SEPARATING INVARIANTS FOR THE BASIC $\mathbb{G}_a$-ACTIONS

JONATHAN ELMER AND MARTIN KOHLS

Abstract. We explicitly construct a finite set of separating invariants for the basic $\mathbb{G}_a$-actions. These are the finite dimensional indecomposable rational linear representations of the additive group $\mathbb{G}_a$ of a field of characteristic zero, and their invariants are the kernel of the Weitzenböck derivation $D_n = x_0\frac{\partial}{\partial x_1} + \ldots + x_{n-1}\frac{\partial}{\partial x_n}$.

Keywords: Invariant theory, separating invariants, binary forms, locally nilpotent derivations, basic $\mathbb{G}_a$-actions, generalized hypergeometric series.
AMS Classification: 13A50, 13N15

1. Introduction

A great many mathematical problems are special cases of the following: let K be a field of arbitrary characterstic and let G be any group. Suppose G acts on the K-vector space V and that v and w are points of V. Is there a $g \in G$ satisfying $gv = w$? In other words, are v and w contained in the same G-orbit? Important examples include the case where $G = \mathrm{GL}_n(K)$ acts on the vector space V of $n \times n$ matrices by conjugation, and the case where $G = \mathrm{SL}_n(K)$ acts on the space V of binary forms of degree n. The classical approach to these problems is to construct "invariant polynomials". These are polynomial functions $f : V \to K$ which satisfy $f(v) = f(gv)$ for all $g \in G$ and $v \in V$, and so are constant on G-orbits. In fact, one can define an action of G on the set of polynomial functions $K[V]$ via $(g \cdot f)(v) := f(g^{-1}v)$ for which the invariant polynomials are the fixed points, $K[V]^G$, and these form a subalgebra of $K[V]$ called the *algebra of invariants*. Ideally, one would like to find a complete set of algebra generators of $K[V]^G$, then use this set to distinguish as many orbits as possible.

This approach is not without its difficulties. For instance, it is not always possible to distinguish all the orbits using invariant polynomials. As an example, consider once more the case where $\mathrm{GL}_n(K)$ acts on the vector space of $n \times n$ matrices over a field K by conjugation. Provided K is an infinite field, the invariants are generated by the coefficients of the characteristic polynomial [5, Example 2.1.3], but it is well known that a pair of matrices with the same characteristic polynomial are not necessarily conjugate. More problematically, if G is not reductive then the algebra $K[V]^G$ may not even be finitely generated. Even if it is, finding a set of generators can be a very difficult problem. If, however, one is only interested in invariants from the point of view of separating orbits, then finding a complete set of generators is not always necessary. It is perhaps surprising, then, that Derksen and Kemper made the following defininition [5, Definition 2.3.8] as recently as 2002.

Definition 1.1. A *separating set* for the ring of invariants $K[V]^G$ is a subset $S \subseteq K[V]^G$ with the following property: given $v, w \in V$, if there exists an invariant f satisfying $f(v) \neq f(w)$, then there also exists $s \in S$ satisfying $s(v) \neq s(w)$.

There are many instances in which separating sets can be seen to have "nicer" properties than generating sets. For example, it is well known that if G is finite and

Date: November 10, 2010.

the characteristic of K does not divide $|G|$, then $K[V]^G$ is generated by elements of degree $\leq |G|$, see [10, 11], but this is not necessarily true in the modular case. On the other hand, the analogue for separating invariants holds in arbitary characteristic [5, Theorem 3.9.13]. Meanwhile, even if $K[V]^G$ is not finitely generated, it is guaranteed to contain a finite separating set [5, Theorem 2.3.15]. The existence proof is non-constructive, which raises the question how to actually construct separating sets. Kemper [15] gives an algorithm for reductive groups, but using Gröbner bases it is only effective for "small" cases. An example of a finite separating set for a non finitely generated invariant ring is given in [8]. For finite groups, a separating set can always be obtained as the coefficients of a rather large polynomial [5, Theorem 3.9.13]. With refined methods, "nicer" separating sets have been obtained for several classes of finite groups and representations, see for example [20]. This paper goes in the same direction: for the basic actions of the additive group in characteristic zero, we present a rather small separating set. See also [6, 7, 9, 16] for a small selection of other recent publications in the area.

From this point onwards, $\Bbbk$ denotes a field of characterstic zero. In this article we will concentrate on linear actions of the additive group $\mathbb{G}_a$ of the ground field $\Bbbk$. The finite dimensional indecomposable rational linear representations of $\mathbb{G}_a$ are called the basic $\mathbb{G}_a$-actions. There is one such action in each dimension, and these are described below:

Definition 1.2. Let $X_n := \langle x_0, \dots, x_n \rangle_{\Bbbk}$ be a vector space of dimension $n+1$. Then $\mathbb{G}_a$ is said to act *basically* on X_n (with respect to the given basis) if the action of $\mathbb{G}_a$ on X_n is given by the formula

$$a * x_i = \sum_{j=0}^{i} \frac{a^j}{j!} x_{i-j}, \qquad \text{for all } a \in \mathbb{G}_a, \;\; i = 0, \dots, n.$$

Note the isomorphisms $X_n \cong X_n^* \cong S^n(X_1)$ for all n, where S^n denotes the nth symmetric power. Let $\{x_0, x_1, \dots, x_n\}$ be the set of coordinate functions on a $n+1$ dimensional vector space V_n, so we consider $X_n = V_n^*$. As $\Bbbk$ is infinite, $\Bbbk[V_n]$ can be viewed as the polynomial ring $R_n := S(X_n) = \Bbbk[x_0, x_1, \dots, x_n]$. If $\mathbb{G}_a$ acts basically on X_n, one can then check that the induced action of $\mathbb{G}_a$ on R_n is given by the formula

$$a * f = \exp(aD_n) f \qquad \text{for all } a \in \mathbb{G}_a, \;\; f \in R_n,$$

where D_n is the *Weitzenböck derivation*

$$D_n := x_0 \frac{\partial}{\partial x_1} + \ldots + x_{n-1} \frac{\partial}{\partial x_n} \quad \text{ on } R_n.$$

Furthermore, the algebra of invariants $\Bbbk[V_n]^{\mathbb{G}_a}$ is precisely the kernel of D_n. We denote this by A_n. The algebras A_n have been objects of intensive study for well over a hundred years, owing to their connection with the classical invariants and covariants of binary forms. While they are known to be finitely generated by the Maurer-Weitzenböck Theorem [24], the number of generators appears to increase rapidly with dimension, and explicit generating sets are (reliably!) known only for $n \leq 7$ (see also the table in section 2). In this article, we shall instead construct explicit separating sets for *all* values of n.

This article is organised as follows: in Section 2 we state our main results, and explain briefly the connection between the algebras A_n and the covariants of binary forms. In Section 3 we prove a crucial lemma on the radical of the Hilbert ideal of A_n which may be of independent interest. Section 4 contains the main body of the proof of our result, while Section 5 is devoted to the proof of a technical lemma which is required in order to construct a separating set for A_n when $n \equiv 0 \mod 4$.

Most of this work was completed during a visit of the first author to TU München in July 2010. We would like to thank Gregor Kemper for making this visit possible.

2. Background and statement of results

Let U_n denote the $\Bbbk$-vector space of binary forms of degree n, which are homogeneous polynomials of the form $\sum_{i=0}^{n} a_i X^i Y^{n-i}$ in the variables X and Y, $a_i \in \Bbbk$. This is a vector space of dimension $n+1$ with basis the set of monomials in X and Y of degree n. The natural action of the group $G := \mathrm{SL}_2(\Bbbk)$ on a two dimensional vector space with basis $\{X, Y\}$ induces an action of G on the vector space U_n. Classically speaking, an invariant is a polynomial in the coefficients a_i which is unchanged under the action of G - in modern notation, an element of $\Bbbk[U_n]^G$. Note that the additive group $\mathbb{G}_a$ is embedded in G as the subgroup of matrices of the form $\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$, and the subgroup $\mathbb{G}_a$ acts basically on U_n (with respect to the basis $\{\frac{1}{k!}X^{n-k}Y^k : k = 0, \dots, n\}$). The pioneers of invariant theory also studied "covariants", which are polynomials in both the coefficients a_i and the variables X and Y themselves which are fixed under the action of G. In modern notation, the algebra of covariants is $\Bbbk[U_n \oplus U_1^*]^G$. There is, in fact, an even stronger connection between covariants and the basic actions of $\mathbb{G}_a$: the algebras $\Bbbk[U_n \oplus U_1^*]^G$ and $\Bbbk[U_n]^{\mathbb{G}_a}$ are actually isomorphic. Let us identify the algebra $\Bbbk[U_n \oplus U_1^*]$ with the polynomial ring $\Bbbk[a_0, a_1, \cdots, a_n, X, Y]$ (we abuse notation by using the same letters a_i for coordinates and coordinate functions). Define a mapping $\Phi : \Bbbk[U_n \oplus U_1^*]^G \to \Bbbk[U_n]^{\mathbb{G}_a}$ by

$$\Phi(f(a_0, a_1, a_2, \dots, a_n, X, Y)) := f(a_0, a_1, a_2, \dots, a_n, 0, 1). \tag{1}$$

The theorem of Roberts [19] states that Φ is an isomorphism. In classical invariant theory one often studies the basic actions of $\mathbb{G}_a$ in order to get a handle on the covariants of binary forms using Roberts' isomorphism. One word of caution is needed at the point. While [16, Proposition 1] implies that a separating set for $\Bbbk[U_n \oplus U_1^*]^{\mathrm{SL}_2(\Bbbk)}$ must be mapped under Φ to a separating set for $\Bbbk[U_n]^{\mathbb{G}_a}$, the converse is not necessarily true, so the separating sets we construct in this paper most likely do not lift to give separating sets for the covariants of binary forms (cf. [16, Remark 3]).

We now state our main results. For any real number x, the symbol $[x]$ denotes the largest integer less than or equal to x. We begin by definining some important invariants, namely

$$f_m := \sum_{k=0}^{m-1} (-1)^k x_k x_{2m-k} + \frac{1}{2}(-1)^m x_m^2 \in \ker D_n \qquad \text{for } m = 1, \dots, [\frac{n}{2}] \tag{2}$$

and $f_0 := x_0$. Further, we define the elements

$$s_m := \sum_{k=0}^{m} (-1)^k \frac{2m+1-2k}{2} x_k x_{2m+1-k} \in R_n \qquad \text{for } m = 1, \dots, [\frac{n-1}{2}] \tag{3}$$

and $s_0 := x_1$, which satisfy

$$D_n s_m = f_m, \text{ for all } m,$$

and in particular $D_n s_m \in A_n \setminus \{0\}$. Elements with this property are called *local slices.*

For any $a \in R_n \setminus \{0\}$, let $\nu(a)$ denote the nilpotency index $\nu(a) := \min\{m \in \mathbb{N} : D_n^{m+1}(a) = 0\}$, and $\nu(0) := -\infty$. If $s \in R_n$ is a local slice, then for any $a \in R_n$ we define

$$\begin{aligned}\epsilon_s(a) &:= (\exp(tD_n)a)|_{t:=-s/D_n s} \cdot (D_n s)^{\nu(a)} \\ &= \sum_{k=0}^{\nu(a)} \frac{(-1)^k}{k!}(D_n^k a)s^k(D_n s)^{\nu(a)-k} \in A_n.\end{aligned}$$

By the Slice Theorem [12, Corollary 1.22], we have

$$A_n \subseteq \Bbbk[\epsilon_s(x_0), \ldots, \epsilon_s(x_n)]_{D_n s}.$$

When $s = x_1$ and so $D_n s = x_0$, this is the first stage in Lin Tan's (and van den Essen's) algorithm for producing a generating set for A_n [22, 23].

Theorem 2.1. *Given n, we define a set E_n consisting of the following elements:*

$$\begin{gathered} f_0, f_1, \ldots, f_{[\frac{n}{2}]}, \\ \epsilon_{s_0}(x_2), \ldots, \epsilon_{s_0}(x_n), \\ \epsilon_{s_1}(x_1), \ldots, \epsilon_{s_1}(x_n), \\ \epsilon_{s_2}(x_2), \ldots, \epsilon_{s_2}(x_n), \\ \epsilon_{s_3}(x_3), \ldots, \epsilon_{s_3}(x_n), \\ \vdots \\ \epsilon_{s_{[\frac{n-1}{2}]}}(x_{[\frac{n-1}{2}]}), \ldots, \epsilon_{s_{[\frac{n-1}{2}]}}(x_n). \end{gathered}$$

If $n \equiv 0 \mod 4$ we also append to E_n an extra invariant w which is defined in Lemma 5.4. Then the set E_n is a separating set for A_n.

Note that $\epsilon_{s_0}(x_0) = f_0$ and $\epsilon_{s_0}(x_1) = 0$. The size of this separating set is about $\frac{3}{8}n^2$. The following table shows its exact size for some values of n. The lower line gives the size c_n of a minimal generating set for A_n, see Olver [17, p. 40]. Olver says this list can not be trusted for $n \geq 7$. For c_7, we use Bedratyuk's value $c_7 = 147$ [3], while in Olver's list values $c_7 = 124$ or $c_7 = 130$ are offered, depending on the source. We also want to remark that for $n \geq 5$, we could save 5 elements by replacing the 10 elements of $E_4 \setminus \{w\}$ appearing in E_n by the 5 generators of A_4. (For $n = 4$ we could save 6 elements). Note that generators for $n \leq 7$ are listed explicitly in [1, 3].

n	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$\lvert E_n \rvert$	11	16	20	28	34	43	49	61	69	82	90	106	116	133	143	163	175
c_n	5	23	26	147	69	415	475	949	?	?	?	?	?	?	?	?	?

It is also worth noting that our separating set consists of invariants whose degree is at most $2n + 1$.

3. The radical of the Hilbert ideal

Let R_n, D_n and A_n be as in the introduction. For any $m < n$ we have the algebra homomorphism

$$\pi_{m,n} : R_n \to R_m, \quad f(x_0, x_1, \ldots, x_n) \mapsto f(\underbrace{0, \ldots, 0}_{n-m \text{ times}}, x_0, \ldots, x_m)$$

which satisfies $\pi_{m,n} \circ D_n = D_m \circ \pi_{m,n}$ and thus induces a map $A_n \to A_m$.

Consider the Hilbert ideal $I_n := A_{n,+}R_n \trianglelefteq R_n$. With the invariants f_m defined in (2), we get the following inclusion for its radical:

$$(x_0, \ldots, x_{[\frac{n}{2}]})R_n = \sqrt{(f_0, f_1, \ldots, f_{[\frac{n}{2}]})R_n} \subseteq \sqrt{I_n}. \tag{4}$$

The main purpose of this section is to prove that the reverse inclusion holds too.

Proposition 3.1. (a) *The radical of the Hilbert ideal is given by*

$$\sqrt{I_n} = (x_0, \ldots, x_{[\frac{n}{2}]})R_n.$$

(b) $\pi_{n-[\frac{n}{2}]-1,n}(A_n) = \Bbbk$.
(c) $\pi_{m,2m}(A_{2m}) = \Bbbk[x_0^2]$ *for m odd.*
(d) $\pi_{m,2m}(A_{2m}) = \Bbbk[x_0^2, x_0^3]$ *for m even.*

Proof. We will make use of Roberts' isomorphism as defined in the previous section, with the only difference that we will choose variables so that the $\mathbb{G}_a$-actions become basic, using the notations of the introduction. Additionally, let $\mathbb{G}_a$ act basically on $\langle y_0, y_1\rangle_\Bbbk$. The action of $\mathbb{G}_a$ on $\tilde{R}_n := R_n[y_0, y_1]$ extends to an action of $\mathrm{SL}_2(\Bbbk)$ on $\tilde{R}_n$ such that the following holds:

(1) for any $a \in \Bbbk\backslash\{0\}$ and $\mu_a := \begin{pmatrix} a^{-1} & 0 \\ 0 & a \end{pmatrix} \in \mathrm{SL}_2$ we have $\mu_a(x_k) = a^{2k-n}x_k$ for $k = 0, \ldots, n$ and $\mu_a(y_k) = a^{2k-1}y_k$ for $k = 0, 1$.
(2) for $\tau := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathrm{SL}_2$ we have $\tau(x_k) = (-1)^k \frac{(n-k)!}{k!} x_{n-k}$ for $k = 0, \ldots, n$ and $\tau(y_k) = (-1)^k y_{1-k}$ for $k = 0, 1$.

Recall from section 2 that the algebra map

$$\Phi : \tilde{R}_n \to R_n, \quad f(x_0, \ldots, x_n, y_0, y_1) \mapsto f(x_0, \ldots, x_n, 0, 1)$$

induces an isomorphism of invariant rings $\tilde{R}_n^{\mathrm{SL}_2} \to R_n^{\mathbb{G}_a} = A_n$. Now let $f \in A_n$ and $F \in \tilde{R}_n^{\mathrm{SL}_2}$ with $\Phi(F) = f$. Then we also have $f = \Phi(\mu_a \cdot F)$ for all $a \in \Bbbk \setminus \{0\}$, i.e.

$$f(x_0, \ldots, x_n) = F(a^{-n}x_0, a^{-n+2}x_1, \ldots, a^n x_n, 0, a) \text{ for all } a \in \Bbbk \setminus \{0\}.$$

Thus,

$$\pi_{n-[\frac{n}{2}]-1,n}(f) = F(0, \ldots, 0, a^{2([\frac{n}{2}]+1)-n}x_0, \ldots, a^n x_{n-[\frac{n}{2}]-1}, 0, a) \text{ for all } a \in \Bbbk\backslash\{0\},$$

and since this equation is polynomial in a and $\Bbbk$ is an infinite field, it also holds for $a = 0$. Therefore, $\pi_{n-[\frac{n}{2}]-1,n}(f) = F(0, \ldots, 0) \in \Bbbk$, which proves (a) and (b). Similarly, for $n = 2m$ we find

$$\pi_{m,2m}(f) = F(0, \ldots, 0, x_0, a^2x_1, \ldots, a^{2m}x_m, 0, a) \text{ for all } a \in \Bbbk \setminus \{0\},$$

which is again polynomial in a. When $a = 0$, we get

$$\pi_{m,2m}(f) = F(0, \ldots, 0, x_0, 0, \ldots, 0) = p(x_0)$$

for some polynomial $p(x_0) \in \Bbbk[x_0]$, so $\pi_{m,2m}(A_{2m}) \subseteq \Bbbk[x_0]$. Since $\pi_{m,2m}(f_m) = \frac{1}{2}(-1)^m x_0^2$, we get the inclusion $\Bbbk[x_0^2] \subseteq \pi_{m,2m}(A_{2m})$. Using that F is also invariant under τ, we find in the same manner as before

$$\pi_{m,2m}(f) = \pi_{m,2m}(\Phi(\tau\mu_{a^{-1}}F))|_{a=0} = F(0, \ldots, 0, (-1)^m x_0, 0, \ldots, 0) = p((-1)^m x_0).$$

Therefore, for m odd we get $\pi_{m,2m}(f) = p(x_0) = p(-x_0) \in \Bbbk[x_0^2]$, which proves (c). For m even, to prove (d), we refer to Lemma 5.4, which gives a $w \in A_{2m}$ with $\pi_{m,2m}(w) = x_0^3$, so $\Bbbk[x_0^2, x_0^3] \subseteq \pi_{m,2m}(A_{2m}) \subseteq \Bbbk[x_0]$. Since x_0 generates the degree one elements of A_{2m} and $\pi_{m,2m}(x_0) = 0$, we are done. □

We want to mention here that the method of proof for Proposition 5.4 (a) also works for decomposable actions. Consider

$$R := \Bbbk[x_{0,1}, \ldots, x_{n_1,1}, \ldots, x_{0,k}, \ldots, x_{n_k,k}]$$

and $D = D_{n_1} + \ldots + D_{n_k}$ with $D_{n_i} = x_{0,i}\frac{\partial}{\partial x_{1,i}} + \ldots + x_{n_i-1,i}\frac{\partial}{\partial x_{n_i,i}}$. Using an algebra homomorphism π which behaves on each subalgebra $\Bbbk[x_{0,i}, \ldots, x_{n_i,i}]$ as $\pi_{n_i-[\frac{n_i}{2}]-1,n_i}$, we get with the same proof

Theorem 3.2. *The radical of the Hilbert ideal of* $\ker D$ *is given by*

$$(x_{0,1},\dots,x_{[\frac{n_1}{2}],1},\dots,x_{0,k},\dots,x_{[\frac{n_k}{2}],k})R.$$

4. Construction of a separating set

In this section, we prove our main result.

Proof of Theorem 2.1. For $V_n=\Bbbk^{n+1}$ with $\Bbbk[V_n]=R_n$, assume there are two elements $a=(a_0,\dots,a_n)$ and $b=(b_0,\dots,b_n)$ of V_n such that $f(a)=f(b)$ for all $f\in E_n$. We have to show that $f(a)=f(b)$ for all $f\in A_n$. As $x_0\in E_n$, we have $a_0=b_0$. Assume first $a_0=b_0\neq 0$. By the Slice Theorem, $A_n\subseteq \Bbbk[E_n]_{x_0}$. Therefore, $f\in A_n$ can be written as $f=\frac{p}{x_0^l}$ with $p\in\Bbbk[E_n]$ and $l\geq 0$. By assumption, $p(a)=p(b)$ and $a_0^l=b_0^l\neq 0$, so $f(a)=p(a)/a_0^l=p(b)/b_0^l=f(b)$. Now assume $a_0=b_0=0$ and let m be maximal such that $a_0=a_1=\ldots=a_m=0$, so $a_{m+1}\neq 0$ (if $m<n$). By induction on k, we shall show that $b_k=0$ for $k=0,\dots,\min\{m,[\frac{n}{2}]\}$. By assumption this holds for $k=0$, so assume it holds for some $k<\min\{m,[\frac{n}{2}]\}$. Then

$$\frac{(-1)^{k+1}}{2}a_{k+1}^2=f_{k+1}(a)=f_{k+1}(b)=\frac{(-1)^{k+1}}{2}b_{k+1}^2, \tag{5}$$

so $b_{k+1}=0$ since $a_{k+1}=0$. If $m\geq[\frac{n}{2}]$, then $f(a)=f(0)=f(b)$ for any $f\in A_n$ by Proposition 3.1, so now assume $0\leq m<[\frac{n}{2}]$. Equation (5) for $k=m$ shows $0\neq a_{m+1}^2=b_{m+1}^2$. We now distinguish different cases.

1st Case: $m<[\frac{n-1}{2}]$. Then s_{m+1} is defined, and by the Slice Theorem

$$A_n\subseteq\Bbbk[\epsilon_{s_{m+1}}(x_0),\dots,\epsilon_{s_{m+1}}(x_n)]_{f_{m+1}}.$$

Applying $\pi:=\pi_{n-m-1,n}$ on both sides yields

$$\pi(A_n)\subseteq\Bbbk[\pi(\epsilon_{s_{m+1}}(x_{m+1})),\dots,\pi(\epsilon_{s_{m+1}}(x_n))]_{\pi(f_{m+1})},$$

where we used that $\pi_{n-m-1,n}(\epsilon_{s_{m+1}}(x_k))=0$ for $k=0,\dots,m$. The right hand side is included in $\Bbbk[\pi(E_n)]_{\pi(f_{m+1})}$. Therefore, for any $f\in A_n$ there is $p\in\Bbbk[E_n]$ and $l\geq 0$ such that $\pi(f)=\frac{\pi(p)}{\pi(f_{m+1})^l}$. Let

$$\gamma: V_n\to V_{n-m-1},\quad (c_0,\dots,c_n)\mapsto(c_{m+1},\dots,c_n).$$

Then

$$\begin{aligned} f(a) &= \pi(f)(\gamma(a))=\frac{\pi(p)}{\pi(f_{m+1})^l}(\gamma(a))=\frac{\pi(p)(\gamma(a))}{(\pi(f_{m+1})(\gamma(a)))^l}=\frac{p(a)}{f_{m+1}(a)^l}\\ &= \frac{p(b)}{f_{m+1}(b)^l}=\frac{\pi(p)(\gamma(b))}{(\pi(f_{m+1})(\gamma(b)))^l}=\pi(f)(\gamma(b))=f(b).\end{aligned}$$

Here we used that the elements p and f_{m+1} of E_n take the same value on a,b by assumption, and $f_{m+1}(a)=f_{m+1}(b)\neq 0$.

2nd Case: $[\frac{n-1}{2}]=m<[\frac{n}{2}]$. In this case, n has to be even, and $n=2m'$ with $m'=m+1$. Let π and γ be as before, so $\pi=\pi_{m',2m'}$ and $\gamma:V_{2m'}\to V_{m'}$. First assume m' is odd. By Proposition 3.1 (c) we have

$$\pi(A_n)=\Bbbk[x_0^2]=\Bbbk[\pi(f_{m'})].$$

If m' is even, by Proposition 3.1 (d) we have

$$\pi(A_n)=\Bbbk[x_0^2,x_0^3]=\Bbbk[\pi(f_{m'}),\pi(w)],$$

with w the element of Lemma 5.4. In both cases, $\pi(A_n)=\Bbbk[\pi(E_n)]$, so for any $f\in A_n$, there exists $p\in\Bbbk[E_n]$ such that $\pi(f)=\pi(p)$. Therefore,

$$f(a)=\pi(f)(\gamma(a))=\pi(p)(\gamma(a))=p(a)=p(b)=\pi(p)(\gamma(b))=\pi(f)(\gamma(b))=f(b).$$

We have shown: for any $f\in A_n$ we have $f(a)=f(b)$, and so we are done. □

5. The existence of w.

In this section we prove Lemma 5.4, which requires some more machinery. Note that we need this Lemma in order to construct a separating set only in the case where $n \equiv 0 \mod 4$ — in the other cases, w is not contained in our separating set. We will make use of semitransvectants, which are the classical transvectants transformed under Roberts' isomorphism, see for example [3, 4, 17]. Recall that for a covariant $F \in \tilde{R}_n^{\mathrm{SL}_2} = R_n[y_0, y_1]^{\mathrm{SL}_2}$, its total degree in y_0, y_1 is called the *order* of F. For covariants F and G of orders l and m respectively, we can construct new covariants given by

$$\langle F, G\rangle^{(r)} := \sum_{k=0}^{r} (-1)^k \begin{pmatrix} r \\ k \end{pmatrix} \frac{\partial^r F}{\partial y_0^{r-k} \partial y_1^k} \frac{\partial^r G}{\partial y_0^k \partial y_1^{r-k}} \qquad r \leq \min(l, m),$$

which is called the rth *transvectant* of F and G (see [17, p. 88]). Transvectants play a key role in Gordan's famous proof of the finite generation of covariants of binary forms [13]. The transformation of this construction under Roberts' isomorphism leads the following definition (see also [2]).

Definition 5.1. Let $\Phi : R_n[y_0, y_1]^{\mathrm{SL}_2} \to R_n^{\mathbb{G}_a} = A_n$ be Roberts' isomorphism, given by substituting $y_0 := 0,\ y_1 := 1$. Let f and g be a pair of invariants in A_n. Then for $r \leq \min(l, m)$, where l and m are the orders of $\Phi^{-1}(f)$ and $\Phi^{-1}(g)$ as above, we define the rth *semitransvectant* of f and g by

$$[f, g]^{(r)} := \Phi(\langle \Phi^{-1}(f), \Phi^{-1}(g)\rangle^{(r)}).$$

In order to get an explicit expression for the semitransvectant, we introduce a second derivation on R_n, which is somewhat inverse to D_n:

$$\Delta_n := \sum_{k=0}^{n} (n-k)(k+1) x_{k+1} \frac{\partial}{\partial x_k}.$$

This derivation comes from the other canonical embedding of $\mathbb{G}_a$ in SL_2, namely for $f \in R_n,\ a \in \Bbbk$ we have $\begin{pmatrix} 1 & \\ a & 1 \end{pmatrix} * f = \exp(a\Delta_n) f$. Let $\operatorname{ord}(f)$ denote the nilpotency index of f with respect to Δ_n. Assume $F \in R_n[y_0, y_1]^{\mathrm{SL}_2}$ is homogeneous of degree m in the variables y_0, y_1, so it can be written in the form $F = f y_1^m + y_0 \cdot (\ldots)$ with $f \in R_n$. Then $\Phi(F) = f$, and invariance of F under the torus action implies all terms $x_0^{a_0} \ldots x_n^{a_n}$ in f satisfy $m = \sum_{k=0}^{n} (n-2k) a_k$. A polynomial $f \in R_n^{\mathbb{G}_a}$ with this property is called *isobaric* of *weight* m, and then we have $m = \operatorname{ord}(f)$. For an isobaric $f \in R_n^{\mathbb{G}_a}$, by [14, p. 43] the inverse of Roberts' isomorphism is given by

$$\Phi^{-1}(f) = \sum_{i=0}^{\operatorname{ord}(f)} (-1)^i \frac{\Delta_n^i(f)}{i!} y_0^i y_1^{\operatorname{ord}(f)-i}.$$

Proposition 5.2. *Let $f, g \in R_n^{\mathbb{G}_a}$ be isobaric. Then for $r \leq \min(\operatorname{ord}(f), \operatorname{ord}(g))$, the rth semitransvectant of f and g is given by the formula*

$$[f, g]^{(r)} = \sum_{k=0}^{r} (-1)^k \begin{pmatrix} r \\ k \end{pmatrix} \Delta_n^k(f) \frac{(\operatorname{ord}(f)-k)!}{(\operatorname{ord}(f)-r)!} \Delta_n^{r-k}(g) \frac{(\operatorname{ord}(g)-r+k)!}{(\operatorname{ord}(g)-r)!}$$

Proof. Let $\frac{\Delta_n^i(f)}{i!} := \lambda_i$ and $\frac{\Delta_n^i(g)}{i!} := \mu_i$. Then

$$\frac{\partial^r \Phi^{-1}(f)}{\partial y_0^{r-k} \partial y_1^k} = \sum_{i=r-k}^{\operatorname{ord}(f)-k} (-1)^i \lambda_i \frac{i!}{(i-r+k)!} \frac{(\operatorname{ord}(f)-i)!}{(\operatorname{ord}(f)-i-k)!} y_0^{i-r+k} y_1^{\operatorname{ord}(f)-i-k}$$

and

$$\frac{\partial^r \Phi^{-1}(g)}{\partial y_0^k \partial y_1^{r-k}} = \sum_{i=k}^{\mathrm{ord}(g)-r+k} (-1)^i \mu_i \frac{i!}{(i-k)!} \frac{(\mathrm{ord}(g)-i)!}{(\mathrm{ord}(g)-i-r+k)!} y_0^{i-k} y_1^{\mathrm{ord}(g)-i-r+k},$$

therefore

$$\Phi\left(\frac{\partial^r \Phi^{-1}(f)}{\partial y_0^{r-k} \partial y_1^k}\right) = (-1)^{r-k} \lambda_{r-k} \frac{(r-k)!(\mathrm{ord}(f)-r+k)!}{(\mathrm{ord}(f)-r)!}$$

and

$$\Phi\left(\frac{\partial^r \Phi^{-1}(g)}{\partial y_0^k \partial y_1^{r-k}}\right) = (-1)^k \mu_k \frac{k!(\mathrm{ord}(g)-k)!}{(\mathrm{ord}(g)-r)!}.$$

Using the fact that Φ is an algebra homomorphism we have

$$\begin{aligned}
[f,g]^{(r)} &= \sum_{k=0}^{r} (-1)^{k+r} \binom{r}{k} \lambda_{r-k} \frac{(r-k)!(\mathrm{ord}(f)-r+k)!}{(\mathrm{ord}(f)-r)!} \mu_k \frac{k!(\mathrm{ord}(g)-k)!}{(\mathrm{ord}(g)-r)!}, \\
&= \sum_{k=0}^{r} (-1)^{k+r} \binom{r}{k} \Delta_n^{r-k}(f) \frac{(\mathrm{ord}(f)-r+k)!}{(\mathrm{ord}(f)-r)!} \Delta_n^k(g) \frac{(\mathrm{ord}(g)-k)!}{(\mathrm{ord}(g)-r)!} \\
&= \sum_{k=0}^{r} (-1)^{k} \binom{r}{k} \Delta_n^{k}(f) \frac{(\mathrm{ord}(f)-k)!}{(\mathrm{ord}(f)-r)!} \Delta_n^{r-k}(g) \frac{(\mathrm{ord}(g)-r+k)!}{(\mathrm{ord}(g)-r)!}
\end{aligned}$$

as required. □

Remark 5.3. This is analogous to [2, Lemma 1], using a different basis.

The formula shows that, up to some scalar factor, f_m (from (2)) equals $[x_0, x_0]^{(2m)}$ (while $[x_0, x_0]^{(r)} = 0$ for r odd), and $\epsilon_{s_m}(x_1)$ equals $[x_0, f_m]^{(1)}$. We wonder whether there is also a connection between $\epsilon_{s_m}(x_j)$ and $[x_0, f_m^j]^{(j)}$.

Lemma 5.4. *Suppose n is divisible by 4, so $n = 2m = 4p$. Then there is an invariant $w \in A_n$ satisfying $\pi_{m,n}(w) = x_0^3$.*

Proof. Throughout the proof we use the shorthand $\pi := \pi_{m,n}$, and we set $f := f_p = \frac{1}{2}\sum_{i=0}^{m}(-1)^i x_i x_{m-i}$ (which is proportional to $[x_0, x_0]^{(m)}$). Obviously, f is isobaric of weight $(n-2i) + (n-2(m-i)) = 2n - 2m = n = \mathrm{ord}(f)$. Thus we may define

$$\bar{w} := [x_0, f]^{(n)} = \sum_{k=0}^{n} (-1)^k \frac{n!^2 k!}{(n-k)!} x_k \Delta_n^{n-k}(f) = \sum_{k=0}^{n} (-1)^k \frac{n!^2 (n-k)!}{k!} x_{n-k} \Delta_n^k(f),$$

where we used Proposition 5.2. Thus,

$$\pi(\bar{w}) = \sum_{k=0}^{m} (-1)^k \frac{n!^2 (n-k)!}{k!} x_{m-k} \pi(\Delta_n^k(f)).$$

Using Leibniz's formula for iterated differentiation of products, we have

$$\pi(\Delta_n^k(x_i x_{m-i})) = \sum_{j=0}^{k} \binom{k}{j} \pi(\Delta_n^j x_i) \pi(\Delta_n^{k-j} x_{m-i})$$

$$\begin{aligned}
&= \sum_{j=m-i}^{k-i} \binom{k}{j} \frac{(i+j)!(n-i)!}{i!(n-i-j)!} \pi(x_{i+j}) \frac{(m-i+k-j)!(m+i)!}{(m-i)!(m+i-k+j)!} \pi(x_{m-i+k-j}) \\
&= \sum_{j=0}^{k-m} \binom{k}{j+m-i} \frac{(m+j)!(n-i)!}{i!(m-j)!} \pi(x_{m+j}) \frac{(k-j)!(m+i)!}{(m-i)!(n-k+j)!} \pi(x_{k-j}).
\end{aligned}$$

In particular, $\pi(\Delta_n^k(x_i x_{m-i})) = 0$ for all $k < m$, and since f is a linear combination of terms of the form $x_i x_{m-i}$, we have $\pi(\Delta_n^k(f)) = 0$ for all $k < m$. From this, remembering m is even, it follows that $\pi(\bar{w}) = n!^2 x_0 \pi(\Delta_n^m(f))$. Therefore, since

$$\pi(\Delta_n^m(x_i x_{m-i})) = \left(\begin{array}{c} m \\ m-i \end{array} \right) x_0^2 \frac{(n-i)!}{i!} \frac{(m+i)!}{(m-i)!},$$

and $f = \frac{1}{2}\sum_{i=0}^m (-1)^i x_i x_{m-i}$, we obtain

$$\begin{aligned} \pi(\Delta_n^m(f)) &= \frac{1}{2} x_0^2 \sum_{i=0}^m (-1)^i \left(\begin{array}{c} m \\ i \end{array} \right) \frac{(n-i)!}{i!} \frac{(m+i)!}{(m-i)!} \\ &= \frac{((2p)!)^2}{2} x_0^2 \sum_{i=0}^{2p} (-1)^i \left(\begin{array}{c} 2p \\ i \end{array} \right) \left(\begin{array}{c} 4p-i \\ 2p \end{array} \right) \left(\begin{array}{c} 2p+i \\ i \end{array} \right). \end{aligned}$$

Thus, $\pi(\bar{w}) = n!^2 x_0 \pi(\Delta_n^m(f))$ is a nonzero multiple of x_0^3 if the sum above is nonzero. This follows from Lemma 5.6. □

Remark 5.5. With $g := \Delta_n^n(f)$, we have $\bar{w} = c \cdot \sum_{k=0}^n (-1)^k x_k D^k g$ with $c \in \Bbbk$.

Lemma 5.6. *For all $p \geq 1$ we have*

$$\sum_{k=0}^{2p} (-1)^k \left(\begin{array}{c} 2p \\ k \end{array} \right) \left(\begin{array}{c} 4p-k \\ 2p \end{array} \right) \left(\begin{array}{c} 2p+k \\ k \end{array} \right) = (-1)^p \frac{(3p)!}{(p!)^3}.$$

Proof. The argument which follows was produced using the implementation of Zeilberger's algorithm [25] in the remarkable EKHAD package for Maple [18]. Let

$$F(p,k) := \begin{cases} (-1)^k \left(\begin{array}{c} 2p \\ k \end{array} \right) \left(\begin{array}{c} 4p-k \\ 2p \end{array} \right) \left(\begin{array}{c} 2p+k \\ k \end{array} \right) & 0 \leq k \leq 2p \\ 0 & \text{Otherwise,} \end{cases}$$

and let $S(p) := \sum_{k=0}^{2p} F(p,k)$. We show that the following recurrence relation holds:

$$6(3p+2)(3p+1)S(p) + 2(p+1)^2 S(p+1) = 0. \tag{6}$$

To do this, we consider the function

$$\begin{aligned} G(p,k) := \frac{1}{2}k^2(&180 - 184k + 1036p + 59k^2 + 2192p^2 - 790pk - 1116p^2k + 168pk^2 \\ &+2024p^3 - 8k^3 + 688p^4 + k^4 - 520p^3k + 120p^2k^2 - 10pk^3)(-4p+k-1)\frac{F(p,k)}{R(p,k)} \end{aligned}$$

where $R(p,k) = (2p+1)(-2p-2+k)^2(-2p-1+k)^2$. For $0 \leq k \leq 2p-1$ it satisfies the relation

$$G(p,k+1) - G(p,k) = 6(3p+2)(3p+1)F(p,k) + 2(p+1)^2 F(p+1,k).$$

Summing both sides over $0 \leq k \leq 2p-1$ (and adding remaining terms) produces (6), and an easy inductive argument then shows that $S(p) = (-1)^p \frac{(3p)!}{(p!)^3}$. □

Remark 5.7. The sum $S(p)$ is the well-poised hypergeometric series

$$\left(\begin{array}{c} 4p \\ 2p \end{array} \right) \sum_{k=0}^{\infty} \frac{(-2p)_k (2p+1)_k (-2p)_k}{(1)_k (-4p)_k k!} = \left(\begin{array}{c} 4p \\ 2p \end{array} \right) {}_3F_2[-2p, 2p+1, -2p; 1, -4p; 1].$$

Surprisingly, the series is not summable by any classical hypergeometric sum theorem (e.g. Dixon's theorem, Watson's theorem) because the series ${}_3F_2[-2p, 2p+1, -2p; 1, -4p; z]$, p not an integer, does not converge when $z = 1$, see [21, Chapter 2]. For this reason, we have to apply Zeilberger's algorithm for partial sums in order to sum the series. In the language of WZ-theory, $\bar{F}, G$ is a WZ-pair, where $\bar{F}(p,k) := (-1)^p \frac{(p!)^3 F(p,k)}{(3p)!}$, and $R(p,k)$ is the corresponding WZ-proof certificate.

References

[1] Leonid Bedratyuk. Casimir elements and kernel of weitzenböck derivation. *arXiv:math/0512520*, 2005.

[2] Leonid Bedratyuk. On complete system of invariants for the binary form of degree 7. *J. Symbolic Comput.*, 42(10):935–947, 2007.

[3] Leonid Bedratyuk. A complete minimal system of covariants for the binary form of degree 7. *J. Symbolic Comput.*, 44(2):211–220, 2009.

[4] A Cayley. *The Collected Mathematical Papers, vol 1.* Cambridge University Press, Cambridge, England, 1889.

[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] Emilie Dufresne and Martin Kohls. A finite separating set for Daigle and Freudenburg's counterexample to Hilbert's Fourteenth Problem. *Communications in Algebra, to appear*, 2010.

[9] Jonathan Elmer. On the depth of separating algebras for finite groups. *Contributions to Algebra and Geometry, to appear*, 2010.

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

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

[12] Gene Freudenburg. *Algebraic theory of locally nilpotent derivations*, volume 136 of *Encyclopaedia of Mathematical Sciences.* Springer-Verlag, Berlin, 2006. Invariant Theory and Algebraic Transformation Groups, VII.

[13] P Gordan. Beweiss, dass jede Covariante und Invariante einer binären Form eine ganz Funktion mit numerischen Coefficienten einer endlichen Anzahl solher Formen ist. *J. Reine. Angew. Math.*, 69:323–354, 1868.

[14] David Hilbert. *Theory of algebraic invariants.* Cambridge University Press, Cambridge, 1993. Translated from the German and with a preface by Reinhard C. Laubenbacher, Edited and with an introduction by Bernd Sturmfels.

[15] Gregor Kemper. Computing invariants of reductive groups in positive characteristic. *Transform. Groups*, 8(2):159–176, 2003.

[16] Martin Kohls and Hanspeter Kraft. Degree bounds for separating invariants. *Mathematical Research Letters, to appear*, 2010.

[17] Peter J. Olver. *Classical invariant theory*, volume 44 of *London Mathematical Society Student Texts.* Cambridge University Press, Cambridge, 1999.

[18] Marko Petkovšek, Herbert S. Wilf, and Doron Zeilberger. $A = B$. A K Peters Ltd., Wellesley, MA, 1996. With a foreword by Donald E. Knuth, With a separately available computer disk.

[19] Michael Roberts. On the Covariants of a Binary Quantic of the n^{th} Degree. *The Quarterly Journal of Pure and Applied Mathematics*, 4:168–178, 1861.

[20] Müfit Sezer. Constructing modular separating invariants. *J. Algebra*, 322(11):4099–4104, 2009.

[21] Lucy Joan Slater. *Generalized hypergeometric functions.* Cambridge University Press, Cambridge, 1966.

[22] Lin Tan. An algorithm for explicit generators of the invariants of the basic G_a-actions. *Comm. Algebra*, 17(3):565–572, 1989.

[23] Arno van den Essen. An algorithm to compute the invariant ring of a $\mathbf{G}_a$-action on an affine variety. *J. Symbolic Comput.*, 16(6):551–555, 1993.

[24] R. Weitzenböck. Über die Invarianten von linearen Gruppen. *Acta. Math.*, 58:231–293, 1932.

[25] Herbert S. Wilf and Doron Zeilberger. An algorithmic proof theory for hypergeometric (ordinary and "q") multisum/integral identities. *Invent. Math.*, 108(3):575–633, 1992.

University of Bristol, University Walk, Bristol, BS8 1TW
E-mail address: `j.elmer@bris.ac.uk`

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