would like to thank the referee for very carefully reading the manuscript.

2. The degree of nilpotence of the null cone

Let A be a commutative ring. For a subset S let $\langle S\rangle$ be the ideal generated by S. Now assume that the finite group G acts on A. For a G-invariant ideal $\mathfrak{m} \subseteq A$ we define

$$(2.1) \qquad \eta_G(\mathfrak{m}) := \inf\{d \in \mathbb{N} \mid \mathfrak{m}^d \subseteq \langle \mathfrak{m}^G \rangle\}.$$

It is well known that for finite groups, the fibers of $\operatorname{Spec} A \to \operatorname{Spec} A^G$ are precisely the G-orbits. Thus $\mathfrak{m}$ and $\langle \mathfrak{m}^G\rangle$ define the same vanishing set in $\operatorname{Spec} A$, which implies $\eta_G(\mathfrak{m}) < \infty$ whenever A is Noetherian (see also Corollary 2.2 below). The most important case will be when A is the ring of polynomial functions on a representation of G and $\mathfrak{m}$ the maximal ideal corresponding to the origin. Then $\langle \mathfrak{m}^G \rangle$ defines the null cone and $\eta_G(\mathfrak{m}) - 1$ is the degree of nilpotence of the maximal ideal of $A/\langle \mathfrak{m}^G \rangle$.

2.1. Theorem. *Let $H \subseteq G$ be a subgroup and assume $[G:H] \in A^\times$. Then*

$$\eta_G(\mathfrak{m}) \le [G:H]\,\eta_H(\mathfrak{m}). \tag{2.2}$$

Proof: Let $d := [G:H]$. Choose d arbitrary elements $a_u \in \mathfrak{m}^H$ which are indexed by $u \in G/H$. In the following, we will regard u as an element of G representing its coset. Consider the expression

$$\Phi := \sum_{u \in G/H} \prod_{v \in G/H} (va_v - ua_v) \tag{2.3}$$

Since every product contains a zero factor, namely the one corresponding to $v = u$, we have $\Phi = 0$. On the other side, we can expand the product and get

$$\Phi = \sum_S (-1)^{|S|} \Phi_S \tag{2.4}$$

where S ranges over all subsets of G/H and

$$\Phi_S = \sum_{u \in G/H} \prod_{v \notin S} (va_v) \prod_{v \in S} (ua_v) = \prod_{v \notin S} (va_v) \left[\sum_{u \in G/H} u\Big(\prod_{v \in S} a_v \Big) \right]. \tag{2.5}$$

Thus, $\Phi_S \in \langle \mathfrak{m}^G \rangle$ unless $S = \emptyset$ when $\Phi_\emptyset = d \prod_{v \in G/H} va_v$. Since the a_v were arbitrary and $d \in A^\times$ we get

$$\prod_{v \in G/H} v\langle \mathfrak{m}^H \rangle \subseteq \langle \mathfrak{m}^G \rangle. \tag{2.6}$$

With $e := \eta_H(\mathfrak{m})$ we have by definition $\mathfrak{m}^e \subseteq \langle \mathfrak{m}^H \rangle$. Thus,

$$\langle \mathfrak{m}^G \rangle \supseteq \prod_{v \in G/H} v\mathfrak{m}^e = \mathfrak{m}^{de}. \tag{2.7}$$

This implies $\eta_G(\mathfrak{m}) \le de = [G:H]\,\eta_H(\mathfrak{m})$. ☐

For $H = 1$ we get:

2.2. Corollary. *Assume $|G| \in A^\times$. Then $\eta_G(\mathfrak{m}) \le |G|$.*

3. The degree of generators

Let $A = \oplus_{d=0}^{\infty} A_d$ be an $\mathbb{N}$-graded commutative ring. We use the notation $A_{\leq n} := \oplus_{d=0}^{n} A_d$ and analogously $A_{\geq n} := \oplus_{d=n}^{\infty} A_d$. Of particular interest is the ideal $A_{\geq 1}$ which we denote by $\mathfrak{m}$. We define

$$\beta(A) := \inf\{n \in \mathbb{N} \mid A \text{ is generated as a ring by } A_{\leq n}\}, \tag{3.1}$$

which is the smallest degree of a generator of A. Now assume that the finite group G acts on A by degree preserving automorphisms.

3.1. Lemma. *Assume $\beta(A) \leq 1$ and $|G| \in A^{\times}$. Then $\beta(A^G) \leq \eta_G(\mathfrak{m})$.*

Proof: Let $e := \eta_G(\mathfrak{m})$. By the classical argument of Hilbert, any set of G-invariant generators of the ideal $\langle \mathfrak{m}^G \rangle$ also generates the ring A^G. Thus, we have to show that $\langle \mathfrak{m}^G \rangle$ is generated by invariants of degree $\leq e$. To that end, it suffices to show that $\langle \mathfrak{m}^G \rangle$ is generated by elements (invariant or not) of degree $\leq e$ since each of them can be expressed in terms of invariants of degree $\leq e$.

By definition, $\mathfrak{m}^e \subseteq \langle \mathfrak{m}^G \rangle$. Since $\beta(A) \leq 1$ we have $\mathfrak{m}^e = A_{\geq e}$ and that ideal is generated by A_e. Thus, $\langle \mathfrak{m}^G \rangle$ is generated by $A^G_{\leq e-1}$ and A_e. This implies that $\langle \mathfrak{m}^G \rangle$ is generated by $A^G_{\leq e}$. □

For the next statement let $A(G)$ be the polynomial ring over A in $|G|$ variables $\{x_g \mid g \in G\}$ with the G-action extended by $gx_h := x_{gh}$. It is also $\mathbb{N}$-graded with $\deg x_g = 1$. We can write $A(G) = A \otimes_{\mathbb{Z}} \mathbb{Z}(G)$ where $\mathbb{Z}(G) = S^*(\mathbb{Z}[G])$ is the symmetric algebra over the (integral) regular representation of G. Using this, we see that $A(G)$ is actually bigraded where $a \in A_d$ and x_g have bidegree $(d, 0)$ and $(0, 1)$, respectively. The following result is a partial converse of Lemma 3.1.

3.2. Lemma. *For any A we have that $\eta_G(\mathfrak{m}) \leq \beta(A(G)^G)$.*

Proof: Let $b := \beta(A(G)^G)$. For $a \in A_e$ with $e \geq b$ let

$$\bar{a} := \sum_{u \in G} u(ax_1) = \sum_{u \in G} (ua)x_u \in A(G)^G. \tag{3.2}$$

Its bidegree is $(e, 1)$. Since $A(G)^G$ is generated by elements of degree $\leq b$ we can write $\bar{a}$ as a sum of products $p \cdot q$ where p and q are bihomogeneous G-invariants of bidegree $(c, 0)$ and $(e-c, 1)$ respectively. We may assume that $\deg q = e - c + 1 \leq b$. This implies $c \geq e - b + 1 \geq 1$. Thus, $p \in \mathfrak{m}^G$. Comparing the coefficient of x_1 in the expression $\bar{a} = \sum pq$, we obtain $a \in \langle \mathfrak{m}^G \rangle$. Thus $\mathfrak{m}^b \subseteq A_{\geq b} \subseteq \langle \mathfrak{m}^G \rangle$. □

Combining these comparison lemmas with Theorem 2.1, we obtain:

3.3. Theorem. *Assume $\beta(A) \le 1$ and $|G| \in A^\times$. Let $H \subseteq G$ be a subgroup. Then*

$$\beta(A^G) \le [G : H]\beta(A(H)^H). \tag{3.3}$$

Proof: Indeed,

$$\begin{aligned} \beta(A^G) &\le \eta_G(\mathfrak{m}) && \text{(by Lemma 3.1)} \\ &\le [G : H]\eta_H(\mathfrak{m}) && \text{(by Theorem 2.1)} \\ &\le [G : H]\beta(A(H)^H) && \text{(by Lemma 3.2)} \end{aligned} \tag{3.4}$$ □

Remark: Since Theorem 2.1 holds under the weaker assumption $[G : H] \in A^\times$ one might wonder whether also Theorem 3.3 holds in that generality.

Taking for H the trivial subgroup, we obtain Fleischmann's and Fogarty's theorem:

3.4. Corollary. *Assume $\beta(A) \le 1$ and $|G| \in A^\times$. Then $\beta(A^G) \le |G|$.*

This suggests a definition. For a fixed ground ring k let $\beta_k(G) := \sup_A \beta(A^G)$ where A runs through all $\mathbb{N}$-graded k-algebras with G-action and with $\beta(A) \le 1$. The following generalizes a theorem of Schmid [Sch] for $k = \mathbb{C}$:

3.5. Corollary. *Assume $|G| \in k^\times$ and let $H \subseteq G$ be a subgroup. Then*

$$\beta_k(G) \le [G : H]\beta_k(H). \tag{3.5}$$

In particular, $\beta_k(G) \le |G|$.

For arbitrary $\mathbb{N}$-graded algebras we have

3.6. Theorem. *Assume $|G| \in k^\times$ and let A be a k-algebra. Then $\beta(A^G) \le \beta_k(G)\beta(A)$.*

Proof: Let $b := \beta(A)$ and let $\overline{A}$ be the free commutative A_0-algebra over the A_0-module $\oplus_{d=1}^b A_d$. Then $\overline{A}$ has two degree functions, deg and deg$'$, where an element of A_d has degree d or 1, respectively. Clearly, we have $\deg \le b \deg'$. If we denote $\overline{A}$ equipped with the grading deg$'$ by $\tilde{A}$ then $\beta(\tilde{A}) \le 1$.

There is a natural map of $\overline{A}$ onto A. Hence $\overline{A}^G \to A^G$ is also onto. Therefore,

$$\beta(A^G) \le \beta(\overline{A}^G) \le b\,\beta(\tilde{A}^G) \le b\,\beta_k(G). \tag{3.6}$$ □

The following consequence is due to Schmid [Sch] for $k = \mathbb{C}$ and to Fleischmann [Fl1] in general:

3.7. Corollary. *Let $N \lhd G$ be a normal subgroup with $[G : N] \in k^\times$ and A a k-algebra. Then*

$$\beta(A^G) \leq \beta_k(G/N)\beta(A^N) \tag{3.7}$$

In particular, $\beta_k(G) \leq \beta_k(G/N)\beta_k(N)$.

Proof: Apply Theorem 3.6 to G/N acting on A^N. □

Now, we can state the following generalization of Schmid's result:

3.8. Theorem. *Let G be a non-cyclic group with $|G| \in k^\times$. Then $\beta_k(G) < |G|$.*

Proof: A close inspection shows that Schmid's proof works now (i.e., after having proved Corollary 3.5) for any $\mathbb{N}$-graded k-algebra. □

For completeness we state two more results without proof. In characteristic zero, Domokos and Hegedűs strengthened Schmid's theorem to:

3.9. Theorem. ([DH]) *Assume $k = \mathbb{C}$. Let G be a non-cyclic finite group. Then*

$$\beta_k(G) \leq \begin{cases} \frac{3}{4}|G| & \text{if } |G| \text{ is even} \\ \frac{5}{8}|G| & \text{if } |G| \text{ is odd} \end{cases} \tag{3.8}$$

The only part of the proof which cannot be easily carried over to positive characteristic is Schmid's calculation of β_k for dihedral groups. This has been done by Sezer who thereby proved

3.10. Theorem. ([Se]) *Theorem 3.9 holds for any field k whose characteristic does not divide $|G|$.*

In analogy to $\beta_k(G)$ we can define $\eta_k(G)$ as $\sup_{A,\mathfrak{m}} \eta_G(\mathfrak{m})$ where A runs through all (not necessarily graded) k-algebras with G-action and $\mathfrak{m}$ through all G-invariant ideals of A. Then Corollary 2.2 says $\eta_k(G) \leq |G|$ whenever $|G| \in k^\times$. In fact, we have:

3.11. Theorem. *For all finite groups G with $|G| \in k^\times$ holds $\eta_k(G) = \beta_k(G)$.*

Proof: Lemma 3.1 shows $\beta_k(G) \leq \eta_k(G)$. For the converse inequality, let $\mathfrak{m}$ be a G-invariant ideal in the G-ring A. Let $\tilde{A}$ be the blow-up algebra of A along $\mathfrak{m}$, i.e., $\tilde{A}$ is the $\mathbb{N}$-graded subring $\oplus_{d=0}^{\infty}\mathfrak{m}^d t^d$ of $A[t]$. Let $\tilde{\mathfrak{m}} := \tilde{A}_{\geq 1}$. Then Lemma 3.2 implies

$$\eta_G(\tilde{\mathfrak{m}}) \leq \beta(\tilde{A}(G)^G) \leq \beta_k(G). \tag{3.9}$$

Let $\pi : \tilde{A} \to A$ be the composition of $\tilde{A} \hookrightarrow A[t] \twoheadrightarrow A$ where t maps to 1. By linear reductivity, we have $\pi(\langle\tilde{\mathfrak{m}}^G\rangle) = \langle\mathfrak{m}^G\rangle$ and (obviously) $\pi(\tilde{\mathfrak{m}}^d) = \mathfrak{m}^d$. This implies $\eta_G(\mathfrak{m}) \leq \eta_G(\tilde{\mathfrak{m}})$, hence $\eta_k(G) \leq \beta_k(G)$. □

Remark: If $\operatorname{char} k$ divides $|G|$ then Richman has shown $\beta_k(G) = \infty$ (see Theorem 4.3 below). It is feasible that this is also true for $\eta_k(G)$. Then one could remove the assumption $|G| \in k^\times$ in Theorem 3.11. As an example, we compute $\eta_k(G)$ where $G = \{1, \sigma\}$ is the group of order 2 and $\operatorname{char} k = 2$. Let $V = k^{2n}$ with coordinates $x_1, \ldots, x_n, y_1, \ldots, y_n$ and assume $\sigma(x_i) = y_i$. Let $A = k[V]$. Then A^G is generated by all $x_i y_i$, $1 \leq i \leq n$ and $x^\alpha + y^\alpha$, $\alpha \in \mathbb{N}^n$. In particular, $\mathfrak{m}^G$ contains $x_i + y_i$. That means $A/\langle \mathfrak{m}^G \rangle \cong k[x_1, \ldots, x_n]/I$ where I is the ideal generated by all $-x_i^2$ and all $(1+(-1)^{|\alpha|})x^\alpha$ with $|\alpha| > 0$. In characteristic two, the second kind of elements are all zero. Therefore, $A/\langle \mathfrak{m}^G \rangle = k[x_1, \ldots, x_n]/(x_1^2, \ldots, x_n^2)$. Since $x_1 \ldots x_n \notin I$ we have $\eta_G(\mathfrak{m}) = n + 1$. Thus, $\eta_k(G) = \infty$.

4. Some functoriality properties

In this section, we develop some machinery to compute $\beta_k(A)$ for arbitrary ground rings k.

4.1. Theorem. *Let $\varphi : k \to K$ be a ring homomorphism. Then $\beta_k(G) \geq \beta_K(G)$ with equality if φ is faithfully flat.*

Proof: Let A be a graded K-algebra with G-action and $\beta(A) \leq 1$. By means of φ, we can consider A as a k-algebra. Since $\beta(A)$ and $\beta(A^G)$ depend only on the underlying ring structure, we get $\beta_k(G) \geq \beta_K(G)$.

Now assume that φ is faithfully flat and let A be a graded k-algebra with G-action and $\beta(A) \leq 1$. Put $B := A^G$, $\overline{A} := A \otimes_k K$ and $\overline{B} := B \otimes_k K$. Then $\overline{B} = \overline{A}^G$. Indeed, by flatness, the exact sequence

$$(4.1) \qquad 0 \longrightarrow B \longrightarrow A \overset{ga-a}{\longrightarrow} \oplus_{g \in G} A$$

stays exact upon tensoring with K. Now we claim $\beta(\overline{B}) \geq \beta(B)$. For $b := \beta(\overline{B})$ and any $d \in \mathbb{N}$ let C be the cokernel of

$$(4.2) \qquad \bigoplus_{\substack{\sum i_j = d \\ 0 \leq i_j \leq b}} B_{i_1} \otimes \ldots \otimes B_{i_s} \longrightarrow B_d.$$

This map becomes surjective after tensoring with K. Thus, $C \otimes_k K = 0$. Faithful flatness implies $C = 0$ and the claim follows.

We conclude $\beta(A^G) \leq \beta(\overline{A}^G) \leq \beta_K(G)$ for arbitrary A, hence $\beta_k(G) \leq \beta_K(G)$. □

4.2. Corollary. *Let k be a field. Then $\beta_k(A)$ depends only on the characteristic of k.*

In view of this result, we write $\beta_p(G) := \beta_k(G)$ where k is any field of characteristic p.

Next, we cite a result of Richman, [R], which shows that in the modular case there is no universal bound on the degree of generators:

4.3. Theorem. *Let k be a field whose characteristic divides $|G|$. Then $\beta(S_k^*(U)^G)$ is unbounded, where U runs through all finitely generated $k[G]$-modules. In particular, $\beta_k(G) = \infty$.*

This is being used in the proof of:

4.4. Proposition. *Let k be an arbitrary ring. Then, to compute $\beta_k(G)$ it suffices to take the supremum over all $\beta(S_k^*(U)^G)$ where U is a finitely generated $k[G]$-module.*

Proof: First assume that $|G|$ is not invertible in k. Then there is a prime divisor p of $|G|$ with $p \notin k^\times$, i.e., $pk \neq k$. Let $\mathfrak{m} \subset k$ be a maximal ideal containing pk and $K := k/\mathfrak{m}$. Then $\operatorname{char} K = p$ and the assertion follows from Theorem 4.3.

Thus, we may assume $|G| \in k^\times$. Let A be any graded k-algebra with G-action and $\beta(A) = 1$. Let $b := \beta(A^G)$. Then there is $f \in A_b^G$ which is not a polynomial of lower degree invariants. Clearly, we may assume $b \geq 1$. Hence, f is a polynomial of finitely many elements of A_1 with coefficients in k. Let U be the $k[G]$-module generated by these elements. Then we get an algebra homomorphism $S^*(U) \to A$ whose image contains f. Since G is linearly reductive, we can lift f to an invariant $\tilde{f} \in S_k^b(U)^G$. Clearly, we can't obtain $\tilde{f}$ as a polynomial of lower degree invariants, either. Hence $\beta(S^*(U)^G) \geq \beta(A^G)$ which shows the assertion. $\square$

Now we can reduce the computation of $\beta_k(G)$ for an arbitrary ring k to fields. For this we define

$$\operatorname{char}(k) := \begin{cases} \{0\} & \text{if } \mathbb{Q} \subseteq k; \\ \{p \mid p \notin k^\times\} & \text{otherwise.} \end{cases} \tag{4.3}$$

4.5. Theorem. *Let k be any ring. Then*

$$\beta_k(G) = \sup\{\beta_p(G) \mid p \in \operatorname{char}(k)\}. \tag{4.4}$$

Proof: If $\mathbb{Q} \subseteq k$ then k is faithfully flat over $\mathbb{Q}$ and the assertion follows from Theorem 4.1. Thus, assume $\mathbb{Q} \not\subseteq k$. If p is not invertible in k then there is a maximal ideal $\mathfrak{m}$ such that $k/\mathfrak{m}$ has characteristic p. Thus, $\beta_k(G)$ is at least as big as the left hand side of (4.4). In particular, we are done if one of the $\beta_p(G)$ is ∞. Thus we may assume that $|G|$ is invertible in k.

Let A be a $k[G]$-algebra which, by Proposition 4.4, we may assume to be of the form $A = S_k^*(U)$ where U is a finitely generated $k[G]$-module. This implies that each homogeneous component A_d is a finitely generated k-module. Being a direct summand the same holds for A_d^G.

Let $\mathfrak{m} \subset k$ be a maximal ideal. Then we claim that

$$A^G \otimes_k k/\mathfrak{m} \to (A \otimes_k k/\mathfrak{m})^G \tag{4.5}$$

is an isomorphism. Indeed, injectivity follows from the fact that A^G is a direct summand of A. Moreover, $A^G \to (A/\mathfrak{m}A)^G = (A \otimes_k k/\mathfrak{m})^G$ is surjective since $|G|$ is invertible in k.

Let b be the right hand side of (4.4) and let $C(b,d)$ be the cokernel of (4.2). Then $C(b,d) \otimes_k k/\mathfrak{m} = 0$ for every maximal ideal. Since $C(b,d)$ is finitely generated this implies $C(b,d) = 0$ (Nakayama Lemma). Thus $\beta(A^G) \leq b$ and therefore $\beta_k(G) \leq b$. $\square$

4.6. Corollary. *For any ring k holds $\beta_k(G) < \infty$ if and only if $|G| \in k^\times$.*

A similar argument shows that $k = \mathbb{Q}$ is the best possible case:

4.7. Theorem. *For all primes p holds $\beta_p(G) \geq \beta_0(G)$. For almost all primes holds equality.*

Proof: If p divides $|G|$ there is nothing to prove since then $\beta_p(G) = \infty$ (Theorem 4.3). Thus assume $p \not| \, |G|$. Let $k = \mathbb{Z}_p$ be the ring of p-adic integers with field of fractions $K = \mathbb{Q}_p$ and residue field $\mathbb{F}_p$. By Proposition 4.4, there is a finite dimensional $K[G]$-module U_K with $\beta(S_K^*(U_K)^G) = \beta_0(G) =: b_0$. Let $U_k \subseteq U_K$ be a G-stable lattice and $A := S_k^*(U_k)$. Let $\overline{A} := A \otimes_k \mathbb{F}_p$. Since $\overline{A}$ is an $\mathbb{F}_p$-algebra, it suffices to prove $\beta(\overline{A}^G) \geq b_0$.

The same argument as for (4.5) shows that

$$A^G \otimes_k \mathbb{F}_p \to (A \otimes_k \mathbb{F}_p)^G \tag{4.6}$$

is an isomorphism. Now let $B := A^G$ and consider for any b again the cokernel $C(b,d)$ of the map (4.2). For any $b < b_0$ there is a $d > 0$ such that $C(b,d) \otimes_k K \neq 0$. But $C(b,d)$ is a finitely generated k-module which implies $C(b,d) \otimes_k \mathbb{F}_p \neq 0$. This implies that $\beta(\overline{A}^G)$ can not be smaller than b.

Now we show that conversely $\beta_p(G) \leq \beta_0(G)$ for almost all primes. Let $k(G)$ be the symmetric algebra over the regular k-representation. Let $f_1, \ldots, f_r$ be a minimal set of generators of $\mathbb{Q}(G)^G$. Let $h_1, \ldots, h_s$ be generators of $\mathbb{Z}[\frac{1}{g}](G)^G$ where $g := |G|$. Then there is a multiple N of g such that the f_i are defined over $k := \mathbb{Z}[\frac{1}{N}]$ and such that all h_j are in the k-algebra generated by the f_i. This implies that $k(G)^G$ is also generated by the f_i. The homomorphism $k(G)^G \to \mathbb{F}_p(G)^G$ is surjective where p is any prime not dividing

N. This implies $\beta(\mathbb{F}_p(G)^G) \leq \beta(\mathbb{Q}(G)^G)$. We conclude with theorems of Schmid [Sch] and Smith [Sm] (see also its generalization, Corollary 7.5, further down) which assert $\beta(\mathbb{Q}(G)^G) = \beta_0(G)$ and $\beta(\mathbb{F}_p(G)^G) = \beta_p(G)$ for $p > |G|$. $\square$

Remark: Presently, no group G and prime p not dividing $|G|$ with $\beta_p(G) > \beta_0(G)$ seems to be known.

5. Polarization in positive characteristic

From now on, k will be a field which, for convenience, we assume to be infinite. We set $p = \operatorname{char} k$ if positive and $p = \infty$ otherwise. Let V be a k-vector space of dimension $\ell < \infty$. Let $k[V]_d = S^d(V^*)$ be the space of polynomial functions which are homogeneous of degree d. We are interested in the space $V^{\oplus n}$ of n-tuples of vectors in V.

There is a natural action of the algebraic group $GL(n)$ on $V^{\oplus n} = V \otimes_k k^n$. Let S be a subset of $k[V^{\oplus n}]$. Then the GL_n-submodule $\langle S \rangle_{GL(n)}$ generated by S is called the *polarization* of S. The action of the Lie algebra is given by the *polarization operators*

$$(5.1) \qquad P_{jj'} := \sum_{i=1}^{\ell} x_{ij} \frac{\partial}{\partial x_{ij'}}, \qquad 1 \leq j, j' \leq n$$

where $x_{1j}, \ldots, x_{\ell j}$ are coordinates of the j-th copy of V in $V^{\oplus n}$. If $m \leq n$ then we consider $V^{\oplus m}$ as a quotient of $V^{\oplus n}$ by forgetting the last $n-m$ components. In characteristic zero, it follows from work by Weyl ([We] II.5) that every function on V^n with $n \geq \ell := \dim V$ can be obtained by polarization from functions on V^ℓ. In positive characteristic this is true in sufficiently small degrees.

5.1. Theorem. *Let $n \geq m \geq \ell := \dim V$ and $d \leq (p-1)m$. Then*

$$(5.2) \qquad k[V^{\oplus n}]_d = \langle k[V^{\oplus m}]_d \rangle_{GL(n)}.$$

Proof: First, by replacing V with $V \oplus k^{m-\ell}$ we may assume that $m = \ell$. Moreover, it clearly suffices to treat the case $n = m+1 = \ell+1$. The proof will proceed by induction on ℓ starting at the trivial case $\ell = 0$.

We label the copies of $V^{\oplus \ell+1}$ by $j = 0, \ldots, \ell$ such that $V^{\oplus \ell}$ corresponds to $j = 1, \ldots, \ell$. The coordinates of the j-th copy are denoted by $x_{1j}, \ldots, x_{\ell j}$. A monomial in the x_{ij} can be represented by an $\mathbb{N}$-valued matrix

$$(5.3) \qquad A = \begin{pmatrix} a_{10} & a_{11} & \ldots & a_{1\ell} \\ \vdots & \vdots & \iddots & \vdots \\ a_{\ell 0} & a_{\ell 1} & \ldots & a_{\ell\ell} \end{pmatrix}$$

where $x^A := \prod_{ij} x_{ij}^{a_{ij}}$. Its degree is $\deg A := \sum_{ij} a_{ij}$.

We introduce a total order on the set of matrices A as follows. Let $r_i = r_i(A) = \sum_j a_{ij}$ and $c_j = c_j(A) = \sum_i a_{ij}$ be the row and column sums, respectively. Then we define the *index* of A as the vector

$$\operatorname{ind} A = (c_0, \ldots, c_\ell, r_1, \ldots, r_\ell, a_{10}, \ldots, a_{1\ell}, a_{20}, \ldots, a_{2\ell}, \ldots, a_{\ell 0}, \ldots, a_{\ell\ell}). \tag{5.4}$$

In words: $\operatorname{ind} A$ starts with the column sums followed by the row sums followed by the coefficients read from left to right, top to bottom. We say $A < B$ if $\operatorname{ind} A$ is lexicographically smaller that $\operatorname{ind} B$. We have to show that $\deg A = d \le (p-1)\ell$ implies $x^A \in X(\ell, d) := \langle k[V^{\oplus \ell}]_d \rangle_{GL(\ell+1)}$. Let A be the smallest counterexample.

First, observe that $X(\ell, d)$ is invariant under both $GL(\ell)$ (acting on V) and $GL(\ell+1)$ (by definition). Thus, the property $x^A \in X(\ell, d)$ is invariant under permutations of the rows or the columns of A. The minimality of A implies

$$c_0(A) \le \ldots \le c_\ell(A) \quad \text{and} \tag{5.5}$$
$$r_1(A) \le \ldots \le r_\ell(A). \tag{5.6}$$

From $d = \sum_i r_i(A) \le (p-1)\ell$ we obtain

$$r_1(A) \le p - 1. \tag{5.7}$$

The matrix A must have a non-zero entry in the first column since otherwise $x^A \in k[V^{\oplus \ell}]_d \subseteq X(\ell, d)$. A fortiori, there is non-zero entry a_{ij} with $i + j \le \ell$ (i.e., which is strictly above the dotted diagonal in (5.3)). Let a_{ij} be the first one, where we read from left to right and top to bottom. Then a_{ij} is not in the last column and the entries above a_{ij+1} are all zero.

Let $\overline{A}$ be the matrix obtained from A by replacing the entries a_{ij} and a_{ij+1} by $a_{ij} - 1$ and $a_{ij+1} + 1$, respectively. Since $c_j(\overline{A}) < c_j(A)$, the minimality of A implies $x^{\overline{A}} \in X(\ell, d)$. Now apply the polarization operator P_{jj+1} to $x^{\overline{A}}$. Then we get

$$P_{jj+1}(x^{\overline{A}}) = (a_{ij+1} + 1)x^A + a_{i+1j+1}x^{A_{i+1}} + \ldots + a_{\ell j+1}x^{A_\ell} \in X(\ell, d). \tag{5.8}$$

Here, A_ν is the matrix

$$\begin{pmatrix} & \vdots & \vdots & \\ \cdots & a_{ij} - 1 & a_{ij+1} + 1 & \cdots \\ & \vdots & \vdots & \\ \cdots & a_{\nu j} + 1 & a_{\nu j+1} - 1 & \cdots \\ & \vdots & \vdots & \end{pmatrix} \tag{5.9}$$

Since all A_ν are smaller than A we obtain[1] $a_{ij+1} + 1 = 0$ in k. In particular, we get

$$a_{ij} \geq 1 \quad \text{and} \quad a_{ij+1} \geq p - 1. \tag{5.10}$$

Let $\ell' := \ell + 1 - i \geq j + 1$. Then $a_{i\ell'}$ is on the diagonal starting from $a_{1\ell}$:

$$A = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & \cdots & 0 & \cdots & a_{1\ell} \\ \vdots & & & & & \vdots & \iddots & \vdots \\ 0 & \cdots & a_{ij} & a_{ij+1} & \cdots & a_{i\ell'} & & \vdots \\ \vdots & & & & \iddots & \vdots & & \vdots \\ \vdots & & & \iddots & & \vdots & & \vdots \\ \vdots & & \iddots & & & \vdots & & \vdots \\ a_{\ell 0} & a_{\ell 1} & \cdots & \cdots & \cdots & a_{\ell\ell'} & \cdots & a_{\ell\ell} \end{pmatrix} \tag{5.11}$$

Let A' be the submatrix of A consisting of columns $0, \ldots, \ell'$. From (5.10) we get $r_i \geq p$ which implies $\ell' < \ell$ by (5.7). Moreover we have $p - 1 \leq c_{j+1} \leq c_{\ell'+1} \leq \ldots \leq c_\ell$. Hence

$$\deg A' = c_0 + \ldots + c_{\ell'} \leq (p-1)\ell - (p-1)(\ell - \ell') = (p-1)\ell'. \tag{5.12}$$

Now observe that the first $i - 1 = \ell - \ell'$ rows of A' are zero. Hence, the monomial $x^{A'}$ "lives" on $(V')^{\oplus \ell'+1}$ where $\dim V' = \ell'$. Thus we can use induction and conclude that $x^{A'}$ can be obtained by polarization from $(V')^{\oplus \ell'}$. But then the same polarization process produces x^A from $V^{\oplus \ell'}$ in contradiction to the choice of A. $\square$

Remarks: 1. The proof shows that under the given conditions it suffices to apply just polarization operators and column permutations.

2. One can show that the monomial $x_{10}(x_{11} \ldots x_{1\ell})^{p-1}$ cannot be obtained from polarization of polynomials on fewer than $\ell + 1$ copies of V. This shows that the given degree bound is optimal.

6. Weyl's theorem

Let k and p be as in the preceding section. We are going to apply polarizations to invariant theory. Let U and V be two finite dimensional representations of a finite group G. Then the $GL(n)$-action on $U \oplus V^{\oplus n}$ commutes with G, hence we get an $GL(n)$-action and a notion of polarization on $k[U \oplus V^{\oplus n}]^G$. If $m \leq n$ we can restrict invariants from $U \oplus V^{\oplus n}$ to $U \oplus V^{\oplus m}$. This process is kind of inverse to polarization and is called *restitution.*

6.1. Theorem. *Let S be a generating set of G-invariants on $U \oplus V^{\oplus n}$ where $n \geq \max(\dim V, \frac{\beta_k(G)}{p-1})$. Then the polarization of S generates the ring of invariants on $U \oplus V^{\oplus m}$ for any $m \geq n$.*

---

[1] In characteristic 0, we are done at this point: A can't exist.

Proof: First, let us remark that G is linearly reductive. Indeed, otherwise $\beta_k(G) = \infty$ and m would not exist.

Let $\mathcal{P}_m \subseteq \operatorname{End} k[U \oplus V^{\oplus m}]$ be the subalgebra generated by all polarization and permutation operators. Let $A(m)$ be the algebra generated by $\mathcal{P}_m \cdot S \subseteq k[U \oplus V^{\oplus m}]$. Since $\mathcal{P}_m$ commutes with G we have $A(m) \subseteq k[U \oplus V^{\oplus m}]^G$. We have to show equality.

The formulas $P(fg) = P(f)g + fP(g)$ for any polarization operator and $\pi(fg) = \pi(f)\pi(g)$ for any permutation imply that

$$(6.1) \qquad \mathcal{P}_m(fg) \subseteq \mathcal{P}_m(f)\mathcal{P}_m(g)$$

for all f, g. This shows that $A(m)$ is a $\mathcal{P}_m$-module. Let $d \leq \beta_k(G)$. Then the assumption on n and Theorem 5.1 imply that the map

$$(6.2) \qquad \mathcal{P}_m \otimes k[U \oplus V^{\oplus n}]_d \to k[U \oplus V^{\oplus m}]_d$$

is surjective. By linear reductivity of G we obtain

$$(6.3) \qquad \mathcal{P}_m \cdot k[U \oplus V^{\oplus n}]_d^G = k[U \oplus V^{\oplus m}]_d^G.$$

On the other hand,

$$(6.4) \qquad \mathcal{P}_m \cdot k[U \oplus V^{\oplus n}]_d^G = \mathcal{P}_m \cdot k[S]_d \subseteq \mathcal{P}_m A(m)_d = A(m)_d.$$

Thus, we have proved that $A(m)$ contains all G-invariants of degree $\leq \beta_k(G)$ which implies that it consists of all invariants. $\square$

This theorem provides a means to construct a representation of G which has the "most general" ring of invariants. To make this more precise, we extend our notion of polarization. Let U be a G-module. For any subset S of $k[U]$ let $\operatorname{Pol}^U(S)$ be the $\operatorname{Aut}^G(U)$-module generated by it. Observe, that if $|G| \in k^\times$ then $\operatorname{Aut}^G(U)$ is just a product of general linear groups corresponding to the isotypic components of U. There are two more operations: If $V \subseteq U$ then $\operatorname{Res}_V^U(S)$ be the set of restrictions of elements of S to V. Conversely, if U is a quotient of V let $\operatorname{Ext}_U^V(S)$ be the set of pullbacks. Now, we define

Definition: A G-module U has *universal invariants* if for every finite dimensional G-module V the set

$$(6.5) \qquad \operatorname{Res}_V \operatorname{Pol}^{U \oplus V} \operatorname{Ext}_U^{U \oplus V}(S)$$

generates $k[V]^G$ where S is any generating set of $k[U]^G$.

In other words, if U has universal invariants then one obtains a generating set of invariants for any other G-module by the following process: 1. Start with generators for $k[U]^G$. 2.

Think of them as invariants on $U \oplus V$. 3. Apply $\mathrm{Aut}^G(U \oplus V)$ to S. 4. Restrict the ensuing invariants to V.

One of the main properties of a module with universal invariants is:

6.2. Theorem. *Assume U has universal invariants. Then $\beta_k(G) = \beta(k[U]^G)$.*

Proof: First, $\beta_k(G)$ is the supremum of the $\beta(k[V]^G$ where V runs through all finite dimensional G-modules (Proposition 4.4). Then the assertion follows from the fact that the process of extension, polarization, and restriction does not change degrees. $\square$

Our main criterion for having universal invariants is:

6.3. Theorem. *Assume $|G| \in k^\times$. Let U be a representation of G in which every irreducible module M appears with multiplicity at least* $\max(\dim M, \frac{\beta_k(G)}{p-1})$. *Then U has universal invariants.*

Proof: Let S be a generating set of $k[U]^G$ and let V be a finite dimensional G-module. We show by induction on the number of isotypic components of V the more general statement that $\mathrm{Aut}^G(U \oplus V) \cdot S$ generates $k[U \oplus V]^G$. We start with $V = 0$ where the assertion is trivial.

Assume that the simple module M appears in V. Write $U = U' \oplus M^{\oplus m}$ and $V = V' \oplus M^{\oplus n}$ such that M is not contained in U' and V'. The set $S' := \mathrm{Aut}^G(U \oplus V') \cdot S$ generates $k[U \oplus V']^G$ by the induction hypothesis. We have $U \oplus V' = (U' \oplus V') \oplus M^{\oplus m}$ while $U \oplus V = (U' \oplus V') \oplus M^{\oplus m+n}$. Theorem 6.1 implies that $k[U \oplus V]^G$ is generated by $Gl(m+n) \cdot S' = Gl(m+n) \cdot \mathrm{Aut}^G(U \oplus V') \cdot S = \mathrm{Aut}^G(U \oplus V) \cdot S$. $\square$

6.4. Corollary. *Assume $|G| \in k^\times$. Then G has a finite dimensional representation with universal invariants. More precisely, assume that U contains every irreducible module with multiplicity at least $|G|$. Then U has universal invariants.*

Proof: Let M be an irreducible representation of G. Since M is a quotient of the regular representation, we have $\dim M \le |G|$. On the other hand, $\frac{\beta_k(G)}{p-1} \le |G|$ by Corollary 3.5. $\square$

6.5. Corollary. *Assume $p > \beta_k(G)$. Then the regular representation of G has universal invariants.*

Proof: We may assume that k is algebraically closed. Then the regular representation contains each simple module W with multiplicity $\dim W$. The assertion follows from $\frac{\beta_k(G)}{p-1} \le 1 \le \dim W$. $\square$

7. A further improvement

One can try to obtain better results by generalizing the polarization process. The point is that in positive characteristic, the commutator algebra of $GL(V)$ in $\operatorname{End}(k[V^{\oplus m}]$ is not generated by the image of $GL(m)$ (not even in a topological sense).

In general, one can replace $\mathcal{P}_m$ by any subalgebra $\overline{\mathcal{P}}_m$ of endomorphisms of $k[V^{\oplus m}]$ which commutes with the $GL(V)$-action and which satisfies the multiplicativity property (6.1). We will not pursue this direction in any detail but illustrate it by the following easy fact:

7.1. Theorem. *Assume* $\dim V = 1$ *and* $\overline{\mathcal{P}}_m := \operatorname{End}^{GL(V)}(k[V^{\oplus m}])$. *Then for any* $d \geq 0$ *holds*

$$k[V^{\oplus m}]_d = \langle k[V]_d \rangle_{\overline{\mathcal{P}}_m}. \tag{7.1}$$

Proof: The group $GL(V) \cong GL(1)$ acts on $W_d := k[V^{\oplus m}]_d$ by the character $t \mapsto t^d$. Hence, $\overline{\mathcal{P}}_m = \prod_{d=0}^{\infty} \operatorname{End}_k W_d$. This implies that any non-zero element generates W_d as a $\overline{\mathcal{P}}_m$-module. ☐

Remark: This theorem also serves as an example that $\overline{\mathcal{P}}_m$ may be bigger than $\mathcal{P}_m$. In fact, $k[V^{\oplus m}]_d$ is a simple $\overline{\mathcal{P}}_m$-module but, in general, it is not a simple $GL(m)$-module.

7.2. Lemma. *Let* $\overline{\mathcal{P}}_m$ *be as in Theorem 7.1. Then for any* $f, g \in k[V^{\oplus m}]$ *holds* $\overline{\mathcal{P}}_m(fg) \subseteq \overline{\mathcal{P}}_m(f)\overline{\mathcal{P}}_m(g)$.

Proof: Let $f = \sum_d f_d$ be the decomposition into homogeneous components. Since $\overline{\mathcal{P}}_m$ contains the projection onto $k[V^{\oplus m}]_d$ we have $f_d \in \overline{\mathcal{P}}_m(f)$. Hence, we may assume that f and g are homogeneous of degree, say, d and e, respectively. But then $\overline{\mathcal{P}}_m(f) = k[V^{\oplus m}]_d$ and $\overline{\mathcal{P}}_m(g) = k[V^{\oplus m}]_e$, hence $\overline{\mathcal{P}}_m(f)\overline{\mathcal{P}}_m(g) = k[V^{\oplus m}]_{d+e} \supseteq \overline{\mathcal{P}}_m(fg)$. ☐

Now assume $|G| \in k^\times$. For an arbitrary G-module V we generalize the polarization process as follows. Let $V = \oplus_i M_i^{m_i}$ be the isotypic decomposition of V. Let $\overline{\mathcal{P}}_V \subseteq \operatorname{End} k[V]$ be the subalgebra which is generated by $GL(m_i)$ if $\dim M_i > 1$ and by $\overline{\mathcal{P}}_{m_i}$ if $\dim M_i = 1$. For a subset $S \subseteq k[V]$ let $\overline{\operatorname{Pol}}^V(S) := \overline{\mathcal{P}}_V(S)$.

We say a module U has *weakly universal invariants* if $k[V]^G$ is generated by

$$\operatorname{Res}_V \overline{\operatorname{Pol}}^{U\oplus V} \operatorname{Ext}_U^{U\oplus V}(S) \tag{7.2}$$

where V is any finite dimensional G-module and S is any generating set of $k[U]^G$. Then we get the following analogue of Theorem 6.3

7.3. Theorem. *Assume* $|G| \in k^\times$*Let* U *be a representation of* G *in which every one-dimensional module appears at least once and every other irreducible module* M *appears with multiplicity at least* $\max(\dim M, \frac{\beta_k(G)}{p-1})$. *Then* U *has weakly universal invariants.*

The analogue of Corollary 6.5 is

7.4. Corollary. *Assume k is algebraically closed. Let*

$$\ell := \inf\{\dim W \mid W \textit{ simple } G\textit{-module, } \dim W > 1\} \in [2, \infty]. \tag{7.3}$$

Assume $p \geq \frac{\beta_k(G)}{\ell} + 1$. Then the regular representation R of G has weakly universal invariants. In particular, $\beta_k(G) = \beta(k[R]^G)$.

If we combine this result with the bound of Domokos-Hegedűs-Sezer we obtain the following strengthening of theorems of Schmid, [Sch], and Smith, [Sm].

7.5. Corollary. *Assume $|G| \in k^\times$ and $p \geq \frac{3}{8}|G| + 1$. Then $\beta_k(G) = \beta(k[R]^G)$.*

Proof: If G is cyclic then $\ell = \infty$. Otherwise $\ell \geq 2$ and $\beta_k(G) \leq \frac{3}{4}|G|$ by Theorem 3.10. □

Remark: The condition in [Sch] is $\operatorname{char} k = 0$. In [Sm] it is $p > |G|$.

8. References

[DH] Domokos, M.; Hegedűs, P.: Noether's bound for polynomial invariants of finite groups. *Arch. Math.* **74** (2000), 161–167

[Fl1] Fleischmann, P.: The Noether bound in invariant theory of finite groups. *Adv. Math.* **156** (23–32), 2000

[Fl2] Fleischmann, P.: On invariant theory of finite groups. *Preprint* (2002), 49 pages

[Fo] Fogarty, J.: On Noether's bound for polynomial invariants of a finite group. *Electron. Res. Announc. Amer. Math. Soc.* **7** (2001), 5–7 (electronic)

[N] Noether, E.: Der Endlichkeitssatz der Invarianten endlicher Gruppen. *Math. Ann.* **77** (1916), 89–92

[R] Richman, D.: Invariants of finite groups over fields of characteristic p. *Adv. Math.* **124** (1996), 25–48

[Sch] Schmid, B.: Finite groups and invariant theory. In: *Topics in invariant theory (Paris, 1989/1990).* (M.-P. Malliavin ed.) Lecture Notes in Mathematics **1478**, Berlin: Springer-Verlag 1991, 35–66

[Se] Sezer, M.: Sharpening the generalized Noether bound in the invariant theory of finite groups. *J. Algebra* **254** (2002), 252–263

[Sm] Smith, L.: On a theorem of Barbara Schmid. *Proc. Amer. Math. Soc.* **128** (2000), 2199-2201

[We] Weyl, Hermann: The Classical Groups. Their Invariants and Representations. (Princeton Mathematical Series **1**) Princeton: Princeton University Press 1939