ZERO-SEPARATING INVARIANTS FOR FINITE GROUPS

JONATHAN ELMER AND MARTIN KOHLS

Abstract. We fix a field $\Bbbk$ of characteristic p. For a finite group G denote by $\delta(G)$ and $\sigma(G)$ respectively the minimal number d, such that for any finite dimensional representation V of G over $\Bbbk$ and any $v \in V^G \setminus \{0\}$ or $v \in V \setminus \{0\}$ respectively, there exists a homogeneous invariant $f \in \Bbbk[V]^G$ of positive degree at most d such that $f(v) \neq 0$. Let P be a Sylow-p-subgroup of G (which we take to be trivial if the group order is not divisble by p). We show that $\delta(G) = |P|$. If $N_G(P)/P$ is cyclic, we show $\sigma(G) \geq |N_G(P)|$. If G is p-nilpotent and P is not normal in G, we show $\sigma(G) \leq \frac{|G|}{l}$, where l is the smallest prime divisor of $|G|$. These results extend known results in the non-modular case to the modular case.

1. Introduction

Let G be a linear algebraic group over an algebraically closed field $\Bbbk$, V a finite dimensional rational representation of G (which we will call a G-module), and denote by $\Bbbk[V]$ the ring of polynomial functions $V \to \Bbbk$. The action of G on V induces an action of G on $\Bbbk[V]$ via $g \cdot f(v) := f(g^{-1}v)$ for $g \in G$, $f \in \Bbbk[V]$ and $v \in V$. The set of G-invariant polynomial functions under this action is denoted by $\Bbbk[V]^G$, and inherits a natural grading from $\Bbbk[V]$, since the given action is degree-preserving. We denote by $\Bbbk[V]^G_d$ the set of polynomial invariants of degree d and the zero-polynomial, and by $\Bbbk[V]^G_{\leq d}$ the set of polynomial invariants of degree at most d. For any subset S of $\Bbbk[V]$, we define S_+ as those elements of S with constant term zero.

A linear algebraic group G is said to be *reductive* if for every G-module V we have that, for all nonzero $v \in V^G$, there exists $f \in \Bbbk[V]^G_+$ such that $f(v) \neq 0$. It is said to be *linearly reductive* if for all nonzero $v \in V^G$ there exists $f \in \Bbbk[V]^G_1$ such that $f(v) \neq 0$. Obviously linear reductivity implies reductivity. Denote by $\mathcal{N}_{G,V}$ the nullcone of V, that is

$$\mathcal{N}_{G,V} = \left\{ v \in V | \quad f(v) = 0 \quad \text{ for all } f \in \Bbbk[V]^G_+ \right\}.$$

Note that the nullcone is the vanishing set of the "Hilbert Ideal" $I_{G,V}$ of $\Bbbk[V]$, which is the ideal of $\Bbbk[V]$ generated by all homogeneous invariants of positive degree. Then G is reductive if for any G-module V, one has that $\mathcal{N}_{G,V} \cap V^G = \{0\}$.

The concept of reductivity is important in both invariant theory and the theory of linear algebraic groups. One of the most celebrated results of 20th century invariant theory is the theorem of Nagata [12] and Popov [13] which states that $\Bbbk[X]^G$ is finitely generated for all affine G-varieties X if and only if G is reductive.

The first part of the present article is motivated by a simple, perhaps even facetious, question: are there any "quadratically reductive" groups? The reader can probably guess the definition, but we explain this in detail, while introducing some useful terminology. Let G be a linear algebraic group over $\Bbbk$ and V a G-module. We shall say a subset $S \subseteq \Bbbk[V]^G$ is a δ-*set* if, for all $v \in V^G \setminus \mathcal{N}_{G,V}$,

Date: August 5, 2013.

there exists an $f \in S_+$ such that $f(v) \neq 0$. We shall call a subalgebra of $\Bbbk[V]^G$ a *δ-subalgebra* if it is a δ-set. The quantity $\delta(G,V)$ is then defined as

$$\delta(G,V) = \min\{d \geq 0| \quad \Bbbk[V]^G_{\leq d} \text{ is a } \delta\text{-set}\}.$$

We will justify below that $\delta(G,V)$ is always a finite number. Finally, we define

$$\delta(G) := \sup\{\delta(G,V)| \quad V \text{ a } G\text{-module}\},$$

where we take the supremum of an unbounded set to be infinity.

Note that if G is reductive over $\Bbbk$, the definitions above simplify: $S \subseteq \Bbbk[V]^G$ is then a δ-set if and only if for all nonzero $v \in V^G$, there is an $f \in S_+$ with $f(v) \neq 0$. Note further, that a reductive group G is linearly reductive if and only if $\delta(G) = 1$. A quadratically reductive group, then, ought be a reductive group G for which $\delta(G) = 2$. There are plenty of examples. We show in Section 2:

Theorem 1.1. *Let G be a finite group, $\Bbbk$ an algebraically closed field of characteristic p, and P a Sylow-p-subgroup of G. Then $\delta(G) = |P|$.*

It is well known that a finite group G is linearly reductive over a field $\Bbbk$ if and only if the order of G is not divisible by the characteristic of $\Bbbk$. The theorem above can be viewed as a generalisation of this result, where we take the Sylow-p-subgroup to be trivial in the non-modular case.

In addition to $\delta(G)$, we also study the closely related quantity $\sigma(G)$. The definition is as follows. We shall say a subset $S \subseteq \Bbbk[V]^G$ is a *σ-set* if, for all $v \in V \setminus \mathcal{N}_{G,V}$, there exists an $f \in S_+$ such that $f(v) \neq 0$.

We shall call a subalgebra of $\Bbbk[V]^G$ a *σ-subalgebra* if it is a σ-set. The quantity $\sigma(G,V)$ is then defined as

$$\sigma(G,V) = \min\{d \geq 0| \quad \Bbbk[V]^G_{\leq d} \text{ is a } \sigma\text{-set}\}.$$

It is clear that a generating set of the Hilbert ideal $I_{G,V}$ which consists of invariants is a σ-set. Therefore, since $\Bbbk[V]$ is Noetherian, $\Bbbk[V]^G$ always contains a finite σ-set and the number $\sigma(G,V)$ is finite. Finally, we define

$$\sigma(G) := \sup\{\sigma(G,V)| \quad V \text{ a } G \text{ -module}\},$$

which can be finite or infinite. It is immediately clear that $\delta(G,V) \leq \sigma(G,V)$ for all G-modules V, and that $\delta(G) \leq \sigma(G)$. It is also well known that $\sigma(G) \leq |G|$, e.g. from Dade's Algorithm [5, Proposition 3.3.2].

Note that $\sigma(G,V)$ can be interpreted in a few different ways. For instance, we see that $\sigma(G,V)$ is the minimal degree d such that there exists a finite set of invariants of degree at most d whose common zero set is $\mathcal{N}_{G,V}$. If G is reductive, then a graded subalgebra $S \subseteq \Bbbk[V]^G$ is a σ-subalgebra if and only if $\Bbbk[V]^G$ is a finitely generated S-module (see [5, Lemma 2.4.5]). So in the case of reductive groups, $\sigma(G,V)$ is the minimal degree d such that there exists a set T of homogeneous invariants of degree at most d such that $\Bbbk[V]^G$ is a finitely generated $\Bbbk[T]$-module. Recall that for reductive groups, $\mathcal{N}_{G,V}$ consists of those $v \in V$ such that $0 \in \overline{G \cdot v}$, where the bar denotes closure in the Zariski topology (see e.g. [5, Lemma 2.4.2]). In particular when G is finite we have that $\mathcal{N}_{G,V} = \{0\}$, so $\sigma(G,V)$ may be defined as the minimal degree d such that there exists a finite set of invariants of degree at most d whose common zero set is $\{0\}$.

For linearly reductive groups in characteristic 0, the σ-number plays an important role in giving upper bounds for the classical Noether number $\beta(G,V) = \beta(\Bbbk[V]^G)$, which is defined as the minimum degree d such that $\Bbbk[V]^G_{\leq d}$ generates $\Bbbk[V]^G$ as an algebra. Again, the "global" value $\beta(G)$ is defined as the supremum of all $\beta(G,V)$. For example, Derksen [4, Theorem 1.1] gives the upper bound

$$\beta(G,V) \leq \max\left\{2\ ,\ \frac{3}{8}\cdot\dim(\Bbbk[V]^G)\cdot\sigma(G,V)^2\right\}.$$

Cziszter and Domokos [3] study $\sigma(G)$ for finite groups over fields of characteristic not dividing $|G|$. In particular, they show

Proposition 1.2 (Cziszter and Domokos [3])**.** *Let G be a finite group, and let $\Bbbk$ be an algebraically closed field of characteristic not dividing $|G|$. Then $\sigma(G) = |G|$ if and only if G is cyclic. More precisely, if G is not cyclic, then $\sigma(G) \leq |G|/l$ where l is the smallest prime dividing $|G|$.*

Proof. [3, Theorem 7.1] states that if G is not cyclic, then $\sigma(G) \leq |G|/l$ where l is the smallest prime dividing $|G|$. In particular $\sigma(G) < |G|$ when G is not cyclic. Conversely, [3, Corollary 5.3] states that if G is abelian, $\sigma(G) = \exp(G)$. In particular, if G is cyclic, $\sigma(G) = |G|$. □

In sections 3 and 4 we generalise some results of Cziszter and Domokos to fields of arbitrary characteristic. In particular, we prove the following version of the above for the modular case (where $N_G(P)$ is the normalizer of the subgroup P in G):

Theorem 1.3. *Suppose G is a finite group, and that $\Bbbk$ is an algebraically closed field of characteristic p, where p divides $|G|$. Let P be a Sylow-p-subgroup of G. Also let l denote the smallest prime divisor of $|G|$. Then the following holds:*

(a) *If $\sigma(G) = |G|$, then $N_G(P)/P$ is a cyclic group. If additionally P is abelian and $G \neq P$, then $N_G(P)/P$ is also non-trivial.*
(b) *If $N_G(P)/P$ is cyclic, then $\sigma(G) \geq |N_G(P)|$. In particular $\sigma(G) = |G|$ when P is normal in G and G/P is cyclic.*
(c) *If G is p-nilpotent and P is not normal, then $\sigma(G) \leq \frac{|G|}{l}$.*

Another quantity associated with $\delta(G,V)$ and $\sigma(G,V)$, which has attracted some attention in recent years, is $\beta_{\mathrm{sep}}(G,V)$. It is defined as follows: a subset $S \subseteq \Bbbk[V]^G$ is called a *separating set* if, for any pair $v, w \in V$ such that there exists $f \in \Bbbk[V]^G$ with $f(v) \neq f(w)$, there exists $s \in S$ with $s(v) \neq s(w)$. Then $\beta_{\mathrm{sep}}(G,V)$ is defined as

$$\beta_{\mathrm{sep}}(G,V) = \min\{d \geq 0 |\quad \Bbbk[V]^G_{\leq d} \text{ is a separating set}\},$$

and once more, the "global" value $\beta_{\mathrm{sep}}(G)$ is defined to be the supremum over all $\beta_{\mathrm{sep}}(G,V)$.

Our point of view is that δ- and σ-sets are "zero-separating" sets. This leads to the following inequalities:

Proposition 1.4. *Let G be a linear algebraic group and V a G-module. Then*

$$\delta(G,V) \leq \sigma(G,V) \leq \beta_{\mathrm{sep}}(G,V) \leq \beta(G,V).$$

Proof. The first and last inequalities are trivial. Assume $S \subseteq \Bbbk[V]^G_+$ is a separating set. It is enough to show that S cuts out the Nullcone. Indeed, if $v \in V \setminus \mathcal{N}_{G,V}$, then there is an $f \in \Bbbk[V]^G_+$ such that $f(v) \neq 0 = f(0)$. Thus there is an $s \in S$ such that $s(v) \neq s(0) = 0$. Consequently, if $S \subseteq \Bbbk[V]^G_+$ is a separating set then it is a σ-set, and we get the second inequality. □

The above implies that one has, for any linear algebraic group G,

$$\delta(G) \leq \sigma(G) \leq \beta_{\mathrm{sep}}(G) \leq \beta(G).$$

For finite groups, Derksen and Kemper [5, Theorem 3.9.13] showed that $\beta_{\mathrm{sep}}(G) \leq |G|$, independently of the characteristic of $\Bbbk$. For this reason we obtain as a consequence of Theorem 1.1, for a finite group G with Sylow-p-subgroup P,

$$|P| = \delta(G) \leq \sigma(G) \leq \beta_{\mathrm{sep}}(G) \leq |G|.$$

Fleischmann [8] and Fogarty [9] proved independently that if p does not divide the order of G, then we have the stronger result that $\beta(G) \leq |G|$ (the result in characteristic zero is due to Emmy Noether, hence the name). In that case we obtain

$$1 = \delta(G) \leq \sigma(G) \leq \beta_{\mathrm{sep}}(G) \leq \beta(G) \leq |G|.$$

In this paper we focus mainly on the case where G is a finite group. However, a subsequent paper dealing with infinite algebraic groups is in preparation. As for some of the results in the present paper the proofs for infinite groups are not more difficult than those for finite groups, we will give the proofs for the most general case.

2. The δ-number for finite groups

The goal of this section is to prove Theorem 1.1, which we do in a series of basic propositions.

Proposition 2.1. *Let G be a reductive group and let U be a G-submodule of V. Then $\delta(G,U) \leq \delta(G,V)$.*

Proof. Let $d = \delta(G,V)$ and take $u \in U^G \setminus \mathcal{N}_{G,U}$. Clearly $u \in V^G \setminus \{0\}$, and reductivity implies $u \notin \mathcal{N}_{G,V}$. It follows that there exists an $f \in \Bbbk[V]^G_{+,\leq d}$ with $f(u) \neq 0$. Now set $g := f|_U$. Then we have $g \in \Bbbk[U]^G_{+,\leq d}$ and $g(u) \neq 0$. This shows that $\delta(G,U) \leq d$. □

Proposition 2.2. *Let G be a reductive group, V_1, V_2 be G modules and $W = V_1 \oplus V_2$. Then $\delta(G,W) = \max\{\delta(G,V_1), \delta(G,V_2)\}$.*

Proof. We have $d := \max\{\delta(G,V_1), \delta(G,V_2)\} \leq \delta(G,W)$ by the previous proposition. Take $w = v_1 + v_2 \in W^G \setminus \mathcal{N}_{G,W}$ with $v_i \in V_i$ for $i = 1,2$. Then $v_i \in V_i^G$ for $i = 1,2$. Clearly $w \neq 0$, hence $v_1 \neq 0$ or $v_2 \neq 0$. Without loss of generality assume $v_1 \neq 0$. Reductivity implies $v_1 \in V_1^G \setminus \mathcal{N}_{G,V_1}$. Hence there exists an $f \in \Bbbk[V_1]^G_{+,\leq \delta(G,V_1)}$ with $f(v_1) \neq 0$. As we have the G-algebra inclusion $\Bbbk[V_1] \subseteq \Bbbk[V_1 \oplus V_2]$, f can be viewed as an element of $\Bbbk[W]^G_{+,\leq d}$ satisfying $0 \neq f(v_1) = f(v_1 + v_2) = f(w)$. This shows that $\delta(G,W) \leq d$ as required. □

Remark 2.3. Using the above and induction, it follows that

$$\delta(G,W) = \max\{\delta(G,V_i)| \quad i = 1, \dots, n\}$$

whenever $W = \bigoplus_{i=1}^n V_i$ is a finite direct sum of G-modules.

Proposition 2.4. *Let G be a finite group. Then $\delta(G) = \delta(G, V_{\mathrm{reg}})$ where $V_{\mathrm{reg}} := \Bbbk G$ denotes the regular representation of G over $\Bbbk$.*

Proof. It is well-known that, given any G-module V, we have an embedding $V \hookrightarrow V_{\mathrm{reg}}^n$ for $n = \dim_{\Bbbk}(V)$ (choosing an arbitrary basis of V^* yields an epimorphism $(\Bbbk G)^n \twoheadrightarrow V^*$, and dualizing yields the desired embedding as $V_{\mathrm{reg}} = \Bbbk G$ is self dual - see also [7, proof of Corollary 3.11]). Now by Proposition 2.1 and Remark 2.3 we obtain

$$\delta(G,V) \leq \delta(G, V_{\mathrm{reg}}^n) = \delta(G, V_{\mathrm{reg}}).$$

The result now follows from the definition of $\delta(G)$. □

The proof of the following Proposition, which is key to proving Theorem 1.1, is similar to [10, Proposition 8], but our point of view is different and we get a new result. Also note that if G is a p-group, Theorem 1.1 and Propositon 1.4 imply $|G| = \delta(G) = \sigma(G) = \beta_{\mathrm{sep}}(G)$, strengthening the result in [10, Proposition 8].

Proposition 2.5. *Let G be a finite group, $\Bbbk$ a field of characteristic p, and let P be a Sylow-p-subgroup of G (if $p = 0$ or does not divide the order of G, take P to be the trivial group). Then $\delta(G, V_{\text{reg}}) = |P|$.*

Proof. Let $\{v_g |\ g \in G\}$ be a $\Bbbk$-basis for $V := V_{\text{reg}}$. The fixed point space V^G of V is 1-dimensional and spanned by $v := \sum_{g\in G} v_g$. Write $\Bbbk[V] = \Bbbk[x_g : g \in G]$ where $\{x_g |\ g \in G\}$ is the basis of V^* dual to $\{v_g |\ g \in G\}$. Since V is a permutation representation, the ring of invariants $\Bbbk[V]^G$ is generated as a vector space by *orbit sums* of monomials, that is, by invariants of the form

$$o_G(m) := \sum_{m'\in G\cdot m} m'$$

where $m := \prod_{g\in G} x_g^{n_g}$ is a monomial in $\Bbbk[x_g : g \in G]$ and $G \cdot m$ denotes the orbit of m. Clearly then for any $g \in G$ we have $x_g(v) = 1$, and therefore for any monomial $m \in \Bbbk[V]$ we have $m(v) = 1$. It follows that for any monomial m, we have

$$o_G(m)(v) = \sum_{m'\in G\cdot m} m'(v) = \sum_{m'\in G\cdot m} 1 = |G \cdot m|.$$

Now let $0 \neq u \in V^G$. Then $u = \lambda v$ for some nonzero $\lambda \in \Bbbk$. Set $m := \prod_{g\in P} x_g$ and $f := o_G(m)$. Note that f is an invariant of degree $|P|$, and that

$$f(u) = \lambda^{|P|}|G \cdot m| = \lambda^{|P|}[G : \operatorname{Stab}_G(m)] = \lambda^{|P|}\frac{|G|}{|P|} \neq 0 \in \Bbbk.$$

This shows that $\delta(G, V) \leq |P|$.

Conversely, any $f \in \Bbbk[V]^G$ can be written as a $\Bbbk$-sum of orbit sums of monomials. Therefore, if $f(v) \neq 0$ for some homogeneous invariant f, for some monomial m of the same degree as f we must have $o_G(m)(v) \neq 0$. This means that $|G \cdot m| = [G : \operatorname{Stab}_G(m)]$ is not divisible by p. Hence, $|P|$ divides $|\operatorname{Stab}_G(m)|$. Therefore $\operatorname{Stab}_G(m)$ contains a Sylow-p-subgroup Q of G. Consequently, if m is divisible by some x_g^k, m must also be divisible by x_{qg}^k for all $q \in Q$. In particular, $\deg(m) \geq |Q| = |P|$. This shows that $\deg(f) \geq |P|$, and hence $\delta(G, V) \geq |P|$. □

Proof of Theorem 1.1. Combine Propositions 2.4 and 2.5. □

3. Relative results for the σ-number

In this section we prove mainly relative results about $\sigma(G)$ for both finite and infinite groups G. Many of these are extensions of results in [3] to fields of arbitrary characteristic and to infinite groups.

Proposition 3.1. *Let G be a reductive group and let U be a G-submodule of V. Then $\sigma(G, U) \leq \sigma(G, V)$.*

Proof. Let $d := \sigma(G, V)$ and take $u \in U \setminus \mathcal{N}_{G,U}$. This implies $0 \notin \overline{G \cdot u}$. As U is a closed subset of V, it does not matter if the closure of $G \cdot u$ is taken in U or in V. Now the reductivity of G implies $u \notin \mathcal{N}_{G,V}$, and therefore there exists an $f \in \Bbbk[V]^G_{+,\leq d}$ with $f(u) \neq 0$. Then $f|_U \in \Bbbk[U]^G_{+,\leq d}$ also separates u from 0, hence $\sigma(G, U) \leq d$. □

Note that for non-reductive groups, it is not always the case that $U \subseteq V$ implies $U \setminus \mathcal{N}_{G,U} \subseteq V \setminus \mathcal{N}_{G,V}$. For example take the action of the additive group $\mathbb{G}_a = (\Bbbk, +)$ on $V = \Bbbk^2$ via $t * (a, b) := (a + tb, b)$ for $t \in \mathbb{G}_a$ and $(a, b) \in V$. We write $\Bbbk[V] = \Bbbk[x, y]$. Take the point $u = (1, 0)$ in the submodule $U := \Bbbk \cdot (1, 0)$. As the action of $\mathbb{G}_a$ on U is trivial, $\Bbbk[U]^{\mathbb{G}_a} = \Bbbk[x|_U]$, so we have $u \in U \setminus \mathcal{N}_{\mathbb{G}_a,U}$. But $\Bbbk[V]^{\mathbb{G}_a} = \Bbbk[y]$, so $u \in \mathcal{N}_{\mathbb{G}_a,V}$.

For arbitrary (even non-reductive) algebraic groups, we have the following result:

Lemma 3.2. *Let G be an arbitrary group and let U and V be G-modules such that U is a direct summand of V. Then $\sigma(G,U) \leq \sigma(G,V)$.*

Proof. Take a $u \in U \setminus \mathcal{N}_{G,U}$ and an $f \in \Bbbk[U]_+^G$ such that $f(u) \neq 0$. As U is a direct summand of V, we have an inclusion of G-algebras $\Bbbk[U] \subseteq \Bbbk[V]$, hence we can view f as an element of $\Bbbk[V]_+^G$. As $f(u) \neq 0$, we have $u \in V \setminus \mathcal{N}_{G,V}$. Therefore there is a $g \in \Bbbk[V]^G_{+,\leq \sigma(G,V)}$ such that $g(u) \neq 0$. Then $g|_U \in \Bbbk[U]^G_{+,\leq\sigma(G,V)}$ satisfies $g|_U(u) \neq 0$, hence $\sigma(G,U) \leq \sigma(G,V)$. □

The following basic result also appears in Cziszter and Domokos [3, Lemma 5.1], but we give a simpler argument here:

Proposition 3.3. *Let G be a finite group and suppose $W = V_1 \oplus V_2$, where V_1, V_2 and W are G-modules. Then $\sigma(G,W) = \max\{\sigma(G,V_1), \sigma(G,V_2)\}$.*

Proof. We have $d := \max\{\sigma(G,V_1), \sigma(G,V_2)\} \leq \sigma(G,W)$ by Proposition 3.1. Conversely, take a nonzero $w = v_1 + v_2 \in W$ with $v_i \in V_i$ for $i = 1,2$. Without loss we can assume $v_1 \neq 0$. Then there is an $f \in \Bbbk[V_1]^G_{+,\leq\sigma(G,V_1)}$ with $f(v_1) \neq 0$. As in the proof of Proposition 2.2, we can view f as an element of $\Bbbk[W]^G_{+,\leq d}$ such that $f(w) = f(v_1+v_2) = f(v_1) \neq 0$. Therefore, $\sigma(G,W) \leq d$. □

Note that the above is not true for reductive algebraic groups in general; a counterexample is provided in [3, Remark 5.2]. However, even for infinite groups the σ-value of vector invariants has an interesting stabilization property, which was observed by Domokos [6, Remark 3.3]. As Domokos only remarks that the proof of the following proposition can be given with the same methods as in his paper [6] (where a similar result for β_{sep} is given), we give the proof here for the sake of completeness.

Proposition 3.4 (Domokos)**.** *Assume G is an arbitrary (possibly infinite) group acting linearly on an n-dimensional vector space V (the action need not even be rational). Then*

$$\sigma(G,V^m) = \sigma(G,V^n) \qquad \textit{for all } \ m \geq n.$$

Note that for finite groups, by Proposition 3.3 we have more precisely $\sigma(G,V^m) = \sigma(G,V)$ for all m.

Under the hypotheses of the theorem, we first show the following:

Lemma 3.5. *Let $v = (v_1,\ldots,v_m)$ and $u = (u_1,\ldots,u_m) \in V^m$ be such that their components span the same $\Bbbk$-vector subspace of V, i.e.*

$$\langle v_1,\ldots,v_m\rangle_\Bbbk = \langle u_1,\ldots,u_m\rangle_\Bbbk.$$

Then we have

$$v \in \mathcal{N}_{G,V^m} \Leftrightarrow u \in \mathcal{N}_{G,V^m}.$$

Proof. Assume $v \notin \mathcal{N}_{G,V^m}$. Then there is an $f \in \Bbbk[V^m]_+^G$ with $f(v) \neq 0$. By assumption, we can find $\alpha_{ij} \in \Bbbk$ such that

$$v_i = \sum_{j=1}^m \alpha_{ij} u_j \qquad \text{for all } i = 1,\ldots,m.$$

Write $f = f(x_1,\ldots,x_m) \in \Bbbk[V^m]^G$, where each x_j belongs to a set of coordinates of an element of V, and set

$$h(x_1,\ldots,x_m) := f\left(\sum_{j=1}^m \alpha_{1,j}x_j, \sum_{j=1}^m \alpha_{2,j}x_j, \ldots, \sum_{j=1}^m \alpha_{m,j}x_j\right) \in \Bbbk[V^m]_+.$$

It is immediately checked that h inherits G-invariance from f, so $h \in \Bbbk[V^m]_+^G$. Now

$$\begin{aligned} h(u) = h(u_1, \ldots, u_m) &= f\left(\sum_{j=1}^m \alpha_{1,j} u_j, \sum_{j=1}^m \alpha_{2,j} u_j, \ldots, \sum_{j=1}^m \alpha_{m,j} u_j\right) \\ &= f(v_1, \ldots, v_m) = f(v) \neq 0, \end{aligned}$$

hence $u \notin \mathcal{N}_{G,V^m}$. We have shown: If $\langle v_1, \ldots, v_m\rangle_\Bbbk \subseteq \langle u_1, \ldots, u_m\rangle_\Bbbk$, then we have the implication: $v \notin \mathcal{N}_{G,V^m} \Rightarrow u \notin \mathcal{N}_{G,V^m}$. The reverse implication follows in the same way, so we are done. □

Proof of Proposition 3.4. By Lemma 3.2 we have $\sigma(G, V^n) \leq \sigma(G, V^m)$, so we have to show the reverse inequality. Take a point $v = (v_1, \ldots, v_m) \in V^m \setminus \mathcal{N}_{G,V^m}$. As $\dim V = n$, we can find a point $u = (u_1, \ldots, u_n, 0, \ldots, 0) \in V^m$ such that $\langle u_1, \ldots, u_n\rangle = \langle v_1, \ldots, v_m\rangle$. By Lemma 3.5 we have $u \notin \mathcal{N}_{G,V^m}$. Hence there is an $f \in \Bbbk[V^m]_+^G$ such that $f(u) \neq 0$. Then $f|_{V^n} \in \Bbbk[V^n]_+^G$ satisfies $f(\tilde{u}) \neq 0$, where $\tilde{u} = (u_1, \ldots, u_n) \in V^n$. Therefore, $\tilde{u} \notin \mathcal{N}_{G,V^n}$, so there is an $\tilde{f} \in \Bbbk[V^n]^G_{+,\leq\sigma(G,V^n)}$ such that $\tilde{f}(\tilde{u}) \neq 0$. As we have a G-algebra inclusion $\Bbbk[V^n] \subseteq \Bbbk[V^m]$, we can take $\tilde{f}$ as an element of $\Bbbk[V^m]^G_{+,\leq\sigma(G,V^n)}$ satisfying $\tilde{f}(u) = \tilde{f}(\tilde{u}) \neq 0$. As in the proof of Lemma 3.5, there is an $h \in \Bbbk[V^m]^G_{+,\leq\sigma(G,V^n)}$ satisfying $h(v) = \tilde{f}(u) \neq 0$. This shows $\sigma(G, V^m) \leq \sigma(G, V^n)$. □

Now we restrict again to finite groups and give two corollaries of Propositions 3.1 and 3.3.

Corollary 3.6. *Let G be a finite group. Then $\sigma(G) = \sigma(G, V_{\mathrm{reg}})$ where V_{reg} denotes the regular representation of G over $\Bbbk$.*

Proof. As the proof of Proposition 2.4. □

Recall that the decomposition of the regular representation into indecomposables gives the complete list of projective indecomposable modules. As a consequence of this, Proposition 3.3 and the above corollary, we have

Corollary 3.7. *Let G be a finite group. Then*

$$\sigma(G) = \max\{\sigma(G, U) \mid U \textit{ is a projective indecomposable } G\textit{-module}\}.$$

Proposition 3.8. *Let G be a group and let N be a normal subgroup of G with finite index. Let V be a G-module. Then*

$$\sigma(G, V) \leq \sigma(N, V)\sigma(G/N) \leq \sigma(N)\sigma(G/N),$$

so particularly we have $\sigma(G) \leq \sigma(N)\sigma(G/N)$.

Proof. Only the first inequality needs to be shown. Choose a finite σ-subset $\{f_1, f_2, \ldots, f_n\}$ of $\Bbbk[V]_+^N$, with $\deg(f_i) \leq \sigma(N, V)$ for all $i = 1, \ldots, n$. Take a left-transversal $\{g_1, g_2, \ldots, g_r\}$ of N in G, that is to say, $G = \bigcup_{i=1}^r g_i N$ where $r = [G : N]$. Let $v \in V \setminus \mathcal{N}_{G,V}$. As a G-invariant separating v from zero is clearly also an N-invariant, we see that $v \in V \setminus \mathcal{N}_{N,V}$. Consequently, the vector

$$(f_1(v), f_2(v), \ldots, f_n(v)) \in \Bbbk^n$$

is not zero, and nor is the vector

$$\hat{v} := (g_1(f_1)(v), g_2(f_1)(v), \ldots, g_r(f_1)(v), \ldots, g_1(f_n)(v), \ldots, g_r(f_n)(v)) \in \Bbbk^{nr}.$$

We may define an action on $\Bbbk^{nr}$ so that it becomes isomorphic to n copies of the regular representation of G/N, i.e. to $V^n_{\mathrm{reg},G/N}$ in such a way that the action of G/N on $\hat{v}$ is given by

$$(1)\quad g^{-1}N \cdot \hat{v} = ((gg_1)(f_1)(v), \ldots, (gg_r)(f_1)(v), \ldots, (gg_1)(f_n)(v), \ldots, (gg_r)(f_n)(v))$$

for all $g \in G$. Since G/N is finite, its nullcone is zero, and as $\hat{v} \neq 0$ we can find an invariant $h \in \Bbbk[V^n_{\mathrm{reg},G/N}]^{G/N}_{+,\leq\sigma(G/N)}$ such that $h(\hat{v}) \neq 0$. Now consider the polynomial

$$\hat{h} := h(g_1(f_1), g_2(f_1), \dots, g_r(f_1), \dots, g_1(f_n), \dots, g_r(f_n)) \in \Bbbk[V]_+.$$

Notice that $\hat{h}(v) = h(\hat{v}) \neq 0$, and that $\deg(\hat{h}) \leq \sigma(N,V)\sigma(G/N)$. It remains to show that $\hat{h}$ is G-invariant. From the definition of the action of G/N on $\Bbbk^{nr}$, we see that for any $g \in G$ and $u \in V$ we have

$$\begin{aligned}
(g\hat{h})(u) &= h(gg_1(f_1), \dots, gg_r(f_1), \dots, gg_1(f_n), \dots, gg_r(f_n))(u) \\
&= h(gg_1(f_1)(u), \dots, gg_r(f_1)(u), \dots, gg_1(f_n)(u), \dots, gg_r(f_n)(u)) \\
&\overset{(1)}{=} h(g^{-1}N \cdot (g_1(f_1)(u), \dots, g_r(f_1)(u), \dots, g_1(f_n)(u), \dots, g_r(f_n)(u))) \\
&= (gN \cdot h)(g_1(f_1)(u), \dots, g_r(f_1)(u), \dots, g_1(f_n)(u), \dots, g_r(f_n)(u)) \\
&\overset{(*)}{=} h(g_1(f_1)(u), \dots, g_r(f_1)(u), \dots, g_1(f_n)(u), \dots, g_r(f_n)(u)) = \hat{h}(u),
\end{aligned}$$

where in $(*)$ we used that h is G/N invariant. Hence, $g(\hat{h}) = \hat{h}$ for all $g \in G$, that is $\hat{h} \in \Bbbk[V]^G$, so we are done. □

Proposition 3.9. *Let G be a group and let H be a subgroup of G with finite index. Let V be a G-module. Then*

$$\sigma(G,V) \leq \sigma(H,V)[G:H] \leq \sigma(H)[G:H],$$

so in particular we have $\sigma(G) \leq \sigma(H)[G:H]$.

Proof. As in Proposition 3.8, we can find a finite σ-subset $\{f_1, f_2, \dots, f_n\} \subseteq \Bbbk[V]^H_+$ with the property that $\deg(f_i) \leq \sigma(H,V)$ for all i. Let $\{g_1, g_2, \dots, g_r\}$ be a left-transversal of H in G. Take a new independent variable T on which G acts trivially, and form the polynomial ring $\Bbbk[V][T]$. As the polynomial $\sum_{j=1}^n f_j T^{j-1}$ is H-invariant, its relative "norm"

$$z(T) := \prod_{i=1}^r \left(g_i \left(\sum_{j=1}^n f_j T^{j-1} \right) \right) = \prod_{i=1}^r \left(\sum_{j=1}^n g_i(f_j) T^{j-1} \right) \in \Bbbk[V]_+[T].$$

is G-invariant, hence the coefficients of z as a polynomial in T are G-invariant. Let $S \subseteq \Bbbk[V]^G_+$ be this set of coefficients of z. We claim that S is a σ-set. Suppose that $v \in V$ is such that $f(v) = 0$ for all $f \in S$. We must show that $v \in \mathcal{N}_{G,V}$. We have that $z(T)(v)$ is the zero polynomial, i.e.

$$\prod_{i=1}^r \left(\sum_{j=1}^n f_j(g_i^{-1}v) T^{j-1} \right) = 0 \in \Bbbk[T].$$

Since $\Bbbk[T]$ is an integral domain, this implies that one of the factors of the above is zero, that is, for some $i \in \{1, \dots, r\}$,

$$\sum_{j=1}^n f_j(g_i^{-1}v) T^{j-1} = 0 \in \Bbbk[T].$$

This implies that $f_j(g_i^{-1}v) = 0$ for all $j = 1, \dots, n$. Since the set $\{f_1, f_2, \dots, f_n\} \subseteq \Bbbk[V]^H_+$ is a σ-set, we deduce that $f(g_i^{-1}v) = 0$ for all $f \in \Bbbk[V]^H_+$. In particular, $f(g_i^{-1}v) = (g_i(f))(v) = 0$ for all $f \in \Bbbk[V]^G_+$. This means that $f(v) = 0$ for all $f \in \Bbbk[V]^G_+$, i.e. that $v \in \mathcal{N}_{G,V}$ as required. □

The following is the first statement of Theorem 1.3 (a):

Corollary 3.10. *Let G be a finite group, and let $\Bbbk$ be an algebraically closed field of characteristic p. Let P be a Sylow-p-subgroup of G (if p does not divide $|G|$, take P the trivial group) and suppose $N_G(P)/P$ is not cyclic. Then $\sigma(G) < |G|$.*

Proof. By Propositions 3.9, 3.8 and 1.2, we have

$$\begin{aligned}\sigma(G) &\leq \sigma(N_G(P))[G : N_G(P)] \leq \sigma(P)\sigma(N_G(P)/P)[G : N_G(P)]\\ &< |P|[N_G(P) : P][G : N_G(P)] = |G|.\end{aligned}$$

□

Lemma 3.11. *Assume G is a reductive group with a closed subgroup H of finite index, and V a G-module. Then $\sigma(H, V) \leq \sigma(G, V)$.*

Proof. Let $v \in V \setminus \mathcal{N}_{H,V}$. Clearly this implies $0 \notin \overline{Hv}$. Let $g_1, \ldots, g_r$ be a left transversal of H in G. Then we have

$$\overline{Gv} = \overline{\left(\bigcup_{i=1}^{r} g_i H\right) \cdot v} = \overline{\bigcup_{i=1}^{r} g_i \cdot (Hv)} = \bigcup_{i=1}^{r} \overline{g_i \cdot (Hv)} = \bigcup_{i=1}^{r} g_i \cdot \overline{Hv}.$$

For the last equation, note that each g_i induces a homeomorphism of topological spaces $V \to V$ with inverse map g_i^{-1}. Also note that in an arbitrary topological space, one has the general rule $\overline{A \cup B} = \overline{A} \cup \overline{B}$ for subsets A and B, which justifies the previous equation. Now assume for a contradiction $0 \in \overline{Gv}$. Then $0 \in g_i \cdot \overline{Hv}$ for some i, hence $0 = g_i^{-1} 0 \in \overline{Hv}$, a contradiction. Therefore, $0 \notin \overline{Gv}$, and as G is reductive there is an $f \in \Bbbk[V]^G_{+,\leq\sigma(G,V)}$ with $f(v) \neq 0$. As f is clearly H-invariant, this shows that $\sigma(H, V) \leq \sigma(G, V)$. □

Note that this lemma does not hold for arbitrary subgroups: Take the action of the multiplicative group $\mathbb{G}_m = (\Bbbk \setminus \{0\}, \cdot)$ on $V = \Bbbk$ by left multiplication and H the subgroup generated by a primitive nth root of unity. Then $\Bbbk[V]^{\mathbb{G}_m} = \Bbbk$ and $\Bbbk[V]^H = \Bbbk[x^n]$, hence $\sigma(\mathbb{G}_m, V) = 0$ while $\sigma(H, V) = n$.

Proposition 3.12. *Let G be a linear algebraic group, $H \subseteq G$ a closed subgroup of finite index and V an H-module. Then*

$$\sigma(H, V) \leq \sigma(G, \mathrm{Ind}_H^G(V)) \leq \sigma(G),$$

so in particular we have $\sigma(H) \leq \sigma(G)$.

Proof. There is a natural H-equivariant embedding $V \hookrightarrow \mathrm{Ind}_H^G(V)$, which turns V into an H-submodule of $\mathrm{Ind}_H^G(V)$. The restriction map $\Phi : \Bbbk[\mathrm{Ind}_H^G(V)]^G \to \Bbbk[V]^H$, $f \mapsto f|_V$ is well defined and decreases degrees. Let $S \subseteq \Bbbk[\mathrm{Ind}_H^G(V)]^G_+$ be a σ-set for G. We will show that $\Phi(S) \subseteq \Bbbk[V]^H_+$ is a σ-set for H, which proves the proposition. Take $v \in V$ such that $\Phi(f)(v) = f(v) = 0$, for all $f \in S$. Since S is a σ-set for G, this means $f(v) = \Phi(f)(v) = 0$ for all $f \in \Bbbk[\mathrm{Ind}_H^G(V)]^G_+$. By the proof of Schmid [15, Proposition 5.1], the map Φ is surjective, so we have $f(v) = 0$ for all $f \in \Bbbk[V]^H_+$. Thus, $\Phi(S) \subseteq \Bbbk[V]^H_+$ is a σ-set for H. □

An immediate consequence of Propositions 3.9 and 3.12 is

Corollary 3.13. *Let G be a linear algebraic group, with G^0 the connected component of G containing the identity. We have the inequalities*

$$\sigma(G) \leq [G : G^0]\sigma(G^0) \quad \textit{and} \quad \sigma(G^0) \leq \sigma(G).$$

In particular, $\sigma(G)$ and $\sigma(G^0)$ are either both finite or infinite.

Remark 3.14. If G is a linear algebraic group and N a closed normal subgroup of G, then we have

$$\sigma(G/N) \leq \sigma(G).$$

This follows from the fact that every G/N module can be turned into a G-module via the canonical map $G \to G/N$.

Propositions 3.8 and 3.9 are proved in [3] under the assumption that $[G : N]$ is not divisible by p. Our proofs are rather similar to [10, Theorem 2]. The proof of Proposition 3.12 is similar to [10, Corollary 1].

4. The σ-number for finite groups

We now specialize to the case of finite groups. Throughout this section we work over an algebraically closed field $\Bbbk$ of characteristic p, which is assumed to divide $|G|$. Our first result is a generalisation of [3, Corollary 5.3] to the modular case.

Theorem 4.1. *Let G be a group of the form $P \times A$, where P is a p-group and A is an abelian group of order not divisible by p. Then $\sigma(G) = |P| \exp(A)$.*

Proof. We have $\sigma(G) \leq [G : A]\sigma(A) = |P|\sigma(A)$ by Proposition 3.9. By [3, Corollary 5.3], $\sigma(A) = \exp(A)$ when $|A|$ is not divisible by p, so we have proved $\sigma(G) \leq |P| \exp(A)$. It remains to show that $\sigma(G) \geq |P| \exp(A)$. To this end, let W be a 1-dimensional $\Bbbk A$-module with character of order $e := \exp(A)$, and consider the $P \times A$-module $V := \Bbbk P \otimes_{\Bbbk} W$, where P acts on only the first factor and A on only the second. We write $\{v_g \mid g \in P\}$ for a basis of V on which P acts via the regular representation, with $\{x_g \mid g \in P\}$ the dual basis. Notice that a homogeneous $f \in \Bbbk[V]$ is A-invariant if and only if $\deg(f)$ is divisible by e. Now consider the point $v := \sum_{g\in P} v_g \in V$. Take a homogeneous $f \in \Bbbk[V]^G_+$ such that $f(v) \neq 0$. As all monomials are eigenvectors under the A-action, every monomial appearing in f is A-invariant. As $\Bbbk[V]^P$ is linearly spanned by orbit sums of monomials $o_P(m)$, $f(v) \neq 0$ implies there exists an A–invariant monomial m of the same degree as f such that $o_P(m)(v) \neq 0$. The proof of Proposition 2.5 implies that $m = (\prod_{g\in P} x_g)^d$ for some $d \in \mathbb{N}$. In order for m to be A-invariant, it follows $e|\deg(m) = d|P|$. Since $|P|$ and e are coprime, this means that $e|d$, i.e. $\deg(f) \geq e|P|$, and so $\sigma(G, V) \geq e|P| = \exp(A)|P|$ as required. □

From this result we obtain immediately:

Corollary 4.2. *For any cyclic group G, we have $\sigma(G) = |G|$.*

The following is part (b) of Theorem 1.3.

Theorem 4.3. *Let G be a finite group. Assume P is a Sylow-p-subgroup of G such that $N_G(P)/P$ is cyclic. Then $\sigma(G) \geq |N_G(P)|$.*

Proof. Firstly, $\sigma(G) \geq \sigma(N_G(P))$ by Proposition 3.12, so we may assume $G = N_G(P)$. It is enough to find a G-module V with $\sigma(G, V) \geq |G|$. Set $r := |G/P|$ and let $\zeta \in \Bbbk$ be a primitive rth root of unity. By the Schur-Zassenhaus Lemma (see [14, Theorem 7.41]), P has a complement H in G. Let t be a generator of H. Define a $\Bbbk G$-module as follows: a $\Bbbk$-basis is given by $\{v_g \mid g \in P\}$, with dual basis $\{x_g \mid g \in P\}$. The action of P on V is via the regular representation, while the action of H is given via $t^i \cdot v_g := \zeta^{-i} v_{t^i g t^{-i}}$ for any $i \in \mathbb{Z}$ and $g \in P$.

Let $v := \sum_{g\in P} v_g$, and let $f \in \Bbbk[V]^G_+$ be homogeneous such that $f(v) \neq 0$. Once more, $\Bbbk[V]^P$ is linearly spanned by orbit sums of monomials $o_P(m)$. As $r = |H|$

is invertible in $\Bbbk$, $\Bbbk[V]^G$ is linearly spanned by elements $s_m := \sum_{i=0}^{r-1}(t^i \cdot o_P(m))$. Therefore, there exists a monomial m of the same degree as f such that

$$0 \neq s_m(v) = \sum_{i=0}^{r-1}(t^i \cdot o_P(m))(v) = \sum_{i=0}^{r-1} \zeta^{\deg(m)i}|Pm|.$$

If the rightmost sum is to be non-zero, we must have again $|Pm| = 1$, that is, m must be P-invariant, i.e. of the form $(\prod_{g \in P} x_g)^d$ for some $d \geq 1$. It follows

$$0 \neq s_m(v) = \sum_{i=0}^{r-1} \zeta^{d|P|i}.$$

However, the sum on the right hand side is non-zero if and only if $r|d$. Therefore, m (hence f) has to be of degree at least $|P|r = |G|$. This shows that $\sigma(G, V) \geq |G|$ as required. □

Corollary 4.4. *Let G be a finite group. Assume P is a Sylow-p-subgroup of G and $\sigma(G) = |P|$. Then $N_G(P) = |P|$.*

Proof. Assume for a contradiction $N_G(P) \supsetneq P$ and take a $g \in N_G(P) \setminus P$. Then the subgroup $H := \langle P, g \rangle$ of G satisfies $N_H(P) = H$ and H/P is cyclic. Hence by Propostion 3.12 and Theorem 4.3 we would have

$$\sigma(G) \geq \sigma(H) = |H| > |P|,$$

a contradiction. □

We will write Z_n for a cyclic group of order n, which if convenient we identify with $\mathbb{Z}/n\mathbb{Z}$. Recall $\mathrm{Aut}(Z_n) \cong (\mathbb{Z}/n\mathbb{Z})^\times$, which is cyclic when n is prime.

Proposition 4.5. *Assume p, q are primes and $d \in \mathbb{N}$ such that $p|d$ and $d|q-1$. Take an embedding $Z_d \hookrightarrow \mathrm{Aut}(Z_q)$ and form the corresponding semidirect product $Z_q \rtimes Z_d$. Then over a field of characteristic p, we have $\sigma(Z_q \rtimes Z_d) = q$.*

Note that over a field of characteristic q, $\sigma(Z_q \rtimes Z_d) = dq$ by Theorem 4.3, and in the non-modular case, $\sigma(Z_q \rtimes Z_d) = q$ by Cziszter and Domokos [3, Proposition 6.2]. This proof here is an adapted version to the modular case of the latter proof by Cziszter and Domokos. We want to thank Cziszter for explaining some details of their exposition to us via eMail. In the proof we will use a decomposition of the regular representation of $G = Z_q \rtimes Z_d$ into a direct sum of (not necessarily indecomposable) smaller modules which we construct below. We write $G = \langle g, h \rangle$ such that

$$\mathrm{ord}(g) = d, \qquad \mathrm{ord}(h) = q,$$

and set $H := \langle h \rangle \cong Z_q$ and $D := \langle g \rangle \cong Z_d$. Then with $k + q\mathbb{Z}$ a suitable element of multiplicative order d in $\mathbb{Z}/q\mathbb{Z}$ we have

$$g^a h^b = h^{k^a b} g^a \qquad \text{for all } a, b \in \mathbb{Z}.$$

For convenience, we will write k^{-a} for a suitable representative of the class $(k + q\mathbb{Z})^{-a}$. Then V_{reg} has a basis $\{v_{g^j h^r} \mid j = 0, \ldots, d-1,\ r = 0, \ldots, q-1\}$. We choose a primitive qth root of unity $\zeta \in \Bbbk$ and define

$$w_{i,j} := \sum_{r=0}^{q-1} \zeta^{-ir} v_{g^j h^r} \in V_{\mathrm{reg}} \qquad \text{for } j = 0, \ldots, d-1,\ i = 0, \ldots, q-1.$$

Lemma 4.6. *For all $i = 0, \ldots, q-1$, the vector space*

$$V_i := \langle w_{i,0}, w_{i,1}, \ldots, w_{i,d-1} \rangle$$

is a G-submodule of V_{reg}, and we have a decomposition

$$V_{\text{reg}} = \bigoplus_{i=0}^{q-1} V_i.$$

The action of G on the summands is given by

$$g^a \cdot w_{i,j} = w_{i,(j+a \mod d)} \quad \text{and} \quad h^b \cdot w_{i,j} = (\zeta^i)^{k^{-j}b} w_{i,j}$$

for $j = 0, \ldots, d-1,\ i = 0, \ldots, q-1$.

Proof. As $(\zeta^{-ir})_{i,r=0,\ldots,q-1} \in \Bbbk^{q \times q}$ is a Vandermonde matrix of full rank, we obtain for any $j = 0, \ldots, d-1$ the equality of vector subspaces

$$\langle w_{0,j}, w_{1,j}, \ldots, w_{q-1,j} \rangle = \langle v_{g^j h^0}, v_{g^j h^1}, \ldots, v_{g^j h^{q-1}} \rangle.$$

Therefore, the set $\{w_{i,j}\}_{i,j}$ is a basis of V_{reg} and we get the desired direct sum decomposition as vector spaces. We also see that

$$g^a \cdot w_{i,j} = g^a \cdot \sum_{r=0}^{q-1} \zeta^{-ir} v_{g^j h^r} = \sum_{r=0}^{q-1} \zeta^{-ir} v_{g^{a+j} h^r} = w_{i,(j+a \mod d)}$$

and

$$\begin{aligned} h^b \cdot w_{i,j} &= h^b \cdot \sum_{r=0}^{q-1} \zeta^{-ir} v_{g^j h^r} = \sum_{r=0}^{q-1} \zeta^{-ir} v_{h^b g^j h^r} = \sum_{r=0}^{q-1} \zeta^{-i(r+k^{-j}b-k^{-j}b)} v_{g^j h^{k^{-j}b+r}} \\ &= \zeta^{ik^{-j}b} \sum_{r=0}^{q-1} \zeta^{-i(r+k^{-j}b)} v_{g^j h^{r+k^{-j}b}} = \zeta^{ik^{-j}b} w_{i,j}, \end{aligned}$$

as desired, and therefore the V_i's are G-submodules. □

Proof of Proposition 4.5. As Z_q is a subgroup of G, we have $q = \sigma(Z_q) \leq \sigma(G)$ by Corollary 4.2 and Proposition 3.12, so it remains to show $\sigma(G) \leq q$. By Corollary 3.6, Lemma 4.6 and Proposition 3.3 we have

$$\sigma(G) = \sigma(G, V_{\text{reg}}) = \sigma(G, \oplus_{i=0}^{q-1} V_i) = \max\{\sigma(G, V_i) \mid i = 0, \ldots, q-1\}.$$

Note that V_0 is obtained from the regular representation of Z_d and the projection $G \to Z_d$. Therefore, $\sigma(G, V_0) \leq \sigma(Z_d) = d < q$. As the ζ^i's for $i = 1, \ldots, q-1$ are just different primitive roots of unity, the modules V_i for $i = 1, \ldots, q-1$ are pairwise isomorphic, so it is enough to show $\sigma(G, V_1) \leq q$. We write $V := V_1$ and $K[V] = K[x_0, \ldots, x_{d-1}]$. The action on $K[V]$ then has the following form: For all $a, b \in \mathbb{Z}$ we have

$$\begin{aligned} g^a \cdot x_j &= x_{(j+a \mod d)}, \\ h^b \cdot x_j &= \zeta^{-k^{-j}b} x_j \quad \text{for all } j = 0, \ldots, d-1. \end{aligned}$$

Note that k^{-j} is understood mod q at all times. From this we see that a monomial

$$x_{j_1}^{\alpha_1} \cdot \ldots \cdot x_{j_s}^{\alpha_s} \text{ is } H\text{-invariant if and only if}$$

$$\alpha_1 \bar{k}^{-j_1} + \ldots + \alpha_s \bar{k}^{-j_s} = \bar{0} \in \mathbb{Z}/q\mathbb{Z}.$$

Now for any non-empty subset $S := \{j_1, \ldots, j_s\} \subseteq \{0, \ldots, d-1\}$, we consider the subset (of same length s) $\{\bar{k}^{-j_1}, \ldots, \bar{k}^{-j_s}\} \subseteq (\mathbb{Z}/q\mathbb{Z})^\times$. By [3, Lemma 6.1] there exist $\alpha_1, \ldots, \alpha_s > 0$ such that $\alpha_1 + \ldots + \alpha_s \leq q$ and

$$\alpha_1 \bar{k}^{-j_1} + \ldots + \alpha_s \bar{k}^{-j_s} = \bar{0} \in \mathbb{Z}/q\mathbb{Z}.$$

We can thus define the monomial

$$m_S := x_{j_1}^{\alpha_1} \cdot \ldots \cdot x_{j_s}^{\alpha_s} \in \Bbbk[V]^H,$$

where $\alpha_1, \ldots, \alpha_s$ are chosen in such a way as to minimise $\alpha_1 + \ldots + \alpha_s$. We now claim that the common zero set of all the orbit sums

$$o_D(m_S) := \sum_{m' \in D \cdot m_S} m' \in \Bbbk[V]^G$$

(for all non-empty subsets S) is 0: otherwise, take $u = (u_0, \ldots, u_{d-1}) \in V \setminus \{0\}$ in the common zero set of all those $o_D(m_S)$. Consider the non-zero coordinates of u,

$$S = \{j_1, \ldots, j_s\} := \{j \in \{0, \ldots, d-1\} \mid u_j \neq 0\} \neq \emptyset.$$

By assumption, $o_D(m_S)(u) = 0$. We show this is a contradiction by considering two cases. Define $m := x_{j_1} \cdot \ldots \cdot x_{j_s}$ (which might be different from m_S).

First, assume the D-stabilizer of m is trivial. Then every monomial in the orbit $D \cdot m_S$ different from m_S contains a variable outside $\{x_{j_1}, \ldots, x_{j_s}\}$, hence its value on u is zero. Therefore,

$$o_D(m_S)(u) = m_S(u) = u_{j_1}^{\alpha_1} \cdot \ldots \cdot u_{j_s}^{\alpha_s} \neq 0,$$

a contradiction.

Second, assume the D-stabilizer of m is non-trivial. So there exists a non-identity element $g^a \in D$ with $g^a \cdot m = m$. This means

$$\{j_1 + d\mathbb{Z}, \ldots, j_s + d\mathbb{Z}\} = \{j_1 + a + d\mathbb{Z}, \ldots, j_s + a + d\mathbb{Z}\},$$

hence

$$\underbrace{\bar{k}^{-j_1} + \ldots + \bar{k}^{-j_s}}_{=:w} = \bar{k}^{-j_1-a} + \ldots + \bar{k}^{-j_s-a} = \bar{k}^{-a} \cdot w \in \mathbb{Z}/q\mathbb{Z}.$$

Hence $(\bar{k}^{-a} - 1)w = 0$. As $\mathbb{Z}/q\mathbb{Z}$ is a field and $\bar{k}^{-a} \neq \bar{1}$, we get

$$0 = w = \bar{k}^{-j_1} + \ldots + \bar{k}^{-j_s},$$

which means the monomial m is H-invariant. By the minimality assumption we have $m = m_S$. Now again, every monomial in the orbit $D \cdot m$ different from m contains a variable outside $\{x_{j_1}, \ldots, x_{j_s}\}$, hence its value on u is zero. So we have

$$o_D(m)(u) = m(u) = u_{j_1} \cdot \ldots \cdot u_{j_s} \neq 0,$$

a contradiction. □

In [3, Theorem 7.1], it is shown that for a non-modular, non-cyclic group G, we have $\sigma(G) \leq \frac{|G|}{l}$, where l denotes the smallest prime divisor of $|G|$. The following, which is part (c) of Theorem 1.3, is an extension of this to the modular case.

Theorem 4.7. *Let G be a finite p-nilpotent group which has a non-normal Sylow-p-subgroup. If l denotes the smallest prime divisor of $|G|$, then*

$$\sigma(G) \leq \frac{|G|}{l}.$$

Proof. Recall that a group G is p-nilpotent if and only if it has a Sylow-p-subgroup P of G with a normal complement, i.e. a normal subgroup $H \trianglelefteq G$ such that $G = PH$ and $P \cap H$ is the trivial group. Let l' denote the smallest prime divisor of $|H|$. In case H is not cyclic, by the aforementioned result of Cziszter and Domokos, we have $\sigma(H) \leq \frac{|H|}{l'}$, thus $\sigma(G) \leq \sigma(H)[G : H] \leq \frac{|H|}{l'}[G : H] = \frac{|G|}{l'} \leq \frac{|G|}{l}$. So we may assume that $H \cong Z_h$ is cyclic of order h. We have a group homomorphism

$$\varphi : P \to \operatorname{Aut}(H), \quad a \mapsto \varphi_a : \left\{ \begin{array}{ccc} H & \to & H \\ h & \mapsto & aha^{-1}. \end{array} \right.$$

As by assumption P is not a normal subgroup, we have $\varphi(P) \neq \{\mathrm{id}_H\}$. Let $U = \ker(\varphi)$, and write $\overline{\varphi} : P/U \to \operatorname{Aut}(H)$ for the induced injective morphism. Note that $U \neq P$ as $\varphi(P) \neq \{\mathrm{id}_H\}$. We first show that U is a normal subgroup of G. By definition, $hu = uh$ for all $u \in U$ and $h \in H$. As U is a normal subgroup of P, for any $a \in P$ and $h \in H$ we hence have $haU = hUa = Uha$, so indeed $U \trianglelefteq G$. The canonical epimorphism $P \twoheadrightarrow P/U$ induces an epimorphism

$$G = HP \cong H \rtimes_\varphi P \twoheadrightarrow H \rtimes_{\overline{\varphi}} (P/U)$$

with kernel U, hence we have $G/U \cong H \rtimes_{\overline{\varphi}} (P/U)$. Let l'' denote the smallest prime-divisor of G/U. If we can show the claim for G/U, i.e. $\sigma(G/U) \le \frac{|G/U|}{l''}$, then $\sigma(G) \le \sigma(G/U)\sigma(U) \le \frac{|G/U|}{l''}|U| = \frac{|G|}{l''} \le \frac{|G|}{l}$, so we are done. Hence we can replace G by G/U, i.e. we will assume that $G \cong H \rtimes_\varphi P$ where $\varphi : P \hookrightarrow \operatorname{Aut}(H)$ is an injective map and P is a non-trivial p-group. We now choose a cyclic subgroup Z_p of order p of P. The restriction of φ to Z_p is of course still injective. By the same argument as before, it is enough to show the claim for the subgroup $H \rtimes Z_p$ of $H \rtimes P$. Thus we now will assume that $G \cong Z_h \rtimes_\varphi Z_p$ where $\varphi : Z_p \hookrightarrow \operatorname{Aut}(Z_h) \cong (\mathbb{Z}/h\mathbb{Z})^\times$ is a monomorphism. Therefore, the element $\varphi(1 + p\mathbb{Z}) = a + h\mathbb{Z}$ is of multiplicative order p in $(\mathbb{Z}/h\mathbb{Z})^\times$. We write $h = q_1^{s_1} \cdot \ldots \cdot q_e^{s_e}$ for the prime factorization of h with different primes $q_1, \ldots, q_e$. The cyclic subgroups $U_{q_i} := \langle \frac{h}{q_i} + h\mathbb{Z} \rangle$ of Z_h of order q_i are characteristic. Therefore for each i, we have an induced homomorphism $\varphi_i : Z_p \to \operatorname{Aut}(U_{q_i})$. As Z_p is of prime order, this homomorphism is either injective or trivial, where it is trivial if and only if

$$a \cdot q_1^{s_1} \cdot \ldots \cdot q_i^{s_i - 1} \cdot \ldots \cdot q_e^{s_e} \equiv q_1^{s_1} \cdot \ldots \cdot q_i^{s_i - 1} \cdot \ldots \cdot q_e^{s_e} \mod q_1^{s_1} \cdot \ldots \cdot q_e^{s_e},$$

i.e. if and only if $a \equiv 1 \mod q_i$. We want to show that at least one of the maps φ_i is injective. For the sake of a proof by contradiction, we therefore assume $a \equiv 1 \mod q_i$ for all $i = 1, \ldots, e$. As a has multiplicative order p modulo h, we have $a^p \equiv 1 \mod q_1^{s_1} \cdot \ldots \cdot q_e^{s_e}$, so particularly $a^p \equiv 1 \mod q_i^{s_i}$ for all $i = 1, \ldots, e$. Lemma 4.8 therefore implies $a \equiv 1 \mod q_i^{s_i}$ for $i = 1, \ldots, e$, hence $a \equiv 1 \mod q_1^{s_1} \cdot \ldots \cdot q_e^{s_e}$, i.e. $a \equiv 1 \mod h$, a contradiction to $a + h\mathbb{Z}$ being of multiplicative order p. So we have that φ_i is injective for at least one i. Then for the subgroup $U_{q_i} \rtimes_{\varphi_i} Z_p$ of $G = Z_h \rtimes_\varphi Z_p$, by Proposition 4.5 we have $\sigma(U_{q_i} \rtimes_{\varphi_i} Z_p) = q_i = \frac{|U_{q_i} \rtimes_{\varphi_i} Z_p|}{p}$, which as before implies $\sigma(G) \le \frac{|G|}{p} \le \frac{|G|}{l}$, so we are done. □

We have used the following number-theoretic lemma in the proof.

Lemma 4.8. *Let $p, q > 0$ be coprime, $s > 0$ and $a \in \mathbb{Z}$. If*

$$a \equiv 1 \mod q \quad \text{and} \quad a^p \equiv 1 \mod q^s,$$

then

$$a \equiv 1 \mod q^s.$$

Proof. We have $a = kq + 1$ for some $k \in \mathbb{Z}$ by the first assumption. Hence by the second assumption,

$$a^p = (1 + kq)^p = 1 + kq \left(\sum_{i=1}^{p} \binom{p}{i} (kq)^{i-1} \right) \equiv 1 \mod q^s.$$

Therefore,

$$kq \left(\sum_{i=1}^{p} \binom{p}{i} (kq)^{i-1} \right) = kq \left(p + kq \left(\sum_{i=2}^{p} \binom{p}{i} (kq)^{i-2} \right) \right) \equiv 0 \mod q^s.$$

As p, q are coprime, the second factor $p + kq(\cdots)$ is coprime to q^s, and hence it follows $kq \equiv 0 \mod q^s$. Thus we have $a = kq + 1 \equiv 1 \mod q^s$, which is what we wanted to prove. □

Proof of Theorem 1.3 (a). It remains only to show the second part of (a). If $\sigma(G) = |G|$, we have already seen in Corollary 3.10 that $N_G(P)/P$ must be cyclic. Now assume additionally P is abelian and $G \neq P$. If $N_G(P) = P$, then Burnside's Theorem (see [14, Theorem 7.50]) implies G is p-nilpotent, hence $\sigma(G) < |G|$ by Theorem 4.7, a contradiction. Therefore, $N_G(P)/P$ must be non-trivial. □

It remains an open question to classify those finite groups which satisfy $\sigma(G) = |G|$. Though we do not have any evidence, the following conjecture was a motivation for many of our results:

Conjecture 4.9. *Suppose G is a finite group. Let P be a Sylow-p-subgroup of G. Then $\sigma(G) = |G|$ implies P is normal in G.*

Note that for p-nilpotent groups, the conjecture follows from Theorem 4.7. From this conjecture, we would get the classification that $\sigma(G) = |G|$ if and only if P is normal in G and G/P is cyclic. Indeed, if P is normal and G/P is cyclic, $\sigma(G) = |G|$ by Theorem 4.3. Conversely, if $\sigma(G) = |G|$ and the conjecture holds, P is normal in G, and then as $|G| = \sigma(G) \leq |P|\sigma(G/P)$ (Proposition 3.8), the result of Cziszter and Domokos, Proposition 1.2, forces G/P to be cyclic.

Also note that whenever G contains a p-nilpotent subquotient with non-normal Sylow-p-subgroup, $\sigma(G) < |G|$ by our relative results. So for the proof of the conjecture, a classification of groups not containing such a p-nilpotent subquotient could be the key.

In [10, Question 1], the authors ask the similar (and also still open) question which finite groups satisfy $\beta_{\text{sep}} = |G|$? At least, as a consequence of Theorem 4.3 and Proposition 1.4, we can add groups G with normal Sylow-p-subgroup P and G/P cyclic to the list.

We conclude with some explicit examples:

Example 4.10. Assume throughout characteristic 2. As $S_3 \cong Z_3 \rtimes Z_2$, $\sigma(S_3) = 3$ by Proposition 4.5. More generally, for D_{2q}, the dihedral group of order $2q$ and q an odd prime, we have $\sigma(D_{2q}) = \sigma(Z_q \rtimes Z_2) = q$ by that proposition. Also note that $\beta_{\text{sep}}(D_{2q}) = q + 1$ by [11, Theorem 8]. So here we have the strict inequalities $\delta(D_{2q}) = 2 < \sigma(D_{2q}) = q < \beta_{\text{sep}}(D_{2q}) = q + 1$. The group A_4 has the normal Sylow-p-subgroup $P = \langle (12)(34), (14)(23) \rangle$ of order 4, and its factor group A_4/P is cyclic of order 3. Hence $\sigma(A_4) = 12$ by Theorem 4.3. As $A_4 \leq S_4$, we have $12 = \sigma(A_4) \leq \sigma(S_4)$ by Proposition 3.12. Also as $S_3 \leq S_4$ we have $\sigma(S_4) \leq \sigma(S_3)[S_4 : S_3] = 3 \cdot 4 = 12$ by Proposition 3.9. This shows $\sigma(S_4) = 12$.

Example 4.11. Assume throughout characteristic 3. Then $\sigma(S_3) = 6$ by Theorem 4.3. Furthermore $\sigma(A_4) = 4$: The projective indecomposable representations of A_4 are obtained by induction of the characters of the Klein four group $H = \langle (12)(34), (13)(24) \rangle \leq A_4$, which leads to a three-dimensional representation. Its matrix group is either the regular representation of Z_3, or conjugate to

$$G = \langle \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \rangle.$$

Then a computation with MAGMA [1] yields that the corresponding invariant ring is minimally generated by

$$x_1^2 + x_2^2 + x_3^2, \quad x_1x_2x_3, \quad x_1^4 + x_2^4 + x_3^4, \quad x_1^4x_2^2 + x_1^2x_3^4 + x_2^4x_3^2.$$

It is easily seen that the first three invariants in that list minimally cut out 0, which shows the claim.

It is also worth mentioning that in the non-modular characteristics (i.e. $p \neq 2, 3$), $\sigma(A_4) = 4$ by [2, Corollary 4.2].

Example 4.12. Assume throughout arbitrary characteristic $p > 0$. If G is a p-group, we have $\sigma(G) = |G|$ either from Theorem 4.3 or Proposition 1.4 and Theorem 1.1. For the dihedral groups D_{2p^n} with $n \geq 0$, we have $\sigma(D_{2p^n}) = 2p^n$, since in characteristic 2 it would be a p-group, and by Theorem 4.3 otherwise. This strengthens the corresponding result on β_{sep}, see [10, Proposition 10].

Acknowledgments. This paper was prepared during visits of the first author to TU München and the second author to University of Aberdeen. The second of these visits was supported by the Edinburgh Mathematical Society's Research Support Fund. We want to thank Gregor Kemper and the Edinburgh Mathematical Society for making these visits possible. We also thank K. Cziszter for some helpful communication via eMail.

References

[1] Wieb Bosma, John Cannon, and Catherine Playoust. The Magma algebra system. I. The user language. *J. Symbolic Comput.*, 24(3-4):235–265, 1997. Computational algebra and number theory (London, 1993).

[2] K. Cziszter. On the noether bound for polynomial invariants of finite groups. *Ph. D. thesis, Central European University. Available from http://mathematics.ceu.hu/theses/2/on-the-noether-bound-for-polynomial-invariants-of-finite-groups*, 2012.

[3] Kálmán Cziszter and Mátyás Domokos. On the generalized Davenport constant and the Noether number. *Cent. Eur. J. Math.*, 11(9):1605–1615, 2013.

[4] Harm Derksen. Polynomial bounds for rings of invariants. *Proc. Amer. Math. Soc.*, 129(4):955–963 (electronic), 2001.

[5] Harm Derksen and Gregor Kemper. *Computational invariant theory*. Invariant Theory and Algebraic Transformation Groups, I. Springer-Verlag, Berlin, 2002. Encyclopaedia of Mathematical Sciences, 130.

[6] M. Domokos. Typical separating invariants. *Transform. Groups*, 12(1):49–63, 2007.

[7] Jan Draisma, Gregor Kemper, and David Wehlau. Polarization of separating invariants. *Canad. J. Math.*, 60(3):556–571, 2008.

[8] Peter Fleischmann. The Noether bound in invariant theory of finite groups. *Adv. Math.*, 156(1):23–32, 2000.

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

[10] Martin Kohls and Hanspeter Kraft. Degree bounds for separating invariants. *Math. Res. Lett.*, 17(6):1171–1182, 2010.

[11] Martin Kohls and Müfit Sezer. Invariants of the dihedral group D_{2p} in characteristic two. *Math. Proc. Cambridge Philos. Soc.*, 152(1):1–7, 2012.

[12] Masayoshi Nagata. On the 14-th problem of Hilbert. *Amer. J. Math.*, 81:766–772, 1959.

[13] V. L. Popov. On Hilbert's theorem on invariants. *Dokl. Akad. Nauk SSSR*, 249(3):551–555, 1979.

[14] Joseph J. Rotman. *An introduction to the theory of groups*, volume 148 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, fourth edition, 1995.

[15] Barbara J. Schmid. Finite groups and invariant theory. In *Topics in invariant theory (Paris, 1989/1990)*, volume 1478 of *Lecture Notes in Math.*, pages 35–66. Springer, Berlin, 1991.

University of Aberdeen, King's College, Aberdeen, AB24 3UE
E-mail address: j.elmer@abdn.ac.uk

Technische Universität München, Zentrum Mathematik-M11, Boltzmannstrasse 3, 85748 Garching, Germany
E-mail address: kohls@ma.tum.de