

WEYL'S POLARIZATION THEOREM IN POSITIVE CHARACTERISTIC

H. DERKSEN*

V. MAKAM**

Department of Mathematics
University of Michigan
Ann Arbor, MI 48104, USA
hderksen@umich.edu

School of Mathematics
Institute for Advanced Study
Princeton, NJ 08540, USA
visu@umich.edu

Abstract. Let V be an n -dimensional algebraic representation over an algebraically closed field K of a group G . For $m > 0$, we study the invariant rings $K[V^m]^G$ for the diagonal action of G on V^m . In characteristic zero, a theorem of Weyl tells us that we can obtain all the invariants in $K[V^m]^G$ by the process of polarization and restitution from $K[V^n]^G$. In particular, this means that if $K[V^n]^G$ is generated in degree $\leq d$, then so is $K[V^m]^G$ no matter how large m is.

There are several explicit counterexamples to Weyl's theorem in positive characteristic. However, when G is a (connected) reductive affine group scheme over \mathbb{Z} and V^* is a good G -module, we show that Weyl's theorem holds in sufficiently large characteristic. As a special case, we consider the ring of invariants $R(n, m)$ for the left-right action of $\mathrm{SL}_n \times \mathrm{SL}_n$ on m -tuples of $n \times n$ matrices. In this case, we show that the invariants of degree $\leq n^6$ suffice to generate $R(n, m)$ if the characteristic is larger than $2n^6 + n^2$.

1. Introduction

Let K be an algebraically closed field. Suppose V is a rational representation of a reductive group G . The ring of invariant polynomials $K[V]^G$ is a finitely generated graded subalgebra of the coordinate ring $K[V]$, see [27], [29], [30], [40]. A long standing theme in invariant theory is to extract a minimal set of generators – apart from a few instances, this is perhaps too ambitious a problem. A more approachable problem is to find upper bounds on the degree of generators.

Definition 1. We define $\beta(K[V]^G)$ to be the smallest integer D such that invariants of degree $\leq D$ form a generating set, i.e.,

$$\beta(K[V]^G) := \min\{D \mid K[V]_{\leq D}^G \text{ is a generating set}\},$$

where $K[V]_{\leq D}^G$ denotes the invariants of degree $\leq D$.

A general bound for $\beta(K[V]^G)$ is given in [3]. In this paper, we will be concerned with the growth of $\beta(K[V^m]^G)$ as m gets large, where V^m denotes the direct sum

DOI: 10.1007/s00031-020-09559-3

*Supported by NSF grant DMS 1601229.

**Supported by NSF grants DMS 1638352 and CCF 1412958.

Received March 18, 2019. Accepted July 19, 2019.

Corresponding Author: V. Makam, e-mail: visu@umich.edu

of m copies of the representation V . It is easy to see that $\beta(K[V^a]^G) \leq \beta(K[V^b]^G)$ if $a \leq b$, and so for fixed G and V , the sequence $\beta(K[V^m]^G)$ is increasing. In characteristic 0, it is a remarkable result due to Weyl (see [50], [36]) that this sequence is actually bounded!

Theorem 1 (Weyl’s polarization theorem — weak form). *Assume $\text{char}(K)$ is 0, and let $\dim V = n$. Then for all m , we have*

$$\beta(K[V^m]^G) \leq \beta(K[V^n]^G).$$

Weyl’s result is actually a little stronger than the version we state above, which we will now discuss. Interpreting V^m as $V \otimes K^m$ illuminates a GL_m action on V^m . Since this GL_m action commutes with the G action, the invariant ring $K[V^m]^G = K[V \otimes K^m]^G$ inherits an action of GL_m . For $a \leq b$, we have the inclusion $K[V^a]^G \subseteq K[V^b]^G$. Suppose S is a generating set for $K[V^a]^G$. Starting with S , we can construct some obvious invariants in $K[V^b]^G$. For example, take any $f \in S$ and $g \in \text{GL}_b$, then $g \cdot f \in K[V^b]^G$. In the same spirit, we can consider the smallest GL_b -stable subspace containing S , i.e., $\langle S \rangle_{\text{GL}_b} \subseteq K[V^b]^G$ (see Section 3 for a more detailed definition). Consider the subalgebra of $K[V^b]^G$ generated by $\langle S \rangle_{\text{GL}_b}$. This subalgebra may not be all of $K[V^b]^G$, and there may be some other ‘genuinely new’ invariants in $K[V^b]^G$. Weyl’s polarization theorem says that there are no genuinely new invariants if we take a at least as big as n .

Theorem 2 (Weyl’s polarization theorem — strong form). *Assume $\text{char}(K)$ is 0, and let $\dim V = n$. Let $S \subseteq K[V^n]^G$ be a generating set for $K[V^n]^G$. Then for all $m \geq n$, the set $\langle S \rangle_{\text{GL}_m}$ is a generating set for $K[V^m]^G$.*

It is easy to see that the weak form of Weyl’s theorem stated before is a consequence of the strong form stated above. In positive characteristic, one does not have to look far to get counterexamples. Suppose $\text{char}(K) = p > 0$. Let C_p denote the cyclic group of order p , and consider the action of C_p on $V = K^2$ where the generator of C_p acts by the matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Weyl’s theorem fails in this case, see [44]. Other examples of failure for finite groups can be found in [48]. We note that finite groups are reductive in arbitrary characteristic. Knop showed in [35] that the strong form of Weyl’s theorem holds for invariant rings of finite groups if the characteristic is large enough.

In this paper, we will restrict ourselves to connected reductive groups. Even in this restricted setting, Weyl’s theorem still fails. For example, in characteristic 2, it fails for the natural action of $G = \text{SO}(V)$ on V , see [12], [11]. An analogue of Weyl’s theorem in positive characteristic was proved for *separating* invariants in [22], see also [26].

1.1. Matrix invariants and semi-invariants

Let $\text{Mat}_{p,q}$ denote the set of $p \times q$ matrices. Consider the group $G = \text{GL}_n$ acting on $V = \text{Mat}_{n,n}^m$ by simultaneous conjugation, i.e.,

$$g \cdot (X_1, \dots, X_m) = (gX_1g^{-1}, \dots, gX_mg^{-1}).$$

We set $S(n, m) = K[V]^G$, the ring of invariants for this action. The ring $S(n, m)$ is often referred to as the ring of matrix invariants. Procesi showed that traces of monomials (in the X_i 's) generate $S(n, m)$ in characteristic 0, see [42]. In [17], [18], Donkin extended this result to all characteristics, by replacing traces with the coefficients of the characteristic polynomial instead.

A bound on the degree of generators in characteristic 0 followed from the work of Razmyslov, see [43].

Theorem 3 (Procesi–Razmyslov). *Assume $\text{char}(K) = 0$. Then we have*

$$\beta(S(n, m)) \leq n^2.$$

One can observe that this bound is independent of m , as predicted by Weyl's theorem. It was pointed out to us by Domokos that the proof of the above result in [23] goes through once characteristic is larger than $n^2 + 1$. In particular, this means that the statement of Weyl's theorem holds for matrix invariants if we assume a modest lower bound on characteristic! However, the techniques used for this are very specific to matrix invariants, and it is not clear if they can be generalized. For example, even in the closely related example of matrix semi-invariants discussed below, such a result was not known prior to this paper.

Consider the left-right action of $G = \text{SL}_n \times \text{SL}_n$ on $V = \text{Mat}_{n,n}^m$, i.e., for $(A, B) \in \text{SL}_n \times \text{SL}_n$ and $(X_1, \dots, X_m) \in \text{Mat}_{n,n}^m$, we have

$$(A, B) \cdot (X_1, \dots, X_m) = (AX_1B^{-1}, \dots, AX_mB^{-1}).$$

We set $R(n, m) = K[V]^G$, the invariant ring in this case. The ring $R(n, m)$ is often referred to as the ring of matrix semi-invariants. In recent times, connections to computational complexity have generated a lot of interest in matrix semi-invariants, see [5], [25], [32], [39].

A determinantal description for the generators follows from results on semi-invariants of quivers, see [8, Cor. 3], [13] and [46]. A polynomial bound on the degree of generators was given in [5], [7].

Theorem 4 ([5], [7]). *Let $n \geq 2$. We have $\beta(R(n, m)) \leq mn^3(n-1)$. If $\text{char}(K) = 0$, then we have $\beta(R(n, m)) \leq n^6$.*

The bound stated in [5], [7] for $\beta(R(n, m))$ was mn^4 , but these slightly stronger bounds are evident in the proof of [5, Thm. 1.2]. The bound in characteristic 0 is once again a consequence of Theorem 1. We prove that the statement of Weyl's theorem holds for matrix semi-invariants with only a modest lower bound on the characteristic.

Theorem 5. *Suppose $\text{char}(K) = p > 2n^6 + n^2$. Then the statement of Weyl's polarization theorem holds for the left-right action of $\text{SL}_n \times \text{SL}_n$ on $\text{Mat}_{n,n}$. In particular, for all $m \in \mathbb{Z}_{>0}$ we have*

$$\beta(R(n, m)) \leq \beta(R(n, n^2)) \leq n^6.$$

Our techniques give similar results for matrix invariants as well, but the lower bound on characteristic we obtain is weaker than the already known $n^2 + 1$.

Remark 1. In small characteristic (i.e., $p \leq n$), the statement of Weyl's theorem is false for matrix invariants, see [12], [9]. By a standard reduction, the same phenomenon holds for matrix semi-invariants as well. However, it remains an open problem to understand whether the statement of Weyl's theorem holds for matrix invariants for $n < p \leq n^2 + 1$ and matrix semi-invariants for $n < p \leq 2n^6 + n^2$.

We can further decrease the lower bound on characteristic if all we want is a bound that doesn't depend on m . However, the degree bound will become a bit worse. For example, the techniques in this paper can be used to show the following:

Proposition 6. *Suppose $\text{char}(K) = p > n^6$, then for all $m \in \mathbb{Z}_{>0}$;*

$$\beta(R(n, m)) \leq \beta(R(n, n^3)) \leq n^7.$$

However, with these techniques, one cannot decrease the lower bound on characteristic to $O(n^{6-\epsilon})$ for any $\epsilon > 0$.

1.2. Main results

We need some technical definitions, for which we follow [45]. An affine group scheme G over $\text{Spec } \mathbb{Z}$ (or simply \mathbb{Z}) is said to be reductive if $G \rightarrow \text{Spec } \mathbb{Z}$ is smooth, and the geometric fibers are connected reductive algebraic groups (in the usual sense). Let G be a reductive group scheme over \mathbb{Z} , and let V be a free \mathbb{Z} -module of finite rank n with a linear action of G . We will call V a free G - \mathbb{Z} -module of rank n . We will denote the ring of invariants by $\mathbb{Z}[V]^G = \text{Sym}(V^*)^G$.

For any algebraically closed field K , the K -points G_K form a connected reductive group over K , and the K -points of V , i.e., $V_K = V \otimes_{\mathbb{Z}} K$ is an n -dimensional representation of G_K . We will write $K[V] = K[V_K]$ and $K[V]^G = K[V_K]^{G_K}$ for simplicity. Note that $K[V]^G$ is not necessarily the same as the base change $\mathbb{Z}[V]^G \otimes_{\mathbb{Z}} K$.

Definition 2. Let $S = \bigoplus_{i \geq 0} S_i$ be a graded R -algebra. Then let $S_{\{d\}}$ denote the R -subalgebra generated by $\bigcup_{i \leq d} S_i$. Further, let $\delta_R(S)$ denote the smallest d such that S is a finite extension over $S_{\{d\}}$.

The following theorem requires the notion of good modules, which we recall in Section 4. A reductive group scheme over \mathbb{Z} is called split if there is a (fiberwise) maximal torus defined over \mathbb{Z} .

Theorem 7. *Let G be a split reductive group scheme over \mathbb{Z} , and let V be a free G - \mathbb{Z} -module of rank n . Suppose V^* is a good G -module. Then, the following statements hold.*

- (1) *The number $\delta_{\mathbb{Z}}(\mathbb{Z}[V^n]^G)$ is finite.*
- (2) *Suppose K is an algebraically closed field such that*

$$\text{char}(K) > 2Q(n+1) + n,$$

where $Q = \max\{2, \frac{3}{8}n(\delta_{\mathbb{Z}}(\mathbb{Z}[V^n]^G))^2\}$. Then the statement of Weyl's polarization theorem holds for the action of G_K on V_K , i.e.,

1. if S is a set of generators for $K[V^n]^G$, then $\langle S \rangle_{\mathrm{GL}_m}$ is a set of generators for $K[V^m]^G$ for all $m \geq n$;
2. we have $\beta(K[V^m]^G) \leq \beta(K[V^n]^G)$ for all $m \geq 1$.

For the first part of the theorem, we will need some results of Seshadri from [45]. The bulk of the paper will go towards proving the second part of the above theorem. The approach is a delicate interplay between combinatorics, representation theory and commutative algebra.

Acknowledgements. We would like to thank Matyas Domokos, David Wehlau, Jerzy Weyman and Jakub Witaszek for helpful discussions.

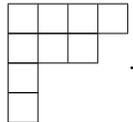
1.3. Organization

In Section 2, we recall some necessary preliminaries. We study polarization in Section 3. Then, in Sections 4 and 5, we discuss good filtrations. In Section 6, we discuss the technical details needed, and prove Theorem 5. Finally, in Section 7, we bring together all the results to prove the main result, i.e., Theorem 7.

2. Preliminaries

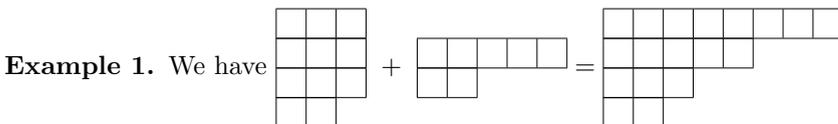
2.1. Partitions

A partition $\lambda = (\lambda_1, \lambda_2, \dots)$ is a (weakly) decreasing sequence of non-negative numbers, such that only finitely many λ_i are non-zero. We often omit writing the trailing zeros. We say λ is a partition of n if $\sum_i \lambda_i = n$, and we write $\lambda \vdash n$. Associated to any partition is its Young diagram. For example, if $\lambda = (4, 3, 1, 1)$, then its Young diagram is



We will not distinguish between a partition and its Young diagram. For a partition λ , we define its size $|\lambda| := \sum_i \lambda_i =$ number of boxes in the Young diagram, and its length $l(\lambda) =$ length of the first column in its Young diagram. For the above example, we have $|\lambda| = 9$ and $l(\lambda) = 4$. We define λ^\dagger to be the conjugate of the partition λ .

Definition 3 (Horizontal concatenation). Given two partitions λ and μ , we define their horizontal concatenation $\lambda + \mu = (\lambda_1 + \mu_1, \lambda_2 + \mu_2, \dots)$. Note that $\lambda + \mu$ is a partition.



2.2. Schur functors

For any commutative ring R , any R -module E and any partition λ , one can construct a Schur module $S_\lambda(E)$ (denoted E^λ in [24, Sect. 8.1]). Let $E^{\times \lambda}$ denote

the direct product of $|\lambda|$ copies of E labelled by boxes in the Young diagram of λ . The Schur module $S_\lambda(E)$ is defined as the universal target for R -module maps from $E^{\times \lambda}$ that are multilinear, alternating along columns, and satisfying some exchange relations. We do not restate the exchange relations, but refer instead to [24, Sect. 8.1] for details.

Let K be an algebraically closed field. For any partition λ , the aforementioned construction gives a polynomial functor $S_\lambda : \text{Vect} \rightarrow \text{Vect}$, where Vect represents the category of finite-dimensional vector spaces (over K). We call S_λ the Schur functor associated to λ . We have $S_{(n)} = \text{Sym}^n$, the n^{th} symmetric power, whereas $S_{(n)^\dagger} = S_{1^n} = \wedge^n$, the n^{th} alternating power. Note that S_λ is denoted by L_{λ^\dagger} in [1], [49].

We require the following result that is well known to experts.

Proposition 8. *Let λ and μ be two partitions. Then, there is a $\text{GL}(V)$ -equivariant surjection $S_\lambda(V) \otimes S_\mu(V) \rightarrow S_{\lambda+\mu}(V)$*

We will discuss a stronger statement, i.e., Corollary 19, later using the theory of good filtrations. Here, we indicate a combinatorial proof of the above proposition for the reader who is more familiar with Young tableaux.

Proof of Proposition 8. One way to construct the partition $\lambda + \mu$ is to take all the columns of (the Young diagrams) of λ and μ and rearrange them in decreasing order. This gives a map $V^{\times \lambda} \times V^{\times \mu} \rightarrow V^{\times (\lambda+\mu)} \rightarrow S_{\lambda+\mu}(V)$. We leave it to the reader to check that this map factors to give a surjective map $S_\lambda(V) \otimes S_\mu(V) \rightarrow S_{\lambda+\mu}(V)$ as required. \square

The above proposition has a version for an arbitrary reductive group. For any reductive group G , fix $T \subseteq B$ a maximal torus and a Borel. Then for each dominant integral weight λ , one defines the K_λ to be the 1-dimensional representation of T corresponding to λ and extended to a representation of B in the standard way. For λ, μ dominant integral weights, the induced modules $\text{ind}_B^G(K_\lambda)$ and $\text{ind}_B^G(K_\mu)$ can be realized inside the coordinate algebra $K[G]$ in the usual way. Multiplication gives a surjective map $\text{ind}_B^G(K_\lambda) \otimes \text{ind}_B^G(K_\mu) \rightarrow \text{ind}_B^G(K_{\lambda+\mu})$. In the case of $G = \text{GL}_n$, these induced modules are the Schur modules, which recover the above result.

2.3. Polynomial representations of GL_m of degree n

We will only need Corollary 10 from this section, but a general reference for the definitions and results in this section is [47]. We first note that $S_\lambda(V)$ is a representation of $\text{GL}(V)$. It is an irreducible representation in characteristic 0, but not necessarily in positive characteristic. We denote by $\text{Rep}^{\text{pol}}(\text{GL}_m)_d$, the category of polynomial representations of GL_m of degree d . This category is a highest weight category, and the costandard objects are precisely the Schur modules $S_\lambda(K^m)$ for $|\lambda| = d$. Totaro was able to give upper bounds on the homological dimension of this category, and compute it precisely under mild assumptions, see [47].

Theorem 9 (Totaro). *Let $\text{char}(K) = p$, and let $\alpha_p(d)$ denote the sum of the digits in the base p expansion of d . The homological dimension of $\text{Rep}^{\text{pol}}(\text{GL}_m)_d$ is $\leq 2(d - \alpha_p(d))$. Further, we have equality if $m \geq d$.*

We refer the reader to [20], [41] for similar results in quantized or generalized settings. For our purposes, Totaro's result in the following special case will suffice.

Corollary 10. *Assume $p > d$. Then $\text{Rep}^{\text{pol}}(\text{GL}_m)_d$ is semi-simple. Further, the Schur modules $S_\lambda(K^m)$ for $\lambda \vdash d$ are irreducible representations of GL_m .*

Proof. The semisimplicity of $\text{Rep}^{\text{pol}}(\text{GL}_m)_d$ follows from the aforementioned Totaro's theorem on homological dimension. The costandard objects in any semisimple highest weight category are irreducible. Hence the Schur modules $S_\lambda(K^m)$'s with $\lambda \vdash d$ are irreducible. \square

3. Polarization

Let E be a $\text{GL}(W)$ representation. For any subset $S \subseteq E$, recall that we define $\langle S \rangle_{\text{GL}(W)}$ to be the smallest $\text{GL}(W)$ stable subspace containing S . This is often referred to as polarization. In more concrete terms $\langle S \rangle_{\text{GL}(W)}$ consists of elements $e \in E$ that can be written as a sum $e = \sum_i g_i s_i$ with $s_i \in S$ and $g_i \in \text{GL}(W)$. Let us note here that the definition of $\langle S \rangle_{\text{GL}(W)}$ depends on the ambient $\text{GL}(W)$ representation E . For our discussion, it will almost always be obvious what the ambient representation is.

Understanding the following special case is the most crucial part of this paper.

Problem 1. For an inclusion of vector spaces $V \subseteq W$, we have $S_\lambda(V) \subseteq S_\lambda(W)$. When is $\langle S_\lambda(V) \rangle_{\text{GL}(W)} = S_\lambda(W)$?

In characteristic 0, this is always true as long as $S_\lambda(V)$ is non-empty, because the module $S_\lambda(W)$ is an irreducible $\text{GL}(W)$ -module. In positive characteristic, this is often not the case as the following example shows:

Example 2. Suppose $\text{char}(K) = 2$, and let $\lambda = (2)$, so $S_\lambda = \text{Sym}^2$. Consider $K^1 \hookrightarrow K^2$, and let x, y be a basis for K^2 with x being a basis for K^1 . Then we have $\text{Sym}^2(K^1) = \text{span}(x^2)$, whereas $\text{Sym}^2(K^2) = \text{span}(x^2, y^2, xy)$. It is easy to see that $\langle \text{Sym}^2(K^1) \rangle_{\text{GL}_2} = \text{span}(x^2, y^2)$ which is a proper subset of $\text{Sym}^2(K^2)$.

Remark 2. If $\dim V \geq |\lambda|$, we will always have $\langle S_\lambda(V) \rangle_{\text{GL}(W)} = S_\lambda(W)$. This is a simple consequence of the description of the Schur module in terms of semistandard Young tableaux. We need a much stronger statement to be of any use for our purposes.

Proposition 11. *Suppose $V = V_1 \oplus V_2 \subseteq W$. Further, suppose $\langle S_\mu(V_1) \rangle_{\text{GL}(W)} = S_\mu(W)$ and $\langle S_\nu(V_2) \rangle_{\text{GL}(W)} = S_\nu(W)$. Then*

$$\langle S_\mu(V_1) \otimes S_\nu(V_2) \rangle_{\text{GL}(W)} = S_\mu(W) \otimes S_\nu(W).$$

Proof. Consider $E \in S_\mu(W)$ and $F \in S_\nu(W)$. We will show that $E \otimes F \in \langle S_\mu(V_1) \otimes S_\nu(V_2) \rangle_{\text{GL}(W)}$. Since we have $\langle S_\mu(V_1) \rangle_{\text{GL}(W)} = S_\mu(W)$, we can write $E = \sum_i g_i e_i$ for some $g_i \in \text{GL}(W)$ and $e_i \in S_\mu(V_1)$. Similarly, we can write $F = \sum_j h_j f_j$ for some $h_j \in \text{GL}(W)$ and $f_j \in S_\nu(V_2)$.

Decompose $W = W_1 \oplus W_2$ with $V_i \subseteq W_i$. Let (w_1, \dots, w_k) be a basis for W_1 and (w'_1, \dots, w'_l) be a basis for W_2 . Let $(w_1, \dots, w_k, w'_1, \dots, w'_l)$ be an ordered basis for W . In this ordered basis, we have a block decomposition $g_i = [A_i \mid B_i]$, where A_i represents the first k columns, and B_i the last l columns. Observe that

since $e_i \in S_\mu(V_1)$, the action of g_i on e_i only depends on A_i . Similarly, write $h_j = [P_j \mid Q_j]$, and the action of h_j on f_j only depends on Q_j . Hence, if we define $\sigma_{ij} = [A_i \mid Q_j]$, we have

$$E \otimes F = \sum_{i,j} \sigma_{ij} \cdot (e_i \otimes f_j).$$

There is a small issue that σ_{ij} may not be invertible, but this is easy to circumvent. For some non-zero constant c_{ij} , we have $c_{ij}I + \sigma_{ij}$ is invertible, where I denotes the identity transformation. Then we can write

$$E \otimes F = \sum_{i,j} (c_{ij}I + \sigma_{ij}) \cdot (e_i \otimes f_j) - (c_{ij}I) \cdot (e_i \otimes f_j).$$

Since $c_{ij}I + \sigma_{ij}$ as well as $c_{ij}I$ are elements of $\text{GL}(W)$, we have that

$$E \otimes F \in \langle S_\mu(V_1) \otimes S_\nu(V_2) \rangle_{\text{GL}(W)}.$$

The proposition follows since elements of the form $E \otimes F$ span $S_\mu(W) \otimes S_\nu(W)$.
□

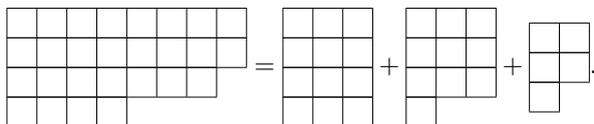
Theorem 12. *Let $\lambda = \mu + \nu$, and let $V_1 \oplus V_2 = V \hookrightarrow W$. Further, suppose we have $\langle S_\mu(V_1) \rangle_{\text{GL}(W)} = S_\mu(W)$ and $\langle S_\nu(V_2) \rangle_{\text{GL}(W)} = S_\nu(W)$. Then we have $\langle S_\lambda(V) \rangle_{\text{GL}(W)} = S_\lambda(W)$.*

Proof. Consider the surjection $\pi : S_\mu(W) \otimes S_\nu(W) \twoheadrightarrow S_\lambda(W)$ from Proposition 8. It suffices to show $\pi(E \otimes F) \in \langle S_\lambda(V) \rangle_{\text{GL}(W)}$ for $E \in S_\mu(W)$ and $F \in S_\nu(W)$. Indeed by the Proposition 11, we have $E \otimes F = \sum_i g_i \cdot (e_i \otimes f_i)$ for $g_i \in \text{GL}(W)$, $e_i \in S_\mu(V_1)$, and $f_i \in S_\nu(V_2)$. Thus we have $\pi(E \otimes F) = \sum_i g_i \cdot (\pi(e_i \otimes f_i)) \in \langle S_\lambda(V) \rangle_{\text{GL}(W)}$.
□

Lemma 13. *Let $\lambda \vdash d$ such that $l(\lambda) \leq n$. Fix $k \geq 2$. Then we can write $\lambda = \mu_1 + \mu_2 + \dots + \mu_s$ for some positive integer s and non-empty partitions μ_i for $i = 1, \dots, s$ such that $n(k-1) < |\mu_i| \leq kn$ for all $i < s$ and $|\mu_s| \leq kn$. Further, we have $l(\mu_i) \leq n$ for all i .*

Proof. Suppose $|\lambda| \leq kn$, then there is nothing to do. So, let us assume $|\lambda| > kn$. The lengths of the columns in λ are given by the conjugate partition $\lambda^\dagger = (\lambda_1^\dagger, \lambda_2^\dagger, \dots)$. Let t be the smallest integer such that $\sum_{i=1}^t \lambda_i^\dagger > kn$. Then let μ_1 be the first $(t-1)$ columns of λ , so that $\lambda = \mu_1 + \nu$, where ν is a partition. We have $kn \geq \sum_{i=1}^{t-1} \lambda_i^\dagger$ by minimality of t , and we have $\sum_{i=1}^{t-1} \lambda_i^\dagger > kn - n$, since $\lambda_t^\dagger \leq n$ by hypothesis. Hence we have $n(k-1) < |\mu_1| \leq kn$. Now, proceed by induction on ν .
□

Example 3. Suppose $n = 4$ and $k = 3$ and $\lambda = (8, 8, 7, 4)$, then the decomposition in the above lemma is best visualized by the following picture.



Corollary 14. *Let $\lambda \vdash d$ such that $l(\lambda) \leq n$, and suppose $\text{char}(K) > kn$ for some $k \geq 2$. Then for $V \hookrightarrow W$, with $\dim V \geq n \lceil d/(n(k-1)) \rceil$, we have $\langle S_\lambda(V) \rangle_{\text{GL}(W)} = S_\lambda(W)$*

Proof. Write $\lambda = \mu_1 + \mu_2 + \cdots + \mu_s$ be the decomposition from the previous lemma. If $s \geq \lceil d/(n(k-1)) \rceil + 1$, then since $|\mu_i| > n(k-1)$ for $i < s$, and $|\mu_s| > 0$, we have $|\lambda| = \sum_i |\mu_i| > (s-1)n(k-1) = \lceil d/(n(k-1)) \rceil n(k-1) \geq d$, which is a contradiction. Hence we have $s \leq \lceil d/(n(k-1)) \rceil$, and consequently, we have $sn \leq n \lceil d/(n(k-1)) \rceil \leq \dim V$. This allows us to choose subspaces $V_1, \dots, V_s \subseteq W$ such that $\dim V_i = n$ and $V_1 \oplus V_2 \oplus \cdots \oplus V_s \subseteq V$.

First observe that $S_{\mu_i}(V_i)$ is non-zero as $l(\mu_i) \leq \dim V_i = n$. Next, we see from Corollary 10 that $S_{\mu_i}(W)$ is an irreducible $\text{GL}(W)$ representation as $|\mu_i| \leq kn < \text{char}(K)$. Hence, we have $\langle S_{\mu_i}(V_i) \rangle_{\text{GL}(W)} = S_{\mu_i}(W)$. The result follows by a repeated application of Theorem 12. \square

Corollary 15. *Suppose $L : \text{Vect} \rightarrow \text{Vect}$ is a functor such that it has a filtration (of functors) whose subquotients are of the form S_λ with $\lambda \vdash d$ and $l(\lambda) \leq n$. Suppose $\text{char}(K) > kn$ with $k \geq 2$, and let $V \in \text{Vect}$ such that $\dim V \geq n \lceil d/(n(k-1)) \rceil$. Then for $V \hookrightarrow W$, we have $\langle L(V) \rangle_{\text{GL}(W)} = L(W)$.*

4. Good filtrations and the Littlewood–Richardson rule

The theory of good filtrations is very powerful in positive characteristic. A comprehensive introduction to this theory can be found in [14] (see also [15], [16], [19], [38]). We also refer the reader to [10], [51] for an exposition with a view of using them for invariant rings coming from quivers including matrix invariants and semi-invariants.

Let G be a connected reductive algebraic group over an algebraically closed field K . Let B be a choice of Borel subgroup of G and let $T \subset B$ be a maximal torus of G . Let Λ^+ denote the set of dominant integral weights. For each $\lambda \in \Lambda^+$, one can associate a one-dimensional representation of B . The corresponding induced G -module is called a dual Weyl module, and denoted $\nabla(\lambda)$. Note that for GL_n and SL_n , Schur modules are dual Weyl modules. There is a partial order \prec on Λ^+ defined by $\lambda \prec \mu$ if $\mu - \lambda$ is a non-negative sum of roots.

Definition 4. A G -module V is called a good G -module if it has a filtration $0 \subseteq V_0 \subseteq V_1 \subseteq \dots$ such that $\bigcup_i V_i = V$ and each subquotient V_i/V_{i-1} is a dual Weyl module. Such a filtration is called a good filtration.

The dual Weyl modules occurring as subquotients (including multiplicities) are independent of the choice of filtration.

Remark 3. For a split reductive group defined over \mathbb{Z} , Weyl modules and dual Weyl modules are defined over \mathbb{Z} , see [33] or [37]. More precisely, for $\lambda \in \Lambda^+$, there is a free G - \mathbb{Z} -module $\nabla_{\mathbb{Z}}(\lambda)$ such that $\nabla_{\mathbb{Z}}(\lambda) \otimes_{\mathbb{Z}} K$ is the dual Weyl module $\nabla(\lambda)$ for G_K for any algebraically closed field K . So, we call a free G - \mathbb{Z} -module a good G module if it has a filtration by the dual Weyl modules $\nabla_{\mathbb{Z}}(\lambda)$'s. Moreover, the characters of dual Weyl modules are given by the Weyl character formula and in particular independent of the characteristic.

The following lemma is straightforward (see for example [14, Prop. 3.1.1]).

Lemma 16. *Suppose a G -module V has a filtration $0 \subseteq V_0 \subseteq V_1 \subseteq \dots$ such that $\bigcup_i V_i = V$ and each subquotient V_i/V_{i-1} is a good G -module, then V is a good G -module.*

Let us recall some well-known properties of good G -modules. They can be found in the standard references mentioned above.

Lemma 17. *Let V and W be good G -modules.*

- (1) *If $V \subseteq W$, then W/V is a good G -module.*
- (2) *$V \otimes W$ is a good G -module.*
- (3) *$\dim(V^G)$ is the multiplicity of the trivial module in any good filtration for G .*

The following result is [14, Prop. 3.2.6].

Lemma 18. *Suppose V is a good G -module. Suppose it has a good filtration $0 = V_0 \subseteq V_1 \subseteq \dots \subseteq V_n = V$ with $V_i/V_{i-1} = \nabla(\lambda_i)$. Let π be a permutation of $\{1, 2, \dots, n\}$ such that whenever $\lambda_{\pi(i)} \succ \lambda_{\pi(j)}$, we have $\pi(i) > \pi(j)$. Then there is a good filtration $0 = V'_0 \subseteq V'_1 \subseteq \dots \subseteq V'_n = V$ such that $V_i/V_{i-1} = \nabla(\lambda_{\pi(i)})$.*

The following result is already evident in the proof of the universal form of the Littlewood–Richardson rule (see [2]). However, we will provide a sketch of the proof. Let us note that the dominance order on partitions agrees with the partial order \prec for $G = \text{GL}(V)$.

Corollary 19. *Suppose λ, μ are two partitions, and V a vector space over an algebraically closed field K . Then we have a surjection $\zeta : S_\lambda(V) \otimes S_\mu(V) \twoheadrightarrow S_{\lambda+\mu}(V)$ such that $\ker(\zeta)$ has a filtration whose subquotients are Schur modules of the form $S_\nu(V)$ with $\nu \prec \lambda + \mu$.*

Proof. Since $S_\lambda(V)$ and $S_\mu(V)$ are good $\text{GL}(V)$ modules, so is $S_\lambda(V) \otimes S_\mu(V)$ by Lemma 17. To understand the multiplicities of dual Weyl modules in any good $\text{GL}(V)$ -filtration for $S_\lambda(V) \otimes S_\mu(V)$, it suffices to write its character as a sum of characters of dual Weyl modules. This is a computation that is independent of characteristic as the dual Weyl modules have the same formal character in any characteristic, see Remark 3.

In characteristic zero, the celebrated Littlewood–Richardson rule describes how $S_\lambda(V) \otimes S_\mu(V)$ decomposes as a sum of Schur modules. Hence, in any characteristic, the Littlewood–Richardson rule describes the subquotients in any good filtration of $S_\lambda(V) \otimes S_\mu(V)$. The Schur module $S_{\lambda+\mu}(V)$ occurs with multiplicity one, and all others are of the form $S_\nu(V)$ with $\nu \prec \lambda + \mu$.

Using the above lemma, we can get a good filtration $0 = V_0 \subseteq \dots \subseteq V_k = S_\lambda(V) \otimes S_\mu(V)$ such that $V_k/V_{k-1} = S_{\lambda+\mu}(V)$. Interpreting this as a map $\zeta : V_k \twoheadrightarrow S_{\lambda+\mu}(V)$ whose kernel is V_{k-1} , we get the required conclusion. \square

5. Good filtrations for invariant rings

For this section, let us assume G is a connected reductive group over an algebraically closed field K whose characteristic is $p > 0$, and V is an n -dimensional good G -module. The following lemma is [51, Lem. 2].

Lemma 20. *The module $\Lambda^i(V)$ is a good G -module for $i < p$.*

We can use the above lemma to prove the more general statement.

Lemma 21. *If $p > n$, then $S_\lambda(V)$ is a good G -module for all partitions λ .*

Proof. The minimal elements in the dominance order on partitions are the partitions of the form 1^t . We note that $S_{1^t}(V) = \Lambda^t(V)$. From the above lemma, we see that all of these are good G -modules. We proceed by induction. Let λ be a partition such that $S_\mu(V)$ is a good G -module for all μ smaller than λ in the dominance order. If $l(\lambda) > n$, then $S_\lambda(V) = 0$, so we can assume $l(\lambda) \leq n < p$. Since $l(\lambda) < p$, we can write $\lambda = \mu + 1^t$ where μ is a partition and $t \leq l(\lambda) < p$.

By the inductive hypothesis, $S_\mu(V)$ is a good G -module. Further, we have already observed that $S_{1^t}(V) = \Lambda^t(V)$ is also a good G -module. Hence, by Lemma 17, $M = S_\mu(V) \otimes S_{1^t}(V)$ is a good G -module. By Corollary 19, we have a surjection $\zeta : M \twoheadrightarrow S_{\mu+1^t}(V) = S_\lambda(V)$ such that $\ker(\zeta)$ has a filtration by Schur modules $S_\nu(V)$ satisfying $\nu \prec \lambda$. By induction, all such $S_\nu(V)$'s are good G -modules. This means that $\ker(\zeta)$ has a filtration by good G -modules, and so by Lemma 16, we conclude that $\ker(\zeta)$ is a good G -module. By Lemma 17, we conclude that $S_\lambda(V) = M/\ker(\zeta)$ is also a good G -module. \square

Corollary 22. *If $p > n$, then $\text{Sym}(V)$ is a good G -module.*

Lemma 23. *If $p > n$, the module $\text{Sym}(V \otimes W)$ is a good G -module for any finite-dimensional W (G acts trivially on W).*

Proof. We have $\text{Sym}(V \otimes W) = \text{Sym}(V)^{\otimes \dim W}$. So, it is a good module by Lemma 17. \square

The following result first appeared in [21], but can also be found in [1].

Theorem 24 (Doubilet–Rota–Stein). *$\text{Sym}^d(V \otimes W)$ has a natural filtration whose associated graded module is*

$$\bigoplus_{\lambda \vdash d} S_\lambda(V) \otimes S_\lambda(W).$$

Corollary 25. *$\text{Sym}^d(V \otimes W)$ has a natural filtration whose associated graded module is*

$$\bigoplus_{\lambda \vdash d, l(\lambda) \leq n} S_\lambda(V) \otimes S_\lambda(W).$$

Proof. This follows from the above theorem, since $S_\lambda(V) = 0$ if $l(\lambda) > n$. \square

Lemma 26. *Suppose $p > n$. Then $\text{Sym}^d(V \otimes W)^G$ has a natural filtration whose associated graded module is*

$$\bigoplus_{\lambda \vdash d, l(\lambda) \leq n} S_\lambda(V)^G \otimes S_\lambda(W).$$

Proof. Let $0 = F_0 \subseteq F_1 \subseteq \dots \subseteq F_m = \text{Sym}^d(V \otimes W)$ denote the $G \times \text{GL}(W)$ -filtration from the above corollary. We know that F_i/F_{i-1} is of the form $S_\lambda(V) \otimes S_\lambda(W)$, so $(F_i/F_{i-1})^G = S_\lambda(V)^G \otimes S_\lambda(W)$

Observe that $\text{Sym}^d(V \otimes W)^G$ has a filtration

$$0 = F_0^G \subseteq F_1^G \subseteq \dots \subseteq F_m^G = \text{Sym}^d(V \otimes W)^G.$$

The associated graded module of this filtration is $\bigoplus_i F_i^G/F_{i-1}^G$. Hence, if we show that $F_i^G/F_{i-1}^G = (F_i/F_{i-1})^G$, we would be done. It is easy to see that we have natural injective maps $\eta_i : F_i^G/F_{i-1}^G \hookrightarrow (F_i/F_{i-1})^G$ for each i . So, it suffices to show that the maps η_i are isomorphisms. Indeed, if $A \subseteq B$ are modules with good filtrations, then the natural surjection $B \rightarrow B/A$ restricts to a surjection $B^G \rightarrow (B/A)^G$ (see for example [9, Prop. 1(ii)]), and hence descends to a surjection $B^G/A^G \rightarrow (B/A)^G$. This shows that the η_i 's are all surjective, and hence isomorphisms. \square

Corollary 27. *Suppose $p > n$, then $\text{Sym}^d(V \otimes W)^G$ has a natural filtration whose associated subquotients are all of the form $S_\lambda(W)$ with $\lambda \vdash d$ and $l(\lambda) \leq n$.*

Corollary 28. *Assume $p > kn$ for some $k \geq 2$. Let W be a vector space with $\dim W \geq n \lceil d/(n(k-1)) \rceil$. Then, for any inclusion of vector spaces $W \hookrightarrow W'$, we have*

$$\langle \text{Sym}^d(V \otimes W)^G \rangle_{\text{GL}(W')} = \text{Sym}^d(V \otimes W')^G.$$

Proof. First observe that $p > kn > n$. Consider the polynomial functor L defined by $L(U) = \text{Sym}(V \otimes U)^G$. L has a filtration by S_λ with $\lambda \vdash d$ and $l(\lambda) \leq n$ by the previous corollary. Hence, by Corollary 15, we have $\langle L(W) \rangle_{\text{GL}(W')} = L(W')$. \square

6. Technical details

We discuss a few elementary results before proceeding to the main technical result.

6.1. Decomposable elements

For a graded ring $R = \bigoplus_{d=0}^{\infty} R_d$, we define the notion of decomposable and indecomposable elements.

Definition 5. A homogeneous element $f \in R_d$ of degree d is called decomposable if it can be written as $f = \sum_{i \in I} g_i h_i$, where g_i, h_i are homogeneous elements of degree $< d$. If a homogeneous element is not decomposable, we call it indecomposable.

Corollary 29. *The set $R_{\leq N} = \bigoplus_{i=1}^N R_i$ is a set of generators for R if and only if for all $d > N$, every element of R_d is decomposable.*

For the rest of the section, let V be a representation of a group G over some algebraically closed field K . The following lemma is straightforward.

Lemma 30. *The set of decomposable invariants in $K[V^m]_d^G = \text{Sym}^d(V^* \otimes K^m)^G$ is GL_m stable.*

Lemma 31. *Assume $a \leq b$ and let S be a set of generators for $K[V^a]^G$. If $\langle K[V^a]_d^G \rangle_{\text{GL}_b} = K[V^b]_d^G$ for all $d \leq \beta(K[V^b]^G)$, then $\langle S \rangle_{\text{GL}_b}$ is a generating set for $K[V^b]^G$.*

Proof. Clearly, it suffices to show that every indecomposable invariant in $K[V^b]^G$ can be generated by $\langle S \rangle_{\text{GL}_b}$. Take an indecomposable invariant $f \in K[V^b]^G$. It has degree $d \leq \beta(K[V^b]^G)$. Thus $f \in \langle K[V^a]_d^G \rangle_{\text{GL}_b}$ by hypothesis. Hence, we have $f = \sum_i g_i \cdot f_i$ with $f_i \in K[V^a]_d^G$ and $g_i \in \text{GL}_b$. But since S is a generating set for $K[V^a]^G$, we can write each $f_i = p_i(s_{i_1}, \dots, s_{i_{r_i}})$ for some polynomial p_i in r_i variables, and $s_{i_j} \in S$. Thus we have

$$f = \sum_i g_i \cdot f_i = \sum_i g_i \cdot p_i(s_{i_1}, \dots, s_{i_{r_i}}) = \sum_i p_i(g_i \cdot s_{i_1}, \dots, g_i \cdot s_{i_{r_i}})$$

But this means that f is generated by $\langle S \rangle_{\text{GL}_b}$. \square

6.2. Main technical result

For this section, let G be a reductive group over an algebraically closed field K of characteristic p , and let V be an n -dimensional representation such that V^* is a good G -module. This section is devoted to proving the following result.

Proposition 32. *Suppose $Q \geq \frac{1}{2}$ is such that $\beta(K[V^m])^G \leq mQ$ for all $m \geq 1$ and $p > 2Q(n+1) + n$. If S is a generating set for $K[V^n]^G$, then $\langle S \rangle_{\text{GL}_m}$ is a generating set for $K[V^m]^G$ for all $m \geq n$.*

To prove this proposition, we need some other results first.

Lemma 33. *Assume $\text{char}(K) > kn$ for some $k \geq 2$. If $m \geq n \lceil d/(n(k-1)) \rceil$, then we have $\langle K[V^m]_d^G \rangle_{\text{GL}_{m+1}} = K[V^{m+1}]_d^G$.*

Proof. This follows from Corollary 28 because we assume that V^* is a good G -module. \square

Corollary 34. *Assume the hypothesis of Proposition 32. Suppose $m \geq n$. Then, we have $\langle K[V^m]_d^G \rangle_{\text{GL}_{m+1}} = K[V^{m+1}]_d^G$ for all $d \leq \beta(K[V^{m+1}]^G)$.*

Proof. It follows from the hypothesis that $d \leq \beta(K[V^{m+1}]^G) \leq (m+1)Q$. Also, observe that by hypothesis, we have $p = \text{char}(K) > kn$, where $k = 2Q(1 + 1/n) + 1 \geq 2$. So, in order to use the above lemma, we only need to show

$$m \geq n \left\lceil \frac{d}{n(k-1)} \right\rceil.$$

We have two cases:

Case 1: $d/(n(k-1)) \leq 1$. In this case, we have $m \geq n = n \lceil d/(n(k-1)) \rceil$ by hypothesis.

Case 2: $d/(n(k-1)) > 1$. In this case, we have $\lceil d/(n(k-1)) \rceil < 2 \frac{d}{n(k-1)}$, hence it suffices to show

$$m \geq 2n \frac{d}{n(k-1)} = \frac{2d}{k-1}.$$

But this is the same as showing that $d \leq m(k-1)/2$. We know that $d \leq (m+1)Q$, so it suffices to show that $(m+1)Q \leq m(k-1)/2$. Rearranging, we need to show that $k \geq 2Q(1+1/m) + 1$. But it is easy to see that $k = 2Q(1+1/n) + 1 \geq 2Q(1+1/m) + 1$ since $m \geq n$. \square

Proof of Proposition 32. This follows from Lemma 31 and a repeated application of the above corollary. \square

Proof of Theorem 5. For $n = 1$, the result is obvious. Assume $n \geq 2$, and now the result follows from Proposition 32 once we observe that the hypothesis is satisfied for $Q = n^3(n-1)$ by Theorem 4. One does need to be a little careful in applying Proposition 32 as $\dim(\text{Mat}_{n,n}) = n^2$ and not n . \square

Proof of Proposition 6. One has to mimic the proof of Proposition 32. First, note that we have to replace n by n^2 because $\dim(\text{Mat}_{n,n}) = n^2$. In Corollary 34, one needs to adjust the assumption to $m \geq n^3$. Accordingly the two cases in the proof should be modified. Let us write $T = d/(n^2(k-1))$. Then the two cases one should use are that either $\lceil T \rceil \leq n$ or $\lceil T \rceil \leq (1+1/n)T$. \square

Remark 4. Similar results can be formulated for various invariant rings and semi-invariant rings associated to quivers, by standard reductions to be reduced to the case of matrix invariants and semi-invariants, see [5], [6], [7] for details.

7. Proof of main theorem

The results of the previous section can be applied in any situation where an explicit Q that satisfies the hypothesis of Proposition 32 is known. The importance of Theorem 7 is to provide a plethora of examples. The only catch, though, is that part (1) of Theorem 7 is an existential result rather than an explicit one. It is an interesting problem to give any kind of explicit bound on the number Q .

In characteristic zero, an approach to upper bounds for the degree of generators of invariant rings was proposed by Popov, and improved by the first author in [3]. For a collection of polynomials T , we denote its zero locus by $\mathbb{V}(T)$.

Definition 6. Let V be a representation of a reductive group G over an algebraically closed field K . The null cone is given by

$$\mathcal{N}(G, V) := \mathbb{V}\left(\bigcup_{d=1}^{\infty} K[V]_d^G\right) \subseteq V.$$

Further, define

$$\gamma_K(G, V) := \min \left\{ D \mid \mathbb{V}\left(\bigcup_{d=1}^D K[V]_d^G\right) = \mathcal{N}(G, V) \right\}.$$

Now, let us study the null cone for the action of G on several copies of V . It is easy to see that $\gamma_K(G, V^a) \leq \gamma_K(G, V^b)$ for $a \leq b$. However, we claim that the sequence stabilizes.

Lemma 35. *Let V be an n -dimensional representation of a reductive group G over an algebraically closed field K . For all $a \geq 0$, we have*

$$\gamma_K(G, V^a) \leq \gamma_K(G, V^n).$$

Proof. Without loss of generality we can assume $a > n$. From the Hilbert–Mumford criterion for the null cone, it is easy to see that whether a tuple $(v_1, \dots, v_a) \in V^a$ is in the null cone or not depends only on the linear subspace $\text{span}(v_1, \dots, v_a) \subseteq V$. So, for any tuple (v_1, \dots, v_a) not in the null cone, let v_{i_1}, \dots, v_{i_r} be a basis for $\text{span}(v_1, \dots, v_a)$. Then $(v_{i_1}, \dots, v_{i_r})$ is not in the null cone for the action of G on V^r , and observe that $r \leq n$. This gives an invariant $f \in K[V^r]^G$ such that $f(v_{i_1}, \dots, v_{i_r}) \neq 0$. Now, define $\tilde{f} \in K[V^a]^G$ by $\tilde{f}(w_1, \dots, w_a) = f(w_{i_1}, \dots, w_{i_r})$. Then $\tilde{f}(v_1, \dots, v_a) = f(v_{i_1}, \dots, v_{i_r}) \neq 0$. This means there is an invariant of degree $\leq \gamma_K(G, V^r) \leq \gamma_K(G, V^n)$ that doesn't vanish on (v_1, \dots, v_a) . This means that the null cone for V^a is cut out by invariants of degree $\leq \gamma_K(G, V^n)$. This proves the lemma. \square

In the above, instead of using the Hilbert–Mumford criterion, one may also use the results from [22] to establish a similar result.

If V is a free G - \mathbb{Z} -module of rank n for some reductive group scheme G over \mathbb{Z} , then we write $\gamma_K(G, V) = \gamma_K(G_K, V_K)$ for simplicity. Recall the definition of $\delta_R(S)$ from Definition 2.

Lemma 36. *Let G be a reductive group scheme over \mathbb{Z} . Suppose V is a free G - \mathbb{Z} -module of rank n . Then $\gamma_K(G, V) \leq \delta_{\mathbb{Z}}(\mathbb{Z}[V]^G)$ for any algebraically closed field K .*

Proof. This follows in two steps. The first is to see that $\gamma_K(G, V) = \delta_K(K[V]^G)$. This follows from [4, Lem. 2.5.5, Rem. 4.7.2]. The second is to check that $\delta_K(K[V]^G) \leq \delta_{\mathbb{Z}}(\mathbb{Z}[V]^G)$. First, we note that $\mathbb{Z}[V]^G$ is finitely generated, see [45, Thm. 2], so in particular, we know that $\delta_{\mathbb{Z}}(\mathbb{Z}[V]^G) < \infty$.

By definition of $\delta_{\mathbb{Z}}(\mathbb{Z}[V]^G)$, we have that $\mathbb{Z}[V]^G$ is a finite extension over the ring $\mathbb{Z}[G_1, \dots, G_m]$ where G_i are homogeneous of degree $\leq \delta_{\mathbb{Z}}(\mathbb{Z}[V]^G)$. But finite extensions are preserved under base change, so $\mathbb{Z}[V]^G \otimes_{\mathbb{Z}} K$ is a finite extension over $\mathbb{Z}[G_1, \dots, G_m] \otimes_{\mathbb{Z}} K = K[G_1, \dots, G_s]$. Hence, if we show that $K[V]^G$ is a finite extension of $\mathbb{Z}[V]^G \otimes_{\mathbb{Z}} K$, then it follows that $K[V]^G$ is a finite extension of $K[G_1, \dots, G_s]$. So, we can conclude that $\delta_K(K[V]^G) \leq \max\{\deg(G_i)\} \leq \delta_{\mathbb{Z}}(\mathbb{Z}[V]^G)$.

So, all that remains is to prove that $K[V]^G$ is a finite extension of $\mathbb{Z}[V]^G \otimes_{\mathbb{Z}} K$. As we noticed before, $\mathbb{Z}[V]^G$ is finitely generated, so $\mathbb{Z}[V]^G = \mathbb{Z}[F_1, \dots, F_r]$, and so $\mathbb{Z}[V]^G \otimes_{\mathbb{Z}} K = K[F_1, \dots, F_r]$. We see from [45, Prop. 6] that $\mathbb{V}(F_1, \dots, F_r) \subseteq V_K$ is the null cone for the action of G_K on V_K . Again, by [4, Lem. 2.5.5, Remark 4.7.2], we see that $K[V]^G$ is a finite extension of $K[F_1, \dots, F_r] = \mathbb{Z}[V]^G \otimes_{\mathbb{Z}} K$ as required. \square

Combining the above two lemmas, we get the following result.

Corollary 37. *Let G be a reductive group scheme over \mathbb{Z} . Suppose V is a free G - \mathbb{Z} -module of rank n . Then we have $\gamma_K(G, V^a) \leq \gamma_{\mathbb{Z}}(\mathbb{Z}[V^n]^G)$ for all algebraically closed fields K and for all a .*

Proof. From the previous two lemmas, we get

$$\gamma_K(G, V^a) \leq \gamma_K(G, V^n) \leq \gamma_{\mathbb{Z}}(\mathbb{Z}[V^n]^G). \quad \square$$

We now discuss how the bounds for invariants defining the null cone translate into bounds for the degrees of generators. Let us state the main result in [3] in a slightly different way.

Theorem 38. *Assume K is an algebraically closed field of characteristic 0. Let V be an n -dimensional representation of a reductive group G over K . Then we have*

$$\beta(K[V]^G) \leq \max\{2, \frac{3}{8}n\gamma_K(G, V)^2\}.$$

There are two reasons that we need characteristic zero in the above statement. The first is that invariant rings for reductive groups are Cohen–Macaulay in characteristic zero, see [31]. The second is Kempf’s result that the Hilbert series is a proper rational function, see [34]. In the situations that we are interested in, these required ingredients are true in positive characteristic as well. The Cohen–Macaulay condition was addressed by Hashimoto in [28].

Theorem 39 ([28]). *Suppose G is a reductive group over an algebraically closed field K . Suppose V is a representation of G such that $K[V] = \text{Sym}(V^*)$ is a good G -module. Then $K[V]^G$ is strongly F -regular, and in particular Cohen–Macaulay.*

To get Kempf’s result on the Hilbert series to arbitrary characteristic, we use a comparison to the characteristic zero case. In order to compare across characteristics, we will need the representation and the reductive group to be defined over \mathbb{Z} . So, we will work under the hypothesis of the Theorem 7 for the following result.

Proposition 40. *Let G be a split reductive group scheme over \mathbb{Z} , and V a free G - \mathbb{Z} -module of rank n such that V^* is a good G -module. Let K be an algebraically closed field such that $K[V]$ is a good G_K -module (for, e.g., if $\text{char}(K) > n$ by Corollary 22). Then the Hilbert series for $K[V]^G$ is the same as the Hilbert series for $\mathbb{C}[V]^G$. In particular, it is a proper rational function.*

Proof. This fact that the Hilbert series of $K[V]^G$ is the same as the Hilbert series for $\mathbb{C}[V]^G$ has been observed before in the context of matrix invariants and matrix semi-invariants in [17], [9], [7]. The same proof works, and we sketch it.

First note that the group $G_{\mathbb{C}}$ is reductive, and hence linearly reductive. In particular, all $G_{\mathbb{C}}$ modules have good filtrations. Since $K[V]_d$ (resp. $\mathbb{C}[V]_d$) has a good filtration, to get $\dim(K[V]^G)$ (resp. $\dim \mathbb{C}[V]^G$), one has to write the character of $K[V]_d$ (resp. $\mathbb{C}[V]_d$) as a sum of characters of dual Weyl modules and read off the coefficient of the trivial character (see Lemma 17). The characters of the dual Weyl modules are independent of characteristic, so the two computations are identical. \square

Thus the degree bound for invariants in Theorem 38 continues to hold in positive characteristic if we add some hypothesis so as to use the above results. So, we get the following:

Corollary 41. *Let G be a split reductive group scheme over \mathbb{Z} , and V a free G - \mathbb{Z} -module of rank n such that V^* is a good G -module. Let K be an algebraically closed field such that $K[V]$ is a good G_K -module. Then we have*

$$\beta(K[V]^G) \leq \max \left\{ 2, \frac{3}{8} n \gamma_K(G, V)^2 \right\}.$$

Proof of Theorem 7. To prove the first part, it suffices to show that $\delta_{\mathbb{Z}}(\mathbb{Z}[V^n]^G)$ is finite. This follows from the fact that $\mathbb{Z}[V^n]^G$ is finitely generated, see [45, Thm. 2].

Let us now turn to the second part. Let K be an algebraically closed field with $\text{char}(K) > 2Q(n+1) + n$. Observe that since $\text{char}(K) > n$ and V^* is a good G -module, we have that $K[V^m]$ is a good G -module for all m by Lemma 23. So the above corollary applies and so for all m , we have

$$\beta(K[V^m]^G) \leq \max \left\{ 2, \frac{3}{8} mn \gamma_K(G, V^m)^2 \right\} \leq \max \left\{ 2, m \left(\frac{3}{8} n \delta_{\mathbb{Z}}(\mathbb{Z}[V^n]^G)^2 \right) \right\} \leq mQ.$$

The first inequality follows from the above corollary since $\dim(V^m) = mn$. The second inequality follows from Corollary 37, and the last inequality follows from the definition of Q . Since $\beta(K[V^m]^G) \leq mQ$, we can apply Proposition 32 and the second part of the theorem follows. \square

To conclude, we wish to emphasize an important future direction of research, namely, to produce a strong upper bound for $\delta_{\mathbb{Z}}(\mathbb{Z}[V^n]^G)$. At the moment, we do not have any kind of explicit bound!

References

- [1] K. Akin, D. Buchsbaum, J. Weyman, *Schur functors and Schur complexes*, Adv. in Math. **44** (1982), 207–278.
- [2] G. Boffi, *The universal form of the Littlewood–Richardson rule*, Adv. in Math. **68** (1988), no. 1, 64–84.
- [3] H. Derksen, *Polynomial bounds for rings of invariants*, Proc. Amer. Math. Soc. **129** (2001), no. 4, 955–963.
- [4] H. Derksen, G. Kemper, *Computational Invariant Theory*, 2nd enlarged ed., with two appendices by Vladimir L. Popov, and an addendum by Norbert A’Campo and Popov, Encyclopaedia of Mathematical Sciences, Vol. 130, Subseries *Invariant Theory and Algebraic Transformation Groups*, Vol. VIII, Springer, Heidelberg, 2015.
- [5] H. Derksen, V. Makam, *Polynomial degree bounds for matrix semi-invariants*, Adv. in Math. **310** (2017), 44–63.
- [6] H. Derksen, V. Makam, *Degree bounds for semi-invariant rings of quivers*, J. Pure Appl. Algebra **222** (2018), no. 10, 3282–3298.
- [7] H. Derksen, V. Makam, *Generating invariant rings of quivers in arbitrary characteristic*, J. Algebra **489** (2017), 435–445.
- [8] H. Derksen, J. Weyman, *Semi-invariants of quivers and saturation of Littlewood–Richardson coefficients*, J. American Math. Soc. **13** (2000), 467–479.

- [9] M. Domokos, *Matrix invariants and the failure of Weyl's theorem*, in: *Polynomial Identities and Combinatorial Methods* (Pantelleria, 2001), Lecture Notes in Pure Appl. Math. **235**, Dekker, New York, 2003, pp. 215–236.
- [10] M. Domokos, *Finite generating system of matrix invariants*, Math. Pannon **13** (2002), 175–181.
- [11] M. Domokos, P. E. Frenkel, *Mod 2 indecomposable orthogonal invariants*, Adv. Math. **192** (2005), 209–217.
- [12] M. Domokos, S. G. Kuzmin, A. N. Zubkov, *Rings of matrix invariants in positive characteristic*, J. Pure Appl. Algebra **176** (2002), no. 1, 61–80.
- [13] M. Domokos, A. N. Zubkov, *Semi-invariants of quivers as determinants*, Transformation Groups **6** (2001), 9–24.
- [14] S. Donkin, *Rational Representations of Algebraic Groups: Tensor Products and Filtrations*, Lecture Notes in Mathematics, Vol. 1140, Springer, Berlin, 1985.
- [15] S. Donkin, *Skew modules for reductive groups*, J. Algebra **113** (1988), 465–479.
- [16] S. Donkin, *The normality of closures of conjugacy classes of matrices*, Inv. Math. **101** (1990), 717–736.
- [17] S. Donkin, *Invariants of several matrices*, Inv. Math. **110** (1992), 389–401.
- [18] S. Donkin, *Invariant functions on matrices*, Math. Proc. Cambridge Math. Soc. **113** (1993), 23–43.
- [19] S. Donkin, *On tilting modules for algebraic groups*, Math. Z. **212** (1993), 39–60.
- [20] S. Donkin, *The q -Schur Algebra*, London Mathematical Society Lecture Note Series, Vol. 253, Cambridge University Press, Cambridge, 1998.
- [21] P. Doubilet, G. C. Rota, J. Stein, *On the foundations of combinatorial theory: IX combinatorial methods in invariant theory*, Studies in Appl. Math. **53** (1974), 185–216.
- [22] J. Draisma, G. Kemper, D. Wehlau, *Polarization of separating invariants*, Canad. J. Math. **60** (2008), 556–571.
- [23] E. Formanek, *Generating the ring of matrix invariants*, in: *Ring Theory*, Lecture Notes in Mathematics, Vol. 1197, Springer, Berlin, 1986, pp. 73–82.
- [24] W. Fulton, *Young Tableaux, With Applications to Representation Theory and Geometry*, London Mathematical Society Student Texts, Vol. 35, Cambridge University Press, Cambridge, 1997.
- [25] A. Garg, L. Gurvits, R. Oliveira, A. Widgerson, *A deterministic polynomial time algorithm for non-commutative rational identity testing*, Foundations of Computer Science (FOCS), IEEE, 2016, 109–117.
- [26] F. D. Grosshans, *Vector invariants in arbitrary characteristic*, Transformation Groups **12** (2007), 499–514.
- [27] W. Haboush, *Reductive groups are geometrically reductive*, Ann. of Math. **102** (1975), 67–85.
- [28] M. Hashimoto, *Good filtrations of symmetric algebras and strong F -regularity of invariant subrings*, Math. Z. **236** (2001), 605–623.
- [29] D. Hilbert, *Über die Theorie der algebraischen Formen*, Math. Ann. **36** (1890), 473–534.

- [30] D. Hilbert, *Über die vollen Invariantensysteme*, Math. Ann. **42** (1893), 313–370.
- [31] M. Hochster, J. L. Roberts. *Rings of invariants of reductive groups acting on regular rings are Cohen-Macaulay*. Adv. in Math. **13** (1974), 115–175.
- [32] G. Ivanyos, Y. Qiao, K.V. Subrahmanyam, *Non-commutative Edmonds' problem and matrix semi-invariants*, Comput. Complexity **26** (2017), no. 3, 717–763.
- [33] J. C. Jantzen, *Representations of Algebraic Groups*, 2nd ed., Mathematical Surveys and Monographs, Vol. 107. American Mathematical Society, Providence, RI, 2003.
- [34] G. Kempf, *The Hochster–Roberts theorem of invariant theory*, Michigan Math. J. **26** (1979), no 1, 19–32.
- [35] F. Knop, *On Noether's and Weyl's bound in positive characteristic*, in: *Invariant Theory in All Characteristics*, CRM Proc. Lecture Notes, Vol. 35, Amer. Math. Soc., Providence, RI, 2004, pp. 175–188.
- [36] H. Kraft, C. Procesi, *Classical Invariant Theory : A primer*, <http://www.unibas.math.ch>.
- [37] U. Kulkarni, *A homological interpretation of Jantzen's sum formula*, Transformation Groups **11** (2006), no. 3, 517–538.
- [38] O. Mathieu, *Filtrations of G -modules*, Ann. Scient. Ec. Norm. Sup (2) **23** (1990), 625–644.
- [39] K. Mulmuley, *Geometric complexity theory V: Efficient algorithms for Noether normalization*, J. Amer. Math. Soc. **30** (2017), no. 1, 225–309.
- [40] M. Nagata, *Invariants of a group in an affine ring*, J. Math. Kyoto Univ. **3** (1963/1964), 369–377.
- [41] A. E. Parker, *The global dimension of Schur algebras for GL_2 and GL_3* , J. Algebra **241** (2001), no. 1, 340–378.
- [42] C. Procesi, *The invariant theory of $n \times n$ matrices*, Adv. in Math. **19** (1976), 306–381.
- [43] Ю. П. Размыслов, *Тождества со следом полных матричных алгебр над полем характеристики нуль*, Изв. АН СССР, Сер. матем. **38** (1974), вып. 4, 723–745. Engl. transl.: Y. Razmyslov, *Trace identities of full matrix algebras over a field of characteristic zero*, Math. USSR Izv. **8** (1974), no. 4, 727–760.
- [44] D. Richman, *On vector invariants over finite fields*, Adv. in Math. **81** (1990), no. 1, 30–65.
- [45] C. S. Seshadri, *Geometric reductivity over arbitrary base*, Adv. in Math. **26** (1977), no. 3, 225–274.
- [46] A. Schofield, M. van der Bergh, *Semi-invariants of quivers for arbitrary dimension vectors*, Indag. Mathem., N.S **12** (2001), 125–138.
- [47] B. Totaro, *Projective resolutions of representations of GL_n* , J. Reine Angew. Math. **482** (1997), 1–13.
- [48] D. Wehlau, *The Noether number in invariant theory*, C. R. Math. Acad. Sci. Soc. R. Can. **28** (2006), no. 2, 39–62
- [49] J. Weyman, *Cohomology of Vector Bundles and Syzygies*, Cambridge Tracts in Mathematics, Vol **149**, Cambridge University Press, Cambridge, 2003.

- [50] H. Weyl, *The Classical Groups, their Invariants and Representations*, Princeton Mathematical Series, Vol. 1, Princeton Univ. Press, Princeton, 1946.
- [51] А. Н. Зубков, *Матричные инварианты над бесконечным полем конечной характеристики*, Сиб. матем. журн. **34** (1993), ном. 6, 68–74. Engl. transl.: A. N. Zubkov, *Matrix invariants over an infinite field of finite characteristic*, Siberian Math. J. **34** (1993), no. 6, 1059–1065.

Publishers Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.